



الجمهورية الجزائرية الديمقراطية الشعبية
République Algérienne Démocratique et Populaire

وزارة التعليم العالي والبحث العلمي
Ministère de L'Enseignement Supérieur et de la Recherche Scientifique

جامعة سعد دحلب البليدة-1-
Université Saad Dahlab Blida -1-



Mémoire de fin D'études

En vue de l'obtention du diplôme Master

Faculté des Sciences
Département : Mathématiques

Spécialité : Modélisation stochastique et statistique

THEME

**Prévision de la Var pour les modèles Garch et
gestion de portefeuille**

Présenté par :
SALMI Razika
Devant le Jury :

Soutenu le : 08/07/2021

Président :	R.FRIHI	MCB	Université de Blida 1
Promoteur:	O.TAMI	MAA	Université de Blida 1
Examineur :	A. RASSOUL	Prof.	ENSH de Blida

JUIN 2021

Remerciement

Avant tout, nous remercions Dieu le Tout Puissant de nous avoir guidés vers les portes du savoir.

Nous tenons à remercier toutes les personnes qui ont contribué de près ou de loin à l'élaboration de ce travail.

Tout d'abord, notre première pensée va à notre encadreur, M^r TAMI Omar, pour avoir bien voulu nous assister et nous conseiller au cours de notre travail.

Nous désirons ensuite remercier l'ensemble des enseignants de département de Mathématiques.

Enfin, nous remercions les membres du jury pour avoir accepté d'évaluer notre travail.

ملخص

بالنسبة الى اي مستثمر من الضروري عليه القيام بتقدير قيمة الخطر و قيمة الخسارة التي يمكن له التعرض لها من خلال وضعيات مختلفة للاستثمار كثير من الطرق العملية متاحة له ولكل منها ايجابيات و مساوئ لذا وجب عليه معرفة اي منها اكثر نجاعة و موضوعية. من خلال هذا العمل دراستنا تتمحور حول مقياسين هما قيمة المخاطرة وقيمة المخاطرة الشرطية التي تقنر قيمتهما من خلال طريقتين التجريبية و GPD وايضا قرش(1.1) بالنسبة لقيمة المخاطرة . هذا النموذج يهدف الى صياغة وقائع لنمذجة ظواهر بهدف التنبؤ بالقيم المستقبلية لقيمة المخاطرة التي تعد مقياس ثابت لتسيير المخاطرة في المجال المالي. و ذلك من خلال معالجة معطيات حقيقية لسلاسل الايراد لأربعة مؤشرات في البرصة (دكس سمني كاك فتنسيو).
كلمات مفتاحية : قيمة المخاطرة, طريقة تجريبية جبدي, قرش, سلسلة زمنية, توقعات.

Résumé :

Pour n'importe quel investisseur, il est primordial d'évaluer le risque et la perte potentielle qu'il pourrait encourir suite à une position d'investissement bien déterminée. Plusieurs mesures s'offrent à lui. Chacune a ses avantages comme elle a ses inconvénients. Il doit savoir alors laquelle est la plus fiable, afin de réduire le risque de son placement. Dans ce travail, nos étude s'articule autour le mesures de risque : la Value at Risk (VaR) qui est estimé par deux méthodes : empirique, GPD et GARCH(1,1) pour la VAR seulement. L'objectif de ce travail est d'étudier les techniques d'analyse et de modélisation des séries chronologiques. On s'intéresse essentiellement aux modèles GARCH(1.1). Ce modèle permettent de représenter une chronique avec comme finalité de prévoir des valeurs futures de la VAR qu'est devenue un standard de la gestion de risque dans le monde financier à partir de traitement d'une application sur des données réelles pour les séries des rendements pour les quatre indices boursiers (DAX, SMI,CAC40 et FTSE).

Mots clés : Var ; ES; Méthode empirique; GPD; GARCH(1,1); Série chronologique; Prévision .

Abstract:

For any investor, it is essential to assess the risk and the potential loss that they could incur as a result of a well-defined investment position. Several measures are available to him. Each has its advantages as it has its disadvantages. He must then know which one is the most reliable, in order to reduce the risk of his investment. In this work, our study revolves around the risk measures: the Value at Risk (VaR) which is estimated by two methods: empirical, GPD and GARCH (1,1) for the VAR only. The objective of this work is to study the techniques of analysis and modeling of time series. We are mainly interested in GARCH models (1.1). This model makes it possible to represent a chronicle with the aim of forecasting future values of the VAR that has become a standard of risk management in the financial world from the processing of an application on real data for the series of returns for the four stock market indices (DAX, SMI, CAC40 and FTSE).

Keywords: Var; ES; Empirical method; GPD; GARCH (1.1); Time series; Forecast

Table des matières

Introduction Générale	1
1 Généralités sur les mesures de risque	4
1.1 Aperçu général sur les mesures de risque	4
1.2 Type du risque	4
1.2.1 Le risque du marché	4
1.2.2 Le risque financier	5
1.2.3 Le risque d'inflation	5
1.2.4 Le risque du taux d'intérêt	5
1.2.5 Le risque de liquidité	5
1.3 Aperçu sur les mesures de risque d'un portefeuille du marché	6
1.4 Définition de la Value-at-Risk	7
1.5 Les principales méthodes de mesure de la Value at Risk	12
1.6 Méthodes non paramétriques	12
1.6.1 Méthode du quantile empirique	12
1.6.2 Méthode du Bootstrap	12
1.7 Méthode semi paramétriques	13
1.7.1 Méthode basée sur la théorie des valeurs extrêmes	13
1.7.2 Simulation historique filtrée	14
1.8 Méthodes paramétriques	14
1.8.1 Méthode de Variance Covariance	14
1.8.2 Simulation Monte Carlo	15
1.8.3 Méthodes basées sur les modèles GARCH	16
1.9 Propriétés de la VaR	17
1.10 Validation d'une prévision de la VaR	18
1.11 Méthodes de mesure de risque cohérentes	21
1.11.1 Définition	21
1.11.2 Propriétés des mesures de risque cohérentes	21
1.12 Expected Shortfall (ES)	22
1.12.1 Définition et méthodes de calcul	22

1.13 Rendements financiers	23
1.13.1 Rendements des titres financiers	23
1.13.2 Caractéristiques des rendements financiers	25
2 Rappel sur les séries chronologiques	27
2.1 Introduction	27
2.2 Définition et concept de base	27
2.3 Domaine d'application	27
2.4 Type de séries chronologiques	28
2.4.1 Série continues	28
2.4.2 Séries discrète	28
2.5 Objectifs de l'analyse d'une chronique	28
2.5.1 Modélisation	28
2.5.2 Prévision	29
2.6 Les composantes d'une série chronologique	29
2.7 Les différents schémas d'une série chronologique	29
2.8 Choix du modèle	30
2.9 Opérateur définis sur une série chronologique	30
2.9.1 Opérateur de retard	30
2.9.2 Opérateur de différence d'ordre d	31
2.9.3 Processus aléatoire	31
2.10 Filtre linéaires et moyennes mobiles	31
2.10.1 Filtre linéaire	32
2.10.2 Moyen mobiles	32
2.10.3 Moyenne mobile et composante saisonnière	32
2.11 Stationnarité	33
2.11.1 Relation entre stationnarité faible et stricte	33
2.11.2 Test de Dickey Fuller Augmenté	34
2.12 Procédure de choix d'un schéma de composition :	35
2.13 Analyse des composantes d'une série chronologique	36
2.13.1 Analyse de la tendance :	36
2.13.2 Analyse de la saisonnalité	37
2.14 Le concept de bruit blanc	38
2.15 Auto-corrélations simple et partielle	38
2.15.1 Fonction d'auto-covariance et auto-corrélation	38
2.15.2 La fonction d'autocorrélation partielle	40
2.16 Série linéaire	41
2.17 Analyse de la tendance	41
2.17.1 Rappels sur la regression linéaire	41

2.17.2	La méthode des moindres carrés	42
2.17.3	Ajustement tendanciel linéaire les moindres carrés	43
2.17.4	Ajustement tendanciel linéaire par points médians	44
2.17.5	Ajustements tendanciels non linéaires	44
2.17.6	Estimation non paramétrique	44
2.17.7	Théorème de Wold	45
2.18	Processus de référence	46
2.18.1	Processus auto régressive (AR)	46
2.18.2	Processus moyenne mobile (MA)	47
2.18.3	Processus auto régressif moyenne mobile (ARMA)	47
2.18.4	Processus ARIMA (auto régressive moyen mobile intégré)	48
2.18.5	Processus auto régressive moyen mobile intégré saisonnier SARIMA(p,d,q)	48
2.18.6	Les Modèles des séries temporelles à variance conditionnelle	48
2.18.7	ARCH(q)	49
2.18.8	Modèle GARCH	50
2.18.8.1	Présentation du Modèle GARCH	50
2.18.8.2	Estimation des Paramètres	51
2.18.8.3	Problèmes de convergence et choix des conditions initiales	52
2.18.9	Modèle EGARCH (Exponential GARCH)	53
2.18.9.1	Présentation du Modèle EGARCH	53
2.18.9.2	Estimation des Paramètres	54
2.18.9.3	Problèmes de convergence et choix des conditions initiales	56
2.18.10	Modèle QGARCH (Quadratic GARCH)	56
2.18.10.1	Présentation du Modèle QGARCH	57
2.18.10.2	Estimation des Paramètres	58
2.18.10.3	Problèmes de convergence et choix des conditions initiales	59
2.18.11	Modèle IGARCH (Integrated GARCH)	60
2.18.11.1	Présentation du Modèle IGARCH	60
2.18.11.2	Estimation des Paramètres	61
2.18.11.3	Problèmes de convergence et choix des conditions initiales	62
2.18.12	Modèle GJR-GARCH	63
2.18.12.1	Présentation du Modèle GJR-GARCH	63
2.18.12.2	Estimation des Paramètres	64

2.18.12.3	Problèmes de convergence et choix des conditions initiales	65
2.18.13	Modèle GARCH-M (GARCH in Mean)	66
2.18.13.1	Présentation du Modèle GARCH-M(GARCH in Mean)	66
2.18.13.2	Estimation des Paramètres	67
2.18.13.3	Problèmes de convergence et choix des conditions initiales	68
2.19	Distributions Conditionnelles - Modèles GARCH	69
2.19.1	Loi Normale	69
2.19.2	Loi de Student	70
2.19.3	Loi GED (General Error Distribution)	71
2.19.4	Loi Skewed Student	72
2.19.5	Loi Skewed GED	73
3	La Modélisation du problème d'optimisation du portefeuille	75
3.1	Introduction	75
3.2	Résolution du problème d'optimisation du portefeuille	76
3.3	Le cadre empirique de l'étude	79
3.3.1	Test de la stationnarité des quatres indices boursier	82
3.3.2	Conclusion	82

Table des figures

1.1	VaR sous distribution normale.	9
3.1	Représentation graphique de la série de rendement de l'indice DAX. . .	91
3.2	Représentation graphique de l'autocorrélation simple et la corrélogramme partiel de la série de rendement de l'indice DAX.	91
3.3	Histogramme et Normal Q-Q-Plot des rendements journaliers de l'indice DAX.	92
3.4	Représentation graphique de la série de rendement de l'indice SMI. . .	92
3.5	Représentation graphique de l'autocorrélation simple et la corrélogramme partiel de la série de rendement de l'indice SMI.	93
3.6	Histogramme et Q-QPlot des rendements journaliers de l'indice SMI. .	93
3.7	Représentation graphique de la série de rendement de la série de rendement de l'indice CAC40.	94
3.8	Représentation graphique de l'autocorrélation simple et la corrélogramme de la série de rendement de l'indice CAC40.	94
3.9	Histogramme et normale Q-QPlot des rendements journaliers de l'indice CAC40.	95
3.10	Représentation graphique de la série de rendement de l'indice FTSE. . .	95
3.11	Représentation graphique de l'autocorrélation simple et la corrélogramme partiel de la série de rendement de l'indice FTSE.	96
3.12	Histogramme et Q-QPlot des rendements journaliers de l'indice FTSE. .	96
3.13	Représentation graphique de la série du rendement et du résidus de l'indices DAX PAR GARCH(1,1).	96
3.14	Histogramme de rendement et du résidus de l'indice DAX par GARCH(1,1). .	97
3.15	Représentation graphique du Q-QPlot de la série de rendement de l'indice DAX etQ-QPlot du résidue par garch(1,1).	98
3.16	Représentation graphique de l'autocorrélation simple de la série de rendement de DAX et du résidue par garch(1,1).	98
3.17	Représentation graphique de la série de rendement de l'indice SMI PAR GARCH (1,1).	99

TABLE DES FIGURES

3.18 Représentation de la corelogramme de la série du rendement et du résidue par garch (1,1).	99
3.19 Représentation graphique de Q-QPlot de la série du rendement et du résidue de l'indice SMI par GARCH(1,1).	100
3.20 Représentation graphique de l'autocorrélation simple et du squared résidue de l'indice de SMI PAR GARCH(1,1).	100
3.21 Représentation graphique de la série de rendement de l'indice CAC 40 PAR garch (1,1).	101
3.22 Histogramme de la série de rendement et du résidue de l'indice CAC 40 par garch (1.1).	101
3.23 Représentation graphique de Q-QPlot et du résidue de l'indice CAC40.	102
3.24 Représentation graphique de squared de la série de rendement et du résidue par GARCH (1.1).	102
3.25 Représentation graphique de la série de rendement de l'indice FTSE PAR GARCH (1.1).	103
3.26 Histogramme de la série de rendement et du résidue de l'indice FTSE par garch(1.1).	103
3.27 Représentation graphique de la série de rendement et du résidue pour l'indice FTSE par GARCH(1.1).	104
3.28 Représentation graphique de squared de la série du rendement et de résidue de l'indice FTSE par GARCH(1.1).	104
3.29 Représentation graphique des valeurs de sigma pendant la période de prévision pour l'indice DAX.	105
3.30 Représentation graphique des valeurs de sigma pendant la période de prévision de l'indice SMI.	105
3.31 Représentation graphique des valeurs de sigma pendant la période de prévision pour l'indice CAC40.	106
3.32 Représentation graphique des valeurs de sigma pendant la période de prévision pour l'indice FTSE.	106

Liste des tableaux

2.1	Table d'un processus centré	35
2.2	Table d'un processus avec un drift égale à 1	35
2.3	Tableau des résultats de l'analyse de la variance	37
3.1	Paramètres de description des quatres indices boursiers	83
3.2	Résultats du tests de la stationnarité des quatres indices	83
3.3	Le montant de la VAR par la méthode mpirique pour un niveau de confiance de 0.95 et 0.99	84
3.4	Le montant de la CVAR par la méthode mpirique pour un niveau de confiance de 0.95 et 0.99	84
3.5	Le montant de la VAR par la méthode GPD pour un niveau de confiance de 0.95 et 0.99	84
3.6	Table de la VaR à 95% et 99% pour les quatres indices boursier	84
3.7	Estimation des paramètre du modèle Garch(1.1) au niveau de confiance 95% pour l'indice DAX	84
3.8	Estimation des paramètre du modèle Garch(1.1) au niveau de confiance 95% pour l'indice SMI	84
3.9	Estimation des paramètre du modèle Garch(1.1) au niveau de confiance 95% pour l'indice CAC	84
3.10	Estimation des paramètre du modèle Garch(1.1) au niveau de confiance 95% pour l'indice FTSE	85
3.11	Test de Jarque Bera pour les quatres indices	85
3.12	les valeur de la VAR estimé au niveau de confiance 95% par garch(1.1)	85
3.13	Montant de la VAR par modèle GARCH(1.1) pendant la période de pré- vision pour les quatres indices boursiers	85

Introduction Générale

La libéralisation des marchés financiers, l'extension considérable des mouvements de capitaux et l'apparition continue de nouveaux produits financiers ont contribué à rendre le paysage financier de plus en plus incertain et instable. Cette situation a provoqué une volatilité importante au niveau des prix et a conduit parfois à des mouvements de prix extrêmes.

Ces dernières décennies ont été marquées par plusieurs crises financières qui ont entravé de graves problèmes économiques non seulement à l'échelle nationale mais aussi à l'échelle mondiale. La crise récente des suprimés (subprime mortgage meltdown), qui a été déclenchée par une crise de marché des prêts hypothécaires à risque aux États Unis, est un exemple. Elle a conduit à une récession économique nationale puis proliféré pour devenir une crise financière mondiale à partir de l'été 2007. Autre exemple est la quasi-faillite du Hudge Fund «Long term Capital Management» qui a entraîné de graves problèmes dans le système bancaire international conduisant à d'importantes perturbations financières mondiales en 1998.

Suite à ce genre d'événements, les autorités sont intervenues à plusieurs reprises afin d'instaurer une réglementation plus rigoureuse aux institutions financières. Elles leur exigeaient de développer des mesures de risque internes plus efficaces dans le but de détecter les défaillances et les ajuster pour éviter les grandes pertes, les faillites et même les crises mondiales.

La notion de mesure de risque n'est pas nouvelle dans la théorie financière. En effet, elle est liée au concept de probabilité apparu au dix-septième siècle par les jeux du hasard. Mais cette approche est mise en évidence par l'apparition de la théorie de Markowitz dans les années cinquante, puis d'autres mesures ont été développées tel que la notion de Beta, le modèle CAPM et le modèle APT.

Néanmoins, toutes ces mesures classiques ont montré leurs limites dans ce nouvel environnement financier très turbulent.

La prolifération des crises financières et la limite des approches de mesure de risque antérieures ont amené les chercheurs et praticiens à développer de nouvelles mesures de risque de marché compatibles avec ce nouveau contexte. D'où l'initiation de la banque JP Morgan par la fameuse méthode Valeur à risque (Value At Risk) par le système Riskmetrics. Il est devenu accessible librement sur internet à partir d'Octobre

1994. Ce système était développé suite à la décision du directeur de la banque JP Morgan. Ce dernier exigeait un rapport journalier sur le risque et les pertes potentielles encourus par la banque dans les prochaines vingt-quatre heures.

Suite au succès de cette méthode, conçue comme une simple mesure de risque interne, la VaR est devenue une approche utilisée par toutes les banques. Elle leur permettait d'évaluer le risque global conformément aux accords de Bâle sur le capital des banques. Elle reflète la perte maximale d'une position d'investissement sur une période bien déterminée et qui ne sera dépassée qu'avec une probabilité fixée auparavant.

Outre le Riskmetrics, d'autres méthodes ont été développées pour estimer la VaR. Trois principales catégories de méthodes sont apparues : méthodes paramétriques (la famille du modèle économétrique GARCH), méthodes semi-paramétriques (CAViaR, théorie des extrêmes) et les méthodes non paramétriques (tel que la simulation historique et la simulation Monte Carlo).

Malgré le grand succès de la VaR, la philosophie de la méthode est souvent mal comprise et donc mal interprétée. Elle est devenue sujet de plusieurs critiques notamment au sujet de ses propriétés. Il s'est avéré qu'elle contredit la notion de la diversification dans la théorie financière puisqu'elle ne vérifie pas la propriété de la « sous additivité » à l'échelle d'un portefeuille agrégé.

L'origine de cette critique est les travaux d'Artzner et al (1997,1999) qui ont montré que la VaR n'est pas une mesure de risque cohérente. D'où l'apparition d'autres mesures dont les propriétés sont plus avantageuses que la VaR.

Notre travail s'accroît sur les deux mesures de risque : la VaR et la VaR conditionnelle (ou appelée Expected Shortfall). Elles sont les plus connues sur le marché financier ; largement utilisées par les gestionnaires grâce à leur simplicité, leur moindre coût ainsi qu'à la diversité de leurs méthodes de calcul. . Le but est de déterminer la méthode qui procure une meilleure prévision de risque du marché par la VaR et la ES.

Afin de répondre à cette question, la démarche suivie par cette étude s'articulera sur trois chapitres :

- Le premier chapitre sera consacré à une analyse théorique qui comportera trois sections. La première section présentera un aperçu général sur les mesures de risques sur le marché financier. La deuxième section portera sur la notion de la VaR et ses propriétés. La troisième section portera sur la notion des mesures de risques cohérentes et ses propriétésles en particulier celle de notre étude qui est l'ES sera développée. Par la suite une autre section sur le rendement financier.
- Dans le second chapitre, on commence par un rappel sur la définition d'une série temporelle, nous présentons ensuite les différentes composantes et les principales caractéristiques des série temporelle ,la présentation des différents tests ,aussi on présente quelques modèles de la série temporelle à variance condition-

nelle et nous nous basons sur cette section sur les modèles GARCH, l'estimation des paramètres, problèmes de convergence et choix des conditions initiales. Nous terminons ce chapitre par les distributions conditionnelles relatives au modèle GARCH.

- Dans le troisième chapitre on élabore le cadre empirique de notre étude avec des hypothèses sur les différents facteurs. Cette étude empirique sera composée de trois parties : Dans la première partie, on applique et on analyse les différentes méthodes d'estimations de la VaR et CVAR cités dans le chapitre précédent. Dans la seconde partie, on applique et on analyse les résultats de notre modèle dans un cadre statique. Les méthodes d'estimation de la VaR et CVAR retenue sont la méthode empirique, (TVE ou aussi GPD en se référant à la distribution de Pareto Généralisé). Dans la troisième partie, on passe à un cadre dynamique. L'objectif est de tester les deux approches d'estimation de la VaR et CVAR dans notre modèle générale. Pour cela, on procède à l'application de ces derniers pour une gestion journalière. On se réfère ensuite à un Backtesting en comparant les pertes effectives des portefeuilles optimaux constitués journalièrement avec la VaR limite prévue par les investisseurs.

Chapitre 1

Généralités sur les mesures de risque

La présente étude traite un concept important de la finance qui est le risque. Dans ce chapitre, une revue de littérature sur la notion de risque sera présentée. Ensuite, les deux méthodes adaptées dans notre recherche, la *VaR* et *CVAR*.

1.1 Aperçu général sur les mesures de risque

Il n'existe pas une définition bien précise du risque, bien que Mc Neil, Frey et Embrechts (2005) le définit comme « chaque événement ou action qui pourrait affecter la capacité d'une organisation à atteindre ses objectifs et exécuter ses stratégies ». Le risque reste un concept difficile à comprendre.

1.2 Type du risque

Il existe plusieurs types de risque. Le type de risque qui nous intéresse présentement est le risque financier. Ce dernier est lié à quatre ingrédients essentiels : événements, décisions, conséquences et incertitude.

Plusieurs catégories de risque sont présentées tel que :

1.2.1 Le risque du marché

C'est la possibilité de fluctuations du prix du placement en raison des facteurs externes, indépendamment de ses caractéristiques.

Parmi ces facteurs on trouve la conjoncture économique générale (exemple : récession économique, hausse des prix de matières premières), événements politiques ou sociaux (Guerre civile, coup d'état..), ou tout simplement l'évolution des préférences des consommateurs et leur utilité.

Tel est le cas de la récession économique récente des États-Unis. En effet, la dépréciation du dollar américain a incité les investisseurs à retirer leur argent des marchés

boursiers puisque la valeur boursière a diminué et investir ainsi dans des actifs moins risqués tel que l'or. Les indices boursiers ont subi une chute contrairement au marché de l'or, qui a connu une hausse de prix.

Le risque du marché reflète alors la volatilité du prix de l'actif financier ; plus le prix est volatile plus l'actif est risqué.

1.2.2 Le risque financier

Il est lié à la structure financière de l'entreprise, c'est à dire sa capacité à honorer ses endettements et ainsi assurer sa survie. Il fait référence au poids des dettes de l'entreprise par rapport à ses capitaux propres. Pour évaluer ce risque, plusieurs indicateurs sont calculés pour déterminer son état financier.

Ainsi, plus les dettes utilisées sont importantes, moins elle est capable de faire face au service des emprunts, plus la probabilité de faillite est élevée. Ce qui engendre des pertes pour ses créanciers, soit ses actionnaires ou détenteurs d'obligations.

1.2.3 Le risque d'inflation

L'évolution de l'indice des prix dans une économie affecte considérablement les investissements. Cette évolution peut être dans les deux sens : une hausse des prix (inflation) au contrairement une baisse (déflation).

En général, les placements qui évoluent dans le même sens que le niveau des prix sont plus rémunérés dans les périodes d'inflation tel que les actions des fabricants de biens durables. Quant aux investissements qui procurent des rendements fixes sont plus vulnérables surtout en période d'inflation, tel que les obligations à taux fixes.

1.2.4 Le risque du taux d'intérêt

La valeur de certains placements est affectée par la variation du taux d'intérêt. Généralement, la variation de la valeur de l'investissement varie en sens inverse à la variation du taux d'intérêt.

Il existe des placements qui sont affectés plus que d'autres, tel que les obligations à taux fixe ou les SICAV, bien qu'ils soient des placements à long terme.

1.2.5 Le risque de liquidité

Un actif financier est dit liquide s'il est échangé sur le marché financier régulièrement, à des intervalles de temps courts sans avoir un impact important sur son prix (hausse importante lors d'un ordre d'achat ou baisse importante lors d'une vente). Plus un actif est liquide, plus il est susceptible de se vendre facilement à un prix raisonnable.

Généralement, les actions de grandes sociétés et des sociétés opérant dans des secteurs prometteurs sont plus liquides que celles des petites entreprises ou des sociétés opérant dans des secteurs en déclin.

Chaque instrument de placement sur le marché financier est assujéti à un certain risque. Cependant, l'investisseur doit en prendre compte et évaluer son impact sur son investissement. Pour cela, il doit bien déterminer quelle est la source du risque qui pourrait induire sa position et le traiter rigoureusement.

Notre étude concerne un seul type de risque : le risque du marché. Un aperçu historique sur ses différentes mesures sera détaillé dans la section suivante.

1.3 Aperçu sur les mesures de risque d'un portefeuille du marché

La volatilité importante du marché financier a mis l'accent sur l'importance de la gestion du risque.

Les mesures de risque du marché ont bien évolué depuis le développement de la théorie de diversification de Markowitz. Ce dernier a avancé en 1959 « une méthode de solution générale du problème de la structure des portefeuilles qui incorpore le traitement quantifié du risque »

Elle repose sur deux critères essentiels : la moyenne pour la rentabilité espérée du portefeuille et la variance comme une mesure du risque, d'où son appellation théorie de la 'moyenne-variance'. Selon ce critère, on a abouti à un ensemble de portefeuilles qui offrent le rendement le plus élevé (moyenne) pour chaque niveau de risque (variance) donné et le risque le plus faible pour chaque niveau de rendement donné. Cet ensemble de portefeuilles constitue la frontière efficiente.

Dans cet ensemble, on ne peut pas espérer d'augmenter le rendement du portefeuille sans augmenter le risque en contre partie. D'où ce critère de moyenne-variance est insuffisant. Cependant, Markowitz intègre la fonction d'utilité; chaque investisseur choisit le portefeuille qui maximise sa fonction d'utilité. Ainsi son portefeuille efficient sera la tangence entre sa fonction d'utilité et la frontière efficiente.

De nombreuses critiques ont été adressées à ce modèle d'optimisation tel que sa sensibilité aux critères adoptés moyenne-variance. Ce qui a incité à développer d'autres modèles : les modèles à facteurs qui s'intéressent aux facteurs influençant sur les fluctuations des cours des actifs. Il y a deux types de facteurs : des facteurs communs qui affectent l'ensemble de marché (risque systématique) et d'autres qui sont spécifiques à chaque actif (risque spécifique). Le modèle le plus connu est le modèle de marché, développé par Sharpe (1964) et qui s'est basé sur les travaux de Markowitz. Bien que ce modèle ait contribué à enrichir la théorie de gestion de risque des portefeuilles, il a

présenté certains inconvénients en pratique. Ce qui a poussé Sharpe à développer un autre modèle : le modèle d'équilibre des actifs financiers MEDAF (ou CAPM) qui lui a valu le prix Nobel. Comme tout autre modèle économique (tel que le modèle Keynésien de la consommation), Le MEDAF repose sous certaines hypothèses. Il propose de déterminer les prix des actifs financiers qui permettent d'ajuster l'offre et la demande sur le marché et ainsi instaurer l'équilibre général du marché.

Au milieu des années soixante-dix, Ross a présenté un autre modèle *APT* (Arbitrage Pricing Theory). C'est un modèle à facteurs où Ross stipule qu'avec un nombre peu élevé d'actifs et un nombre moins élevé de facteurs, on peut construire un portefeuille où le risque diversifiable de chaque composante de ce portefeuille est négligeable. Et dans un marché efficient en absence d'arbitrage, son rendement espéré est alors la combinaison linéaire des primes de risques des facteurs.

Toutefois, l'évolution du marché financier et l'apparition continue de nouveaux aspects des actifs financiers ont manifesté les limites de ces approches de mesure de risque ; elles ne permettent pas d'évaluer le risque de type asymétrique (tel que associé aux options) puisque les critères bêta et l'écart type sont insuffisants. D'autres mesures se sont développées. D'où l'apparition de la mesure de risque « Valeur à Risque » ou en anglais « Value at Risk ».

Selon Dowd, à la fin des années soixante-dix, la majorité des institutions financières ont commencé à développer des modèles internes de mesure de risque agrégé afin de les appliquer dans leur gestion mais aussi à le commercialiser pour d'autres institutions financières incapables de les développer par elles même.

Le système le plus connu était alors le « Riskmetrics » développé par la banque JP Morgan. Il était en exclusivité pour la banque JP Morgan mais à partir de 1994, il est devenu accessible par internet.

Cette nécessité d'évaluer le risque interne des institutions financières a encouragé à développer d'autres mesures de la *VaR* plus précises que le système Riskmetrics.

Dans la section suivante, nous présenterons l'approche *VaR* et les différentes méthodes de son estimation ainsi que ces principales avantages et limites.

1.4 Définition de la Value-at-Risk

La notion de Value-at-Risk (*VaR*) est apparue pour la première fois dans le secteur de l'assurance. A la fin des années 1980, la banque Bankers Trust fut l'une des premières institutions à utiliser cette notion sur les marchés financiers aux Etats-Unis, mais c'est principalement la banque JP Morgan qui dans les années 90 a popularisée ce concept notamment grâce à son système RiskMetrics (pour un historique complet de la notion de Value-at-Risk et de sa diffusion se reporter au livre de Dowd, 2005). La Value-at-Risk est ensuite devenue, en moins d'une dizaine d'années, une mesure de

référence du risque sur les marchés financiers, consacrée notamment par la réglementation prudentielle définie dans le cadre des accords de Bâle II.

De façon générale, la Value-at-Risk est définie comme la perte maximale potentielle qui ne devrait être atteinte qu'avec une probabilité donnée sur un horizon temporel donné (Engle et Manganelli, 2001). La Value at Risk est donc la pire perte attendue sur un horizon de temps donné pour un niveau de confiance donné. Cette définition très simple constitue l'un des principaux attraits de la Value-at-Risk : il est en effet très facile de communiquer sur la VaR et de ainsi proposer une mesure homogène et générale (quelque soit la nature de l'actif, la composition du portefeuille etc.) de l'exposition au risque.

Ainsi, la Value-at-Risk n'est rien d'autre qu'un fractile de la distribution de perte et profit associée à la détention d'un actif ou d'un portefeuille d'actifs sur une période donnée. La mesure de Value-at-Risk ne fait que refléter l'information contenue dans la queue gauche (associée aux pertes) de la distribution des rendements d'un actif. Si l'on considère un taux de couverture de $\alpha\%$ ou de façon équivalente un niveau de confiance de $(1 - \alpha\%)$ la Value-at-Risk correspond tout simplement au fractile de niveau $\alpha\%$ de la distribution de perte et profit valable sur la période de détention de l'actif :

$$VaR(\alpha) = F^{-1}(\alpha) \quad (1.1)$$

où $F(\cdot)$ désigne la fonction de répartition associée à la distribution de perte et profit. De cette définition générale découlent plusieurs définitions techniques tout aussi simples.

Ainsi, la Value-at-Risk dépend de trois éléments : (i) la distribution des pertes et profits du portefeuille valable pour la période de détention (ii) le niveau de confiance (ou de façon équivalente le taux de couverture égal à un moins le niveau de confiance) et (iii) la période de détention de l'actif.

— Taux de couverture et Niveau de Confiance :

Le niveau de confiance choisi est un paramètre compris entre 0 et 1 (95% ou 99% en général) qui permet de contrôler la probabilité que l'on obtienne un rendement supérieur ou égale à la Value-at-Risk. Supposons que la distribution des pertes et profits associée à la détention d'un actif sur une période corresponde à une distribution normale standard. Sur la Figure 1 est reproduite cette distribution de perte et profit supposée normale : sur la partie gauche de l'axe des abscisses figurent les rendements négatifs (pertes) tandis qu'à droite figure les rendements positifs (profits). Dans ce cas, la Value-at-Risk définie pour un niveau de confiance de 95% ($\alpha = 5\%$) est égale tout simplement à -1.645 . Dit autrement, dans cet exemple il y a 95% de chances que le rendement de l'actif, noté r , soit au moins égal à -1.645 sur la période de détention.

Ainsi, on aura :

$$VaR = \sqrt{SCS^T} \quad (1.13)$$

Avec :

S = le vecteur des VaR pour chaque position ou facteur de risque = $[...q, \delta, \omega, ...]$

C = la matrice des corrélations entre les facteurs de risque.

Les calculs utilisés dans la méthode RiskMetrics sont rapides et simples, et requièrent uniquement la connaissance de la matrice des Variances Covariances des rendements du portefeuille. Néanmoins, cette méthode s'avère être inadaptée aux portefeuilles non linéaires (instruments optionnels), et théoriquement peu adaptée aux queues de distribution épaisses et aux distributions non normales des rendements.

1.8.2 Simulation Monte Carlo

La méthode de Monte Carlo consiste à simuler un grand nombre de fois les comportements futurs possibles des facteurs de risque selon un certain nombre d'hypothèses, et d'en déduire une distribution des pertes et profits à partir de laquelle on estime finalement un fractile. Plus précisément, on peut considérer l'exemple de l'hypothèse de normalité. La méthode Monte Carlo s'applique en trois étapes :

- La première étape consiste à simuler N scénario de l'évolution des facteurs de risque. Un scénario est obtenu à travers la formule suivante :

$$\Delta S = tZ \quad (1.14)$$

Avec :

S = le vecteur des facteurs de risque

t = la matrice résultant de la décomposition de Cholesky de la matrice des Variances Covariances des facteurs de risque ($\sum S = tt'$)

Z = un vecteur de variables aléatoires indépendantes de loi normale centrée et réduite. On obtient suite à cette étape la série $(\Delta S^1 \dots \dots \Delta S^N)$

- La deuxième étape consiste à déterminer les N variations respectives du portefeuille. On obtient donc la suite $(\Delta L^1 \dots \dots \Delta L^N)$ des évolutions de la valeur de la position initiale.

- La troisième étape consiste à déterminer le quantile de la même façon que pour la simulation historique à partir de la distribution simulée.

Si cette approche peut s'appliquer, en théorie, quelles que soient les lois de probabilité suivies par les facteurs de risque, elle est couramment utilisée en pratique, pour

$$Pr[r < Var(0.05)] = Pr[r < -1.645] = 0.05$$

De la même façon, la Value-at-Risk définie pour un niveau de confiance de 99% ($\alpha = 1\%$) est égale à -2.326 .

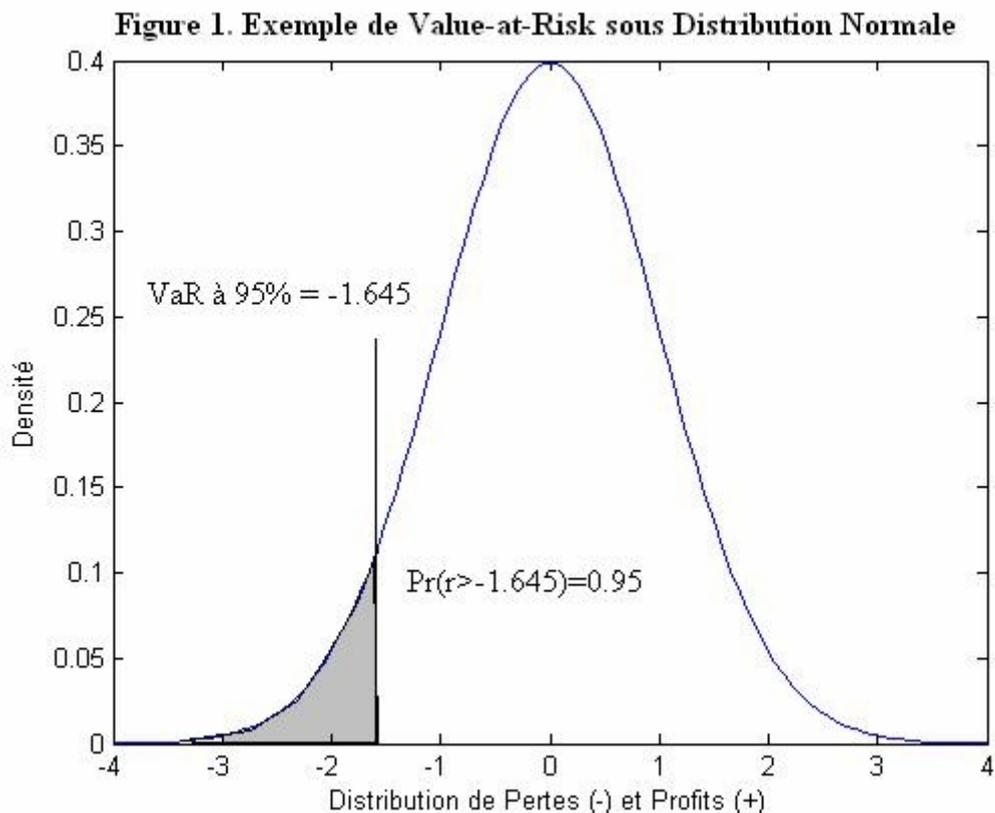


FIGURE 1.1 – VaR sous distribution normale.

Ainsi, la Value-at-Risk correspond généralement à une perte (valeur négative). Toutefois, on trouve souvent une Value-at-Risk définie non pas à partir de la distribution de perte (-) et profit (+), mais à partir au contraire d'une distribution de profit(-) et perte(+). Dit autrement, une telle définition revient à omettre le signe moins devant la perte et donc à afficher une Value-at-Risk positive. Dans ce cas, la définition de la Value-at-Risk correspond à l'opposé du fractile de la distribution de perte et profit :

$$VaR(\alpha) = -F^{-1}(\alpha) \quad (1.2)$$

Si l'on reprend notre exemple de distribution normale, on affiche alors une Value-at-Risk pour un niveau de confiance de 95% ($\alpha = 5\%$) égale à 1.645. Cela signifie qu'il y a 95% de chances que la perte associée à la détention de l'actif n'excède pas 1.645.

— Horizon de détention

Le deuxième élément fondamental dans le calcul de la Value-at-Risk est la période de détention de l'actif ou du portefeuille d'actifs. La formule de calcul de la Value-at-Risk doit alors être ajustée de façon à tenir compte de la composition des rendements. Il n'existe aucune règle quant au choix de la période de détention dans le calcul de la Value-at-Risk puisque ce choix dépend fondamentalement de l'horizon de reporting ou d'investissement des opérateurs.

Toutefois, les autorités de régulation peuvent spécifier des horizons de détention spécifiques notamment dans le cadre des procédures de validation de la Value-at-Risk.

Dans le cadre de ce site, tous les calculs proposés portent sur une Value-at-Risk définie à partir de la distribution de pertes (-) et profits(+)(ce qui implique une valeur négative de la Value-at-Risk) et pour un horizon de détention d'une période (une journée, un mois etc. suivant les données utilisées pour les calculs).

— La distribution des pertes et profits du portefeuille

C'est le paramètre le plus important mais aussi le plus difficile à déterminer. Une question fondamentale se pose : comment peut-on choisir une distribution pour la variable de marché? Idéalement, il faudrait un modèle qui soit simple et qui convienne le plus aux observations empiriques. La méthode de calcul est déterminée par la distribution choisie pour modéliser les pertes et profits du portefeuille. Empiriquement, trois principaux problèmes dans les séries temporelles financières sont détectés. Le premier est la non stationnarité des séries manifesté par la non stabilité des paramètres de la loi régissant les variables du processus temporel. Le deuxième problème est le caractère leptokurtique de la distribution des données qui consiste à des queues empiriques plus épaisses que celles considérées par la loi mise en hypothèse, ceci a pour conséquence la sous estimation de la Value at Risk et donc du risque assumé par l'investisseur. Le troisième problème est le phénomène de la dépendance de la volatilité appelé la persistance de la volatilité (clustering) issu du constat en pratique du fait que les volatilités élevées sont souvent suivies par des volatilités élevées et les volatilités faibles sont souvent suivis par des volatilités faibles. Ce phénomène peut être pris en compte par la famille des modèles *ARCH* – *GARCH* sur de courtes périodes d'estimation. La détection de l'effet *ARCH* dans la série des observations reste à vérifier pour justifier ce type de modélisation.

En théorie financière, une hypothèse de normalité des rendements est souvent adoptée comme réponse à la problématique de la distribution des rendements des actifs financiers. En effet, cette hypothèse accélère considérablement les calculs. Elle est bien adaptée dans l'application de la méthode RiskMetricsTM. Un test dit test de Jarque et Bera (*JB*) permet de valider ou non la normalité de la distribution. Il tient compte implicitement de deux paramètres essentiels : le coefficient d'asymétrie (*Skewness*) et le coefficient d'aplatissement (*Kurtosis*).

Le premier paramètre (*Skewness*) est le moment d'ordre 3 et mesure l'asymétrie du

comportement des rendements autour de leur moyenne empirique. Son expression est la suivante :

$$S = \frac{1}{T} \left[\frac{\sum_{t=1}^T (X_t - \bar{X})^3}{(\delta^2)^{(3/2)}} \right] \quad (1.3)$$

Si $S = 0$ la distribution est dite symétrique à l'instar de la loi normale. Si $S > 0$ alors la densité de la distribution s'étale vers la droite et on a une asymétrie positive. Si $S < 0$ alors la densité de la distribution s'étale vers la gauche et on est en présence d'une asymétrie négative.

Le deuxième paramètre, le kurtosis, est le moment d'ordre 4 et représente une mesure associée à l'épaisseur des queues de la distribution. Son expression est la suivante :

$$K = \frac{1}{T} \left[\frac{\sum_{t=1}^T (X_t - \bar{X})^4}{(\delta^2)^{(2)}} \right] \quad (1.4)$$

Si $K = 3$ on dit que la distribution est mésokurtique comme c'est le cas pour la distribution normale qui sert de point de référence. Si $K > 3$ la distribution présente des queues épaisses (fat tails). Elle est dite leptokurtique. Si $K < 3$ la distribution présente des queues minces (thins tails). Elle est dite platikurtique.

En ce qui concerne le test de Jarque et Bera, il est basé sur la statistique suivante :

$$JB = (TS^2|6) + [T(K - 3)^2|24] \neq \chi^2(2) \quad (1.5)$$

Cette statistique suit asymptotiquement la loi χ^2 avec deux degré de liberté lorsque T (nombre des observations) est assez grande.

Le test d'hypothèse est le suivant :

H_0 : la distribution est normale

H_1 : la distribution n'est pas normale

La règle consiste à rejeter H_0 si la statistique JB est plus grande que χ^2 avec deux degrés de liberté au seuil de signification choisie (95% par exemple).

Notons que la stationnarité des rendements du portefeuille d'actions est une condition nécessaire pour appliquer la méthode de simulation historique . Le test de racine unitaire ADF (augmented Dickey-Fuller) consiste à tester l'hypothèse nulle :

H_0 existence d'une racine unitaire série non stationnaire

Ce test consiste à rejeter H_0 si la statistique obtenue est inférieure à une valeur critique dite valeur de MacKinnon.

1.5 Les principaux méthodes de mesure de la Value at Risk

Mathématiquement, la notion de la Value-at-Risk se traduit ainsi :

$$Pr(\Delta V < Var) = 1 - c \quad (1.6)$$

Avec : ΔV = la variation de la valeur V du portefeuille sur la période de détention.
 c = le niveau de confiance

Plusieurs modèles ont été présentés pour l'estimation de la Value-at-Risk (Manganelli et Engle (2001)). L'élément clé qui distingue ces modèles est l'existence ou non d'une hypothèse de para métrisation de la distribution des pertes et des profits. Ainsi on classera ces méthodes en trois classes : les méthodes non paramétriques, les méthodes semi paramétriques et les méthodes paramétriques.

1.6 Méthodes non paramétriques

1.6.1 Méthode du quantile empirique

La méthode du quantile empirique (ou Historical Simulation) est une méthode très simple d'estimation des mesures de risque fondée sur la distribution empirique des données historiques de rendements. Formellement, la VaR est estimée simplement par la lecture directe des fractiles empiriques des rendements passés. Si l'on considère par exemple un niveau de confiance de 95% et que l'on dispose d'un échantillon de 1000 observations historiques de rendements, la VaR est donnée par la valeur du rendement qui correspond à la 50^{ème} forte perte.

1.6.2 Méthode du Bootstrap

Une amélioration simple de la méthode de la simulation historique consiste à estimer la VaR à partir de données simulées par Bootstrap. Le Bootstrap consiste à ré échantillonner les données historiques de rendements avec remise. Plus précisément, dans notre contexte, la procédure consiste à créer un grand nombre d'échantillons de rendements simulés, où chaque observation est obtenue par tirage au hasard à partir de l'échantillon original. Chaque nouvel échantillon constitué de la sorte permet d'obtenir une estimation de la VaR par la méthode HS standard, et l'on définit au final une estimation en faisant la moyenne de ces estimations basées sur les ré échantillonnages.

1.7 Méthode semi paramétriques

1.7.1 Méthode basée sur la théorie des valeurs extrêmes

Parmi les méthodes semi paramétriques figurent tout d'abord l'ensemble des méthodes et approches qui relèvent de la théorie des extrêmes (*TVE*) qui diffère de la théorie statistique habituelle fondée pour l'essentiel sur des raisonnements de type tendance centrale. Les extrêmes sont en effet gouvernés par des théorèmes spécifiques qui permettent d'établir sous différentes hypothèses la distribution suivie par ces extrêmes. Il existe deux principales branches de la théorie des valeurs extrêmes : la théorie des valeurs extrêmes généralisée et l'approche Peaks Over Threshold (*POT*) basée sur la loi de Pareto généralisée. L'approche *POT* permet l'étude de la distribution des pertes excessives au dessus d'un seuil (élevé), tandis que la théorie des valeurs extrêmes généralisée permet de modéliser la loi du maximum ou du minimum d'un très grand échantillon. Dans ce qui suit, on procèdera à l'application de cette approche. Pour cela, on définit la moyenne en excédent pour une distribution F par :

$$e(u) = E(X - u \mid X > u) \quad (1.7)$$

C'est simplement une fonction de u qui s'exprime à l'aide de la fonction de survie de F . Plus les queues de distribution sont épaisses, plus cette fonction a tendance à tendre vite vers l'infini.

En pratique, si n est le nombre total de l'échantillon et si est le nombre d'observations au dessus du seuil u , on a :

$$\widehat{e}(u) = \frac{1}{N_u} \sum_{j=1}^n (x_j - u) 1_{\{x_j > u\}}(x_j), u > 0 \quad (1.8)$$

Le problème du choix de u reste entier. Usuellement, on trace cette fonction Mean Excess pour différents niveaux du seuil u . Le bon seuil est celui à partir duquel $e(u)$ est approximativement linéaire. Graphiquement, cela se traduit par un changement de la pente de la courbe qui ensuite reste stable. Ce résultat provient de la remarque que pour la distribution de Pareto généralisée, $e(u)$ est linéaire en u . Une fois le seuil optimal choisi, on construit une nouvelle série d'observations au dessus de ce seuil, et la distribution de ces données suit une distribution généralisée de Pareto, qui se définit comme suit :

$$G_{\xi, \beta}(x) = \begin{cases} 1 - \left(1 + \frac{\xi x}{\beta}\right)^{\frac{1}{\xi}} & \text{si } \xi \neq 0 \\ 1 - \exp\left(-\frac{x}{\beta}\right) & \text{si } \xi = 0, \text{ avec } \beta > 0. \end{cases} \quad (1.9)$$

ξ est appelée l'indice de queue. Le paramètre β est un indicateur de la taille de la

quee à une distance finie. L'estimation des paramètres ξ et β se fait par le maximum de vraisemblance.

La densité de la distribution *GPD* s'écrit :

$$g_{\xi,B}(x) = \begin{cases} \beta^{\frac{1}{\xi}} (\beta + \xi x)^{-\frac{1}{\xi}-1} & \text{si } \xi \neq 0 \\ \beta^{-1} \exp\left(-\frac{x}{\beta}\right) & \text{si } \xi = 0 \end{cases} \quad (1.10)$$

Et la log vraisemblance que nous maximisons est de la forme :

$$\ln \ell(\xi, \beta) = \sum_{i=1}^n \ln \left[g_{\xi,B}(x_t) 1_{\{\xi > 0\}}(x_t) \right] \quad (1.11)$$

Une fois l'estimation terminée, on peut vérifier graphiquement la pertinence des estimations en comparant la distribution *GPD* estimée avec la distribution empirique des observations au dessus du seuil. La Value-at-Risk pour un niveau de confiance c est obtenue par la formule :

$$VaR_c = u + \frac{\widehat{\beta}}{\widehat{\xi}} \left(\left(\frac{n}{N_n} (1-c) \right)^{-\widehat{\xi}} - 1 \right) \quad (1.12)$$

1.7.2 Simulation historique filtrée

La méthode de la simulation historique filtrée est une forme de Bootstrap semi-paramétrique qui vise à combiner les avantages de la simulation historique avec la puissance et la flexibilité des modèles à volatilité conditionnelle tel que le modèle *GARCH*. Elle consiste à faire un Bootstrap sur les rendements dans un cadre de volatilité conditionnelle, le Bootstrap préservant la nature non paramétrique de la simulation historique, et le modèle à volatilité conditionnelle donnant un traitement sophistiqué de la volatilité.

1.8 Méthodes paramétriques

1.8.1 Méthode de Variance Covariance

Cette méthode connu aussi sous le nom de méthode Riskmetrics. Les principales hypothèses simplificatrices consistent à supposer, d'une part, que les lois de probabilité qui régissent les distributions des variations des prix de marché sont normales et, d'autre part, que les instruments présentent un profil de risque linéaire. Sous ces hypothèses, la matrice de Variances Covariances peut être appliquée assez directement aux positions détenues pour calculer la *VaR*.

des raisons techniques, en supposant que les variations relatives des paramètres de marché suivent des lois normales. Cette méthode convient également à tous les types d'instruments, y compris optionnels, et permet de tester de nombreux scénarios et d'y inclure explicitement des queues de distribution épaisses (événements extrêmes pris en compte dans une certaine mesure) (voir Glasserman et al. (2001))

1.8.3 Méthodes basées sur les modèles GARCH

La méthodologie d'estimation de la VaR en se basant sur la modélisation $GARCH$ est largement étudiée en littérature (voir par exemple Christoffersen et al. (2001), Engle (2001)). En ce qui concerne la détection de l'effet $ARCH$ dans la série des observations, deux principaux tests complémentaires peuvent être effectué.

- ⊙ Le premier s'intéresse au phénomène d'auto corrélation entre les termes d'erreur au carré du modèle : Test $Q(p)$ de Ljung-Box (1978). Si le processus est $ARCH$, les résidus au carré doivent être corrélés. L'hypothèse nulle est l'absence d'auto corrélation d'ordre p . La statistique du test est supposée suivre une loi χ^2 avec p degrés de liberté.
- ⊙ Le deuxième test d'intéresse plutôt au phénomène d'homoscédasticité (constance de la volatilité des termes d'erreur) : Test $ARCH(p)$ d'Engle (1982). Ce test vérifie l'absence d'hétéroscédasticité autorégressive conditionnelle d'ordre p . Si le processus est $ARCH$, les résidus au carré doivent être hétéroscédastiques. L'hypothèse nulle est celle de l'homoscédasticité. La statistique du test est supposée aussi suivre une loi χ^2 avec p degré de liberté. La règle de décision est la même pour les deux tests : accepter H_0 si la statistique du test est inférieure à la valeur critique de la loi χ^2 avec p degrés de liberté à un niveau de confiance donné. Notons que pour tester l'effet $GARCH(p, q)$, il suffit de procéder à un test d'effet $ARCH(p + q)$.

La prévision de la Value-at-Risk à partir d'un modèle $GARCH$ est effectué selon une démarche indirecte : dans un premier temps, on fait une hypothèse sur la distribution conditionnelle des rendements de l'actif, puis l'on estime les paramètres du modèle $GARCH$ sur les observations de la période 1 à T , généralement par une procédure de type maximum de vraisemblance. Dans une seconde étape, on déduit du modèle $GARCH$ estimé une prévision de la variance conditionnelle, qui couplée à l'hypothèse retenue sur la distribution des rendements, permet de construire une prévision sur le fractile de la distribution de pertes et profits valable pour $T + 1$.

Considérons l'exemple d'un modèle $GARCH$ sous hypothèse d'une distribution quelconque (normale, student...) de paramètre v .

On suppose ainsi que les rendements d'un actif, notés r_t , satisfont le modèle suivant :

$$r_t = c + \varepsilon_t$$

$$\varepsilon_t = Z_t \delta_t$$

$$\delta_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \delta_{t-1}^2$$

Les z_t sont indépendantes identiquement distribués selon la loi mise en hypothèse. Les paramètres $\alpha_0, \alpha_1, \beta_1, c$ sont des réels à estimer vérifiant les contraintes suivantes : $\alpha_0 > 0, \alpha_1 \geq 0, \beta_1 \geq 0$ (v peut aussi faire partir des paramètres à estimer comme dans le cas de la distribution de Student).

Le terme $\delta_t^2 = E(\varepsilon_t | \varepsilon_{t-1})$ désigne la variance conditionnelle du résidu ε_t et donc des rendements r_t . Une fois les variables sont estimées (par la méthode de maximum de vraisemblance par exemple), on obtient l'expression suivante :

$$\widehat{\delta}_{t+1}^2 = \widehat{\alpha}_0 + \widehat{\alpha}_1 \varepsilon_t^2 + \widehat{\beta}_1 \delta_t^2$$

avec δ_1^2 donné et donc :

$$\widehat{\delta}_{t+1}^2 = \widehat{\alpha}_0 + \widehat{\alpha}_1 (r_t - \widehat{c})^2 + \widehat{\beta}_1 \widehat{\delta}_t^2$$

Soit $G^{-1}(\alpha, v)$ la fonction de répartition de la loi de Z_t . La Value-at-Risk pour $t + 1$ et pour un niveau de confiance $1 - \alpha$ obtenue par la formule suivante :

$$\widehat{vaR}_{t+1} = \widehat{\sigma}_{t+1} G^{-1}(\alpha, \widehat{v}) + \widehat{c} \tag{1.15}$$

1.9 Propriétés de la VaR

Les sections précédentes explicitent la notion de la *VaR* ainsi que les différentes méthodes pour la calculer. Cette section explicite les propriétés de la *VaR*.

Comme toute approche dans la théorie de portefeuille, la méthode *VaR* a contribué au développement des mesures de risque du marché. En effet :

- Elle offre une mesure qui est valide à différentes positions d'investissement et en présence de différents facteurs de risque. Ce qui permet aux institutions financières de mieux gérer le risque de leurs placements dans divers actifs avec différentes manières qui n'existent pas auparavant.
- Les mesures de risque traditionnelles ne tiennent pas compte de l'envergure de la perte encourue par l'investisseur. L'écart type du rendement d'un actif, par exemple, reflète seulement la déviation absolue possible par rapport à la

moyenne. Il n'indique même pas le sens de la déviation vers la hausse ou la baisse. Quant au facteur bêta (β), il reflète le degré d'interaction de l'actif avec le marché. Par contre la *VaR* procure une estimation en valeur de la perte maximale d'une position.

Malgré ses nombreux avantages, la *VaR* reflète certains inconvénients :

- La non sous-additivité : la *VaR* a été largement critiquée au niveau de ses propriétés agrégées. Selon les études d'Artzner et al (1997,1999), elle est constatée comme une mesure de risque non cohérente puisqu'elle n'obéit pas à un critère très important : la sous-additivité.

La non sous-additivité signifie que si on a deux portefeuilles F_{L1} et F_{L2} , et la distribution du portefeuille agrégé $L = L_1 + L_2$ est F_L , alors on n'obtient pas nécessairement

$$VaR(\alpha)_L \leq VaR(\alpha)_{L1} + VaR(\alpha)_{L2}.$$

Ce qui contredit la théorie de la diversification qui stipule que l'agrégation de différents actifs dans un portefeuille permet de réduire le risque. Ce qui signifie aussi que la décentralisation du Risk Management est difficile puisque l'addition des *VaR* de différents portefeuilles ne reflète pas le risque global encouru par l'entreprise.

- Liquidité du marché :

La liquidité du marché est une source de risque pour l'investisseur qu'il doit en tenir compte. Selon Lawrence et Robinson (1995), la mesure de *VaR* doit inclure le coût de liquidité en cas de perte prévue. En réalité, cette mesure néglige cet aspect.

- Choix des paramètres de la *VaR* :

Dans la section précédente au niveau de la définition de la *VaR*, on a spécifié les critères qu'elle requiert. Parmi elles, on a noté l'horizon de temps du calcul. Il peut être influencé par plusieurs facteurs : liquidité du marché, contraintes légales ou contractuelles. Cependant, il est difficile de choisir l'horizon approprié pour l'estimation de la *VaR*. Quant au niveau de confiance, il est aussi difficile de choisir le plus approprié. La *VaR* peut être calculée à différents niveaux de confiance simultanément afin de comparer les résultats fournis. De plus, le choix de la distribution de perte et profit de l'actif. Si elle est mal déterminée, la *VaR* sera biaisée.

Suite aux limites de la *VaR*, d'autres mesures sont apparues. D'où les mesures de risque cohérentes développées par Artzner et al. La section suivante traitera ces mesures.

1.10 Validation d'une prévision de la VaR

Il est nécessaire de tester la validité d'une prévision de la *VaR* en mettant en place une procédure de backtesting des modèles estimés.

Selon Jorion (2007) : Le backtesting est un ensemble de procédures statistiques dont le but est de vérifier que les pertes réelles observées ex-post sont en adéquation avec les pertes prévues. Cela implique de comparer systématiquement l'historique des prévisions de la Value-at-Risk aux rendements observés du portefeuille (Jorion, 2007, page 139).

Notons qu'un test de validation des modèles de VaR doit être de type model free, c'est-à-dire s'applique à n'importe quelle prévision de VaR provenant de n'importe quel modèle. Pour tester la validité de la prévision de la VaR , on compare sa réalisation ex-post à la valeur prédite ex-ante. La comparaison entre plusieurs modèles de prévision de la VaR peut se faire via le critère Mean Squared Error (MSE) qui est fondé sur l'écart entre la vraie valeur réalisée et la valeur prédite ou via les critères d'information standards (AIC, BIC, \dots). La réalisation ex-post de la VaR n'est pas observable. Pour cela la validation de la VaR se fait par des tests sur le processus de violation de la VaR .

On note r_t la rentabilité d'un actif ou d'un portefeuille d'actifs à la date t . La valeur ex-ante de la VaR pour un taux de couverture β notée $VaR_t|_{t-1}(\beta)$, anticipée conditionnellement à un ensemble d'informations, noté Ω_{t-1} disponible à la date $t-1$ est définie par la relation suivante :

$$\rho(r_t < VaR_t|_{t-1}(\beta)) = \beta$$

Par définition, cette quantité correspond au quantile d'ordre β de la distribution conditionnelle des rendements r_t de l'actif ou du portefeuille d'actif considéré. On appelle violation (ou hit, ou exception) une situation dans laquelle à la date t la perte observée excède la VaR anticipée, c'est à dire lorsque la valeur r_t du rendement est inférieur à la valeur prévue ex-ante de la VaR . On note $I_t(\beta)$ la variable associée à une éventuelle violation de la VaR à la date t pour un taux de couverture β .

$$I_t(\beta) = \begin{cases} 1, & \text{si } r_t < VaR_t|_{t-1}(\beta) \\ 0, & \text{sinon} \end{cases} . \quad (1.16)$$

Selon Christoffersen (1998) une prévision de VaR est valide sur une période T si et seulement si la séquence des violations $\{I_t(\beta)\}_{t=1}^T$ satisfait les deux hypothèses de couverture non conditionnelle (UC) et d'indépendance (IND). Ces deux tests peuvent être regroupés dans un seul test qui est celui de la couverture conditionnelle (CC). Selon l'hypothèse (UC) la probabilité que se réalise ex-post une perte (r_t) en excès de la VaR anticipée ex-ante ($VaR_t|_{t-1}$) doit précisément être égale au taux de couverture β , en d'autre terme $E(I_t(\beta)) = \beta$. Selon l'hypothèse (IND), les violations de la VaR pour un même taux de couverture à deux dates différentes doivent être indépendamment distribuées. La couverture conditionnelle (CC) caractérise la situation où la VaR satisfait

les deux hypothèses précédentes. Hurlin & Tokpavi (2008) donnent une description bien détaillée de ces tests.

Pour choisir le modèle le plus adéquat pour la prévision de la VaR parmi plusieurs modèles concurrents, on procède à l'évaluation de la fonction de perte proposée par Sarma et al. (2003) et ceci pour les modèles validés par les tests précédents.

On trouve l'application de cette fonction de perte et de plusieurs autres fonctions utilisées pour tester la performance des modèles de prévisions de la VaR dans Angelidis & Degiannakis (2007), Abad & Benito (2013), Abad et al. (2015) et plusieurs autres. Pour un modèle de prédiction donné, à l'instant t , la fonction de perte de Sarma et al. (2003) est définie par :

$$LF_t(\alpha) = \begin{cases} (VaR_t(\alpha) - r_t)^2, & \text{si } r_t < VaR_t(\alpha) \\ 0 & \text{sinon} \end{cases}. \quad (1.17)$$

On combine cette fonction de perte avec le test de Diebold-Mariano (DM) (1995) pour déterminer si un modèle prédit mieux qu'un autre modèle. Soit deux modèles de prédictions A et B d'erreurs de prédiction respectives $LF_t^A(\alpha)$ et $LF_t^B(\alpha)$. On veut tester l'hypothèse nulle suivante

$$H_0 : E(LF_t^A(\alpha)) = E(LF_t^B(\alpha))$$

contre l'hypothèse alternative

$$H_1 : E(LF_t^A(\alpha)) \neq E(LF_t^B(\alpha)).$$

Soit la fonction de perte différentielle donnée par

$$d_t = LF_t^A(\alpha) - LF_t^B(\alpha).$$

Le test DM teste l'hypothèse nulle d'égalité du pouvoir prédictif suivante

$$H_0 : E(d_t) = 0.$$

Soit $d = (d_1, \dots, d_T)$. La statistique du test est donnée par

$$DMS = \frac{\bar{d}}{\sqrt{T \times var(\bar{d})}}. \quad (1.18)$$

Sous H_0 , DMS suit une normale centrée réduite. On rejette H_0 au seuil de 5% si $DMS < -1.645$.

1.11 Méthodes de mesure de risque cohérentes

Cette section développera les mesures de risques cohérentes et leurs propriétés.

Elle s'intéressera spécialement à la mesure de risque appelée Expected Shortfall (ES).

1.11.1 Définition

Après avoir échoué par son critère de non sous-additivité, il s'est avéré que la mesure de risque VaR comprend certaines lacunes et qu'elle n'est pas une mesure exacte qui reflète le risque d'un portefeuille, d'où le développement par Artzner, Delbaen, Eber et Heath de développer d'autres mesures de risque dites « mesure de risque cohérente » à la fin des années quatre-vingt dix. C'est une mesure de risque d'un portefeuille à une période déterminée. Elle est similaire à la VaR de point de vue conceptuel mais elle diffère au niveau du calcul.

1.11.2 Propriétés des mesures de risque cohérentes

Selon Szegö (2002), mesurer le risque d'un actif correspond à établir une fonctionnelle $\rho : X \rightarrow R$ où X est la variable aléatoire (exemple : le rendement de l'actif) et R est un réel non négatif. Cette valeur de risque permet alors de comparer les différents actifs.

Certaines conditions sont considérées pour cette fonctionnelle.

- Homogénéité positive : cette propriété stipule que $\rho(\lambda x) = \lambda \rho(x)$ pour toute variable aléatoire x et tout réel positif λ . C'est-à-dire que la hausse de la taille d'un portefeuille x par une quantité positive λ ne fait que multiplier le risque de ce portefeuille par ce facteur λ .
- Sous additivité : c'est la propriété que la VaR ne respecte pas. Elle montre que pour toute variable aléatoire x et y ,

$$\rho(x + y) \leq \rho(x) + \rho(y),$$

c'est à dire que le risque agrégé d'un portefeuille doit être moins élevé que la somme du risque individuel des actifs composant ce portefeuille. Ce qui prouve qu'une fonction homogène positive ρ est convexe si et seulement si elle est sous additive.

- Monotonie : cette propriété montre que

$$\text{si } x \leq y \text{ alors } \rho(x) \leq \rho(y)$$

pour toute variable aléatoire x et y . Ce critère montre que plus la perte d'un portefeuille est faible, son risque sera nécessairement moins élevé.

C'est-à-dire que si le portefeuille x a une perte moins grande que le portefeuille y , alors ce dernier est plus risqué.

- Invariance transitionnelle :
- Si $x \leq y$ implique que

$$\rho(x + \alpha r_0) = \rho(x) - \alpha,$$

pour chaque variable aléatoire x , pour chaque réel α et pour chaque actif sans risquer $_0$. Cette propriété explique que si on ajoute à un portefeuille risqué une quantité α d'actif sans risque, alors le risque du portefeuille se réduit par α .

Une mesure de risque qui respecte ces quatre propriétés est appelée une mesure de risque cohérente au sens d'Artzner et al. Il existe plusieurs mesures de risque cohérentes qui seront énumérées dans la section suivante.

1.12 Expected Shortfall (ES)

1.12.1 Définition et méthodes de calcul

Parmi les mesures de risque cohérentes développés : Expected regret (ER), Expected shortfall (ES) appelée aussi tail conditional expectation (TCE), worst conditional expectation (WCE) et spectral risk measures. Cette section s'intéressera à une seule méthode qui est l'expected shortfall (ES), sujet de notre recherche.

La valeur de la VaR reflète le montant de la perte que l'investisseur ne pourrait pas dépasser à une certaine probabilité et sur un horizon de temps bien déterminé. Cette approximation ne tient pas compte des événements extrêmes qui pourraient se produire et qui pourraient engendrer des pertes plus graves. Ce qui amène l'investisseur à prendre des décisions biaisées sur la base de la VaR en sous-estimant les pertes.

Néanmoins, l' ES évite cette lacune de la VaR . En effet, elle prend en compte des événements extrêmes qui pourraient se produire. Elle est définie comme la moyenne des pertes du portefeuille qui sont au delà du niveau de la VaR .

Comme dans le cas de la VaR , l' ES dispose de différentes méthodes de calcul.

Pour les méthodes paramétriques, certains auteurs tel que Giot et Laurent (2003), Bali et Theodossiou (2007) proposent l'estimation d' ES comme étant la moyenne des rendements observés du portefeuille étudié qui sont au delà du niveau de la VaR . Cette dernière est calculée à partir de l'une des méthodes paramétriques décrites au dessus. Analytiquement, elle est calculée comme suit :

$$ES(\alpha) = \frac{1}{\alpha} \int_0^\alpha F^{-1}(\rho) d\rho \quad (1.19)$$

$$ES(\alpha) = E(r|r \leq VaR(\alpha)) \quad (1.20)$$

F est la fonction de répartition des pertes, où α est le taux de couverture de la VaR . Acerbi et Tasche (2002) proposent une méthode non paramétrique qui est la simulation historique. Elle demeure la plus répandue et la plus facile.

L' ES est calculé à partir d'une série de rendements historiques du portefeuille en considération. En effet, si la distribution de ce portefeuille est connue, alors l' ES est la moyenne des $100\alpha\%$ rendements les plus petits

$$ES(\alpha) = \frac{\sum_{i=1}^N r_i}{N} \quad (1.21)$$

avec N est le rang de la $100\alpha\%$ rendement de la distribution

Cette méthode est généralement utilisée lorsque la distribution de l'innovation suit une loi Normale ou Student (détaillée dans la partie empirique).

Il existe une autre méthode non paramétrique pour le calcul d' ES qui est la simulation Monte Carlo. Le processus adapté est le même qu'avec le calcul de la VaR sauf que l'évaluation de l' ES est la moyenne des rendements au delà du niveau de la VaR .

Par exemple, pour l'évaluation de la VaR à 1% et en présence d'un échantillon de 1000 rendements de la distribution des pertes et profits, la VaR est estimée comme étant le 10^{ème} plus petit rendement de la distribution. L' ES est alors la moyenne des rendements qui sont au dessous du niveau de la VaR , c'est-à-dire, la moyenne des dix derniers rendements classés par ordre décroissant.

1.13 Rendements financiers

La première attention des investisseurs porte sur l'analyse des données relatives aux titres financiers. Certainement, les données auxquelles les investisseurs ont accès sont les historiques des prix. Mais, le résultat futur des décisions d'investissement prises aujourd'hui dépend des niveaux futurs des prix qui sont incertain. Il est donc nécessaire d'analyser cette incertitude à travers l'étude des caractéristiques des séries passées qui sont observées. Ceci nous aidera à une meilleure compréhension du comportement des prix, une bonne quantification des risques et à proposer un modèle efficace pour l'évolution des prix pour assurer une prévision rigoureuse des résultats futurs.

On commence par donner les définitions de rentabilités arithmétiques et logarithmiques par la suite on présente quelques caractéristiques des séries financières.

1.13.1 Rendements des titres financiers

Il y a deux approches pour calculer les rendements des titres financiers. Les praticiens utilisent généralement les rendements discrets dit arithmétiques qui reflètent

le vrai change- ment relatif des prix et qui sont utilisés pour mesurer la performance des titres. Alors que les chercheurs utilisent plutôt les rendements continus qui représentent une approximation convenable pour des investigations empiriques ou analytiques. Soit P_t et P_{t-1} les prix des actifs à la fin de la période t et $t - 1$, respectivement. Sans prise en compte des dividendes, le rendement arithmétique durant la période $[t - 1, t]$, note par R_t est donné par

$$R_T(=) \frac{P_t - P_{t-1}}{P_{t-1}} \quad (1.22)$$

et le rendement géométrique, dit aussi la log-rentabilité ou rendement continu, est donné par

$$r_t = \ln(P_t) - \ln(P_{t-1}) = \ln\left(\frac{P_t}{P_{t-1}}\right) = \ln(R_t + 1) \quad (1.23)$$

Lorsque le pas de temps est petit ou lorsque la rentabilité est proche de zéro, les rendements discrets peuvent être approximes par des rendements continus i.e.

$$r_t = \ln(R_t + 1) \approx R_t$$

Soit $R_t(k), r_t(k)$ la rentabilité arithmétique, respectivement logarithmique, entre la date t et la date $t-k$. Si on ne tient pas compte des dividendes, elles sont données par

$$r_t(k) = (1 + r_t) \times \dots \times (1 + r_{t-k+1}) - 1,$$

et

$$R_T(k) = R_T + \dots + R_{T-K+1}.$$

Dans Solvabilité II, on considéré des horizons de longueur une année, pour cela on utilise les rendements arithmétiques pour notre analyse empirique, en effet, une approximation continue peut ne pas donner des bons résultats.

On a la relation suivante : $1 + R_t = \exp(r_t)$ qui lie le rendement arithmétique au rendement logarithmique.

L'utilisation des observations du passé pour modéliser les réalisations du futur est fondée sur la principale idée que certaines propriétés de ce processus aléatoire demeurent constantes dans le temps, on parle ici de la notion statistique de stationnarité. Ceci signifie que le processus générant les prix reste invariable d'une période à l'autre. Une autre notion fondamentale dans l'analyse des séries financières est l'autocorrélation des données qui décrit la manière avec laquelle une observation dépend des réalisations passées. L'analyse directe des séries de prix se révèle difficile car elles ne sont pas stationnaires et exhibent une forte corrélation. Les séries des rentabilités ont en général de meilleures propriétés statistiques. C'est pour ces raisons que l'ana-

lyse empirique porte généralement sur le processus des rendements et non pas sur le prix des titres.

1.13.2 Caractéristiques des rendements financiers

De nombreuses études empiriques e.g. Rama.Cont (2001) et Chakraborti et al. (2011), indiquent que les séries financières sont caractérisées par des faits stylisés tels que les queues épaisses, les phénomènes de clustering de volatilité et d'asymétrie. Plusieurs études se sont concentrées sur la modélisation des distributions non conditionnelles des rendements. Cette modélisation permet de prendre compte certains faits stylisés sans pour autant pouvoir reproduire la fonction d'autocorrélation des rendements. D'autres modèles sont à variance conditionnelles, permettent non seulement de prendre en compte les faits stylisés mais aussi de reproduire la structure de corrélation. Ces modèles seront présentes plus tard dans ce mémoire.

Dans ce qui suit on présente quelque phénomènes caractérisant les séries des rendements financiers

- A court terme, les rendements oscillent autour de zéro ceci, n'est pas le cas terme.
- a moyenne est nettement plus petite que la volatilité. Par conséquent une simple connaissance de la moyenne ne permet pas de prédire avec précision le rendement futur.
- Mandelbrot (1963) a montré que la distribution des rendements n'est pas normale puisque les queues de distribution sont caractérisées par des queues épaisses. Depuis, un nombre important de modélisation a été proposé. L'identification des queues épaisses peut se faire à travers le moment d'ordre 4 normalisé ou encore le Kurtosis. Pour une distribution normale, le Kurtosis est égal à 3, si le Kurtosis est supérieur à 3, la distribution est dite leptokurtique.
- Une autre caractéristiques des séries de rendements est la présence d'asymétrie négative qui se traduit par une proportion de rendements extrêmes négatifs qui est plus élevée que celle de rendements extrêmes positifs : il y a plus de mouvement fort à la baisse que à la hausse.
- Le clustering de volatilité selon lequel des périodes de forte volatilité alternent avec des périodes de faible volatilité. Les grands rendements (positifs ou négatifs) ont tendance à se suivre, et les petits rendements également. La volatilité se produit ainsi en grappe et les extrêmes sont regroupés en cluster ou paquets de volatilité
- Le processus du prix des actifs financiers est généralement non stationnaire au sens faible (ou au second ordre) alors que le processus des séries de rendements est stationnaire.

- Auto-corrélation des carrés des rendements : cela se traduit par une fonction d'autocorrélation (ACF) des carrés des rendements qui décroît lentement à partir d'une valeur faible mais positive. Cette forme typique de ACF est plus prononcée pour une fréquence élevée que faible. Cette propriété implique que la volatilité est prévisible dans une certaine mesure (ce qui n'est pas le cas pour les rendements). Mandelbrot (1963) indique que les grands changements sont suivis par des grands changements d'autres signes et les petits changements sont suivis par des petits changements. Ce fait stylisé est connu par le phénomène de mémoire longue dans les séries de rendements. Ce fait stylisé est important puisqu'il montre que les incréments des prix ne sont pas indépendamment distribués. On note que pour la série des rendements les autocorrélations sont généralement très faibles.
- La baisse des cours a tendance à engendrer une augmentation de la volatilité supérieure à celle engendrée par une hausse des cours, c'est ce qu'on appelle l'effet de levier.
- Les rendements présentent de nombreux phénomènes de saisonnalité on cite l'effet week-end étudié par French (1980) et Gibbon et Hess (1981) et l'effet de premier janvier par exemple

Chapitre 2

Rappel sur les séries chronologiques

2.1 Introduction

La théorie des séries chronologiques(ou temporelles) est appliquée de nos jours dans des domaines aussi variés que l'économétrie, la médecine ou la démographie. Dans ce chapitre, plusieurs concepts importants liés à l'analyse des séries chronologique seront abordés. Parmi ceux-ci, on trouve les notions de bruit blanc, d'autocorrelation, de la stationnarité et les opérateurs définis sur une série chronologique.

2.2 Définition et concept de base

Une série chronologique est une série des valeur que prend une variable étudiée en fonction du temps, on c'est une suite d'observation d'un phénomène en fonction du temps.

l'analyse d'une série chronologique consiste à identifier et classer les facteurs qui expliquent les variations parmi les valeurs prises par la variable observée sur des intervalles régulières.

Les intervalles sont généralement mesurés en années , mois, semaine,....

2.3 Domaine d'application

On trouve des exemples de séries chronologiques univariées dans de très nombreux domaines.

La liste suivante n'est qu'un échantillon :

- finance et économétrie : évolution des indices boursiers, des prix, des données économiques des entreprises, des ventes et achats de biens, des productions agricoles ou industrielles,
- assurance : analyse des sinistres,

- médecine / biologie : suivi des évolutions des pathologies, analyse d'électroencéphalogrammes et d'électrocardiogrammes,
- sciences de la Terre et de l'Espace : indices de marées, variations des phénomènes physiques (météorologie), évolution des taches solaires, phénomènes d'avalanches,
- traitement du signal : signaux de communications, de radars, de sonars, analyse de la parole,
- traitement des données : mesures successives de position ou de direction d'un objet mobile (trajectographie),
- métrologie : variation de phase ou de fréquence des oscillateurs (où l'on voit que dans un laser, un bruit stationnaire de position sur un miroir conduit à un bruit de phase stationnaire sur l'onde produite, alors qu'un bruit stationnaire sur la longueur de la cavité laser se traduit par un bruit stationnaire sur la fréquence de c'est-à-dire une marche au hasard sur sa phase), dérive et bruit des capteurs inertiels.

2.4 Type de séries chronologiques

On distingue deux types de séries chronologique :

2.4.1 Série continues

C'est une série où l'observation est faite d'une manière continue dans le temps. Exemple : résultat d'un électrocardiogramme.

2.4.2 Séries discrète

C'est une série où l'observation est faite sur des intervalles de temps fixés à priori.

Dans ce cas, $T \in \mathbb{Z}$ et les dates d'observations sont le plus souvent équidistantes par exemple relevés mensuels, trimestriels,.....

Exemple 2.1 *Evolution du nombre de voyageurs utilisant le train.*

2.5 Objectifs de l'analyse d'une chronique

les objectifs de l'analyse d'une chronique sont :

2.5.1 Modélisation

Elle consiste à :

- Développer des modèles permettant de décrire le comportement d'une ou plusieurs séries chronologiques.
- Mettre au point une méthodologie pour spécifier, estimer, valider (juger) un modèle approprié pour des données particulières.

2.5.2 Prévision

Étant données des observations X_1, \dots, X_t , la prévision consiste à évaluer une valeur non observée, X_{t+h} . la prévision peut être ponctuelle, ou prendre la forme d'un intervalle de prévision.

2.6 Les composantes d'une série chronologique

On distingue les composantes suivantes :

- La composante tendancielle (Trend) $T(t)$: La tendance représente l'évolution à long terme du phénomène lié à la croissance générale de l'économie. Il traduit l'allure d'ensemble des phénomènes, il décrit le mouvement de fond de la série qui est la marque de la tendance à la hausse ou à la baisse.
- La composante saisonnière $S(t)$: Les variations saisonnières sont les fluctuations périodiques plus ou moins régulières qui se superposent au mouvement extra-saisonnier. Leurs périodes peuvent être journalières, hebdomadaires ou annuelles. Parmi les facteurs responsables de saisonnalité : les facteurs climatiques, l'inégalité des différents mois, les congés,.....etc.
- La composante cyclique $C(t)$: Elle représente des fluctuations qui apparaissent d'une période à une autre, la durée de ces périodes est généralement irrégulière, elle peut être la cause de changements économiques.
- La composante aléatoire $\varepsilon(t)$:
On retrouve souvent dans les séries chronologiques des mouvements qui apparaissent irréguliers et généralement durant de courtes périodes. Les mouvements ne suivent pas de modèle particulier.

Remarque : la tendance F_t prend différentes formes :

- linéaire : $F_t = a + bt$.
- Quadratique : $F_t = a + bt + Ct^2$.
- Exponentielle : $F_t = ae^{bt}$.

2.7 Les différents schémas d'une série chronologique

Si T est l'observation obtenue à l'instant t , on peut admettre la relation combinant, les quatre composantes peuvent s'exprimer par deux types de schémas de composi-

tion :

- Un schéma additif :

$$T = T_1 + S_t + C_t + \varepsilon_t$$

Dans ce schéma , la saisonnalité est rigide en amplitude et en période.

- Un schéma multiplicatif :

$$T = T_1 \times S_t \times C_t \times \varepsilon_t$$

La composante saisonnière est liée à l'extra saisonnier.

$$T = T_1 \times S_t \times C_t + \varepsilon_t,$$

il est actuellement le plus utilisé en économie.

Il est commabdé puisque le logarithme de la chronique conduit au schéma additif.

2.8 Choix du modèle

Graphiquement : Constatation graphique du parallélisme ou non des droites de minima et des maxima :

- Droites parallèles ; modèle additif.
- Droite non parallèle : modèle multiplicatif.

Algébriquement : On recherche le coefficient directeur de ces deux droites par une méthode analytique (deux points)

- Coefficients directeurs très proches : le modèle choisi sera additif.
- Coefficient directeurs très différents : le modèle choisi sera multiplicatif.

2.9 Opérateur définis sur une série chronologique

Si la série X_t n'est pas stationnaire l'opérateur suivant rendra cette série stationnaire.

2.9.1 Opérateur de retard

Définition 2.1 On note indifféremment β , l'opérateur qui fait passer de X_t à X_{t-1} :

$$\beta X_t = X_{t-1} \tag{2.1}$$

Remarque : L'opérateur β est linéaire et inversible .son inverse $\beta^{-1} = F$ est définit par

$$\forall t \in \mathbb{Z}; FX_t = X_{t+1}$$

L'opérateur F est appelé opérateur d'avance .

Si on compose β avec lui meme on obtient $\beta^2 = \beta \circ \beta$ tel que :

$$\forall t \in \mathbb{Z}; \beta^2 X_t = \beta(\beta X_t) = \beta X_{t-1} = X_{t-2}$$

On peut itérer cette application et définir par récurrence :

$$\beta^k X_t = X_{t-k}; k \in \mathbb{N}$$

Par convention, β_0 est l'opérateur identité I

2.9.2 Opérateur de différence d'ordre d

Définition 2.2 On définit l'opérateur Δ^d

de différence d'ordre d comme l'opérateur linéaire tel que :

$$\Delta^d X_t = (1 - \beta)^d X_t. \tag{2.2}$$

On peut aussi prendre l'opérateur d'ordre 1 est :

$$\Delta X_t = (1 - \beta)X_t = X_t - \beta X_t = X_t - X_{t-1},$$

C'est la série des accroissements, alors que la différence seconde donne la série Δ^2 des accroissements, on a :

$$\Delta^2 X_t = \Delta(\Delta X_t) = (1 - \beta)^2 X_t = (1 - 2\beta + \beta^2)X_t = X_t - 2X_{t-1} + X_{t-2}$$

2.9.3 Processus aléatoire

Pour chaque instant du temps la valeur de la quantité étudiée x_t est appelée variable aléatoire, l'ensemble des valeurs x_t quand "t" varie est appelé processus aléatoire.

2.10 Filtre linéaires et moyennes mobiles

Nous donnons ici deux définitions importantes , pour l'étude des série chronologiques

2.10.1 Filtre linéaire

Un filtre linéaire d'ordre $m = p_1 + p_2 + 1$ est défini par

$$FL = \sum_{j=p_1}^{p_2} \omega_j \beta^{-j} \quad (2.3)$$

$$= \omega_{-p_1} \beta^{p_1} + \omega_{-p_1+1} \beta^{p_1+1} + \dots + \omega_{-1} \beta + \omega_0 I + \omega_1 F + \dots + \omega_{p_2-1} F^{p_2-1} + \omega_{p_2} F^{p_2}.$$

où $p_1, p_2 \in \mathbb{N}$ et $\omega_j \in \mathbb{R}$.

2.10.2 Moyen mobiles

Une moyenne mobile d'ordre $m = p_1 + p_2 + 1$ est un filtre linéaire tel que :

$$\sum_{j=p_1}^{p_2} \omega_j = 1, \text{ pour tout } j = -p_1, \dots, p_2 \quad (2.4)$$

-Beaucoup de moyennes mobiles ont des poids ω_j positifs, mais pas toutes.

-Une moyenne mobile est symétrique si $p_1 = p_2 = p$ et $\omega_j = \omega_{-j}$, pour tout $j = 1, \dots, p$.

-Une moyenne mobile symétrique est dite non pondérée si $\omega_j = cst$ pour tout $j = -p_1, \dots, p_2$

2.10.3 Moyenne mobile et composante saisonnière

Une moyenne mobile est un outil intéressant pour lisser une série temporelle et donc pour enlever une composante saisonnière. On utilise de préférence des moyennes mobiles non pondérées d'ordre égal à la période, par exemple d'ordre 7 pour des données journalières, d'ordre 12 pour des données mensuelles. Par exemple, pour enlever la composante saisonnière due au jour de la semaine, on peut appliquer une moyenne mobile non pondérée d'ordre 7.

$$MM_7 = \frac{1}{7}(\beta^3 + \beta^2 + \beta + 1 + F + F^2 + F^3)$$

Cette moyenne mobile accorde le même poids à chaque jour de la semaine. En effet :

$$MM_7(X_t) = \frac{1}{7}(X_{t-3} + X_{t-2} + X_{t-1} + X_t + X_{t+1} + X_{t+2} + X_{t+3}).$$

Remarque :

-Pour les composantes saisonnières d'une période paire, il n'existe pas de moyennes mobiles centrées non-pondérées.

-Si la période est paire et égale à m , on utilise une moyenne mobile accordant un demi-poids aux deux extrémités.

-Par exemple , pour des données trimestrielles ($m=4$), la moyenne mobile est définie par :

$$MM_4 = \frac{1}{4} \left(\frac{1}{2} \beta^2 + \beta + 1 + F + \frac{1}{2} F^2 \right)$$

En effet

$$MM_4(X_t) = \frac{1}{4} \left(\frac{1}{2} X_{t-2} + X_{t-1} + X_t + X_{t+1} + \frac{1}{2} X_{t+2} \right).$$

2.11 Stationnarité

La stationnarité est une caractéristique d'une série chronologique qui implique que le comportement de la série ne dépend pas du temps. En particulier, on dit qu'une série X_t est stable si elle ne comporte pas de tendance à la hausse ou à la baisse.

Plus formellement , on distingue deux types de stationnarité, à savoir stricte et faible.

Définition 2.3 : (*Stationnarité faible*) Un processus $(X_t)_{t \in \mathbb{Z}}$ est dit stationnaire ou faiblement stationnaire ou stationnaire au second ordre si :

- $E(X_t^2) < \infty$
- $E(X_t) = \mu; \forall t \in \mathbb{Z}$ (constante indépendante de t)
- $Cov(X_t, X_{t-h}) = E(X_t X_{t-h}) - E(X_t)E(X_{t-h}) = \gamma(h), \forall t \in \mathbb{Z}, \forall h \in \mathbb{R}$ (ne dépend pas du temps).

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

Définition 2.4 (*Stationnarité stricte*) Un processus $(X_t)_{t \in \mathbb{Z}}$ est dit strictement stationnaire

ou stationnaire au sens strict si les lois jointes de $(X_{t_1}, \dots, X_{t_k})$ et de $(X_{t_1+h}, \dots, X_{t_k+h})$ sont identiques pour tout entier positif k et pour tout $t_1, \dots, t_k \forall h \in \mathbb{R}$

Intuitivement, une série chronologique strictement stationnaire doit avoir le même comportement statique sur des intervalles de temps égaux.

2.11.1 Relation entre stationnarité faible et stricte

Un processus strictement stationnaire du second ordre est faiblement stationnaire. La réciproque n'est pas vraie en général.

Contre - exemple :

Soit (X_t) une suite de variables aléatoires indépendantes telle que :

- $X_t \rightarrow \exp(1)$ lorsque t est pair,

$-X_t \rightarrow N(1, 1)$ lorsque t est impair,

Alors (X_t) est stationnaire avec $\gamma_x(0) = 1$ et $\gamma_x(h) = 0$ lorsque $h \neq 0$. Cependant X_1 et X_2 n'ont pas la même loi donc (X_t) n'est pas strictement stationnaire.

2.11.2 Test de Dickey Fuller Augmenté

Les tests de stationnarité permettent de vérifier si une série est stationnaire ou non. le test de Dickey Fuller, le test augmenté de Dickey Fuller, ou encore le test de Phillips - Perron (pp), sont des tests de racine unitaire, pour lesquels l'hypothèse nulle est que la série a été générée par un processus présentant une racine unitaire, et donc, qu'elle n'est pas stationnaire. On donne ici le test de ADF .

Les hypothèses du test ADF se définissent de la façon suivante; Un processus non stationnaire, il correspond à une de ces formes de non stationnarité :

$$\Delta X_t = \phi X_{t-1} + \sum_{i=1}^p \alpha_i \Delta X_{t-1} + \varepsilon_t \text{ avec } \phi = p - 1 \text{ Modèle(1)}$$

$$\Delta X_t = \alpha + \phi X_{t-1} + \sum_{i=1}^p \alpha_i \Delta X_{t-1} + \varepsilon_t \quad \text{Modèle(2)}$$

$$\Delta X_t = \alpha + \beta t + \sum_{i=1}^p \alpha_i \Delta X_{t-i} + \varepsilon_t \quad \text{Modèle (3)}$$

1. Test de Dickey et Fuller augmenté : Robuste à l'autocorrélation par rapport au test de Dickey - Fuller.

$$H_0 : X_t \rightarrow I(1)$$

$$H_1 : X_t \text{ n'est pas } I(1) \quad \text{tel que : } I(1) \text{ est un racine unitaire.}$$

Et

$$ADF : \frac{T\widehat{\phi}}{(1 - \widehat{a}_1 - \dots - \widehat{a}_p)} \quad (2.5)$$

2. Test de Phillips-Perron : Robuste à l'hétéroscédasticité

$$pp = T\widehat{\phi} - \frac{1}{2} T^2 \frac{SE(\widehat{\phi})}{\widehat{\delta}_2} (\widehat{\lambda}_2 - \widehat{\delta}_2) \quad (2.6)$$

$$\widehat{\delta}_2 = \frac{1}{T-K} \sum_{t=1}^T (\widehat{\varepsilon}_t^2 - \widehat{\varepsilon}_t) :$$

résidus de la régression,

$$SE(\widehat{\phi}) = \text{écart - type de } \phi,$$

$$\lambda^2 = \delta^2 + 2 \sum_{j=1}^q \left(1 - \frac{j}{q+1}\right) \gamma(j)$$

$$\gamma(j) = \frac{1}{T} \sum_{t=j+1}^T \varepsilon_t \varepsilon_{t-j}$$

Valeurs critiques de ADF et PP : $P(\text{ADF} < c) = \alpha$

- $d_t = 0$ (processus centré)

Table 2.1 – Table d'un processus centré

α T	25	50	100	250	500	∞
0.05	-7.3	-7.7	-7.9	-8	-8	-8.1

- $d_t = 1$ processus avec drift)

Table 2.2 – Table d'un processus avec un drift égale à 1

α T	25	50	100	250	500	∞
0.05	-12.5	-13.3	-13.7	-14	-14	-14.1

2.12 Procédure de choix d'un schéma de composition :

Il existe plusieurs méthodes qui permettent de déterminer le schéma qui peut prendre une série chronologique (graphique ou statistique), parmi les méthodes statistiques, on a le test de régression.

— Test de régression : test de Bays Ballot.

Il se base sur l'estimation de l'équation suivante :

$$\delta_i = b_0 + b_1 \bar{X}_i$$

Avec

$$\bar{X}_i = \frac{1}{n} \sum_{j=1}^m X_{ij} \quad \begin{matrix} i = 1, \dots, n \\ j = 1, \dots, m \end{matrix}$$

$$\delta_i = \frac{1}{n} \sum_{j=1}^m X_{ij} - \bar{X}_i^2 \quad 1/2$$

m : nombre de mois.

n : nombre d'années.

Nous utilisons la méthode de moindre carrée ordinaire pour estimer le paramètre b .

$$\widehat{b} = \frac{\sum_{i=1}^n \delta_i \bar{X}_i - n \cdot \overline{\delta \bar{X}}}{\sum_{i=1}^n \bar{X}_i^2 - n \overline{\bar{X}}^2}$$

$$\overline{\bar{X}} = \frac{\sum_{i=1}^n \bar{X}_i}{n}$$

Si $0.05 \leq \widehat{b} \leq 0$ alors la série suit un schéma mixte.

Si $\widehat{b} < 0$ la série suit un schéma additif.

Si $\widehat{b} > 0.10$ la série suit un schéma multiplicatif.

2.13 Analyse des composantes d'une série chronologique

En général, dans la prévision à court et moyen terme, on se base sur l'identification de la tendance et de la saisonnalité.

2.13.1 Analyse de la tendance :

D'après l'analyse des graphes, on peut distinguer l'existence ou l'absence de cette composante .

Pour plus certitude, on peut aussi utiliser les tests statistiques, à savoir le test de différences. Test de différences : C'est un test non paramétrique qu'on peut présenter comme suit :

H_0 : tendance significative

H_1 : tendance nulle.

Soit T l'observation d'une série chronologique homogène à l'instant t . On calcule "S" qui représente le nombre des différences positives.

$$\nabla X_t = X_t - X_{t-1}$$

Si $N > 12$; alors "S" suit la loi normale d'espérance $E(S)$ et de variance $V(S)$

Avec N : le nombre d'observation :

$$E(s) = \frac{1}{2} (N - 1)$$

$$V(s) = \frac{1}{2} (N + 1)$$

Après avoir calculé S , on calcule pour un risque d'erreur $\alpha = 5\%$ la statistique Z.

$|Z| < U_\alpha$: On accepte H_0 (tendance significative)

$|Z| > U_\alpha$: On rejette H_0 (tendance significative)

U_α : le fractile de la loi normale.

2.13.2 Analyse de la saisonnalité

Pour s'assurer de l'existence de la composante saisonnière , on utilise les tests statistiques suivants :

- Le test graphique : Cette méthode consiste en une présentation graphique de la variable étudiée, si l'allure générale du graphe est pratiquement similaire pour chaque année, alors il existe une saisonnalité, et on peut l'analyser par observation visuelle.
- Test par l'analyse de la variance et test de Fisher : C'est un test statistique qui suppose la chronique sans tendance ou sans saisonnalité.

Dans le cas contraire , cette composante sera éliminée pour une régression sur le temps (extra saisonnalité déterministe), ou par une procédure de filtrage (extra saisonnalité aléatoire).

On pose : $X_t = X_{ij}$

n le nombre d'années, $i = 1, \dots, n$

m le nombre de mois, $j = 1, \dots, m$

Donc le déroulement de test sera :

Z ; moyenne arithmétique des $n.m$ observation

$$Z = \frac{1}{n.m} \sum_{i=1}^n \sum_{j=1}^m X_{ij}$$

Z_i : moyenne arithmétique des années : $Z_i = \frac{1}{n} \sum_{i=1}^n X$

Z_j : moyenne arithmétique des mois : $Z_j = \frac{1}{m} \sum_{j=1}^m X_{ij}$

Ensuite , on calcule les sommes des carrées : posons : S_a : somme des carrées des années , $s_a = m \sum_{i=1}^n (Z_i - Z)^2$

S_m : somme des carrées des mois , $S_m = n \sum_{j=1}^m (Z_j - Z)^2$

S_r : somme des carrées des résidus : $S_r = \sum_{j=1}^n \sum_{i=1}^m (X_{ij} - Z_i - Z_j + Z)^2$

Enfin, on a le tableau de l'analyse de variance suivant : Où

TABLE 2.3 – Tableau des résultats de l'analyse de la variance

Sommes des carrées	Degré de liberté	Variance	F_c	F_t
S_m	$m - 1$	$V_m = \frac{S_m}{m-1}$	$F_c \text{ mois} = \frac{V_a}{V_n}$	$F_{0.5\%}(m-1), (m-1)(n-1)$
S_a	$n - 1$	$V_a = \frac{S_a}{n-1}$	$F_c \text{ année} = \frac{V_a}{V_n}$	$F_{0.5\%}(m-1), (m-1)(n-1)$
S_r	$(m-1)(n-1)$	$V_r = \frac{S_r}{(m-1)(n-1)}$		

V_m : variance due au acteur mois.

V_a : variance due au acteur années.

V_r : variance résiduelle.

Si $F_c \text{ mois} > F^{5\%}((N-1), (m-1)(n-1))$; alors effet mois existe

Donc : L'existence de la saisonnalité

$F_c \text{ année} > F^{5\%}((N-1), (m-1)(n-1))$; alors effet année existe

Donc : L'existence de la tendance.

2.14 Le concept de bruit blanc

Définition 2.5 Un bruit blanc est un processus de moyenne nulle, de variance constante et non autocorrélé.

Définition 2.6 Un processus stationnaire ε_t centré est un bruit blanc si :

- $E(\varepsilon_t) = 0, \forall t$
- $V(\varepsilon_t) = \delta_\varepsilon^2, \forall t$
- $COV(\varepsilon_t, \varepsilon_{t'}) = 0$ pour $t \neq t'$

On note $\varepsilon_t \rightarrow BB(0, \delta_\varepsilon^2)$

2.15 Auto-corrélations simple et partielle

Pour juger de la stationnarité d'un processus, on utilise les fonctions d'auto corrélation simple (FAC) et partielle (FACP).

2.15.1 Fonction d'auto-covariance et auto-corrélation

Soit (X_t) une série temporelle, stationnaire.

Définition 2.7 La fonction d'auto-covariance $\gamma(h)_{h \in \mathbb{Z}}$ mesure la covariance entre une variable et cette même variable à des date différentes, pour un horizon h :

$$\gamma(h) = cov(X_t, X_{t-h}) = E[(X_t - E(X_t))(X_{t-h} - E(X_{t-h}))] \quad (2.7)$$

Ainsi

$$\gamma(0) = Var(X_t) = E[(X_t - E(X_t))^2] = \delta_x^2$$

Elle fournit une information sur la variabilité de la série et sur les liaisons temporelles qui existent entre les diverse composantes de la série x_t

Proposition la fonction d'auto-covariance vérifie les propriétés suivantes :

- $\gamma(0) = Var(X_t) \geq 0$;

- $|\gamma(h)| \leq \gamma_0 \forall h$; la fonction auto-covariance est donc bornée.
- $\gamma(h) = \gamma(-h) \forall h$; autrement dit l'auto-covariance est une fonction paire.

Cette fonction étant paire, on ne la représente que pour $h = 0, 1, 2, \dots$. On a également :

La fonction d'auto-covariance d'une série X_t , stationnaire est de type positif. Cette

propriété exprime le fait que la variance d'une combinaison linéaire de n v.a X_{t_1}, \dots, X_{t_n} . La

fonction d'auto-corrélation est défini par :

$$\rho(h) = \frac{\text{cov}(x_t, x_{t-h})}{\sqrt{\text{var}(x_t)\text{var}(x_{t-h})}} = \frac{\text{cov}(x_t, x_{t-h})}{\text{var}(x_t)} = \frac{\gamma(h)}{\gamma(0)}, h \in \mathbb{Z}. \quad (2.8)$$

Avec $\gamma(0) = 1$ et $|\gamma(h)| < 1$:

La dernière égalité tient car

$$\text{var}(x_{t-h}) = \text{var}(X_t) = \gamma(0)$$

Enfin, en notant que par la stationnarité $E(x_t) = u$, indépendant de t , on a en terme d'espérance mathématique :

$$\rho(h) = \frac{E[((X_t - u)(X_{t-h} - u))]}{E[(X_t - u)^2]}$$

L'équivalent empirique de la fonction d'autocorrélation, notée $\widehat{\rho}(h)$; est obtenu à partir de l'estimateur suivant pour l'autovariance $\widehat{\gamma}(h)$ à l'ordre h : Etant donné une série observée $x_t, t = 1, \dots, T$

Notons

$$\bar{X} = \frac{1}{T} \sum_{t=1}^T X_t.$$

L'autocovariance empirique d'ordre h est :

$$\widehat{\gamma}(h) = \frac{1}{T} \sum_{t=h+1}^T (X_t - \bar{X})(X_{t-h} - \bar{X}), 0 \leq h \leq T - 1. \quad (2.9)$$

Le coefficient d'autocorrélation empirique d'ordre h est :

$$\widehat{\rho}(h) = \frac{\sum_{t=h+1}^T (X_t - \bar{X})(X_{t-h} - \bar{X})}{\sum_{t=1}^T (X_t - \bar{X})^2} = \frac{\widehat{\gamma}(h)}{S_T^2}, 0 \leq h \leq T - 1. \quad (2.10)$$

On définit la matrice de corrélation (de dimension m) de la manière suivante :

$$r(m) = \begin{bmatrix} 1 & \rho(1) & \rho(2) & \dots & \rho(m-1) \\ \rho(1) & 1 & \rho(1) & \dots & \rho(m-2) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \rho(m-1) & \rho(m-2) & \dots & \rho(1) & 1 \end{bmatrix}$$

2.15.2 La fonction d'autocorrélation partielle

Elle mesure la liaison (linéaire) entre X_t et X_{t-h} une fois retiré les liens transitant par les variables intermédiaires $X_{t-1}, \dots, X_{t-h+1}$.

Définition 2.8 Le coefficient d'autocovariance partielle d'ordre h , noté $r(h)$, est défini par :

$$r(h) = \text{corr}(X_t, X_{t-h} | X_{t-1}, \dots, X_{t-h+1}) \quad (2.11)$$

Définition 2.9 Le coefficient d'autocorrélation partielle d'ordre h d'un processus stationnaire se calcule de la manière suivante :

$$r(h) = \frac{|r(h)^*|}{|r(h)|}$$

avec

$$r(h) = \begin{bmatrix} 1 & p(1) & \dots & p(h-1) \\ p(1) & 1 & \dots & p(h-2) \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ p(h-1) & p(h-2) & \dots & 1 \end{bmatrix}$$

Et $R(h)^*$ la matrice $R(h)$ dans laquelle on a remplacé la colonne h par :

$$(p(1) \ p(2) \ p(h))'$$

Soit

$$r(h)^* = \begin{bmatrix} 1 & p(1) & \dots & p(1) \\ p(1) & 1 & \dots & p(2) \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ p(h-1) & p(h-2) & \dots & p(h) \end{bmatrix}$$

Ainsi,

$$r(1) = p(1), r(2) = \frac{p(2) - p(1)^2}{1 - p(1)^2}, \dots$$

De manière empirique, les autocorrélations partielles s'estiment par l'estimation des au-tocorrélations simples et en calculant \widehat{r} à partir de la formule ci-dessus.

2.16 Série linéaire

Définition 2.10 Une série (X_t) est dite linéaire si elle peut s'écrire :

$$X_t = \mu + \sum_{i=-\infty}^{+\infty} \psi_i \varepsilon_{t-i} \quad (2.12)$$

Où $\varepsilon_t \rightarrow \beta\beta N(0, \delta_\varepsilon^2)$, $\psi_0 = 1$ et la suite (ψ_i) est absolument sommable, c'est -à-dire $\sum_i |\psi_i| < \infty$.

Définition 2.11 Une série (X_t) est dite linéaire et causale si elle est linéaire avec $\psi_i = 0, i < 0$:

$$X_t = \mu + \sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i}. \quad (2.13)$$

On admettra qu'une série linéaire est stationnaire. L'étude des série non causales conduit à des résultats non intuitifs difficilement utilisables, aussi nous ne considérons parmi les séries linéaires que des séries causales. L'écriture de X_t comme somme de v.a. non corrélées d'obtenir facilement :

$$\begin{aligned} E(X_t) &= \mu, \text{var}(X_t) = \delta_\varepsilon^2 \left(1 + \sum_{i=1}^{\infty} \psi_i^2\right). \\ \gamma(h) &= \text{cov}(X_t, X_{t-h}) = \delta_\varepsilon^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+h}. \end{aligned}$$

2.17 Analyse de la tendance

Dans ce section, nous nous plaçons dans le cadre d'un modèle composé uniquement d'une tendance et de fluctuations irrégulières et donnons différentes méthodes permettant d'estimer la tendance.

2.17.1 Rappels sur la regression linéaire

Lorsqu'une liaison linéaire forte entre deux variable X et Y semble raisonnable au vu du nuage de points, on a alors une relation du type :

$$y = aX + b.$$

Où les coefficients a et b sont inconnus.

Le problème est que les points du nuage sont rarement (parfaitement) alignés : ils sont proches d'une droite.

Nous cherchons maintenant la droite qui passe au plus près des points du nuage. Pour cela, il faut donc mesurer l'éloignement des points du nuage par rapport à une droite D d'équation $\gamma = ax + b$ puis minimiser un critère d'erreur donné. On peut envisager de minimiser.

-La somme des erreurs en valeur absolue :

$$\min_{a,b} \sum_{i=1}^n |\gamma_i - aX_i - b|.$$

-La somme des erreurs au carré :

$$\min_{a,b} \sum_{i=1}^n (\gamma_i - aX_i - b)^2.$$

La méthode des moindres carrés minimisant le second critère est la plus usité.

2.17.2 La méthode des moindres carrés

On démontre en minimisant la fonction de deux variables

$$g(a; b) = \sum_{i=1}^n (\gamma_i - aX_i - b)^2.$$

Que le couple solution $(\widehat{a}; \widehat{b})$ est donné par :

$$\widehat{a} = \frac{\frac{1}{n} \sum_{i=1}^n \gamma_i X_i - (\frac{1}{n} \sum_{i=1}^n \gamma_i)(\frac{1}{n} \sum_{i=1}^n X_i)}{\frac{1}{n} \sum_{i=1}^n X_i^2 - (\frac{1}{n} \sum_{i=1}^n X_i)^2} = \frac{cov(X; y)}{var(X)}.$$

Et

$$\widehat{b} = \frac{1}{n} \sum_{i=1}^n \gamma_i - \widehat{a} \frac{1}{n} \sum_{i=1}^n X_i = \bar{y} - \widehat{a} \bar{X}.$$

La droite d'équation $y = \widehat{a}x + \widehat{b}$ est appelée droite de régression de γ en X et est notée : $\Delta_{y/x}$.

1. Cette droite passe par le point moyen $M(\bar{y}; \bar{x})$.

2. Le coefficient directeur a de $\Delta_{y/x}$; $cov(X; Y)$ et $r(X; Y)$ sont de même signe :

-Lorsqu'ils sont positifs, on parle de corrélation positive (y augmente quand x augmente).

-Lorsqu'ils sont négatifs, on parle de corrélation négative (γ diminue quand x augmente).

Afin de confirmer qu'il est raisonnable d'approximer le nuage de points par une

$$r(x; y) = \frac{Cov(X; Y)}{\delta_x \delta_y}.$$

a. Le coefficient de corrélation linéaire est symétrique :

$$r(x; y) = r(y; x).$$

b. L'inégalité de Cauchy - Schwarz donne :

$$-1 \leq r(x; y) \leq 1.$$

En pratique, il faut commencer par tracer le nuage de points puis calculer $r(x; y)$ et ce est que si la corrélation linéaire est assez forte que l'on cherchera la droite de régression de y en x .

2.17.3 Ajustement tendanciel linéaire les moindres carrés

Supposons que l'on observe T valeurs d'une série dont la tendance semble être linéaire. La méthode des moindres carrés décrite au paragraphe précédent consiste à estimer la tendance par une fonction linéaire

$$\widehat{Z} = \widehat{a}t + \widehat{b}$$

Le couple solution $(\widehat{a}; \widehat{b})$ est donné par :

$$\widehat{a} = \frac{Cov(t; Z_t)}{var(t)} \text{ et } \widehat{b} = \bar{Z} - \widehat{a}\bar{t}.$$

En posant $\bar{t} = \frac{1}{T} \sum_{i=1}^T t$, $\bar{Z} = \frac{1}{T} \sum_{i=1}^T Z_t$ et

$$Cov(t; Z_t) = \frac{1}{T} \sum_{i=1}^T (t - \bar{t})(Z_t - \bar{Z}).$$

$$var(t) = \frac{1}{T} \sum_{i=1}^T (t - \bar{t})^2.$$

Remarque 1.5 Le coefficient de corrélation linéaire est défini par :

$$r = \frac{Cov(t; Z_t)}{\sqrt{var(t)var(Z_t)}}$$

La corrélation linéaire entre la date t et la variable Z_t est d'autant plus importante que $|r|$ est proche de 1.

2.17.4 Ajustement tendanciel linéaire par points médians

On suppose ici aussi que la tendance est linéaire. Cette méthode est empirique et ne repose sur aucun critère d'erreur à minimiser. Elle peut cependant s'avérer efficace en présence de valeurs aberrantes. On choisit deux points de coordonnées $(t_\alpha; Z_\alpha)$ et $(t_\beta; Z_\beta)$

Les coefficients $(\tilde{a}; \tilde{b})$ vérifiant :

$$\begin{aligned} Z_\alpha &= \tilde{a}t_\alpha + \tilde{b}. \\ Z_\beta &= \tilde{a}t_\beta + \tilde{b}. \end{aligned} \quad \text{soit} \quad \begin{aligned} \tilde{a} &= \frac{Z_\alpha - Z_\beta}{t_\alpha - t_\beta} \\ \tilde{b} &= Z_\alpha - \frac{Z_\alpha - Z_\beta}{t_\alpha - t_\beta} t_\alpha. \end{aligned}$$

Pour choisir les deux points, on constitue deux sous-séries d'observations en général d'effectifs égaux. Puis on prend les points médians de chaque sous-série.

2.17.5 Ajustements tendanciels non linéaires

Lorsque la tendance n'est pas linéaire, une technique simple consiste à se ramener à un ajustement linéaire après un changement de variable approprié.

Exemple 2.2 Si $Z_t = at^2 + b$, en posant $\gamma_t = t^2$, on se ramène à $Z_t = a\gamma_t + b$ et on peut faire un ajustement linéaire après entre γ_t et Z_t .

Si $Z_t = b \exp(at)$, en posant $\gamma_t = \ln(Z_t)$, on se ramène à $\gamma_t = at + \ln(b)$ et on peut faire un ajustement linéaire entre γ_t et t .

2.17.6 Estimation non paramétrique

Dans certaines situations, il n'est pas facile de trouver le degré du polynôme d'ajustement pour Z_t ou de changement de variable adéquat. On pourrait utiliser un polynôme avec un degré élevé mais le nombre de paramètres à estimer serait important et rendrait les calculs fastidieux. Par ailleurs, on ne sait pas non plus déterminer l'allure de cette fonction.

Dans cette situation, on a recours à la théorie non paramétrique de l'estimation de la tendance qui ne suppose rien sur celle-ci a priori et on approxime la tendance par la moyenne mobile arithmétique d'ordre k .

$$Z_t^* = M_k(Z_t)$$

Alors la tendance à la date t peut être estimée par la moyenne mobile (centrée) d'ordre K à la date t .

2.17.7 Théorème de Wold

Le théorème de Wold est fondamental pour l'analyse de séries temporelles stationnaires.

Théorème 2.1 *Tout processus $(X_t)_{t \in \mathbb{Z}}$ faiblement stationnaire peut s'écrire sous la forme :*

$$X_t = \sum_{j=0}^{\infty} \psi_j \varepsilon_{t-j} + K_t \quad (2.14)$$

Où les paramètres ψ_1, ψ_2, \dots sont des réels tels que $\psi_0 = 1$ et $\sum_{j=0}^{\infty} \psi_j^2 < +\infty$, $(\varepsilon_t)_{t \in \mathbb{Z}}$ est un bruit blanc et K_t est une composante linéaire telle que :

$$\text{Cov}(K_t, \varepsilon_t) = 0 \text{ pour tout } j \in \mathbb{Z}.$$

Selon le Théorème de Wold, tout processus stationnaire d'ordre 2 peut être représenté comme une somme pondérée infinie de chocs passés, caractérisés par un bruit blanc.

Dans ce qui suit, deux exemples de séries chronologiques stationnaires sont présentés.

On obtiendra, dans chaque cas, leur représentation selon la formule du théorème de Wold.

Exemple 2.3 *Soit le processus :*

$$X_t = \sum_{j=0}^{\infty} \left(\frac{1}{2}\right)^j v_{t-j+1}$$

Où $(v_t)_{t \in \mathbb{Z}}$ est un bruit blanc gaussien de variance unitaire. Afin que la condition $\psi_0 = 1$ du théorème de Wold soit satisfaite, il suffit de poser :

$$X_t = \frac{1}{2} v_t$$

On a alors :

$$X_t = \sum_{j=0}^{+\infty} \left(\frac{1}{2}\right)^{j-1} \varepsilon_{t-j+1} = \varepsilon_t + \sum_{j=0}^{+\infty} \left(\frac{1}{2}\right)^j \varepsilon_{t-j}$$

Où $(\varepsilon_t)_{t \in \mathbb{Z}}$ est un bruit blanc de variance

$$\text{Var}(\varepsilon_t) = \text{Var}\left(\frac{v_t}{2}\right) = \frac{1}{4}$$

Ainsi, $\psi_0 = 1$ et $\psi_j = \left(\frac{1}{2}\right)^j$. On note que la condition $\sum_{j=0}^{+\infty} \psi_j^2 < +\infty$.

Est également satisfaite car :

$$\sum_{j=0}^{\infty} \psi_j^2 = \sum_{j=0}^{+\infty} \left(\frac{1}{2}\right)^{2j} = \sum_{j=0}^{+\infty} \left(\frac{1}{4}\right)^j = \frac{4}{3}$$

Où le résultat :

$$\sum_{j=0}^{+\infty} \alpha^j = \frac{1}{1-\alpha}, |r| < 1.$$

a été utilisé.

2.18 Processus de référence

2.18.1 Processus auto régressive (AR)

Dans le processus auto régressive AR d'ordre (p) l'observation X_t est générées par une moyenne pondérée des observation passées jusqu'à la pieme période suivant la forme :

$$AR(1) : X_t = \phi_1 X_{t-1} + \varepsilon_t$$

$$AR(1) : X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \varepsilon_t$$

$$AR(p) : X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \varepsilon_t$$

où les $\phi_i, i = 1, \dots, p$ sont les paramètre à estimer positifs.

ε_t est un aléa gaussien : $\varepsilon_t \rightarrow N(0, \delta^2)$.

On peut ajouter à ce processus une constante qui ne modifie en rien les propriétés stochastiques. L'équation précédente peut s'écrire à l'aide de l'opérateur retard(B) :

$$(1 - \phi_1 B^1 - \dots - \phi_p B^p) X_t = \varepsilon_t \iff \phi(B) X_t = \varepsilon_t \quad (2.15)$$

ou B défini par :

$$B^p X_t = X_{t-p}, p \geq 0$$

$$\phi(B) = 1 - \phi_1 B^1 - \dots - \phi_p B^p$$

$\phi(B)$ est appelé polynôme caractéristique de processus X_t .

Le processus X_t sera stationnaire si toutes les racines du polynôme caractéristique $\phi(B)$ sont strictement à l'extérieur du disque unité.

2.18.2 Processus moyenne mobile (MA)

Dans le processus moyenne mobile d'ordre (q) chaque observation X_t est générée par une moyenne pondérée d'aléas jusqu'à la $q^{ième}$ périodes suivant la forme :

$$MA(1) : X_t = \varepsilon_t - \vartheta_1 \varepsilon_{t-1}$$

$$MA(1) : X_t = \varepsilon_t - \vartheta_1 \varepsilon_{t-1} - \vartheta_2 \varepsilon_{t-2}$$

$$MA(q) : X_t = \varepsilon_t - \vartheta_1 \varepsilon_{t-1} - \vartheta_2 \varepsilon_{t-2} - \dots - \vartheta_q \varepsilon_{t-q}$$

ou :

$\vartheta_i, i = 1, \dots, q$ sont les paramètres à estimer peuvent être >0 ou <0 et ε_t est un aléa gaussien $\varepsilon_t \rightarrow N(0, \delta^2)$.

L'équation précédente peut s'inscrire :

$$X_t = (1 - \vartheta_1 B^1 - \dots - \vartheta_q B^q) \varepsilon_t \iff X_t = \vartheta(B) \varepsilon_t \quad (2.16)$$

où

$$X_t = 1 - \phi_1 B^1 - \dots - \phi_p B^p$$

Remarque : Tous les processus MA (q), sont stationnaires par définition :

$$X_t = \sum_{i=1}^q \vartheta_i \varepsilon_{t-i} \quad (2.17)$$

avec $|\vartheta_i| < \infty$ et $\varepsilon_t, t \in Z$

2.18.3 Processus auto régressif moyenne mobile (ARMA)

Les modèle sont représentatif d'un processus généré par un combinaison valeurs passées et des erreurs passées.

Le processus $X_t, t \in Z$ est dit auto régressive à moyenne mobile d'ordre p,q noté ARMA(p,q) s'il satisfait l'équation suivante :

$$X_t - \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} = \varepsilon_t - \vartheta_1 \varepsilon_{t-1} - \dots - \vartheta_p \varepsilon_{t-p}$$

L'équation précédente peut encore s'écrire :

$$(1 - \phi_1 B^1 - \dots - \phi_p B^p) X_t = (1 - \vartheta_1 B^1 - \dots - \vartheta_q B^q) \varepsilon_t \implies \phi(B) X_t = \vartheta(B) \varepsilon_t \quad (2.18)$$

où :

$$\phi(B) = 1 - \phi_1 B^1 - \dots - \phi_P B^P$$

$$\vartheta(B) = 1 - \phi_1 B^1 - \dots - \phi_P B^P$$

Un processus $ARMA(p, q)$ est stationnaire si et seulement si sa partie auto régressive est stationnaire.

2.18.4 Processus ARIMA (auto régressive moyen mobile intégré)

l'existence d'une racine unitaire parmi les racines de $\phi_P(B)$ dans le modèle ARMA permet de définir un processus ARIMA noté $ARIMA(p, d, q)$

$$\phi_P(B)(1 - B)X_t = \vartheta_q(B)\varepsilon_t \quad (2.19)$$

En expliquant l'opération de différentiation de fois en aura :

$$\phi_P(B)\nabla^d X_t = \vartheta_q(B)\varepsilon_t \quad (2.20)$$

le processus ainsi défini est non stationnaire en X_t mais il est stationnaire pour la série différenciée $\nabla^d X_t$

2.18.5 Processus auto régressive moyen mobile intégré saisonnier SARIMA(p,d,q)

Un processus $(X_t, t \in Z)$ est un ARIMA saisonnier pur s'il satisfait une formule de type :

$$\phi_P(B^s)(1 - B^s)^d X_t = \vartheta_q(B)\varepsilon_t \quad (2.21)$$

avec :

d : l'ordre d'intégration

s : la période de la saisonnalité.

2.18.6 Les Modèles des séries temporelles à variance conditionnelle

Dans ce qui suit on présente quelques modèles de type $GARCH$ qui sont très utilisés pour modéliser la volatilité conditionnelle des séries des rendements. Ces modèles permettent d'expliquer les changements dans la variance conditionnelle et sont connus

par des modèles avec hétéroscédasticité conditionnelle. Les modèles de volatilité souvent utilisés sont les modèles autorégressifs avec hétéroscédasticité conditionnelle développés par Engle (1982) connus par les modèles ARCH, les modèles ARCH généralisés connus par les modèles de type GARCH de Bollerslev (1982), les modèles GARCH asymétriques comme le modèle exponentiel GARCH (EGARCH) de Nelson (1991), le modèle QGARCH (Quadratic GARCH), le modèle IGARCH (Integrated GARCH), le modèle GJR – GARCH de Glosten (1993), le modèle GARCH - M (GARCH in Mean)

2.18.7 ARCH(q)

la classe des modèles autorégressifs avec hétéroscédasticité conditionnelle ARCH de paramètre q a été introduite par Engle (1982) pour capturer le phénomène de clustering de volatilité des séries financières. Dans le modèle basique ARCH, la variance conditionnelle qui apparait à la date t est une fonction linéaire des carrés des chocs passés ε tels que le processus ARCH(q) satisfait l'équation suivante :

$$\varepsilon_t = \sigma_t Z_t; \tag{2.22}$$

Et

$$X_t = E(X_t | X_{t-1}) + \varepsilon_t$$

$$Var(\varepsilon_t | \underline{\varepsilon}_{t-1}) = \sigma_t^2 = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-1}^2, \tag{2.23}$$

avec

$$\underline{\varepsilon}_{t-1} = \{\varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-j}, \dots\}.$$

$(Z_t)_{t \geq 1}$ est un processus de bruit blanc, c'est à dire un processus de *v.a.i.i.d* de moyenne zéro et de variance unitaire, $\alpha_0 > 0$ est la moyenne inconditionnelle et $\alpha_i \geq 0$ pour tout $q \geq 1$, pour assurer une variance positive.

Considérons une série financière $(X_t)_{t \geq 1}$ qui modélise les rendements des indices boursiers avec résidus. Soit

$$X_t = E(X_t | X_{t-1}) + \varepsilon_t$$

avec $\underline{X}_{t-1} = \{X_{t-1}, X_{t-2}, \dots, X_{t-j}, \dots\}$ et ε_t un bruit blanc faible tel que $E(\varepsilon_t) = 0$, $E(\varepsilon_t \varepsilon_s) = 0$ pour tout $s \neq t$ satisfaisant la condition de différence de martingale $E(\varepsilon_t | \underline{\varepsilon}_{t-1}) = 0$. Le modèle (1.57) décrit à la fois l'évolution dans le temps de l'espérance conditionnelle du processus $(X_t)_{t \geq 1}$ via le terme $E(X_t | X_{t-1})$ et la variance conditionnelle via le terme ε_t . Par exemple le modèle AR(1) - ARCH(1) est donné par

$$X_t = \mu + \rho X_{t-1} + \varepsilon_t$$

avec

$$\varepsilon_t = Z_t \sqrt{\alpha_0 + \alpha_1 + \varepsilon_{t-1}^2},$$

où $|\rho| < 1$.

Remarque : un processus ARCH est conditionnellement hétéroscédastique mais inconditionnellement homoscedastique

Condition suffisante de stationnarité : $\sum_{i=1}^{\rho} \alpha_i < 1$

On peut montrer également, ce qui peut être intéressant pour détecter un ARCH en pratique, que la distribution d'un processus ARCH a un skewness nul (moment centré d'ordre 3) : la distribution est donc symétrique, un kurtosis (moment centré d'ordre 4) supérieur à 3 : la distribution est donc plus aplatie qu'une gaussienne.

2.18.8 Modèle GARCH

Le modèle GARCH (General Autoregressive Conditional Heteroskedasticity) a été introduit par Bollerslev (1986). C'est une extension du modèle ARCH initialement développé par Engle (2002). Ces modèles permettent une représentation autogressive de la variance conditionnelle d'un processus, ce qui permet de les utiliser notamment à des fins de prévisions de la volatilité sur les marchés financiers.

2.18.8.1 Présentation du Modèle GARCH

L'écriture du modèle GARCH porte sur la variance conditionnelle du processus considéré. Soit un processus y_t , d'espérance $E(y_t) = 0$, satisfaisant une représentation de type GARCH (p, q) . Ce processus s'écrit sous la forme suivante :

$$y_t = c + \varepsilon_t \tag{2.24}$$

$$\varepsilon_t = z_t \sqrt{h_t} \tag{2.25}$$

$$h_t = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{\rho} \beta_i h_{t-i} \tag{2.26}$$

où z_t désigne un bruit blanc faible homoscedastique tel que $E(z_t) = 0$ et $Var(z_t) = \sigma_z^2$ et où les

paramètres α_i, β_i sont des réels. De façon usuelle, la quantité h_t désigne la variance conditionnelle

du processus y_t telle que $V(y_t/y_{t-1}) = V(\varepsilon_t/\varepsilon_{t-1}) = h_t$ où y_{t-1} désigne l'ensemble des valeurs passées $\{y_{t-1}, \dots, y_0\}$. Afin de garantir la positivité de la variance conditionnelle, on suppose que $\alpha_0 > 0, \alpha_i \geq 0, i = 1, \dots, q, \beta_i \geq 0, i = 1, \dots, \rho$

2.18.8.2 Estimation des Paramètres

Les paramètres du modèle GARCH peuvent être estimés selon différentes méthodes : maximum de vraisemblance, pseudo maximum de vraisemblance, méthode des moments etc. (pour plus de détails, voir Gouriéroux, 1997). Les méthodes généralement retenues sont celles du maximum de vraisemblance (MV) ou du pseudo maximum de vraisemblance (PMV). L'avantage du PMV réside dans le fait que l'estimateur obtenu converge malgré une mauvaise spécification (supposée normale) de la distribution conditionnelle des résidus, à condition que la loi spécifiée appartienne à la famille des lois exponentielles (Gouriéroux et Montfort, 1989). Ainsi, l'estimateur du MV obtenu sous l'hypothèse de normalité des résidus et l'estimateur du PMV sont identiques, seules leurs lois asymptotiques respectives diffèrent. Toutefois dans les deux cas (MV ou PMV), sous les hypothèses standards, l'estimateur est asymptotiquement convergent et asymptotiquement normal.

Considérons le cas d'un modèle GARCH(1,1) sous l'hypothèse de normalité des innovations.

$$y_t = c + \varepsilon_t$$

$$\varepsilon_t = z_t \sqrt{h_t}$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}$$

$$z_t \rightsquigarrow N(0, 1)$$

où les variables z_t sont identiquement et indépendamment distribuées (*i.i.d.*) selon une loi normale centrée et réduite. La fonction de log-vraisemblance associée à un échantillon de T observations $\{y_1, \dots, y_T\}$ obtenue sous l'hypothèse de normalité de la loi conditionnelle de y_t sachant son propre passé s'écrit :

$$\log L(\theta) = -\frac{T}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^T \log(h_t(\theta)) - \frac{1}{2} \sum_{t=1}^T \frac{[Y_t - m_t(\theta)]^2}{h_t(\theta)}$$

où θ désigne l'ensemble des paramètres du modèle, $m_t(\theta)$ désigne l'espérance conditionnelle et $h_t(\theta)$ désigne la variance conditionnelle. Dans le cas du modèle GARCH(1,1) présenté ci-dessus :

$$m_t(\theta) = c$$

$$ht(\theta) = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}$$

Les estimateurs du maximum de vraisemblance sont alors obtenus par résolution analytique d'un système de $K = p + q + 2$ (nombre de paramètres à estimer) équations non linéaires :

$$\frac{\partial \log L(\theta)}{\partial \theta} \Big|_{\theta = \hat{\theta}} = 0$$

$$\frac{\partial^2 \log L(\theta)}{\partial \theta^2} \Big|_{\theta = \hat{\theta}} < 0$$

Dans le cas général du PMV, l'estimateur du PMV est asymptotiquement convergent et normal (Gouriéroux, 1997).

$$\sqrt{T}(\hat{\theta} - \theta) \xrightarrow[T \rightarrow \infty]{d} N(0, J^{-1} I J^{-1})$$

$$J = E_0 \left[-\frac{\partial^2 \log L(\theta)}{\partial \theta \partial \theta'} \right]$$

$$I = E_0 \left[\frac{\partial \log L(\theta)}{\partial \theta} \frac{\partial \log L(\theta)}{\partial \theta'} \right]$$

où E_0 désigne l'espérance prise par rapport à la vraie loi. Si la vraie distribution des erreurs est une loi normale (cas du MV) alors $I = J$.

2.18.8.3 Problèmes de convergence et choix des conditions initiales

Il n'existe pas de solution analytique à la maximisation de la vraisemblance dans le cas général des modèles GARCH. On doit donc recourir à une optimisation numérique de la vraisemblance ou de la pseudo vraisemblance suivant les cas. Dans la pratique, se pose alors le problème de la convergence de l'algorithme d'optimisation retenu et plus spécifiquement du problème du choix de conditions initiales.

En ce qui concerne l'algorithme d'optimisation, deux choix sont disponibles dans le cadre de la procédure MODEL sous SAS utilisée pour estimer les modèles GARCH : la méthode de Gauss Newton Raphson et la méthode de Marquardt –Levenberg. C'est cette dernière qui est utilisée dans nos programmes. Pour cet algorithme, nous avons retenu pour conditions initiales sur les paramètres α_i et β_i les valeurs estimées dans le cas d'un modèle GARCH sous hypothèse de normalité.

Propriété : Soit X_t un processus GARCH(p, q), et soit $m = \max(p, q)$. Le processus X_t^2 admet une représentation

ARMA(m, q).

Ainsi, pour identifier un GARCH $_{p,q}$, on identifiera tout d'abord le processus ARMA(m, q) qui modélise X_t^2 . Pour identifier ρ dans le cas où $m = q$ ($p \leq q$), il faut effectuer des tests de significativité des coefficients $\alpha_q, \dots, \alpha_1$ du processus ARMA(m, q) (sont-ils significativement non nuls?).

2.18.9 Modèle EGARCH (Exponential GARCH)

Le modèle EGARCH (Exponential General Autoregressive Conditional Heteroscedasticity) a été introduit par Nelson (1991). C'est un modèle ARCH non linéaire qui permet de rendre compte de l'asymétrie dans la réponse de la variance conditionnelle à une innovation. Plus précisément, le modèle EGARCH autorise une forme d'asymétrie qui dépend non seulement du signe positif ou négatif de l'innovation, mais aussi de l'amplitude de ce choc. Par ailleurs le modèle EGARCH présente en outre l'avantage, par rapport au modèle GARCH standard, de ne nécessiter aucune restriction de non négativité sur les paramètres afin de garantir la positivité de la variance conditionnelle.

2.18.9.1 Présentation du Modèle EGARCH

L'écriture du modèle EGARCH porte sur le logarithme de la variance conditionnelle du processus considéré. Soit un processus y_t , d'espérance $E(y_t) = c$, satisfaisant une représentation de type EGARCH $_{(p,q)}$. Ce processus s'écrit sous la forme suivante :

$$y_t = c + \varepsilon_t \quad (2.27)$$

$$\varepsilon_t = z_t \sqrt{h_t} \quad (2.28)$$

$$\log(h_t) = \alpha_0 + \sum_{i=1}^q \alpha_i g(z_{t-i}) + \sum_{i=1}^p \beta_i \log(h_{t-i}) \quad (2.29)$$

$$g(z_{t-i}) = \theta z_{t-i} + \gamma (|z_{t-i}| - E|z_{t-i}|) \quad (2.30)$$

où z_t désigne un bruit blanc faible homoscédastique tel que $E(z_t) = 0$ et $\text{Var}(z_t) = \delta_z^2$ et où les paramètres $\alpha_i, \beta_i, \gamma$ et θ sont des réels. De façon usuelle, la quantité h_t désigne la variance conditionnelle du processus y_t telle que

$$V(y_t | y_{t-1}) = V(\varepsilon_t | \varepsilon_{t-1}) = h_t$$

où désigne l'ensemble des valeurs passées $\{y_{t-1}, \dots, y_0\}$.

La valeur de $t_i E|z_{t-i}|$ dépend naturellement de la loi supposée de z_t . Pour les lois usuelles retenues dans nos applications, cette quantité est définie de la façon suivante :

Loi Normale :

$$E|z_t| = \sqrt{2/\pi} \quad (2.31)$$

Loi de Student(ν) :

$$E|z_t| = 2 \frac{\Gamma(\nu/2) \sqrt{\nu-2}}{\sqrt{\pi}(\nu-1)\Gamma(\nu/2)} \quad (2.32)$$

Loi GED

$$(\nu) : E|z_t| = \frac{\Gamma(2/\nu)}{\sqrt{\Gamma(1/\nu)\Gamma(3/\nu)}} \quad (2.33)$$

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

Dans le cas du modèle EGARCH, il est inutile d'imposer des contraintes sur les paramètres pour satisfaire la condition de positivité de la variance conditionnelle h_t en raison de son écriture logarithmique. Le paramètre θ permet de modéliser un effet asymétrique lié au signe de l'innovation z_t . Si $\theta > 0$ (respectivement si $\theta < 0$), un choc positif sur la variance conditionnelle à la date t se traduira à la date $t + 1$ par une augmentation (respectivement une diminution) de la variance conditionnelle, c'est-à-dire de la volatilité, du processus y_t . Le paramètre γ permet de prendre en compte une asymétrie liée à l'amplitude de l'innovation z_t mesurée par l'écart $|z_t| - E|z_t|$. Si $\gamma = 0$, alors une innovation positive aura le même effet (en valeur absolue) sur la variance conditionnelle qu'une innovation négative. En revanche, si $\gamma > 0$ un choc de forte d'amplitude aura relativement plus d'effet (en valeur absolue) sur la variance conditionnelle qu'un choc de faible ampleur.

2.18.9.2 Estimation des Paramètres

Les paramètres du modèle EGARCH peuvent être estimés selon différentes méthodes :

maximum de vraisemblance, pseudo maximum de vraisemblance, méthode des moments etc. (pour plus de détails, voir Gouriéroux, 1997). Les méthodes généralement retenues sont celles du maximum de vraisemblance (MV) ou du pseudo maximum de vraisemblance (PMV). L'avantage du PMV réside dans le fait que l'estimateur obtenu converge malgré une mauvaise spécification (supposée normale) de la distribution conditionnelle des résidus, à condition que la loi spécifiée appartienne à la famille des lois exponentielles (Gouriéroux et Montfort, 1989). Ainsi, l'estimateur du MV obtenu sous l'hypothèse de normalité des résidus et l'estimateur du PMV sont identiques, seules leurs lois asymptotiques respectives diffèrent. Toutefois dans les deux cas (MV ou PMV), sous les hypothèses standards, l'estimateur est asymptotiquement convergent et asymptotiquement normal. Considérons le cas d'un modèle

EGARCH(1,1) sous l'hypothèse de normalité des innovations.

$$y_t = c + \varepsilon_t$$

$$\varepsilon_t = z_t \sqrt{h_t}$$

$$h_t = \exp\left(\alpha_0 + \alpha_1 \theta z_{t-1} + \alpha_1 (|z_{t-1}| - \sqrt{2/\pi}) + \beta_1 \log(h_{t-1})\right)$$

$$z_t \rightsquigarrow N(0, 1)$$

où les variables z_t sont identiquement et indépendamment distribuées (*i.i.d.*) selon une loi normale centrée et réduite. Le paramètre γ a été normalisé à 1 puisque dans le cas du EGARCH(1,1) il y a 4 régresseurs et 5 paramètres à estimer. La fonction de log-vraisemblance associée à un échantillon de T observations $\{y_1, \dots, y_T\}$ obtenue sous l'hypothèse de normalité de la loi conditionnelle de y_t sachant son propre passé s'écrit :

$$\log L(p) = -\frac{T}{2} \log(2\pi) - \frac{1}{2} \sum_{i=1}^T \log(h_t(p)) - \frac{1}{2} \sum_{i=1}^T \frac{[y_t - m_t(p)]^2}{h_t(p)}$$

où ρ désigne l'ensemble des paramètres du modèle, $E(\cdot | \rho)$ désigne l'espérance conditionnelle et $h_t(p)$ désigne la variance conditionnelle. Dans le cas du modèle EGARCH(1,1) présenté ci-dessus :

$$m_t(\rho) = c$$

$$h_t(p) = \exp\left(\alpha_0 + \alpha_1 z_{t-1} + \alpha_1 (|z_{t-1}| - \sqrt{2/\pi}) + \beta_1 \log(h_{t-1})\right)$$

Les estimateurs du maximum de vraisemblance sont alors obtenus par résolution analytique d'un système de $K = p + q + 2$ (nombre de paramètres à estimer) équations non linéaires :

$$\frac{\partial \log L(p)}{\partial p} \Big|_{p=\hat{p}} = 0$$

$$\frac{\partial^2 \log L(p)}{\partial p^2} \Big|_{p=\hat{p}} < 0$$

Dans le cas général du PMV, l'estimateur du PMV est asymptotiquement convergent et normal (Gouriéroux, 1997).

$$\sqrt{T}(\widehat{p} - p) \xrightarrow{d}_{t \rightarrow \infty} N(0, J^{-1} I J^{-1})$$

$$J = E_0 \left[\frac{\partial^2 \log L(p)}{\partial p \partial p'} \right]$$

$$I = E_0 \left[\frac{\partial \log L(p)}{\partial p} \frac{\partial \log L(p)}{p'} \right]$$

où E_0 désigne l'espérance prise par rapport à la vraie loi. Si la vraie distribution des erreurs est une loi normale (cas du MV) alors $I = J$.

2.18.9.3 Problèmes de convergence et choix des conditions initiales

Il n'existe pas de solution analytique à la maximisation de la vraisemblance dans le cas général des modèles GARCH. On doit donc recourir à une optimisation numérique de la vraisemblance ou de la pseudo vraisemblance suivant les cas. Dans la pratique, se pose alors le problème de la convergence de l'algorithme d'optimisation retenu et plus spécifiquement du problème du choix des conditions initiales.

En ce qui concerne l'algorithme d'optimisation, deux choix sont disponibles dans le cadre de la procédure MODEL sous SAS utilisée pour estimer les modèles GARCH : la méthode de Gauss Newton Raphson et la méthode de Marquardt –Levenberg. C'est cette dernière qui est utilisée dans nos programmes. Pour cet algorithme, nous avons retenu pour conditions initiales sur les paramètres α_i, β_i les valeurs estimées dans le cas d'un modèle GARCH sous hypothèse de normalité.

Concernant le paramètre θ , non définis dans un modèle GARCH, nous considérons une condition initiale fixée à 0, initialisant ainsi l'algorithme d'optimisation sur un modèle symétrique.

2.18.10 Modèle QGARCH (Quadratic GARCH)

Le processus QGARCH introduit par Engle et Ng (1993) et Sentana (1995) permet de prendre en compte des asymétries dans la réponse de la volatilité conditionnelle aux innovations. Sentana (1995) a montré que les propriétés des modèles GARCH(1, 1) et QGARCH (1, 1) sont très similaires :

les conditions de stationnarité sont identiques à celles dérivées dans le cas du modèle GARCH et comme le résidu est un processus centré, l'expression de son espérance non conditionnelle est également identique à celle obtenue avec un GARCH. En revanche, la Kurtosis dans les modèles QGARCH est toujours plus grande que dans les modèles GARCH symétriques correspondants, ce qui explique que le QGARCH domine souvent empiriquement le GARCH.

2.18.10.1 Présentation du Modèle QGARCH

L'écriture du modèle QGARCH porte sur la variance conditionnelle du processus considéré. Soit un processus y_t , d'espérance $E(y_t) = c$, satisfaisant une représentation de type QGARCH(1, 1). Ce processus s'écrit sous la forme suivante :

$$y_t = c + \varepsilon_t \quad (2.34)$$

$$\varepsilon_t = z_t \sqrt{h_t} \quad (2.35)$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1} + \alpha_2 \varepsilon_{t-1}^2 + \beta_1 h_{t-1} \quad (2.36)$$

où z_t désigne un bruit blanc faible homoscédastique tel que $E(z_t) = 0$ et $Var(z_t) = \delta_z^2$ et où les paramètres α_i, β_i sont des réels. De façon suelle, la quantité h_t désigne la variance conditionnelle du processus y_t telle que $V(y_t | \underline{y_{t-1}}) = V(\varepsilon_t | \underline{\varepsilon_{t-1}}) = h_t$ où $\underline{y_{t-1}}$ désigne l'ensemble des valeurs passées $\{y_{t-1}, \dots, y_1\}$. Le paramètre α_2 permet de modéliser l'effet asymétrique.

Dans le cas du modèle QGARCH, il est utile d'imposer des contraintes sur les paramètres pour satisfaire la condition de positivité de la variance conditionnelle h_t . Ces contraintes se ramènent à $\alpha_0, \alpha_2, \beta_1$ et $\alpha_1^2 \leq 4\alpha_2 \alpha_0$. La stationnarité est assurée si $\alpha_2 + \beta_1 < 1$. Notons que la forme quadratique $f(\varepsilon_{t-1}) = \alpha_1 \varepsilon_{t-1} + \alpha_2 \varepsilon_{t-1}^2$ est minimale en $-\alpha_1 / (2\alpha_2)$. Par conséquent, la symétrie de la réponse n'est pas obtenue en zéro mais en ce point : à amplitude donnée de l'innovation passée, on a bien un impact sur h_t différent selon le signe de ε_{t-1} . Sentana (1995) propose une généralisation à un modèle QGARCH(p, q) telle que :

$$h_t = \theta + \sum_{i=1}^q \Psi_i \varepsilon_{t-i} + \sum_{i=1}^q a_{ii} \varepsilon_{t-i}^2 + 2 \sum_{i=1}^q \sum_{j=i+1}^q a_{ij} \varepsilon_{t-i} \varepsilon_{t-j} \quad (2.37)$$

Ce modèle permet de tester simplement les asymétries dynamiques dans la fonction de variance conditionnelle de certaines séries financières. De plus, la formulation du QGARCH peut aussi être interprétée comme une approximation de Taylor au second ordre de la fonction de variance ou comme la projection quadratique de l'innovation au carré sur l'ensemble d'information. En fait, cette projection coïncide avec la projection de la vraie variance conditionnelle sur un « ensemble d'information quadratique », donc nous pouvons ainsi comprendre le modèle QGARCH comme produisant des estimateurs filtrés et lissés de la variance conditionnelle.

2.18.10.2 Estimation des Paramètres

Les paramètres du modèle QGARCH peuvent être estimés selon différentes méthodes : maximum de vraisemblance, pseudo maximum de vraisemblance, méthode des moments etc. (pour plus de détails, voir Gouriéroux, 1997). Les méthodes généralement retenues sont celles du maximum de vraisemblance (MV) ou du pseudo maximum de vraisemblance (PMV). L'avantage du PMV réside dans le fait que l'estimateur obtenu converge malgré une mauvaise spécification (supposée normale) de la distribution conditionnelle des résidus, à condition que la loi spécifiée appartienne à la famille des lois exponentielles (Gouriéroux et Montfort, 1989). Ainsi, l'estimateur du MV obtenu sous l'hypothèse de normalité des résidus et l'estimateur du PMV sont identiques, seules leurs lois asymptotiques respectives diffèrent. Toutefois dans les deux cas (MV ou PMV), sous les hypothèses standards, l'estimateur est asymptotiquement convergent et asymptotiquement normal.

Considérons le cas d'un modèle QGARCH(1,1) sous l'hypothèse de normalité des innovations.

$$y_t = c + \varepsilon_t$$

$$\varepsilon_t = z_t \sqrt{h_t}$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \gamma_1 \varepsilon_{t-1} + \beta_1 h_{t-1}$$

$$z_t \rightsquigarrow N(0, 1)$$

où les variables z_t sont identiquement et indépendamment distribuées (*i.i.d.*) selon une loi normale centrée et réduite.

La fonction de log-vraisemblance associée à un échantillon de T observations $\{y_1, \dots, y_T\}$ obtenue sous l'hypothèse de normalité de la loi conditionnelle de y_t sachant son propre passé s'écrit :

$$\log L(\theta) = -\frac{T}{2} \log(2\pi) - \frac{1}{2} \sum_{i=1}^T \log[h_t(\theta)] - \frac{1}{2} \sum_{i=1}^T \frac{[y_t - m_t(\theta)]^2}{h_t(\theta)}$$

où θ désigne l'ensemble des paramètres du modèle, $m_t(\theta)$ désigne l'espérance conditionnelle et $h_t(\theta)$ désigne la variance conditionnelle. Dans le cas du modèle QGARCH(1,1) présenté ci-dessus :

$$m_t(\theta) = c$$

$$h_t(\theta) = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \gamma_1 \varepsilon_{t-1} + \beta_1 h_{t-1}$$

Les estimateurs du maximum de vraisemblance sont alors obtenus par résolution analytique d'un système de $K = p + q + 2$ (nombre de paramètres à estimer) équations non linéaires :

$$\frac{\partial \log L(\theta)}{\partial \theta} \Big|_{\theta = \hat{\theta}} = 0$$

$$\frac{\partial^2 \log L(\theta)}{\partial \theta^2} \Big|_{\theta = \hat{\theta}} \prec 0$$

Dans le cas général du PMV, l'estimateur du PMV est asymptotiquement convergent et normal (Gouriéroux, 1997).

$$\sqrt{T}(\theta - \hat{\theta}) \xrightarrow[T \rightarrow \infty]{d} N(0, J^{-1} I J^{-1})$$

$$J = E_0 \left[\frac{\partial^2 \log L(\theta)}{\partial \theta \partial \theta'} \right]$$

$$I = E_0 \left[\frac{\partial \log L(\theta)}{\partial \theta} \frac{\partial \log L(\theta)}{\partial \theta'} \right]$$

où E_0 désigne l'espérance prise par rapport à la vraie loi. Si la vraie distribution des erreurs est une

loi normale (cas du MV) alors $I = J$.

2.18.10.3 Problèmes de convergence et choix des conditions initiales

Il n'existe pas de solution analytique à la maximisation de la vraisemblance dans le cas général des modèles GARCH. On doit donc recourir à une optimisation numérique de la vraisemblance ou de la pseudo vraisemblance suivant les cas. Dans la pratique, se pose alors le problème de la convergence de l'algorithme d'optimisation retenu et plus spécifiquement du problème du choix de conditions initiales.

En ce qui concerne l'algorithme d'optimisation, deux choix sont disponibles dans le cadre de la procédure MODEL sous SAS utilisée pour estimer les modèles GARCH : la méthode de Gauss Newton Raphson et la méthode de Marquardt –Levenberg. C'est cette dernière qui est utilisée dans nos programmes. Pour cet algorithme, nous avons retenu pour conditions initiales sur les paramètres α_i et β_i

les valeurs estimées dans le cas d'un modèle GARCH sous hypothèse de normalité.

Concernant le paramètre γ , non défini dans un modèle GARCH, nous retenons une condition initiale fixée à 0, initialisant ainsi l'algorithme d'optimisation sur un modèle

symétrique.

2.18.11 Modèle IGARCH (Integrated GARCH)

Le modèle IGARCH (Integrated General Autoregressive Conditional Heteroskedasticity) correspond au cas d'une racine unitaire dans la variance conditionnelle. C'est un modèle ARCH non linéaire caractérisé par un effet de persistance dans la variance. C'est-à-dire qu'un choc sur la variance conditionnelle

actuelle se répercute sur toutes les valeurs futures prévues.

2.18.11.1 Présentation du Modèle IGARCH

L'écriture du modèle IGARCH porte sur la non stationnarité de son processus de variance

conditionnelle et par une variance non conditionnelle infinie. Soit un processus y_t , d'espérance $E(y_t) = 0$, satisfaisant une représentation de type IGARCH(p, q). Ce processus s'écrit sous la forme suivante :

$$y_t = z_t \sqrt{h_t} \quad (2.38)$$

$$\varepsilon_t = z_t \sqrt{h_t} \quad (2.39)$$

$$\beta_t = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^p (1 - \alpha_i) \beta_{t-i} \quad (2.40)$$

où z_t désigne un bruit blanc faible homoscédastique tel que $E(z_t) = 0$ $Var(z_t) = 1$ et où les paramètres α_0, α_1 sont des réels tels que :

$$(2.41)$$

De façon usuelle, la quantité h_t désigne la variance conditionnelle du processus y_t telle que $V(y_t | \underline{y}_{t-1}) = V(\varepsilon_t | \underline{\varepsilon}_{t-1}) = h_t$ où \underline{y}_{t-1} désigne l'ensemble des valeurs passées $\{y_{t-1}, \dots, y_0\}$.

Les prévisions de la variance conditionnelle aux différents horizons k sont de la forme :

$$E(h_{t+k} | \varepsilon_t) = (\alpha_1 + B_1)^k h_t + \alpha_0 \sum_{i=0}^{k-1} (\alpha_1 + B_1)^i \quad (2.42)$$

Lorsque $\alpha_i + \beta_i < 1$, le processus ε_t est stationnaire et l'influence du choc sur la variance conditionnelle est décroissante et asymptotiquement négligeable sur b_{t+k} .

Dans ce cas, on a $\beta_1 \neq (1 - \alpha_1)$. Lorsque $\alpha_1 + \beta_1 = 1$ avec $\beta_1 = (1 - \alpha_1)$ on a :

$$E(b_{t+k}/\varepsilon_t) = b_t + \alpha_0 k \quad (2.43)$$

en présence d'un terme constant, $E(b_{t+k}|\varepsilon_t)$ diverge avec k . Ce qui traduit alors la persistance des chocs passés.

2.18.11.2 Estimation des Paramètres

Les paramètres du modèle GARCH peuvent être estimés selon différentes méthodes : maximum de vraisemblance, pseudo maximum de vraisemblance, méthode des moments etc. (pour plus de détails, voir Gouriéroux, 1997). Les méthodes généralement retenues sont celles du maximum de vraisemblance (MV) ou du pseudo maximum de vraisemblance (PMV). L'avantage du PMV réside dans le fait que l'estimateur obtenu converge malgré une mauvaise spécification (supposée normale) de la distribution conditionnelle des résidus, à condition que la loi spécifiée appartienne à la famille des lois exponentielles (Gouriéroux et Montfort, 1989). Ainsi, l'estimateur du MV obtenu sous l'hypothèse de normalité des résidus et l'estimateur du PMV sont identiques, seules leurs lois asymptotiques respectives diffèrent. Toutefois dans les deux cas (MV ou PMV), sous les hypothèses standards, l'estimateur est asymptotiquement convergent et asymptotiquement normal. Considérons le cas d'un modèle IGARCH(1,1) sous l'hypothèse de normalité des innovations.

$$y_t = c + \varepsilon_t$$

$$\varepsilon_t = z_t \sqrt{h_t}$$

$$\beta_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + (1 - \alpha_1) \beta_{t-1}$$

$$z_t \rightsquigarrow N(0, 1)$$

où les variables z_t sont identiquement et indépendamment distribuées (*i.i.d.*) selon une loi normale centrée et réduite. La fonction de log-vraisemblance associée à un échantillon de T observations $\{y_1, \dots, y_T\}$ obtenue sous l'hypothèse de normalité de la loi conditionnelle de y_t sachant son propre passé s'écrit :

$$\log L(\theta) = -\frac{T}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^T \log(h_t, \theta) - \frac{1}{2} \sum_{t=1}^T \frac{[y_t - m_t(\theta)]^2}{h_t(\theta)}$$

où θ désigne l'ensemble des paramètres du modèle, $m_t(\theta)$ désigne l'espérance conditionnelle et $h_t(\theta)$ désigne la variance conditionnelle. Dans le cas du modèle IGARCH(1, 1) présenté ci-dessus :

$$m_t(\theta) = c$$

$$b_t = \alpha_0 + 1\alpha_1\varepsilon_{t-1}^2 + (1 - \alpha_1)b_{t-1}$$

Les estimateurs du maximum de vraisemblance sont alors obtenus par résolution analytique d'un système de $K = p + q + 2$ (nombre de paramètres à estimer) équations non linéaires :

$$\frac{\partial \log L(\theta)}{\partial \theta} \Big|_{\theta=\hat{\theta}} = 0$$

$$\frac{\partial^2 \log L(\theta)}{\partial \theta^2} \Big|_{\theta=\hat{\theta}} \prec 0$$

Dans le cas général du PMV, l'estimateur du PMV est asymptotiquement convergent et normal (Gouriéroux, 1997).

$$\sqrt{T}(\hat{\theta} - \theta) \rightarrow_{T \rightarrow \infty}^d N(0, J^{-1} I J^{-1})$$

$$J = E_0 \left[\frac{\partial^2 \log L(\theta)}{\partial \theta \partial \theta'} \right]$$

$$I = E_0 \left[\frac{\partial \log L(\theta)}{\partial \theta} \frac{\partial \log L(\theta)}{\partial \theta'} \right]$$

où E_0 désigne l'espérance prise par rapport à la vraie loi. Si la vraie distribution des erreurs est une loi normale (cas du MV) alors $I = J$.

2.18.11.3 Problèmes de convergence et choix des conditions initiales

Il n'existe pas de solution analytique à la maximisation de la vraisemblance dans le cas général des modèles GARCH. On doit donc recourir à une optimisation numérique de la vraisemblance ou de la pseudo vraisemblance suivant les cas. Dans la pratique, se pose alors le problème de la convergence de l'algorithme d'optimisation retenu et plus spécifiquement du problème du choix de conditions initiales.

En ce qui concerne l'algorithme d'optimisation, deux choix sont disponibles dans le cadre de la procédure AUTOREG sous SAS utilisée pour estimer les modèles GARCH : la méthode de Gauss Newton Raphson et la méthode de Marquardt –Levenberg. Pour cet algorithme, nous avons retenu pour conditions initiales sur les paramètres α_i, B_i

les valeurs estimées dans le cas d'un modèle IGARCH sous hypothèse de normalité.

2.18.12 Modèle GJR-GARCH

Le modèle GJR-GARCH a été introduit par Glosten, Jagannathan et Runkle (1993). C'est un modèle GARCH non linéaire qui permet de rendre compte de l'asymétrie dans la réponse de la variance conditionnelle à une innovation. La logique de ce modèle est similaire à celle des modèles à changements de régimes et plus spécifiquement des modèles à seuils (Tong, 1990). Le principe du modèle GJR-GARCH est que la dynamique de la variance conditionnelle admet un changement de régime qui dépend du signe de l'innovation passée.

2.18.12.1 Présentation du Modèle GJR-GARCH

Soit un processus y_t , d'espérance $E(y_t) = c$, satisfaisant une représentation de type GJRGARCH(p,q). Ce processus s'écrit sous la forme suivante :

$$y_t = c + \varepsilon_t \quad (2.44)$$

$$\varepsilon_t = z_t \sqrt{h_t} \quad (2.45)$$

$$h_t = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \phi_i I_{\varepsilon_{t-1} < 0} \varepsilon_{t-i}^2 + \sum_{i=1}^p \beta_i h_{t-i} \quad (2.46)$$

$$z_t \rightsquigarrow N(0, 1)$$

Où z_t désigne un bruit blanc faible homoscedastique tel que $E(z_t) = 0$ et $V(z_t) = \delta_z^2$, où les paramètres α_i , β_i et ϕ_i sont des réels et où $I_{\varepsilon_{t-1} < 0}$ désigne la fonction indicatrice telle que $I_{\varepsilon_{t-1} < 0} = 1$ si $\varepsilon_{t-1} < 0$ et $I_{\varepsilon_{t-1} < 0} = 0$ sinon. De façon usuelle, la quantité h_t désigne la variance conditionnelle du processus y_t telle que

$$V(y_t | \underline{y}_{t-1}) = V(\varepsilon_t | \underline{\varepsilon}_{t-1}) = h_t$$

où \underline{y}_{t-1} désigne l'ensemble des valeurs passées $\{y_{t-1}, \dots, y_0\}$.

Dans le cas du modèle GJR-GARCH, il est utile d'imposer des contraintes sur les paramètres pour satisfaire la condition de positivité de la variance conditionnelle h_t :

$$\alpha_0 > 0, \alpha_i \geq 0, i = 1, \dots, q \text{ et } \beta_i \geq 0, i = 1, \dots, p.$$

Les paramètres α_i , β_i permettent de modéliser un effet asymétrique lié aux signes des

innovations passées ε_{t-i} . Si $\Phi_i > 0$ (respectivement si $\Phi_i < 0$), un choc positif sur les innovations passées à la $t - i$ se traduira à la date t par une augmentation (respectivement une diminution) de la variance conditionnelle, c'est-à-dire de la volatilité, du processus y_t .

2.18.12.2 Estimation des Paramètres

Les paramètres du modèle GJR-GARCH peuvent être estimés selon différentes méthodes : maximum de vraisemblance, pseudo maximum de vraisemblance, méthode des moments etc. (pour plus de détails, voir Gouriéroux, 1997). Les méthodes généralement retenues sont celles du maximum de vraisemblance (MV) ou du pseudo maximum de vraisemblance (PMV). L'avantage du PMV réside dans le fait que l'estimateur obtenu converge malgré une mauvaise spécification (supposée normale) de la distribution conditionnelle des résidus, à condition que la loi spécifiée appartienne à la famille des lois exponentielles (Gouriéroux et Montfort, 1989). Ainsi, l'estimateur du MV obtenu sous l'hypothèse de normalité des résidus et l'estimateur du PMV sont identiques, seules leurs lois asymptotiques respectives diffèrent. Toutefois dans les deux cas (MV ou PMV), sous les hypothèses standards, l'estimateur est asymptotiquement convergent et asymptotiquement normal. Considérons le cas d'un modèle GJR-GARCH(1,1) sous l'hypothèse de normalité des innovations.

$$y_t = c + \varepsilon_t$$

$$\varepsilon_t = z_t \sqrt{h_t}$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \Phi_i I_{\varepsilon_{t-i} < 0} \varepsilon_{t-1}^2 + \sum_{i=1}^p \beta_i h_{t-1}$$

$$z_t \rightsquigarrow N(0, 1)$$

où les variables z_t sont identiquement et indépendamment distribuées (*i.i.d.*) selon une loi normale centrée et réduite. La fonction de log-vraisemblance associée à un échantillon de T observations $\{y_1, \dots, y_T\}$ obtenue sous l'hypothèse de normalité de la loi conditionnelle de t y sachant son propre passé s'écrit :

$$\log L(p) = -\frac{T}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^T \log(h_t(p)) - \frac{1}{2} \sum_{t=1}^T \frac{[y_t - m_t(p)]^2}{h_t(p)}$$

où ρ désigne l'ensemble des paramètres du modèle, $m_t(\rho)$ désigne l'espérance conditionnelle et $h_t(\rho)$ désigne la variance conditionnelle. Dans le cas du modèle GJR-GARCH(1, 1) présenté cidessus

$$m_t(p) = c$$

$$h_t(p) = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \Phi I_{\varepsilon_{t-1} < 0} \varepsilon_{t-1}^2 + \beta_1 h_{t-1}$$

Les estimateurs du maximum de vraisemblance sont alors obtenus par résolution analytique d'un système de $K = p + q + 2$ (nombre de paramètres à estimer) équations non linéaires :

$$\frac{\partial \log L(\theta)}{\partial \theta} \Big|_{\theta = \hat{\theta}} = 0$$

$$\frac{\partial^2 \log L(\theta)}{\partial \theta^2} \Big|_{\theta = \hat{\theta}} < 0$$

Dans le cas général du PMV, l'estimateur du PMV est asymptotiquement convergent et normal (Gouriéroux, 1997).

$$\sqrt{T}(\hat{\theta} - \theta) \xrightarrow[T \rightarrow \infty]{d} N(0, J^{-1} I J^{-1})$$

$$J = E_0 \left[\frac{\partial^2 \log L(\theta)}{\partial \theta \partial \theta'} \right]$$

$$I = E_0 \left[\frac{\partial \log L(\theta)}{\partial \theta} \frac{\partial \log L(\theta)}{\partial \theta'} \right]$$

où E_0 désigne l'espérance prise par rapport à la vraie loi. Si la vraie distribution des erreurs est une loi normale (cas du MV) alors $I = J$.

2.18.12.3 Problèmes de convergence et choix des conditions initiales

Il n'existe pas de solution analytique à la maximisation de la vraisemblance dans le cas général des modèles GARCH. On doit donc recourir à une optimisation numérique de la vraisemblance ou de la pseudo vraisemblance suivant les cas. Dans la pratique, se pose alors le problème de la convergence de l'algorithme d'optimisation retenu et plus spécifiquement du problème du choix de conditions initiales.

En ce qui concerne l'algorithme d'optimisation, deux choix sont disponibles dans le cadre de la procédure MODEL sous SAS utilisée pour estimer les modèles GARCH : la méthode de Gauss Newton Raphson et la méthode de Marquardt –Levenberg. C'est cette dernière qui est utilisée dans nos programmes. Pour cet algorithme, nous avons

retenu pour conditions initiales sur les paramètres α_i, β_i les valeurs estimées dans le cas d'un modèle GARCH sous hypothèse de normalité.

Concernant les paramètres ϕ_i , non définis dans un modèle GARCH, nous considérons des conditions initiales fixées à 0.0001 permettant ainsi d'évaluer la pertinence de notre modèle asymétrique.

2.18.13 Modèle GARCH-M (GARCH in Mean)

Le modèle GARCH-M (General Autoregressive Conditional Heteroskedasticity in Mean) a été introduit par Engle-Lilien-Robbins (1987). C'est un modèle ARCH qui permet de mesurer l'influence du rendement des titres sur la volatilité conditionnelle. Comme les modèles ARCH linéaires, le modèle GARCH-M repose sur une spécification quadratique de la variance conditionnelle des perturbations.

2.18.13.1 Présentation du Modèle GARCH-M (GARCH in Mean)

L'écriture du modèle GARCH-M porte sur la non stationnarité de son processus de variance conditionnelle et par une variance non conditionnelle infinie. Soit un processus y_t d'espérance $E(y_t) = 0$, satisfaisant une représentation de type GARCH-M (p,q). Ce processus s'écrit sous la forme suivante :

$$y_t = x_t b + \delta h_t + \varepsilon_t \quad (2.47)$$

$$\varepsilon_t = z_t \sqrt{h_t} \quad (2.48)$$

$$h_t = \alpha_0 + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^p \beta_i h_{t-i} \quad (2.49)$$

où z_t désigne un bruit blanc faible homoscedastique tel que $E(z_t) = 0$ et $Var(z_t) = \delta_z^2$ et où les paramètres α_i, β_i sont des réels. De façon usuelle, la quantité h_t désigne la variance conditionnelle du processus y_t telle que $V(y_t | \underline{y}_{t-1}) = V(\varepsilon_t | \underline{\varepsilon}_{t-1}) = h_t$ où \underline{y}_{t-1} désigne l'ensemble des valeurs passées $\{y_{t-1}, \dots, y_0\}$. Afin de garantir la positivité de la variance conditionnelle, les conditions de positivité des paramètres suffisent, on suppose : $\alpha_0 > 0, \alpha_i \geq 0, i = 1, \dots, q, \beta_i \geq 0, i = 1, \dots, p$.

Le paramètre δ mesure l'impact de la variance conditionnelle sur le rendement ; il correspond également au coefficient de l'aversion relative au risque. Les variables z_t sont indépendamment et identiquement distribuées (*i.i.d.*).

En plus de la forme linéaire de l'écriture de y_t ci-dessus des variantes sont possibles ; c'est-à-dire les formes log-linéaire et racine carré.

$$y_t = x_t b + \delta \log(h_t) + \varepsilon_t$$

Forme Log-Linéaire

$$y_t = x_t b + \delta \sqrt{h_t} + \varepsilon_t$$

Forme Racine Carré

2.18.13.2 Estimation des Paramètres

Les paramètres du modèle GARCH peuvent être estimés selon différentes méthodes : maximum de vraisemblance, pseudo maximum de vraisemblance, méthode des moments etc. (pour plus de détails, voir Gouriéroux, 1997). Les méthodes généralement retenues sont celles du maximum de vraisemblance (MV) ou du pseudo maximum de vraisemblance (PMV). L'avantage du PMV réside dans le fait que l'estimateur obtenu converge malgré une mauvaise spécification (supposée normale) de la distribution conditionnelle des résidus, à condition que la loi spécifiée appartienne à la famille des lois exponentielles (Gouriéroux et Montfort, 1989). Ainsi, l'estimateur du MV obtenu sous l'hypothèse de normalité des résidus et l'estimateur du PMV sont identiques, seules leurs lois asymptotiques respectives diffèrent. Toutefois dans les deux cas (MV ou PMV), sous les hypothèses standards, l'estimateur est asymptotiquement convergent et asymptotiquement normal.

Considérons le cas d'un modèle GARCH-M (1,1) sous l'hypothèse de normalité des innovations.

$$y_t = c + \varepsilon_t$$

$$\varepsilon_t = z_t \sqrt{h_t}$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}$$

$$z_t \text{ i.i.d. } N(0, 1)$$

où les variables z_t sont identiquement et indépendamment distribuées (i.i.d.) selon une loi normale centrée et réduite. La fonction de log-vraisemblance associée à un échantillon de T observations $\{y_1, \dots, y_T\}$ obtenue sous l'hypothèse de normalité de la loi conditionnelle de y_t sachant son propre passé s'écrit :

$$\log L(\theta) = -\frac{T}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^T \log(h_t(\theta)) - \frac{1}{2} \sum_{t=1}^T \frac{[y_t - m_t(\theta)]^2}{h_t(\theta)}$$

où θ désigne l'ensemble des paramètres du modèle, $m_t(\theta)$ désigne l'espérance conditionnelle et $h_t(\theta)$ désigne la variance conditionnelle. Dans le cas du modèle GARCH-M (1,1) :

$$m_t(\theta) = c$$

$$h_t(\theta) = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1}$$

Les estimateurs du maximum de vraisemblance sont alors obtenus par résolution analytique d'un système de $K = p + q + 2$ (nombre de paramètres à estimer) équations non linéaires :

$$\frac{\partial \log L(\theta)}{\partial \theta} \Big|_{\theta = \hat{\theta}} = 0$$

$$\frac{\partial^2 \log L(\theta)}{\partial \theta^2} \Big|_{\theta = \hat{\theta}} < 0$$

Dans le cas général du PMV, l'estimateur du PMV est asymptotiquement convergent et normal (Gouriéroux, 1997).

$$\sqrt{T}(\hat{\theta} - \theta) \xrightarrow[T \rightarrow \infty]{d} N(0, J^{-1} I J^{-1})$$

$$J = E_0 \left[-\frac{\partial^2 \log L(\theta)}{\partial \theta \partial \theta'} \right]$$

$$I = E_0 \left[\frac{\partial \log L(\theta)}{\partial \theta} \frac{\partial \log L(\theta)}{\partial \theta'} \right]$$

où E_0 désigne l'espérance prise par rapport à la vraie loi. Si la vraie distribution des erreurs est une loi normale (cas du MV) alors $I = J$.

2.18.13.3 Problèmes de convergence et choix des conditions initiales

Il n'existe pas de solution analytique à la maximisation de la vraisemblance dans le cas général des modèles GARCH. On doit donc recourir à une optimisation numérique de la vraisemblance ou de la pseudo vraisemblance suivant les cas. Dans la pratique, se pose alors le problème de la convergence de l'algorithme d'optimisation retenu et plus spécifiquement du problème du choix de conditions initiales.

En ce qui concerne l'algorithme d'optimisation, deux choix sont disponibles dans le

dans le cadre de la procédure AUTOREG (permet l'estimation et la prévision de modèle linéaire avec les erreurs, autocorrélées ou hétéroscédastique) sous SAS (logiciel fonctionne sur différents systèmes d'exploitation) utilisée pour estimer les modèles GARCH : la méthode de Gauss Newton Raphson et la méthode de Marquardt–Levenberg. C'est cette dernière qui est utilisée dans nos programmes. Pour cet algorithme, nous avons retenu pour conditions initiales sur les paramètres α_i et β_i les valeurs estimées dans le cas d'un modèle GARCH-M sous hypothèse de normalité.

2.19 Distributions Conditionnelles - Modèles GARCH

Différentes lois peuvent être utilisées dans le cadre des procédures de maximum de vraisemblance pour l'estimation des paramètres du modèle GARCH. Nous proposons ici une présentation des propriétés des lois les plus généralement utilisées à savoir :

- la distribution Normale
- la distribution de Student
- la distribution General Error Distribution (GED)
- la distribution Skewed Student
- la distribution Skewed GED

2.19.1 Loi Normale

La densité normale a longtemps été utilisée dans la littérature du fait de sa simplicité. Pour autant, il est aujourd'hui reconnu que les propriétés de la loi normale, symétrique et mésokurtique, ne sont pas compatibles avec les faits stylisés (distribution conditionnelle leptokurtique et asymétrique) observés généralement sur les séries de rendements des actifs financiers (Cont, 2001). Notons toutefois, que supposer à tort la normalité peut tout de même conduire à des estimations convergentes des paramètres d'un modèle GARCH (principe du pseudo-maximum de vraisemblance), bien que non efficaces. Mais, dans le cas d'une application à la VaR, la spécification de la loi conditionnelle ne concerne pas uniquement le problème de l'estimation des paramètres du modèle GARCH et affecte plus directement la détermination du fractile de la distribution conditionnelle. Le choix à tort d'une spécification normale peut dès lors avoir des conséquences notables sur les estimations et les prévisions de VaR.

Les formules utiles dans une perspective de calcul de la VaR associées à la loi normale centrée réduite sont les suivantes :

$$f = \phi(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right)$$

$$F^{-1}(\alpha) = \phi^{-1}(\alpha) \quad (2.51)$$

$$E|z| = \sqrt{2/\pi} \quad (2.52)$$

$$\gamma_3 = E|(z - E(z))^3| = 0 \quad (2.53)$$

$$\gamma_4 = E|(z - E(z))^4| = 3 \quad (2.54)$$

Où γ_3 et γ_4 désignent respectivement la Skewness et la Kurtosis.

La log-vraisemblance associée à une observation z_t sous l'hypothèse de normalité s'écrit :

$$\log L(\theta) = -\frac{T}{2} \log(2\pi) - \frac{1}{2} \sum_{i=1}^T \log[h_t(\theta)] - \frac{1}{2} \sum_{i=1}^T \frac{[z_t - m_t(\theta)]^2}{h_t(\theta)} \quad (2.55)$$

où $h_t(\theta)$ désigne la variance conditionnelle d'un processus GARCH et $m_t(\theta)$ fait référence à l'espérance conditionnelle.

2.19.2 Loi de Student

La distribution de Student permet de modéliser des queues de distribution plus épaisses que celles de la loi normale (distribution leptokurtique). Plus précisément, la Kurtosis de la distribution de Student est déterminée par le degré liberté ν . Dès lors, dans le cadre des modèles GARCH, ce paramètre estimé permet de capturer l'excès de Kurtosis qui ne peut pas être expliqué par le modèle GARCH lui-même. La distribution de Student standardisée est symétrique et la Skewness est nulle à condition que $\nu > 3$. La distribution est leptokurtique dès lors que $\nu > 4$.

Les formules utiles dans une perspective de calcul de la VaR associées à une loi de Student à ν degrés de liberté sont les suivantes (Bao, Lee et Saltoglu, 2004).

$$f(z; \nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \frac{1}{\sqrt{\pi\nu}} \left(1 + \frac{z^2}{\nu}\right)^{-\frac{\nu+1}{2}}, \nu > 2 \quad (2.56)$$

$$F(z; \nu) = G(z; \nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \frac{1}{\sqrt{\pi\nu}} \int_{-\infty}^x \left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}} dt, \nu > 2 \quad (2.57)$$

$$F^{-1}(\alpha) = \sqrt{\frac{(\nu-2)}{\nu}} G^{-1}(\alpha; \nu) \quad (2.58)$$

$$E|z| = \frac{4\Gamma\left(\frac{1+\nu}{2}\right)\sqrt{\nu-2}}{1 + \sqrt{\pi}\Gamma\left(\frac{\nu}{2}\right)(\nu-1)} \quad (2.59)$$

$$\gamma_3 = 0, \nu > 3 \quad (2.60)$$

$$\gamma_4 = \frac{3\nu^2}{(\nu-2)(\nu-4)}, \nu > 4 \quad (2.61)$$

où γ_3 et γ_4 désignent respectivement la Skewness et la Kurtosis.

La log-vraisemblance associée à une observation z_t sous l'hypothèse de Student s'écrit :

$$\log L(\theta, \varepsilon_t) = \log \left[\Gamma \left(\frac{\nu+1}{2} \right) \right] - \log \left[\Gamma \left(\frac{\nu}{2} \right) \right] \quad (2.62)$$

$$-0.5 \left\{ \left[\log [\pi(\nu-2)] + \log(h_t) + (1+\nu) \log \left(1 + \frac{z_t^2}{\nu-2} \right) \right] \right\} \quad (2.63)$$

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

2.19.3 Loi GED (General Error Distribution)

La distribution à erreurs généralisées (GED) est également symétrique mais est assez flexible dans les queues grâce au paramètre ν . En effet, quand $\nu = 2$, nous tombons sur une distribution normale standard. Lorsque $\nu < 2$, cette distribution a des queues plus épaisses que celle de la normale (distribution leptokurtique). Pour $\nu = 1$, la distribution GED dégénère vers une distribution double exponentielle. Lorsque ce paramètre ν est supérieur à 2, la distribution GED a des queues plus fines que celles de la normale (distribution platykurtique). Enfin, quand ν tend vers l'infini, on obtient une distribution uniforme. L'excès de Kurtosis est donné par $\Gamma(1|\nu)\Gamma(5|\nu)/[\Gamma(3|\nu)]^2$, (Bao, Lee et Saltoglu, 2004).

Les formules utiles dans une perspective de calcul de la VaR associées à une loi GED à ν degrés de liberté sont les suivantes.

$$f(z; \nu) = \frac{\nu \exp(-0.5|z|\lambda^\nu)}{\lambda 2^{(1+1/\nu)} \Gamma(1|\nu)}, \nu > 0 \quad (2.64)$$

$$F(z; \nu) = \begin{cases} \Gamma(1|\nu, \frac{1}{2}(-\frac{z}{\lambda})^\nu), & \text{si } z \leq 0 \\ 2\Gamma(1|\nu) \frac{1}{2} + \frac{\gamma(1|\nu, \frac{1}{2}(\frac{z}{\lambda})^\nu)}{2\Gamma(1|\nu)}, & \text{si } z > 0 \end{cases} \quad (2.65)$$

$$F^{-1}(\alpha) = \begin{cases} -2\lambda\Gamma^{-1}(1|\nu, 2\alpha\Gamma(1|\nu)) & \text{si } \alpha \leq 0.5 \\ 2\lambda\gamma^{-1}(1|\nu, 2(\alpha - \frac{1}{2})\Gamma(1|\nu)) & \text{si } \alpha > 0.5 \end{cases} \quad (2.66)$$

$$E|z| = \lambda 2^{1/\nu} \frac{\Gamma(2|\nu)}{\Gamma(1|\nu)} \quad (2.67)$$

$$\gamma_3 = 0 \quad (2.68)$$

$$\gamma_4 = \frac{\Gamma(1|\nu)\Gamma(5|\nu)}{[\Gamma(3|\nu)]^2} \quad (2.69)$$

γ_3, γ_4 désignent respectivement la SKewness et la Kurtosis et λ une constante telle que :

$$\lambda = \sqrt{2^{(-2|\nu)}\Gamma(1|\nu)/\Gamma(3|\nu)}$$

la log-vraisemblance associée à une observation z_t sous l'hypothèse GED s'écrit :

$$\log L(\theta, z_t) = \log(\nu|\lambda) - 0.5 \left| \frac{z_t}{\lambda} \right|^\nu - (1 + \nu^{-1}) \log(2) - \log \left[\Gamma\left(\frac{1}{\nu}\right) \right] - 0.5 \log(h_t) \quad (2.70)$$

2.19.4 Loi Skewed Student

Bien que la GED permette une flexibilité dans la modélisation de la Kurtosis, elle ne peut pas modéliser la Skewness observée dans un grand nombre de séries financières. Or, négliger la Skewness peut conduire à une inférence biaisée du risque.

Fernandez et Steel (1998) ont été les premiers à proposer la Skewed Student comme une distribution à quatre paramètres où les paramètres régissant la location, la dispersion, la Skewness et la Kurtosis ont une interprétation intuitive.

Les formules utiles dans une perspective de calcul de la VaR associées à une loi de Skewed Student à ν degrés de liberté sont les suivantes.

$$f(z, \varepsilon, \nu) = \begin{cases} \frac{2s}{\varepsilon + \varepsilon^{-1}} \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})} \frac{1}{\sqrt{\pi(\nu-2)}} \left(1 + \frac{\varepsilon^2(sz+m)^2}{\nu-2} \right)^{\frac{\nu+1}{2}} \text{ si } z \leq -\frac{m}{s} \\ \frac{2s}{\varepsilon + \varepsilon^{-1}} \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})} \frac{1}{\sqrt{\pi(\nu-2)}} \left(1 + \frac{\varepsilon^{-2}(sz+m)^2}{\nu-2} \right)^{\frac{\nu+1}{2}} \text{ si } z > -\frac{m}{s} \end{cases} \quad (2.71)$$

$$F(z; \varepsilon, \nu) = \begin{cases} \frac{2}{\varepsilon^2+1} G(\varepsilon(sz+m)\sqrt{\frac{\nu}{\nu-2}}; \nu) \text{ si } z \leq -\frac{m}{s} \\ 1 - \frac{2}{\varepsilon^{-2}+1} G(-\varepsilon^{-1}(sz+m)\sqrt{\frac{\nu}{\nu-2}}; \nu) \text{ si } z > -\frac{m}{s} \end{cases} \quad (2.72)$$

$$F^{-1}(\alpha) = \begin{cases} \frac{\frac{1}{\varepsilon} G^{-1}(\frac{\alpha}{2}(1+\varepsilon^2); \nu) - m}{s} \text{ si } \alpha \leq \frac{1}{1+\varepsilon^2} \\ \frac{-\varepsilon G^{-1}(\frac{1-\alpha}{2}(1+\varepsilon^{-2}); \nu) - m}{s} \text{ si } \alpha > \frac{1}{1+\varepsilon^2} \end{cases} \quad (2.73)$$

$$E|z| = \frac{4\varepsilon^2 \Gamma\left(\frac{1+\nu}{2}\right) \sqrt{\nu-2}}{\varepsilon + \frac{1}{\varepsilon} \sqrt{\pi} \Gamma\left(\frac{\nu}{2}\right) (\nu-1)} \quad (2.74)$$

$$\gamma_3 = \frac{(\nu - 2)^{3/2} (\varepsilon^2 - 1) (\varepsilon^4 + 1) \Gamma[(\nu - 3)|2]}{\sqrt{\pi} \varepsilon^3 s^3 \Gamma(\nu|2)} - \frac{m(m^2 + 3s)}{s^3} \quad (2.75)$$

$$\gamma_4 = \frac{3(\varepsilon^5 - \varepsilon^{-5})(\nu - 2)}{(\varepsilon + \varepsilon^{-1})(\nu - 4)s^4} + \frac{3m^2(m^2 + 2s)}{s^4} - \frac{4m(\nu - 2)^{3/2} (\varepsilon^2 - 1) (\varepsilon^4 + 1) \Gamma[(\nu - 3)|2]}{\sqrt{\pi} \varepsilon^3 s^4 \Gamma(\nu|2)} \quad (2.76)$$

où

$$\varepsilon > 0, \nu > 2, m = \Gamma((\nu - 1)/2) \sqrt{(\nu - 2)/\pi} (\varepsilon - 1/\varepsilon) / \Gamma(\nu/2), s = \sqrt{(\varepsilon^2 + 1/\varepsilon^2 - 1) - m^2} \quad (2.77)$$

γ_3 et γ_4 désignent respectivement la Skewness et la Kurtosis si ν est respectivement supérieur à 3 et 4. $G(\cdot; \nu)$ dénote la fonction de répartition d'une distribution de Student standard. ε est un indicateur d'asymétrie tel que lorsque $\varepsilon = 1$, la distribution Skewed Student est égale à la distribution de Student.

La log-vraisemblance associée à une observation z_t sous l'hypothèse Skewed Student s'écrit :

$$\begin{aligned} \log(\theta, z_t) = & T \left[\log \Gamma\left(\frac{\nu + 1}{2}\right) - \log\left(\frac{\nu}{2}\right) - \frac{1}{2} \log(\pi(\nu - 2)) + \log\left[\frac{2}{\varepsilon + \frac{1}{\varepsilon}}\right] + \log(s) \right] \\ & - \frac{1}{2} \sum_{i=1}^T \left[\log(\partial_t^2) + (1 + \nu) \log\left[1 + \frac{(sz_t + m)^2}{(\nu - 2)} \varepsilon^{-2t}\right] \right] \end{aligned} \quad (2.78)$$

2.19.5 Loi Skewed GED

Theodossiou (2000) donne la version Skewed de la GED avec deux paramètres contrôlant respectivement la Skewness et la Kurtosis. La Skewed GED était utilisé pour examiner la distribution empirique des rentabilités de plusieurs actifs financiers à des fréquences différentes et les résultats ont montré qu'elle est appropriée (Bao, Lee et Saltogu, 2004).

Les formules utiles dans une perspective de calcul de la VaR associées à une loi de Skewed GED à ν degrés de liberté sont les suivantes.

$$f(z; k, \lambda) = \begin{cases} c \exp(-w_1 |z - \delta|^k) & \text{si } z \leq \delta \\ c \exp(-w_2 |z - \delta|^k) & \text{si } z > \delta \end{cases} \quad (2.79)$$

$$F(z; k, \lambda) = \begin{cases} \frac{(1-\lambda)\Gamma(1/k, w_1(\delta-z)^k)}{2\Gamma(1/k)} & \text{si } z \leq \delta \\ \frac{1-\lambda}{2} + \frac{(1+\lambda)\gamma(1/k, w_2(z-\delta)^k)}{2\Gamma(1/k)} & \text{si } z > \delta \end{cases} \quad (2.80)$$

$$F^{-1}(\alpha) = \begin{cases} \delta - \frac{1-\lambda}{\phi} \left[\Gamma^{-1} \left(\frac{1}{k}, \frac{2\alpha\Gamma(1|k)}{1-\lambda} \right) \right]^{1|k} & \text{si } \alpha \leq \frac{1-\lambda}{2} \\ \delta + \frac{1+\lambda}{\phi} \left[\gamma^{-1} \left(\frac{1}{k}, \frac{2(\alpha - \frac{1-\lambda}{2})\Gamma(1|k)}{1+\lambda} \right) \right] & \text{si } \alpha > \frac{1-\lambda}{2} \end{cases} \quad (2.81)$$

$$E|z| = \begin{cases} c(-A^1 - B^1 + C^1) & \text{si } \delta \leq 0 \\ c(-A^2 - B^2 + C^2) & \text{si } \delta > 0 \end{cases} \quad (2.82)$$

$$\gamma^3 = \frac{4\lambda(1 + \lambda^2)\Gamma(4|k)}{\phi^3\Gamma(1|k)} + 3\delta + \delta^3 \quad (2.83)$$

$$\gamma^4 = \frac{(1 + 10\lambda^2 + 5\lambda^4)\Gamma(5|k)}{\phi\Gamma(1|k)} + \frac{16\lambda\delta(1 + \lambda^2)\Gamma(4|k)}{\phi^3\Gamma(1|k)} + 6\delta^2 + 3\delta^4 - 3 \quad (2.84)$$

où $k > 0$, $1 > \lambda > -1$, et où γ_3 et γ_4 désignent respectivement la Skewness et la Kurtosis.

$$c = \frac{k\phi}{2\Gamma(1|k)}, \phi = s\sqrt{\frac{\Gamma(3|k)}{\Gamma(1|k)}}, \delta = -2\lambda A s^{-1} \quad (2.85)$$

$$s = \sqrt{1 + 3\lambda^2 - 4A^2\lambda^2}, A = \frac{\Gamma(2|k)}{\sqrt{T(1|k)\Gamma(3|k)}} \quad (2.86)$$

Le paramètre k contrôle l'épaisseur des queues tandis que λ modélise l'asymétrie quand il est différent de 0.

Par ailleurs, il est possible de retrouver plusieurs densités connues comme des cas particuliers. Ainsi, lorsque $k = 2$ et $\lambda = 0$, nous obtenons la densité de la loi Normal standard. Pour $k = 2$ et $\lambda \neq 0$, nous obtenons la Skewed normal d' Azzalini (1985). Enfin pour lorsque le paramètre $\lambda = 0$, nous retombons sur une distribution GED (Cappuccio, Lubian et Raggi, 2004)

Chapitre 3

La Modélisation du problème d'optimisation du portefeuille

3.1 Introduction

Dans un cadre statique d'étude, notre objectif sera principalement de déterminer deux éléments de la gestion de portefeuille. Dans un premier temps, on cherchera les proportions optimales de chaque actif risqué dans le portefeuille. Dans un deuxième temps, on déterminera le montant B de liquidité à prêter ou à emprunter de façon à constituer un portefeuille avec une VaR qui correspond à la VaR préfixé par l'investisseur. Ce niveau de VaR^* reflètera le degré de l'aversion au risque de l'investisseur. Bien évidemment, le portefeuille constitué maximise l'espérance de rendement que peut obtenir l'investisseur sous la contrainte d'une VaR recherché. Dans ce qui suit, l'objectif est de présenter un modèle traduisant la problématique.

On suppose que l'on dispose d'un montant $W(0)$ à investir sur un horizon de temps T . On rappelle qu'on cherche à investir de manière à avoir un niveau de VaR bien définie de notre portefeuille. Ce niveau peut être fixé par le gestionnaire de risque dans les institutions financières de sorte qu'il correspond aux exigences des autorités réglementaires, ou fixé par un investisseur particulier relativement à son degré d'aversion au risque. Ce montant peut être investie avec un autre montant B qui représente un prêt si $B < 0$ et un emprunt si $B > 0$. r_f est le taux d'intérêt sans risque pour lequel l'investisseur peut prêter ou emprunter pendant la période T . On a n actifs disponibles sur le marché. $\gamma(i)$ indique la fraction investie dans l'actif risqué i ainsi la somme des $\gamma(i)$ doit être égale à 1. Soit aussi $P(i, t)$ le prix de l'actif i au temps t (le présent correspond à $t = 0$).

La valeur initiale du portefeuille représente la contrainte budgétaire :

$$w(0) + \beta + \sum_{I=1}^n \gamma(i)P(i, t) \tag{3.1}$$

Le problème fondamental sera ainsi de déterminer les fractions $\gamma(i)$ ainsi que le montant initial β à emprunter ou à prêter.

En choisissant le niveau désiré de la VaR comme VaR^* (exprimé en valeur absolue), on peut formuler la contrainte de perte potentielle de valeur comme suit :

$$pr(w(0) - w(T) \geq VaR^*) \leq 1 - c \quad (3.2)$$

Avec $W(T)$ est la richesse final de l'investisseur compte tenu de son remboursement de l'emprunt ou le cas éventuel de son recouvrement du prêt avec les intérêts y associés, c est le niveau de confiance. Ceci donne :

$$pr(w(T) \leq w(0) - VaR^*) \leq 1 - c \quad (3.3)$$

Du fait que la VaR est la perte maximale, sur l'horizon de temps T , qui peut avoir lieu avec un niveau de confiance c , on constate que le degré d'aversion au risque de l'investisseur est reflété à la fois par le niveau de VaR désiré et par le niveau de confiance associé. L'investisseur est intéressé par la maximisation de la richesse à la fin de la période T . Soit $r(p)$ le rendement total espéré sur le portefeuille p sur cette période. La richesse finale espérée de l'investissement dans le portefeuille p peut s'écrire :

$$E_0(w(T)) = (w(0) + \beta)(1 + r_p) - \beta(1 + r_f) \quad (3.4)$$

3.2 Résolution du problème d'optimisation du portefeuille

A partir de l'équation (3.1), on détermine l'expression de B :

$$\beta = \sum_{I=1}^n \gamma(i)P(i, 0) - w(0) \quad (3.5)$$

Si on remplace cette expression dans l'équation (3.4), on obtient :

$$E_0(w(T)) = (w(0) + (1 + r_f) + \sum_{I=1}^n \gamma(i)P(i, 0)(r_p - r_f)) \quad (3.6)$$

On suppose pour simplifier que $E_0(w(T)) = w(T)$, la valeur de r_p est ainsi donné par :

$$r_p = \frac{(w(T) - w(0)) - w(0)r_f}{\sum_{I=1}^n \gamma(i)P(i, 0)} + r_f \quad (3.7)$$

A partir de la contrainte de la perte de la valeur de l'équation (3.2), on essaiera d'introduire r_p dans l'inéquation. On a ainsi :

$$Pr(w(0) - w(T) \geq VaR^*) \leq 1 - c \quad (3.8)$$

Donc :

$$Pr(w(0) - w(T) \leq -VaR^*) \leq 1 - c$$

Donc :

$$Pr(r_p \leq r_f - \frac{VaR^* + w(0)r_f}{\sum_{I=1}^n \gamma(i)P(i,0)}) \leq 1 - c \quad (3.9)$$

Introduisons maintenant le terme $q(c, p)$ qui représente le quantile correspondant à un niveau de confiance c dans la distribution des rendements du portefeuille. En effet, à partir de la dernière équation on peut obtenir les deux résultats souhaités : d'une part l'expression de la valeur espéré de la richesse finale en fonction du quantile $q(c, p)$ et d'autre part l'expression de β .

Ceci passe par les étapes suivantes :

$$q(c, p) = r_f - \frac{VaR^* + w(0)r_f}{\sum_{I=1}^n \gamma(i)P(i,0)} \quad (3.10)$$

ce qui donne :

$$\sum_{I=1}^n \gamma(i)P(i,0) = \frac{VaR^* + w(0)r_f}{r_f - q(c, p)} \quad (3.11)$$

Une fois remplacé dans la dernière équation exprimant $E_0(W(T, p))$, on obtient :

$$E_0(W(T)) = W(0)(1 + r_f) + \frac{r_p - r_f}{r_f - q(c, p)}(VaR^* + W(0)r_f)$$

En divisant par $W(0)$ on obtient :

$$E_0\left(\frac{W(T)}{W(0)}\right) = (1 + r_f) + \frac{r_p - r_f}{W(0)r_f - W(0)q(c, p)}(VaR^* + w(0)r_f) \quad (6) \quad (3.12)$$

Cette dernière équation implique que la maximisation de l'espérance de rendement de l'investisseur passe à travers la maximisation de l'expression $M(p)$ suivante :

$$M(p) = \frac{r_p - r_f}{W(0)r_f - W(0)q(c, p)} \quad (3.13)$$

On constate que la richesse initiale $W(0)$ n'affecte pas le choix du portefeuille optimal puisque elle est considérée comme une constante dans l'expression $M(p)$ à maximiser. Le processus d'allocation d'actif est ainsi indépendant de la richesse. Cependant, l'avantage d'avoir la richesse initiale dans le dénominateur est son interprétation. En

effet, $M(p)$ est égale au ratio de prime de risque espéré du portefeuille par rapport au risque assumé. Ce dernier est reflété à travers une perte potentielle maximale relativement à une référence (le rendement au taux sans risque). Vu que le produit du quantile négatif par la richesse initiale constitue la VaR du portefeuille pour un niveau donné de confiance, on pourra trouver une nouvelle expression $\varphi(c, p)$ pour le risque.

En notons $VaR(c, p)$ la VaR du portefeuille (avec un signe négatif vu que $q(c, p)$ est un quantile négatif), le dénominateur devient :

$$\varphi(c, p) = W(0)r_f - Var(c, p) \quad (3.14)$$

Cette mesure du risque correspond au profil des investisseurs considérant le taux de rendement sans risque comme un benchmark pour le rendement de leur portefeuille et souhaitant en même temps que l'expression du risque soit en terme de perte potentielle. $M(p)$ est ainsi une mesure de performance comme l'indice de Sharpe et peut être utilisé pour évaluer l'efficacité de portefeuille (voir Sharpe (1994)). En plus sous l'hypothèse que l'espérance de rendement du portefeuille est normalement distribuée et que le taux sans risque est nul, $M(p)$ converge vers un multiple de l'indice de Sharpe. Dans ce cas, les portefeuilles pour lesquelles ces deux indices sont maximisés sont les mêmes.

On constate aussi que le portefeuille optimal qui maximise $M(p)$ est choisi indépendamment du niveau de la richesse initial ainsi que du niveau de VaR désiré (VaR^*). En effet, la mesure de risque $\varphi(c, p)$ pour les différents portefeuilles dépend de la VaR estimé du portefeuille et non de celui désiré. Les investisseurs débutent par la détermination de l'allocation optimale entre les actifs risqués, l'intervention ensuite du montant β vient pour montrer la différence entre la VaR estimé du portefeuille et la VaR désiré. Deux étapes séparées caractérisent le processus de décision comme dans le cas de l'approche de Moyenne Variance.

Afin de déterminer la valeur de β on combine l'équation?? et l'équation?? Ceci donne enfin :

$$\beta = \frac{W(0) * (Var^* + Var(c, p'))}{\varphi(c, p')} \quad (3.15)$$

On note que dans cette dernière expression, la Var^* est exprimé en valeur absolue et que la VaR est de signe négative. On remarque aussi le fait que ce modèle est indépendant des hypothèses de distribution de sorte que le modèle est dérivé dans le cadre de la maximisation de l'espérance de rendement sous la contrainte de perte de valeur désiré.

3.3 Le cadre empirique de l'étude

Afin de déterminer l'effet de la déviation de l'hypothèse de normalité et de la variation de l'horizon d'estimation sur les portefeuilles optimaux à construire, nous avons choisi de se référer à quatre indices du marché des capitaux : le CAC 40, le SMI, le DAX et le Le FTSE 100. Ces quatre actifs représentent dans notre travail les actifs risqués. Rappelons que l'indice CAC 40 est le principal indice boursier de la place de Paris et voit le jour le 15 juin 1988.

- L'indice **CAC 40** est déterminé à partir des cours des 40 plus importantes capitalisations boursières françaises, correspondant aux 40 entreprises multinationales françaises les plus importantes cotées en continu et en temps réel à la bourse de Paris. Le CAC 40 fait partie du groupe de cotation Euronext et son code mnémonique est PX1. Depuis le 1er décembre 2003, le CAC 40 a adopté le système de capitalisation boursière flottante pour s'aligner sur le mode de fonctionnement des grands indices mondiaux. Cela veut dire que, depuis cette date, le nombre de titres disponibles à l'achat sur le marché pour une société est pris en compte dans le calcul de l'indice.
- Le **SMI** signifie Swiss Market Index, indice du marché de la Bourse de Zurich. Il s'agit du plus important marché suisse, puisqu'il regroupe 85 % du montant total des transactions. Il a été créé en 1988 sur une base de 1 500 points, et traite les fluctuations des 20 principales capitalisations boursières. Il est géré et publié par le Six Group. À eux seuls, trois groupes comptent pour la moitié de la capitalisation : Nestlé, Novartis, et Hoffmann-La Roche. Du fait de sa représentativité, l'indice SMI est utilisé comme sous-jacent à plusieurs produits financiers. Sa composition est révisée en septembre.
- Le **DAX** ou Deutscher Aktienindex est le principal indice boursier allemand. Sa valeur est fondée sur le cours des actions des 30 plus importantes entreprises cotées à la Bourse de Francfort (Francfort-sur-le-Main). Le cours des actions est tiré du système de trading électronique Xetra. Selon Deutsche Börse, opérateur de Xetra, le DAX mesure la performance des trente plus importantes sociétés allemandes du Prime Standard en matière de volume de transactions et de capitalisation boursière. L'indice est équivalent au FT 30, au Dow Jones Industrial Average ou au CAC 40 et, en raison du nombre limité de titres qu'il contient, ne représente pas nécessairement la vitalité de l'économie allemande dans son ensemble.
- Et Le **FTSE 100** (Financial Times Stock Exchange), surnommé "le Footsie", est un indice-phare de la Bourse de Londres. Il est apparu en 1984, sur une base de 1 000 points, comme ses équivalents européens. Il est constitué des 100 valeurs les plus représentatives et les plus actives du marché. Ces dernières sont aussi

sélectionnées en fonction de l'importance de leur capitalisation. Très suivi par les analystes financiers, le FTSE 100 sert de baromètre pour mesurer la santé de l'économie britannique. Cet indice boursier appartient au FTSE Group.

Le logiciel utilisé pour l'analyse des données relatives à ces quatre indices et l'implémentation du modèle sera le R(4.0.3). On suppose que les coûts de transactions sont négligeables et que les actifs financiers sont divisibles.

La période d'étude sera à partir du 1991 jusqu'au 1998 et donc s'étale sur sept ans. Le nombre des observations des rendements journaliers est de 1860.

Le rendement journalier de l'indice est obtenu par la formule suivante :

$$R_j = \frac{C_j - C_{j-1}}{C_{j-1}}$$

Avec :

- C_j = valeur de l'indice pour le jour j
- C_{j-1} = valeur de l'indice pour le jour $j - 1$

- La richesse initiale de l'investisseur est supposée égale à (1628.75\$, 1678.1\$, 1772.8\$, 2443.6\$) respectivement aux indices boursiers (DAX, SMI, CAC et FTSE). Le horizon de détention considéré est journalière. On se réfère à la moyenne géométrique pour calculer les moyennes de rendement des indices sur les différents horizons. L'expression générale pour l'obtenir est la suivante :

$$R_m = \left[\prod_{i=1}^n (1 + R_i^{(h)}) \right]^{(1/n)} - 1$$

Avec :

- h = durée de la sous période de détention (soit le jour, la semaine ou dix jours)
- n = nombre de sous période de durée h dans la période totale m
- $R_i^{(h)}$ = rendement de la $i^{\text{ème}}$ sous période de durée h
- R_m = rendement sur la période tota

Par exemple, le rendement journalier moyen de l'indice SMI sur la période d'étude est de 0.077% est le plus élevé par rapport à l'indice de DAX, CAC, et FTSE.

L'écart type journalier du rendement de l'indice DAX et CAC40 sont supérieur aux autres indices boursiers par contre le FTSE est le moins élevé.

La volatilité est presque doublée 10 fois puisque celle du deuxième indice (SMI) atteint 0.92%, et le quatrième (FTSE) atteint 0.79% alors que celle du première et du troisième sont de 10% et 11% successivement, la nouvelle mesure de risque présenté

dans la section précédente est calculée pour les différentes de rendement selon une approche empirique d'estimation de la var et à niveau de confiance de 99%. Rappelons qu'elle est obtenue par la formule suivant :

$$\varphi = W_0 r_f - VaR_{estimé}$$

De même, la VaR à un niveau de confiance de 95% dans la table 2 relative de l'indice de CAC40 reste toujours supérieure aux autres indices boursiers pour la même période de détention. Ceci est en conformité avec le fait que le premier indice offre un rendement espéré plus élevé et donc c'est évident qu'il fait supporter l'investisseur plus de risque. et aussi pour la $CVAR$ l'indice de CAC40 reste toujours supérieure aux autres indices boursiers

On constate aussi que pour la fréquences de données, les valeurs du coefficient d'asymétrie (skewness) et du coefficient d'aplatissement (kurtosis) sont différentes de celles données par une distribution normale. Pour le skewness, la seule valeur positive est celle du cas de rendement journalier du FTSE ce qui indique une distribution asymétrique à droite. Dans tous les autres cas, la distribution est asymétrique à gauche. Pour le kurtosis les valeurs relatives aux indices DAX et SMIs s'éloignent de 3. Il est plus élevé que cette valeur ce qui signifie de l'existence des queues épaisses, aussi les valeurs des indices CAC40, DAX et sont moins de 3. Ceci témoigne à priori de la non normalité des distributions.

La visualisation des corrélogramme simple et partiel (figure 9-15) pour chaque indice boursier montre une stationnarité propre, L'ACF est coupée.

On utilise aussi le test de racine unitaire (test ADF) pour confirmer la stationnarité des séries de rendements (table-2). toutes les p-values pour chaque indice boursier < 0.05 donc les séries sont stationnaires.

Notons enfin que les estimations de la Value-at-Risk et le $CVAR$ à partir de la méthode historique et GPD requièrent théoriquement la stationnarité des séries des rendements. Un élément fondamental du modèle présenté dans cette étude est l'estimation de la Value-at-Risk et $CVAR$. Pour cela, nous avons choisies d'appliquer et de comparer les résultats de deux méthodes (historique et GPD). Les résultats obtenus pour les différentes méthodes et pour les différents niveaux de confiance sont donnés dans les tables 3, 4, 5 et 6. On constate que la méthode historique a tendance à donner les valeurs de la VaR les plus élevées (en valeurs absolues) pour le niveau de confiance de 95%. Dans le même cas au niveau de 99%, la méthode GPD donne la valeur la plus élevée de la VaR qui est égale à 0.02761396. Aussi on remarque que la méthode historique donne les valeurs plus élevées de la $CVAR$ pour les différents niveaux de confiance 99% et 95%, aussi on remarque que dans les mêmes niveaux de confiance, les deux méthodes convergent vers les mêmes valeurs estimées pour la $VAR(0.014, 0.017$ et $0.022)$ et pour $CVAR(0.023$ et $0.029)$. Ces résultats s'interprètent comme suit : pour un investisseur qui cherche à constituer un portefeuille maximisant le rendement espéré sous la contrainte d'une VaR et $CVAR$ limite journalière à un niveau de confiance donné et en

se basant sur une estimation empirique et *GPD* de la *Var* et *CVAR*.

- -Pour la *VAR* à un niveau de confiance 95%
il aura intérêt à constituer un portefeuille dont la part de l'indice DAX est 16% , SMI est 14% ,CAC40 est 17% et FTSE est de12%.
- Pour la *VAR* à un niveau de confiance99%
il aura intérêt à constituer un portefeuille dont la part de l'indice DAX est 26% , SMI est 22% ,CAC40 est27% et FTSE est de19% .
- Pour la *CVAR* à un niveau de confiance 95%
il aura intérêt à constituer un portefeuille dont la part de l'indice DAX est22% , SMI est 20% ,CAC40 est 23% et FTSE est de17% .
- Pour la *CVAR* à un niveau de confiance 99%
il aura intérêt à constituer un portefeuille dont la part de l'indice DAX est 33% , SMI est 29% ,CAC40 est 34% et FTSE est de26% .

Dans ce qui suit, l'objectif est de comparer et de valider les méthodes d'estimation de la *Var* dans le modèle d'allocation proposé à savoir : la méthode Empirique, la méthode *GPD*. Ceci est effectué en passant à un cadre dynamique d'étude. Chaque méthode permet d'avoir un modèle de gestion dynamique issu du modèle général proposé.

aussi, le *VAR* des série de rendement de chaque indice convergent vers la même valeur estimée(1.625773).

Nous nous intéressons à la gestion optimale quotidienne sur une période appelée période de prévision (out-of-sample). L'hypothèse de normalité sera donc présenté à travers le modèle dynamique de prévision de la variance conditionnelle : *GARCH* avec des innovations de loi normale.

On se réfère aux mêmes données journalières des sections précédentes sur l'indice DAX, SMI, CAC40 et sur l'indice FTSE. La période totale de l'étude s'étale donc du 1991 jusqu'au 1998 (1860 observations).et la période de prévision comme celle allant du 1999 jusqu'au 2019 , les résultat sont présenté sur la table (3.12)

3.3.1 Test de la stationnarité des quatres indices boursier

L'application du test de Duckey-Feller sur la stationnarité des quatres indices boursiers est donné dans le tableau 3.2

3.3.2 Conclusion

Les limites de la variance comme mesure de risque ont mené les financiers à chercher d'autres mesures plus précises et plus pratiques. Parmi ces mesures, la Value-at-Risk s'est imposé comme un standard et une référence règlementaire incontournable

aussi bien dans les milieux financiers que dans les milieux d'assurance. Dans ce travail, on a montré comment on peut intégrer cette notion dans le processus de sélection optimale de portefeuille dans une logique de contrôle de risque. Ainsi, à partir d'une Value-at-Risk souhaitée (et ainsi d'un niveau de risque assumé) on présente un modèle général permettant l'allocation optimale de la richesse entre les actifs risqués et les actifs non risqués. Ceci passe par la maximisation de l'espérance de rendement du portefeuille sous la contrainte de la *VaR* limite. Ce processus d'optimisation tient compte de l'estimation de la *VaR*. Ainsi, trois modèles sont choisis : le modèle empirique, le modèle *GARCH* et le modèle *GPD*. L'étude empirique se base sur quatre actifs risqués : le DAX, SMI, CAC40 et le FTSE et ce pour une période de sept ans (1991-1998). Les résultats montrent la supériorité du modèle Empirique par rapport le modèle *GPD* et ces derniers remarque est implicable par le *CVAR*. En ce qui concerne l'étape de prévision, nous avons appliqué le processus *GARCH*(1.1) données par logiciel *R*.

TABLE 3.1 – Paramètres de description des quatres indices boursiers

les variables	DAX	SMI	CAC	FTSE
min	-0.1010640	-0.0874386	-0.0786963	-4.227e-02
1stQu	-0.0046920	-0.0037919	-0.0060780	-4.327e-03
Median	0.0004599	0.0008807	0.0000000	7.946e-05
Mean	0.0005984	0.0007743	0.0003759	4.000e-04
Max.	0.0494933	0.0484659	0.0591554	5.294e-02
3rdQu.	0.0063330	0.0060517	0.0070708	5.239e-03
Var	0.01649141	0.01474110	0.01747416	0.01272508
Cvar	0.02254103	0.02019416	0.02365147	0.01744039
SD	0.01032355	0.009268919	0.01103582	0.007949451
Skewness	-0.68101324	-0.73617803	-0.25003168	
Kurtosis	7.124981	6.370026	2.537210	2.540582

TABLE 3.2 – Résultats du tests de la stationnarité des quatres indices

L'indice	DAX	SMI	CAC	FTSE
Dickey-Fuller	-9.9997	-10.769	-11.447	-10.838
Lag order	12	12	12	12
p-value	0.01	0.01	0.01	0.01

TABLE 3.3 – Le montant de la VAR par la méthode empirique pour un niveau de confiance de 0.95 et 0.99

L'indice	DAX	SMI	CAC	FTSE
95%	0.01649141	0.01474110	0.01747416	0.01272508
99%	0.02607200	0.02245677	0.02650777	0.01912951

TABLE 3.4 – Le montant de la CVAR par la méthode empirique pour un niveau de confiance de 0.95 et 0.99

L'indice	DAX	SMI	CAC	FTSE
95%	0.02254103	0.02019416	0.02365147	0.01744039
99%	0.03385807	0.02905632	0.03325433	0.02681036

TABLE 3.5 – Le montant de la VAR par la méthode GPD pour un niveau de confiance de 0.95 et 0.99

L'indice	DAX	SMI	CAC	FTSE
95%	0.0005450111	0.01424072	0.01700661	2.186099e-05
99%	0.0008390274	0.02282645	0.02761396	3.361033e-05

TABLE 3.6 – Table de la VaR à 95% et 99% pour les quatre indices boursiers

L'indice	DAX	SMI	CAC	FTSE
95%	0.007999325	0.01970236	0.02309391	0.006300607
99%	0.008219286	0.02981998	0.03495835	0.006313999

TABLE 3.7 – Estimation des paramètres du modèle Garch(1.1) au niveau de confiance 95% pour l'indice DAX

coefficients	a_0	a_1	b_1
Estimate	5.165e-06	7.038e-02	8.826e-01
Std. Error	8.041e-07	1.169e-02	1.727e-02
t value	6.423	6.021	51.102
Pr(>t)	1.34e-10	1.74e-09	< 2e-16

TABLE 3.8 – Estimation des paramètres du modèle Garch(1.1) au niveau de confiance 95% pour l'indice SMI

coefficients	a_0	a_1	b_1
Estimate	1.245e-05	1.170e-01	7.419e-01
Std. Error	1.441e-06	1.968e-02	3.248e-02
t value	8.637	5.945	22.841
Pr(>t)	< 2e-16	2.77e-09	< 2e-16

TABLE 3.9 – Estimation des paramètres du modèle Garch(1.1) au niveau de confiance 95% pour l'indice CAC

coefficients	a_0	a_1	b_1
Estimate	9.142e-06	5.087e-02	8.743e-01
Std. Error	2.105e-06	9.804e-03	2.584e-02
t value	4.343	5.188	2.12e-07
Pr(>t)	8.743e-01	2.584e-02	33.836

TABLE 3.10 – Estimation des paramètre du modèle Garch(1.1) au niveau de confiance 95% pour l'indice FTSE

coefficients	a ₀	a ₁	b ₁
Estimate	8.419e-07	4.424e-02	9.434e-01
Std. Error	3.009e-07	6.810e-03	1.012e-02
t value	2.798	6.495	93.180
Pr(>t)	0.00515	8.28e-11	< 2e-16

TABLE 3.11 – Test de Jarque Bera pour les quatres indices

Les indices	X-squared	df	p-value
DAX	17289	2	< 2.2e-16
SMI	9180.2	2	2.2e-16
CAC	789.78	2	2.2e-16
FTSE	199.09	2	2.2e-16

TABLE 3.12 – les valeur de la VAR estimé au niveau de confiance 95% par garch(1.1)

Les indices	DAX	SMI	CAC40	FTSE
VAR	1.655542	1.623542	1.589656	1.556358

TABLE 3.13 – Montant de la VAR par modèle GARCH(1.1) pendant la période de pré-
vision pour les quatres indices boursiers

Indice	DAX	SMI	CAC40	FTSE
1999	0,015162056	0.015159091	0.01253	0,011504976
2000	0,014964376	0,014422733	0.01251	0,011475324
2001	0,01477658	0,013766435	0.01249	0,011435788
2002	0,014598668	0,013185256	0.01247	0,011406136
2003	0,014420756	0,012671288	0.01245	0,011376484
2004	0,014252728	0,012218601	0.01243	0,011336948
2005	0,014094584	0,011821264	0.01242	0,011307296
2006	0,01393644	0,011473347	0.01240	0,011277644
2007	0,01378818	0,011169908	0.01238	0,011247992
2008	0,01363992	0,010906006	0.01236	0,011218340
2009	0,013501544	0,010676697	0.01235	0,011188688
2010	0,013363168	0,010479017	0.01233	0,011149152
2011	0,013234676	0,010308024	0.01231	0,011119500
2012	0,013116068	0,010160752	0.01230	0,011089848
2013	0,01299746	0,010034237	0.01228	0,011060196
2014	0,012878852	0,009925513	0.01227	0,011030544
2015	0,012770128	0,009832603	0.01225	0,011000892
2016	0,012661404	0,009753531	0.01224	0,010971240
2017	0,012562564	0,009685332	0.01222	0,010951472
2018	0,01246372	0,009628004	0.01220	0,010921820

Bibliographie

- [1] Acerbi C., Tasche D. (2001) Expected Shortfall: a Natural Coherent Alternative to Value at Risk. Working Paper.
- [2] Acerbi C., Tasche D. (2002) On the Coherence of Expected Shortfall. *Journal of Banking and Finance*, vol. 26(7), 1487-1503.
- [3] Alexander G., Baptista A. (2001) A VaR-Constrained Mean-Variance Model: implications for Portfolio Selection and the Basle Capital Accord. Working Paper, University of Minnesota.
- [4] Artzner P., Delbaen F., Eber J.M., Heath D. (2000) Coherent measures of risk, *Mathematical Finance*, Vol 9 (3), 203-228.
- [5] Aussenegg W., Pichler S. (1997) Empirical Evaluation of Simple Models to Calculate Value-at-Risk of Fixed Income Instruments. Working Paper, Technische Universität Wien.
- [6] Black F., Litterman R. (1992) Global Portfolio Optimization. *Financial Analysts Journal*, vol. 48(5) 28-43.
- [7] Bouchaud J.P., Potters M. (2000) *Theory of Financial Risks*. Cambridge University Press.
- [8] Bredin D., Hyde S., (2001) FOREX Risk: Measurement and Evaluation using Value-at-Risk. Central Bank of Ireland, November.
- [9] Calvet A. L. (2000) La gestion globale des risques du marché : de la Value at Risk à Corporate Metrics™. *Revue Internationale de Gestion*, volume 25, numéro 3, automne.
- [10] Campbell R., Huisman R., Koedijk K. (2001) Optimal Portfolio Selection in a Value-at-Risk Framework. *Journal of Banking and Finance*, 25, 1789–1804.
- [11] Chabaane A., Laurent J.P., Malevergne Y., Turpin F. (2006) Alternative Risk Measures for alternative Investments. *Journal of Risk* 8, n°4, 1-32.
- [12] Christoffersen P., Hahn J., Inoue A. (2001) Testing and comparing Value-at-Risk measures. *Journal of Empirical Finance*, 8, Number 3, July, pp. 325-342(18).
- [13] Crnkovic C., Drachman J. (1996) A Universal Tool to Discriminate Among Risk Measurement Techniques. Working Paper, J. P. Morgan, corporation.

- [14] Denneberg D. (1990) Premium Calculation: Why Standard Deviation Should Be Replaced By Absolute Deviation. *ASTIN Bulletin*, vol. 20, 181-190.
- [15] De Roon F., Nijman T., Werker B. (2003) Currency Hedging for International Stock Portfolios: the Usefulness of Mean-Variance Analysis. *Journal of Banking and Finance*, 27, 327–349.
- [16] Engle R. (1982) Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, 50, p 987-1007, 1982.
- [17] Engle R. (2000) The Use of ARCH/GARCH Models in Applied Econometrics. *Journal of Economic Perspectives*, 15(4), 157-168.
- [18] Fischer T. (2001) Coherent Risk Measures Depending on Higher Moments, Working Paper.
- [19] Fleming J., Kirby C., Ostdiek B. (2001) The Economic Value of Volatility Timing. *Journal of Finance*, (1), 329–352.
- [20] Gaivoronski A., Pflug G. (1999) Finding Optimal Portfolios with Constraints on Value at Risk. Working paper, University of Vienna.
- [21] Glasserman P., Heidelberger P., Shahabuddin P. (2001) Efficient Monte Carlo Methods for Value-at-Risk. *Mastering Risk*, Vol 2, published by Financial Times-Prentice Hall.
- [22] Glasserman P., Shahabuddin P., Heidelberger P. (2000) Variance Reduction Technics for Estimating Value at Risk. *Management Sciences*, Vol.46 No.10, October.
- [23] Hendricks D. (1996) Evaluation of Value at Risk Models Using Historical data. Federal Reserve Bank of New York. *Economic Policy Review* 2, pp 39-70, Avril.
- [24] Jackson P., Maude D., Perraudin W. (1997) Bank Capital and Value-at-Risk. *Journal of Derivatives* 4, pp73-90, springer
- [25] Jorion P. (1985) International Portfolio Diversification with Estimation Risk. *Journal of Busines.* vol. 58,no. 3.
- [26] Jorion P. (2001) *Value at Risk: The New Benchmark for Managing Financial Risk*. second edition, McGraw- Hill, New York.
- [27] Kahneman D., Knetsch J.L., Thaler R.H. (1990) Experimental tests of the endowment effect and the coax theorem. *Journal of public Economics*, 98, 1325-1350.
- [28] Kaplanski G., Kroll Y. (2002) VaR risk measure vs. traditional risk measures: An analysis and survey . *Journal of Risk*, 4, Printemps , 1-27.

- [29] Konno H., Waki H., Yuuki A. (2002) Portfolio Optimization under Lower Partial Risk Measures. Working paper, Kyoto University.
- [30] Krokmal P., Uryasev S., Zrazhevsky G. (2002) Risk Management for Hedge Fund Portfolios. *The Journal of Alternative Investments*, vol. 5(1), 10-29.
- [31] Kupiec P.H. (1995) Techniques for verifying the Accuracy of Risk Measurement Models. Board of Governors of the Federal Reserve System. Finance and Economics Discussion Series, 95/24 May.
- [32] Leibowitz M.L., Kogelman S. (1991) Asset Allocation under Shortfall Constraints. *Journal of Portfolio Management*, Winter, 18-23.
- [33] Linsmeier T. J., Neil D. P. (1996) Risk Measurement: An Introduction to Value at Risk. ACE OFOR 9604, University of Illinois at Urbana-Champaign Juillet.
- [34] Ljung G.M., Box G.E.P. (1978) On a measure of a lack of fit in Time Series Models. *Biometrika*, 65, p 297-303, 1978.
- [35] Lopez J. A. (1996) Regulatory Evaluation of Value at Risk Models. Working Paper, Federal Reserve Bank of New York.
- [36] Lucas A., Klaassen P. (1998) Extreme Returns, Downside Risk, and Optimal Asset Allocation. *Journal of Portfolio Management*, Fall, 71-79.
- [37] Manganelli,S., Engel R.F. (2001) Value at Risk models in finance. Working Paper n°75.
- [38] Markowitz H. (1952) Portfolio Selection. *Journal of Finance*, vol. 7(1), 77-91.
- [39] Marshall C., Siegel M. (1997) Value at Risk: Implementing a Risk Management Standard. *Journal of Derivatives* 4, .p73-84, spring.
- [40] Pflug G. (2000) Some remarks on the Value-at-Risk and the conditional Value-at-Risk in Probabilistic Constrained Optimization: Methodology and Applications, Kluwer Academic Publishers. S. Uryasev (Ed.), Dordrecht.
- [41] Powell A., Balzarotti V. (1996) Capital requirement for Latin American Banks in Relation to their Market Risks: The Relevance of the Basle 1996 Amendment to Latin America. Working Paper, series 347, Washington, D.C.
- [42] Pritsker M. (1997) Evaluating Value at Risk Methodologies: Accuracy versus Computational Time. *Journal of Financial Services Research*, 12 (2/3), October/ December, 201-242.
- [43] Rengifo E.W., Rombouts J.V.K. (2004) Optimal portfolio selection in a VaR framework. Working Paper.

- [44] Rockafellar R.T., Uryasev S. (2002) Conditional Value-at-Risk for General Loss Distributions. *Journal of Banking and Finance*, vol. 26(7), 1443-1471.
- [45] Roy A.D. (1952) Safety-First and the Holding of Assets. *Econometrica*, 20, 431-449.
- [46] Rockafellar R.T., Uryasev S. (2000) Optimization of Conditional Value-at-Risk. *The Journal of Risk*, vol 2(3), 21-41.
- [47] Sharpe W. (1994) The Sharpe Ratio. *Journal of Portfolio Management*, 21, 49-58.
- [48] Tobin J. (1958) Liquidity Preference as Behavior Towards Risk. *Review of Economic Studies*, vol. 67, pp 65-86.
- [49] ARTZNER, P., DELBAEN, F., EBER, J-M ET HEATH, D., "Coherent Mesures of Risk", *Mathematical Finance*, Vol 9, No 3, Juillet 1999, 203-228.
- [50] ARTZNER, P., DELBAEN, F., EBER, J-M. ET HEATH, D., "A Characterization of Measures of Risk", *Cornell University Operations Research and Industrial Engineering*, Février 1997
- [51] SZEGO, "Measures of Risk", *Journal of Banking & Finance*, Vol 26, 2002, 1253- 1272
- [52] Bali, T.G., Theodossiou, P., 2007. A conditional-SGT-VaR approach with alternative GARCH models. *Ann. Oper. Res.* 151 (1), 241–267
- [53] Acerbi C., Tasche D. (2002) On the Coherence of Expected Shortfall. *Journal of Banking and Finance*, vol. 26(7), 1487-150
- [54] Bollerslev, Tim, 1986, "A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return", *Review of Economics and Statistics*, 69, 542-547
- [55] Gouriéroux C. (1997). ARCH models and financial applications. Springer Series in Statistics. Springer- Verlag, New York.
- [56] NELSON D.B. et D.P. FOSTER, 1991a, Filtering and Forecasting with Misspecified Arch models II: Making the Right Forecast with the Wrong Model, D.P. University of Chicago
- [57] Gouriéroux, C. et Monfort, A. (1989). *Statistique et modèles économétriques*, Economica, Paris
- [58] Engle R. F. et V. K. Ng (1993), « Measuring and Testing the Impact of News on Volatility », *Journal of Finance* 48, 1749-1778.
- [59] Sentana, E., 1995. Quadratic ARCH models. *Review of Economic Studies* 62, 639}661. Taylor, S., 1986. *Modelling Financial Time Series*. Wiley, New York.

- [62] Engle, R. F., D. M. Lilien and R. P. Robins, 'Estimating Time Varying Risk Premia in the Term Structure: The ARCH-M Model', *Econometrica*, 55, 391- 407, (1987).
- [63] Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues. *Quantitative Finance*, 1 , 223–236. URL: <http://www.tandfonline.com/doi/abs/10.1080/713665670>. doi:10.1080/713665670.
- [64] Fernandez, C. and Steel, M.F.J. (1998) On Bayesian modeling of fat tails and skewness. *Journal of the American Statistical Association* 93, 359-371.
- [65] Theodossiou, P., 2000. Skewed generalized error distribution of financial assets and option pricing. SSRN Working Paper.
- [66] Cappuccio, N., D. Lubian, and R. Raggi (2004). MCMC Bayesian estimation of a skew-GED stochastic volatility model. *Studies in Nonlinear Dynamics & Econometrics* 8(2).
- [67] Dowd, K. (2005). *Measuring market risk*. John Wiley & Sons, Lt
- [68] Giot et Laurent (2003) Value –at –Risk for long and short trading positions. *Journal of applied Econometrics* 18-641-663-663.

Annexes

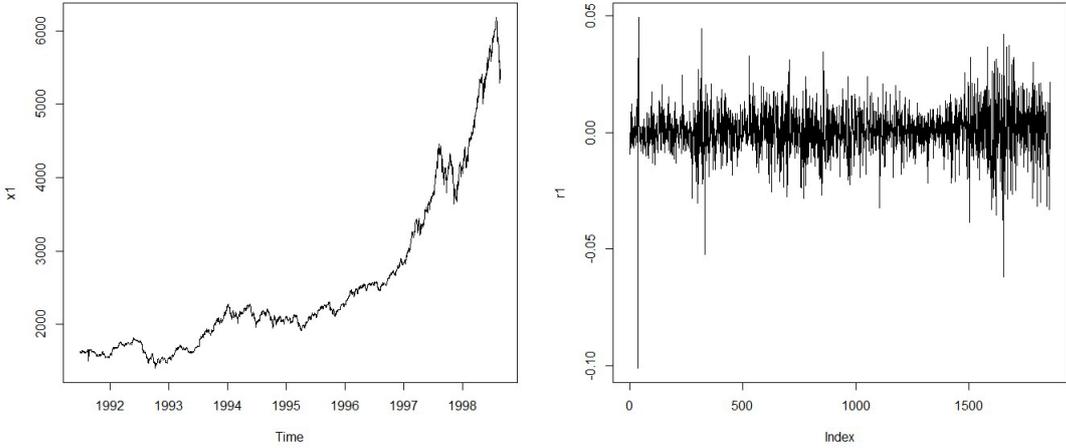


Figure 3.1 – Représentation graphique de la série de rendement de l'indice DAX.

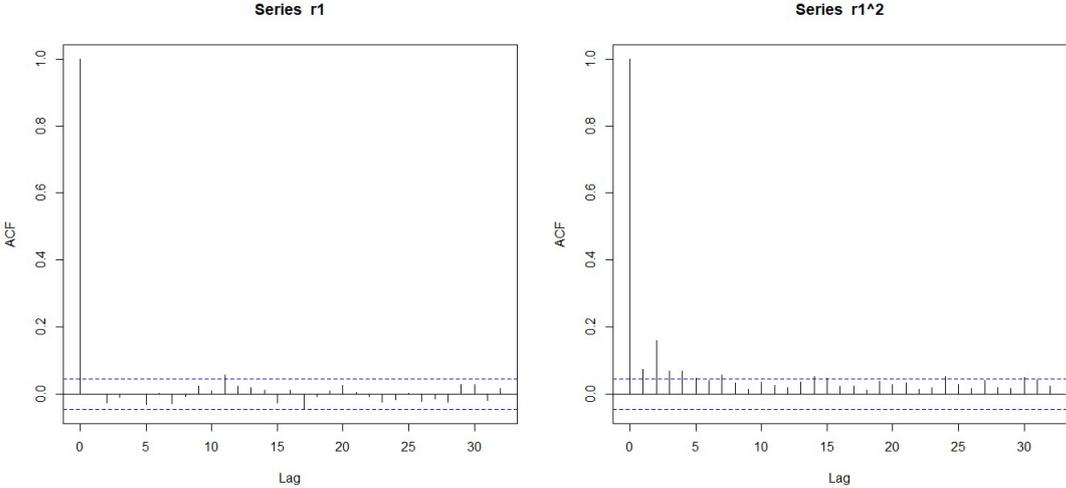


Figure 3.2 – Représentation graphique de l'autocorrélation simple et la corrélogramme partiel de la série de rendement de l'indice DAX.

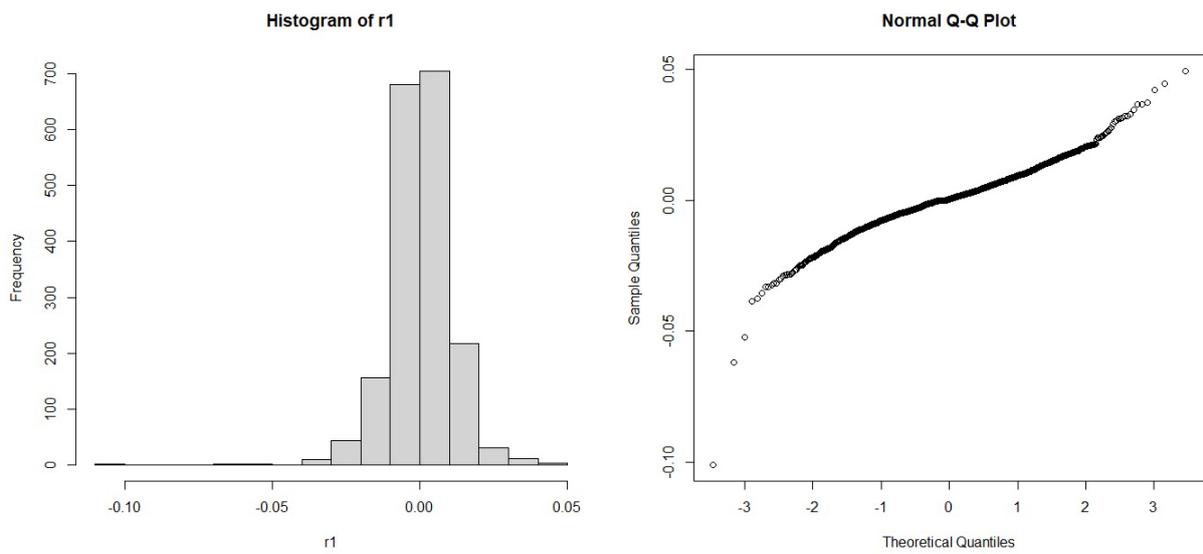


Figure 3.3 – Histogramme et Normal Q-Q-Plot des rendements journaliers de l'indice DAX.

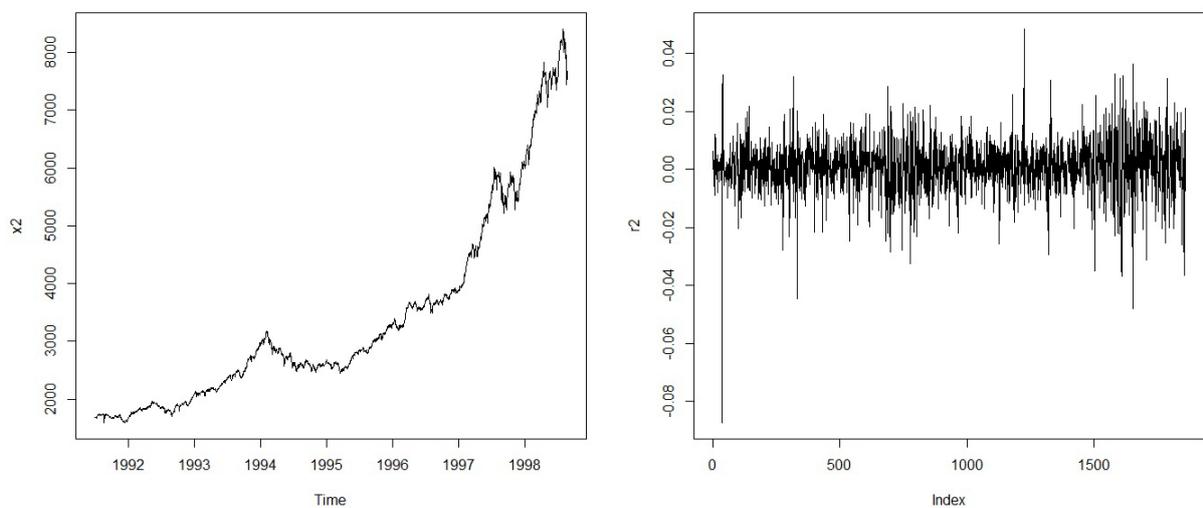


Figure 3.4 – Représentation graphique de la série de rendement de l'indice SMI.

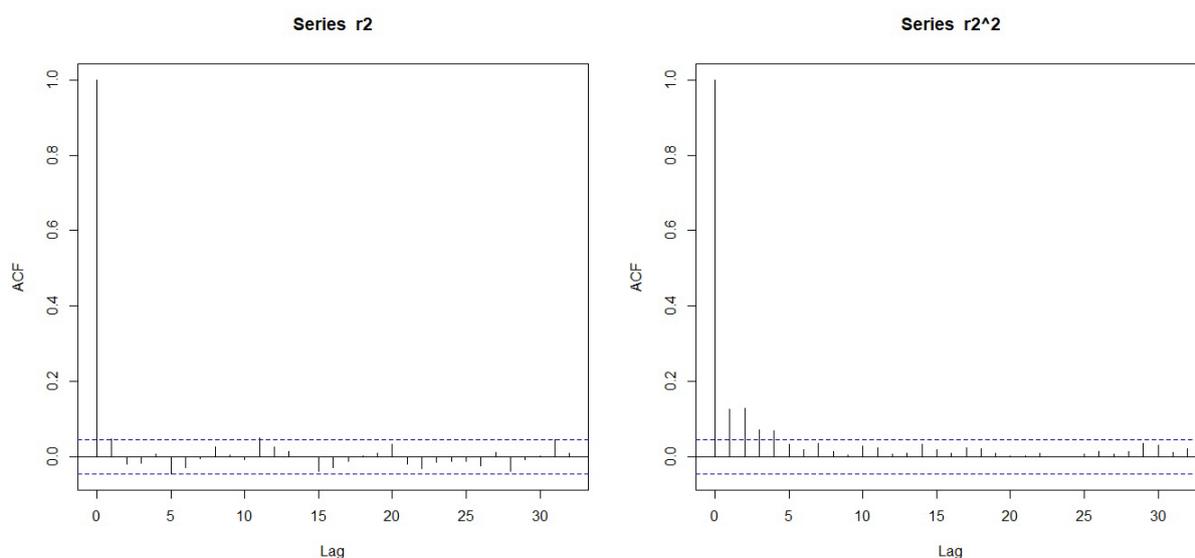


Figure 3.5 – Représentation graphique de l'autocorrélation simple et la corrélogramme partiel de la série de rendement de l'indice SMI.

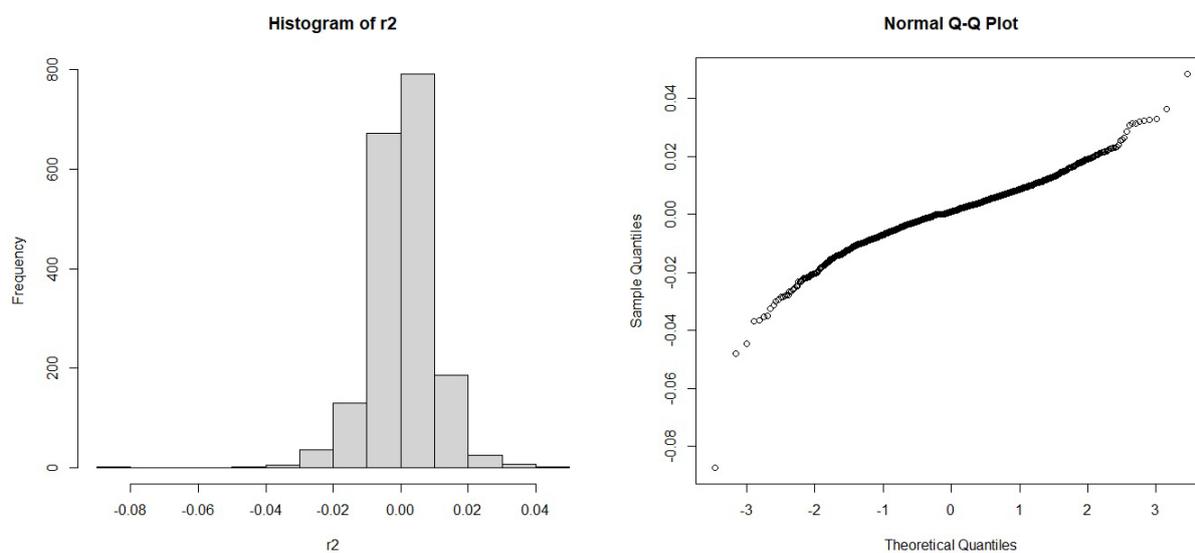


Figure 3.6 – Histogramme et Q-QPlot des rendements journaliers de l'indice SMI.

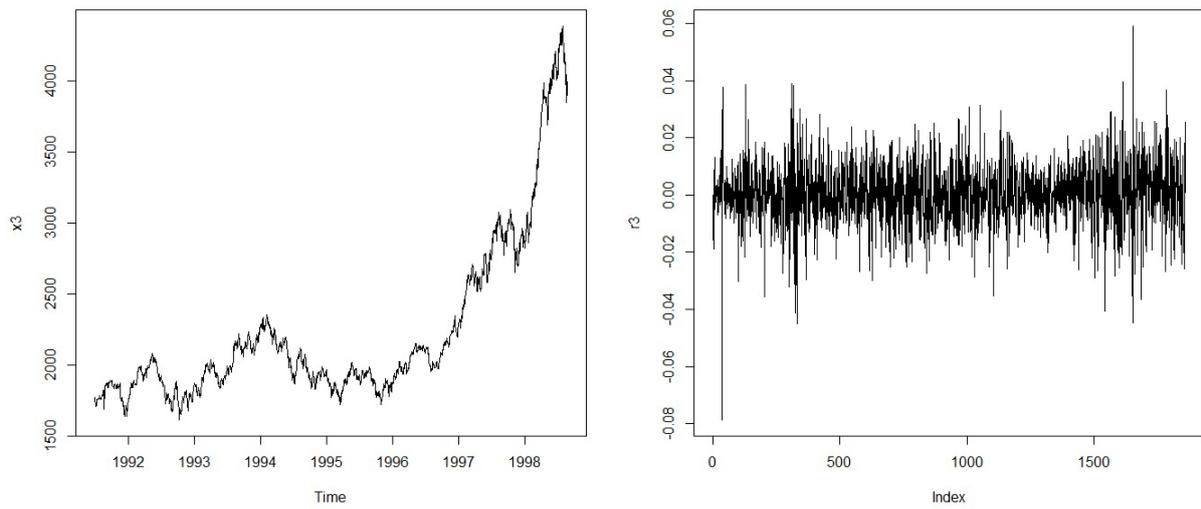


Figure 3.7 – Représentation graphique de la série de rendement de la série de rendement de l'indice CAC40.

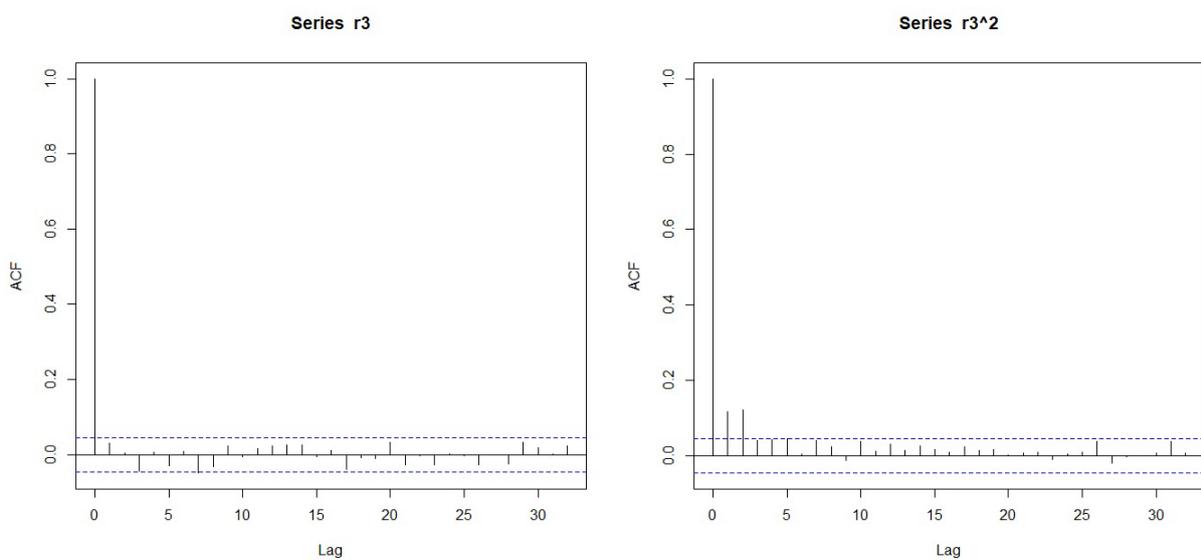


Figure 3.8 – Représentation graphique de l'autocorrélation simple et la corrélogramme de la série de rendement de l'indice CAC40.

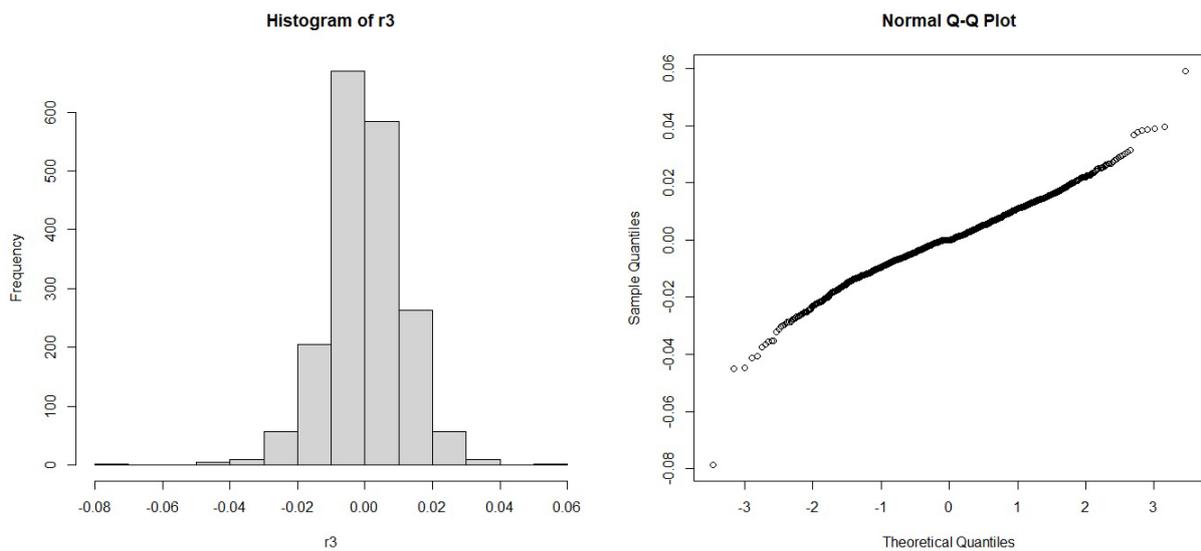


Figure 3.9 – Histogramme et normale Q-QPlot des rendements journaliers de l'indice CAC40.

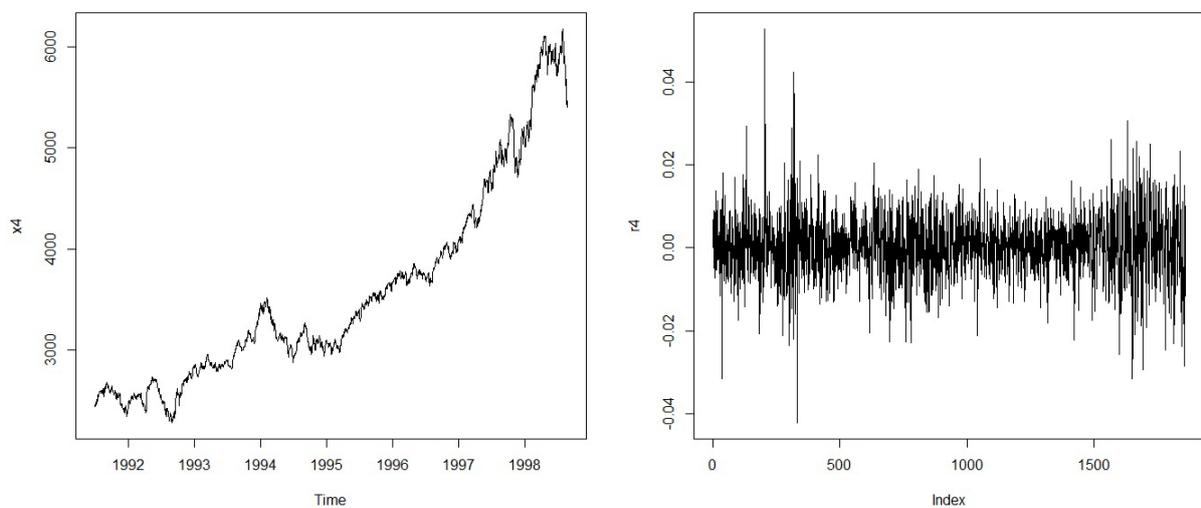


Figure 3.10 – Représentation graphique de la série de rendement de l'indice FTSE.

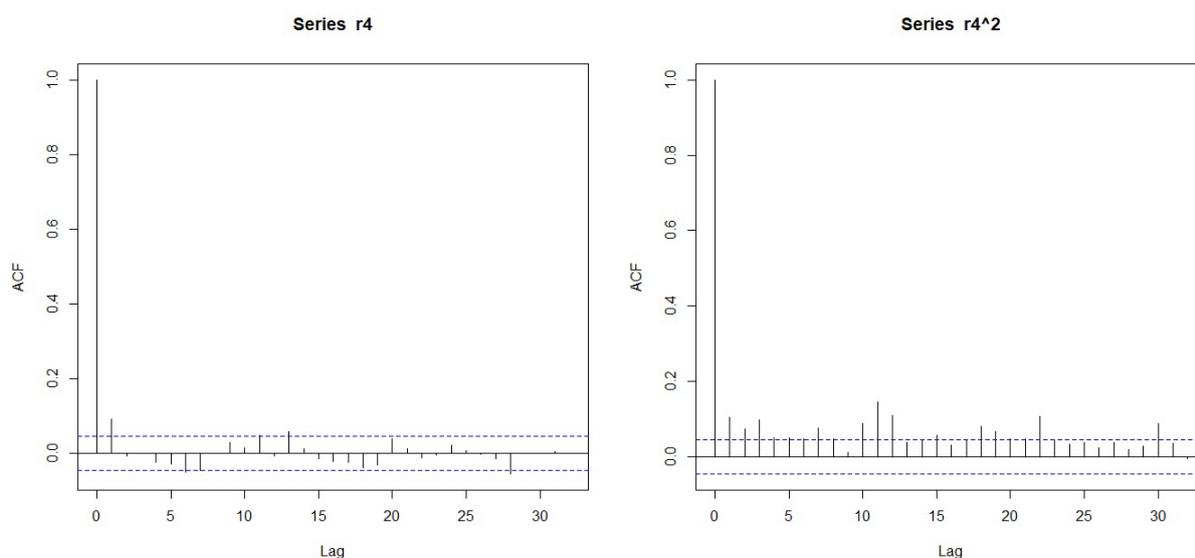


Figure 3.11 – Représentation graphique de l'autocorrélation simple et la corrélogramme partiel de la série de rendement de l'indice FTSE.

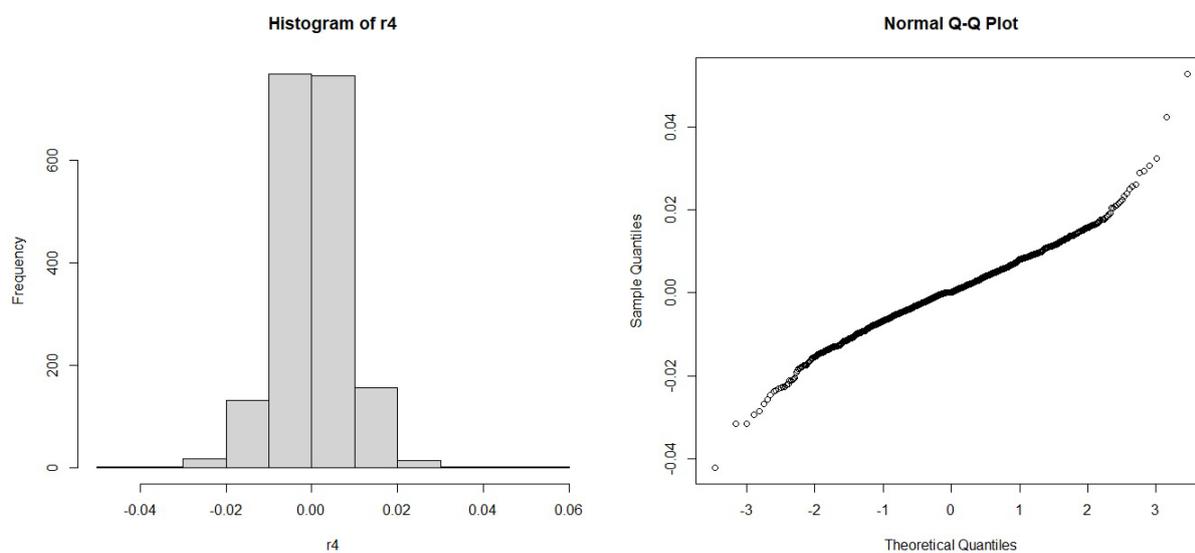


Figure 3.12 – Histogramme et Q-QPlot des rendements journaliers de l'indice FTSE.

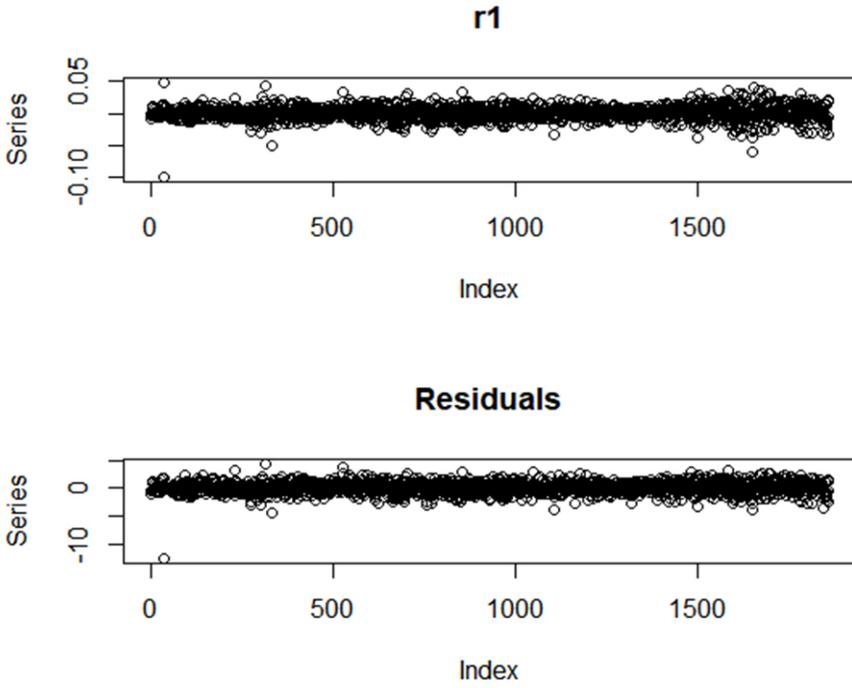


FIGURE 3.13 – Représentation graphique de la série du rendement et du résidu de l'indices DAX PAR GARCH(1.1).

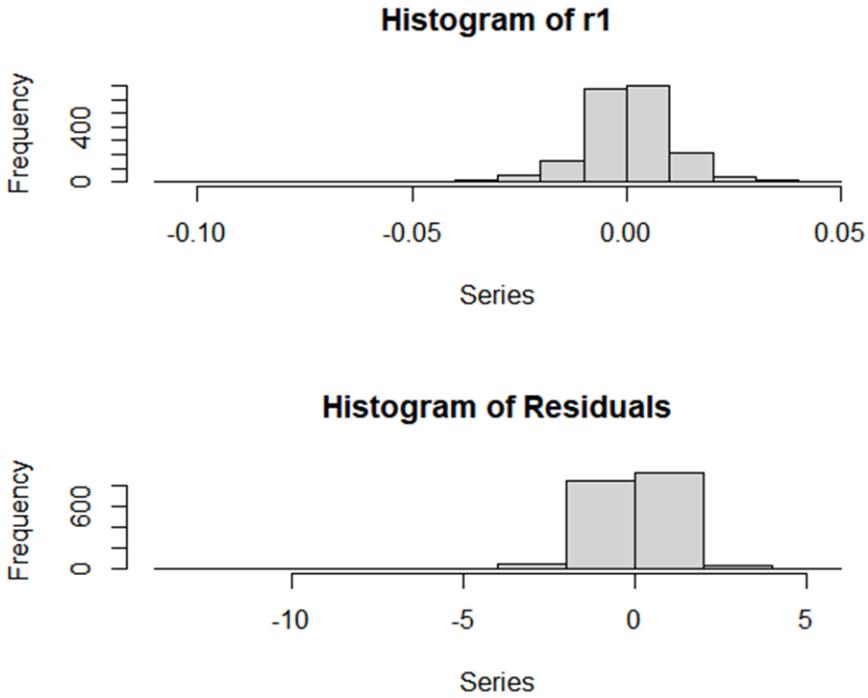


FIGURE 3.14 – Histogramme de rendement et du résidus de l'indice DAX par GARCH(1.1).

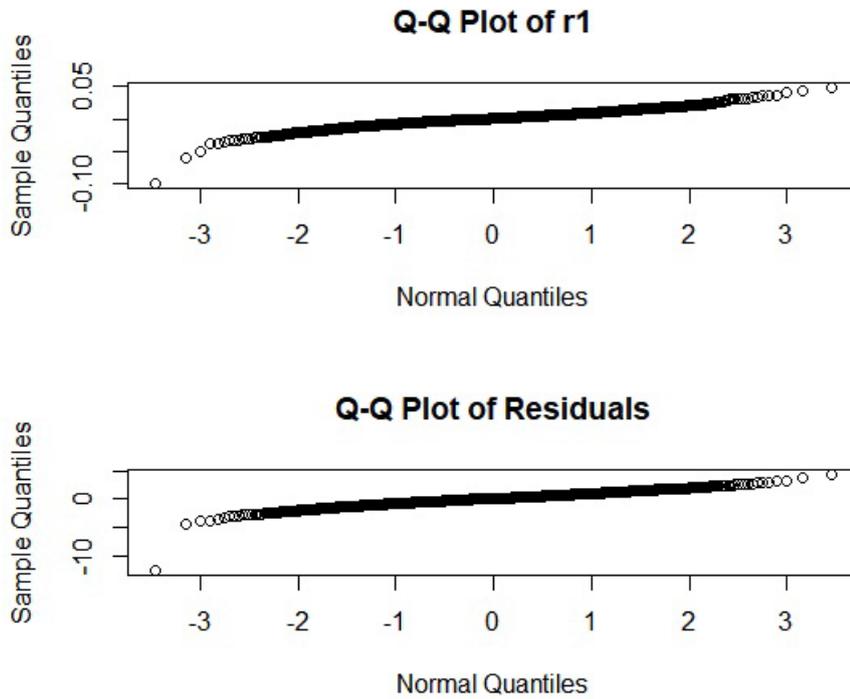


Figure 3.15 – Représentation graphique du Q-QPlot de la série de rendement de l'indice DAX et Q-QPlot du résidu par garch(1,1).

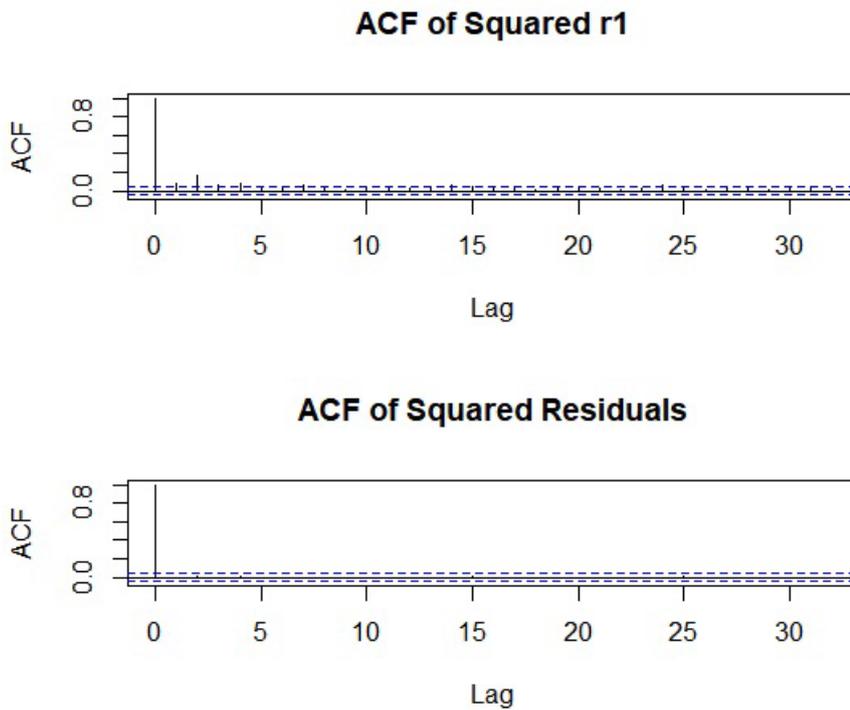


Figure 3.16 – Représentation graphique de l'autocorrelation simple de la série de rendement de DAX et du résidu par garch(1,1).

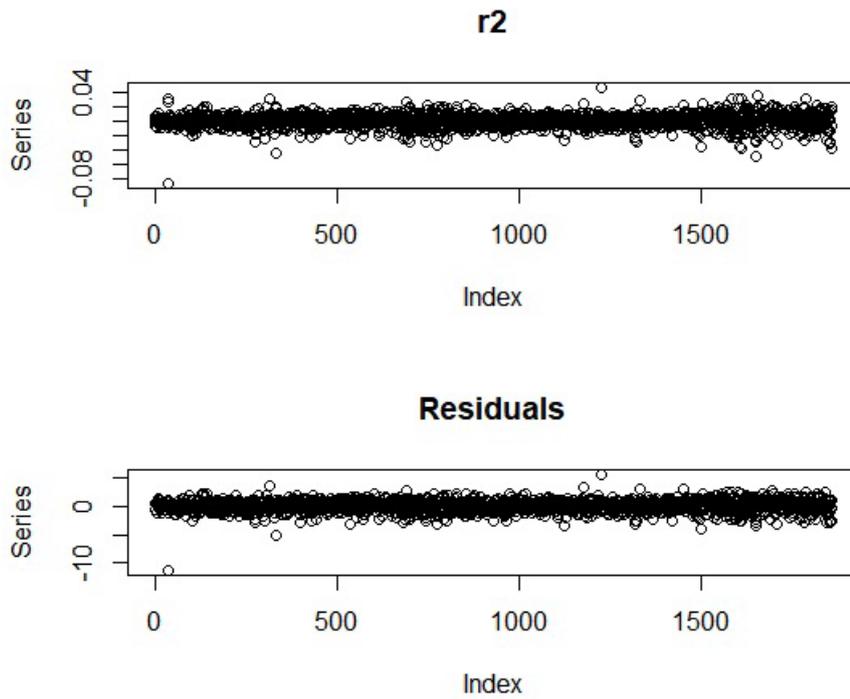


Figure 3.17 – Représentation graphique de la série de rendement de l'indice SMI PAR GARCH (1,1).

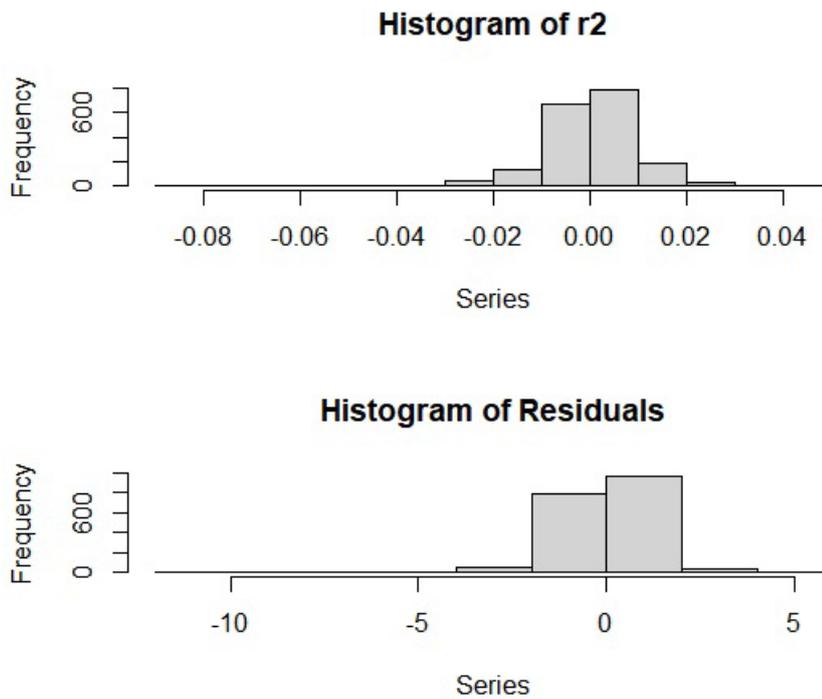


Figure 3.18 – Représentation de la corelogramme de la série du rendement et du résidu par garch (1,1).

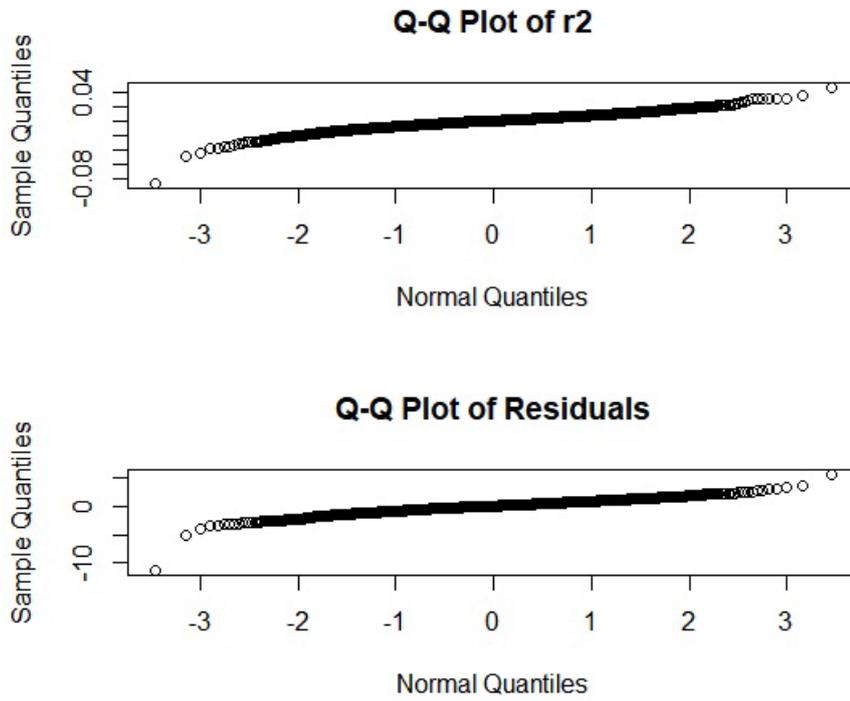


Figure 3.19 – Représentation graphique de Q-QPlot de la série du rendement et du résidue de l'indice SMI par GARCH(1,1).

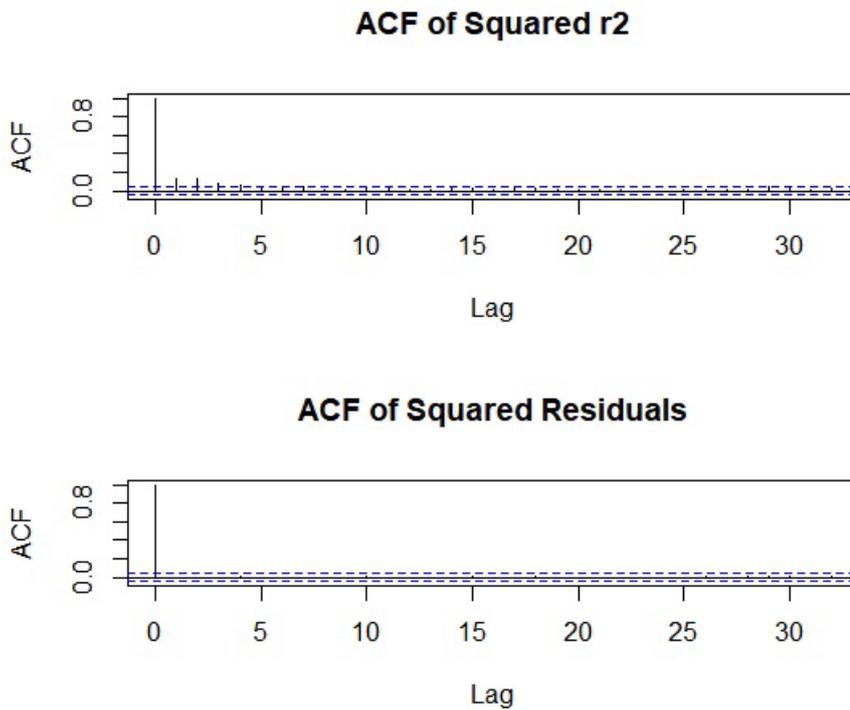


Figure 3.20 – Représentation graphique de l'autocorrélation simple et du squared résidue de l'indice de SMI PAR GARCH(1,1).

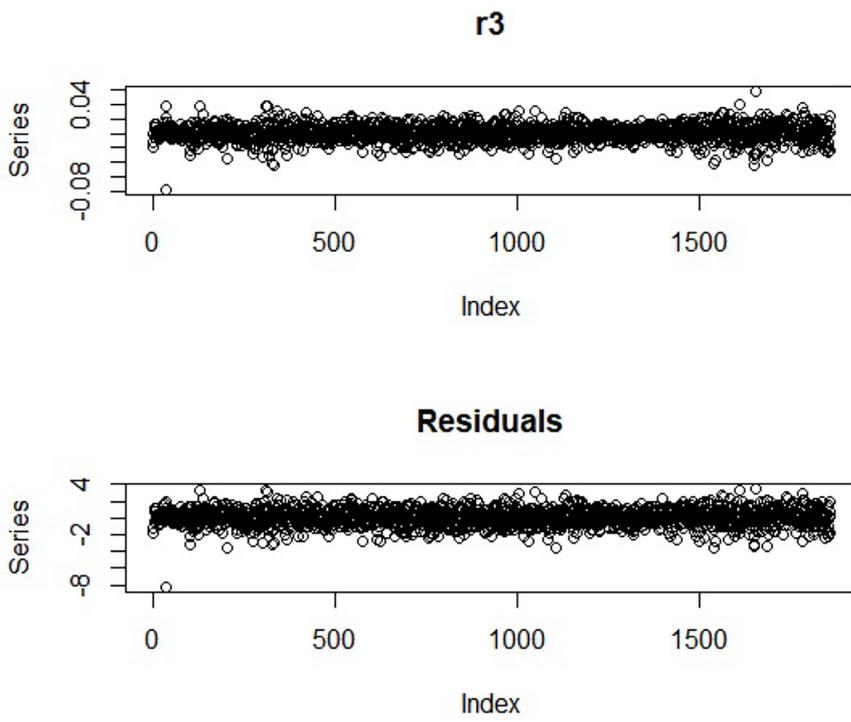


FIGURE 3.21 – Représentation graphique de la série de rendement de l'indice CAC 40 PAR garch (1,1).

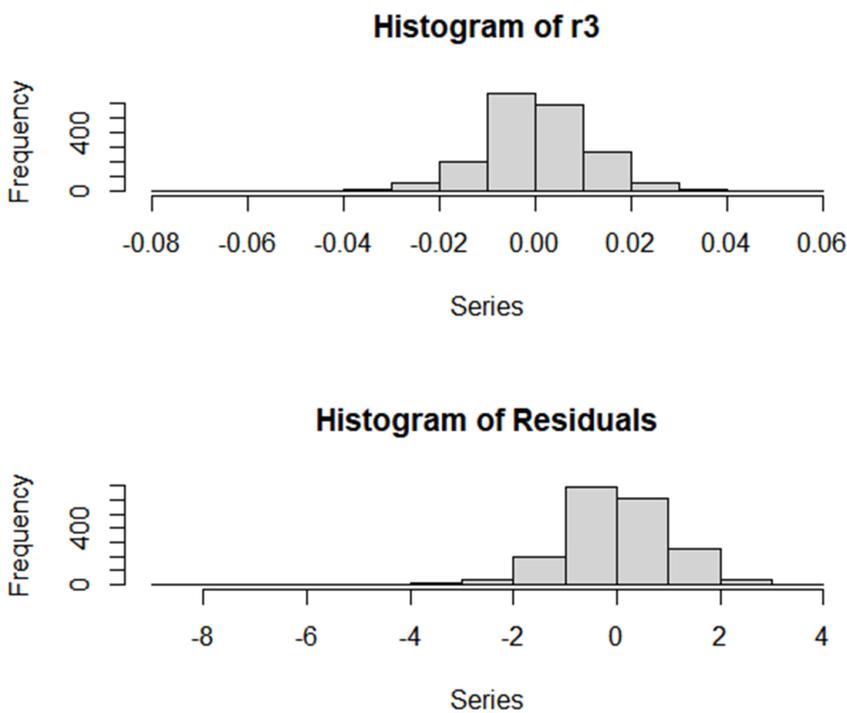


FIGURE 3.22 – Histogramme de la série de rendement et du résidue de l'indice CAC 40 par garch (1.1).

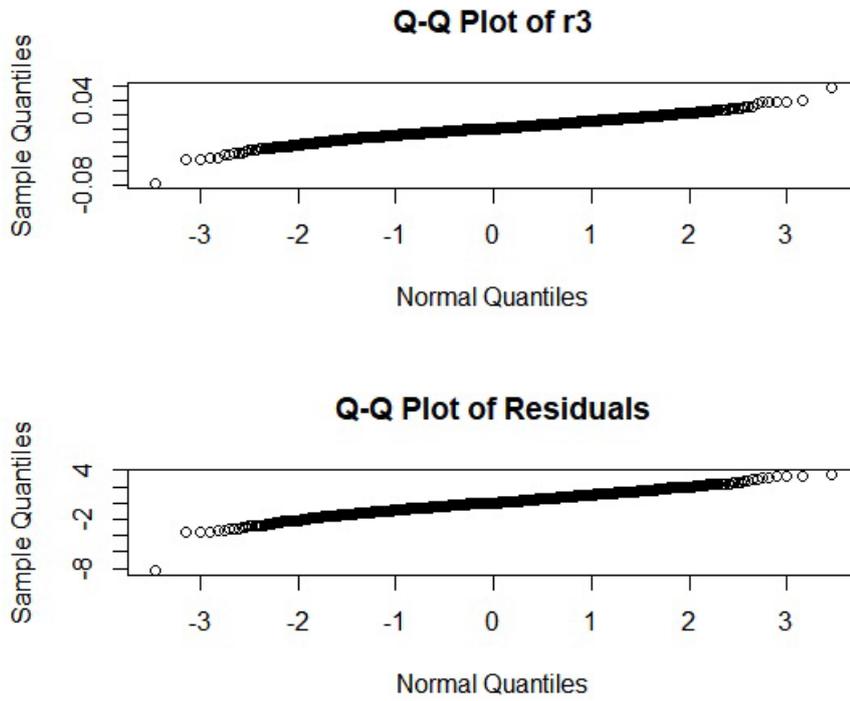


Figure 3.23 – Représentation graphique de Q-QPlot et du résidu de l'indice CAC40.

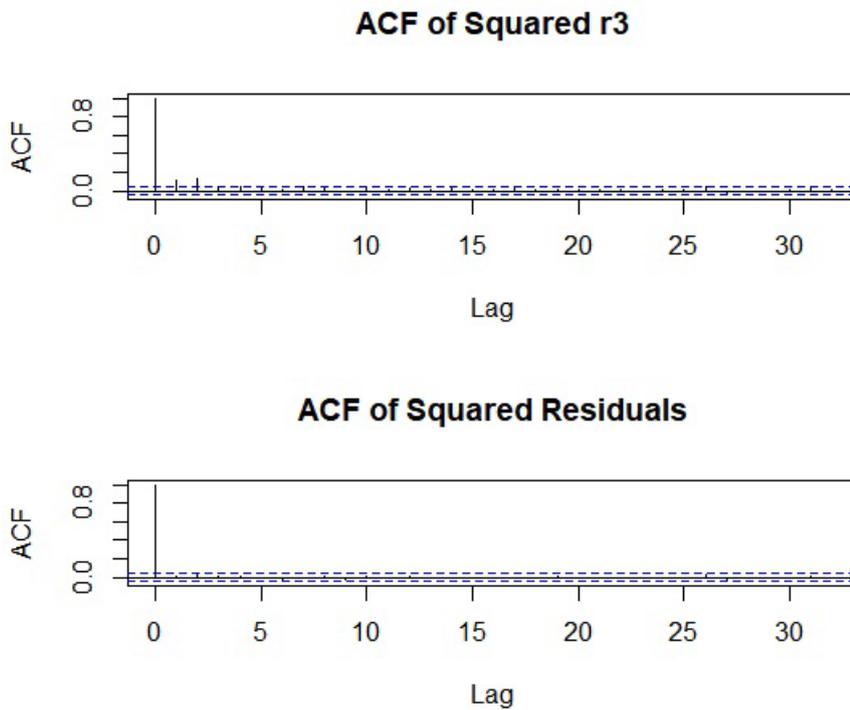


Figure 3.24 – Représentation graphique de squared de la série de rendement et du résidu par GARCH (1.1).

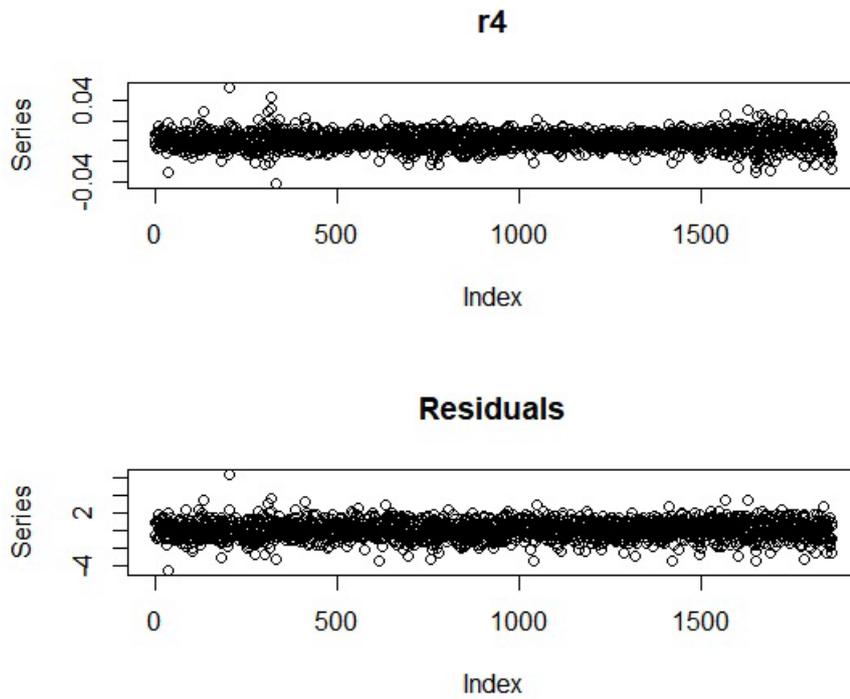


Figure 3.25 – Représentation graphique de la série de rendement de l'indice FTSE PAR GARCH (1.1).

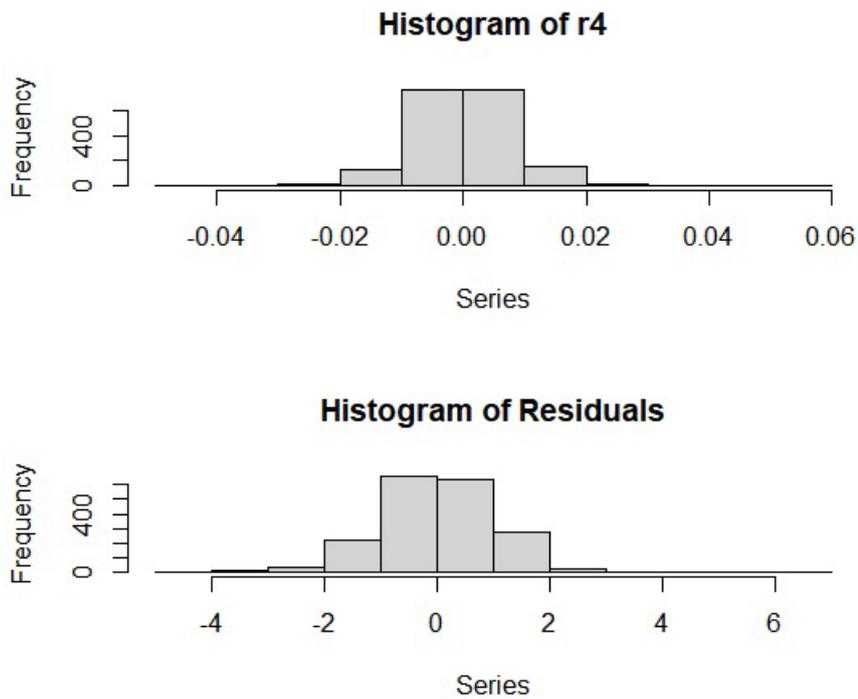


Figure 3.26 – Histogramme de la série de rendement et du résidue de l'indice FTSE par garch(1.1).

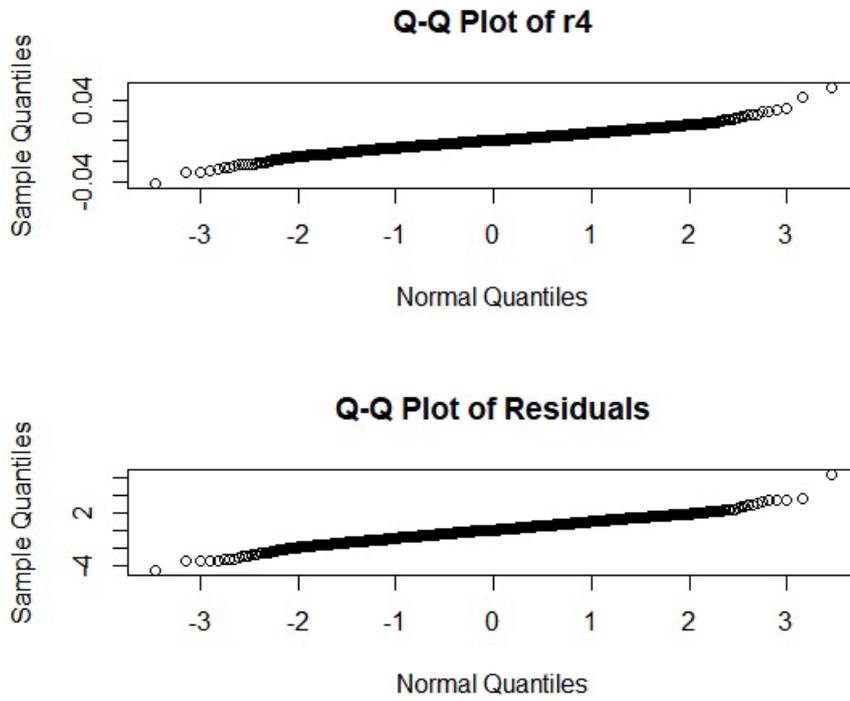


Figure 3.27 – Représentation graphique de la série de rendement et du résidue pour l'indice FTSE par GARCH(1.1).

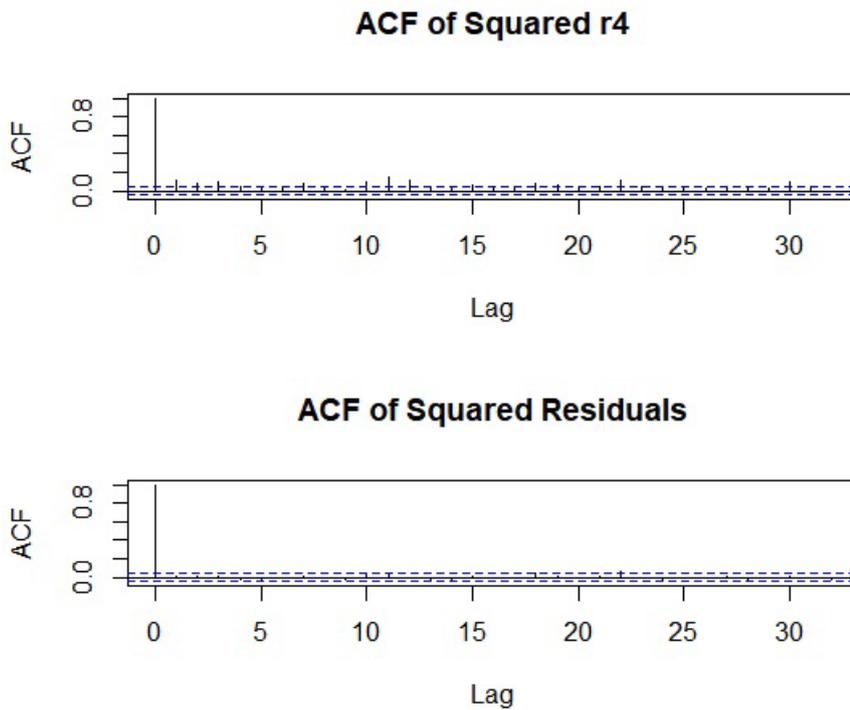


Figure 3.28 – Représentation graphique de squared de la série du rendement et de résidue de l'indice FTSE par GARCH(1.1).

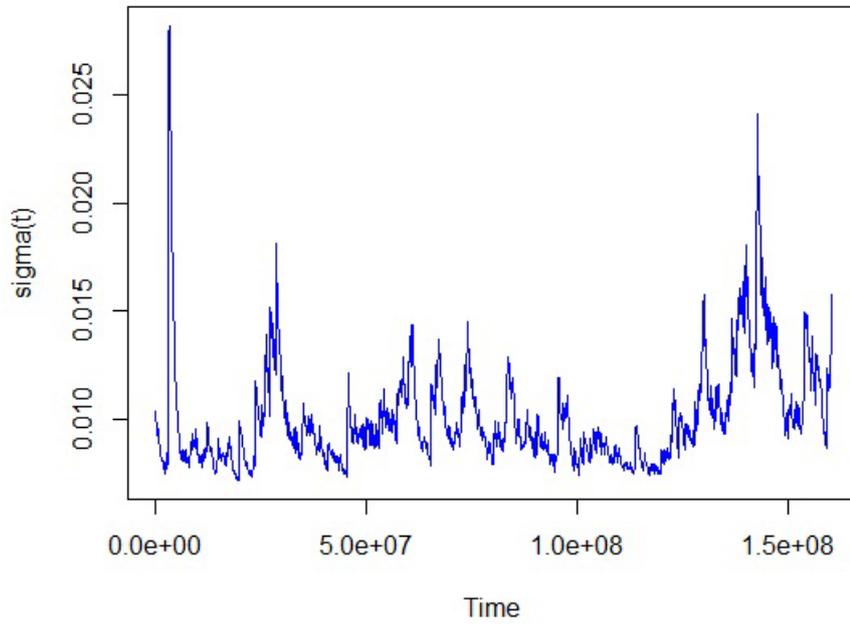


Figure 3.29 – Représentation graphique des valeurs de sigma pendant la période de prévision pour l'indice DAX.

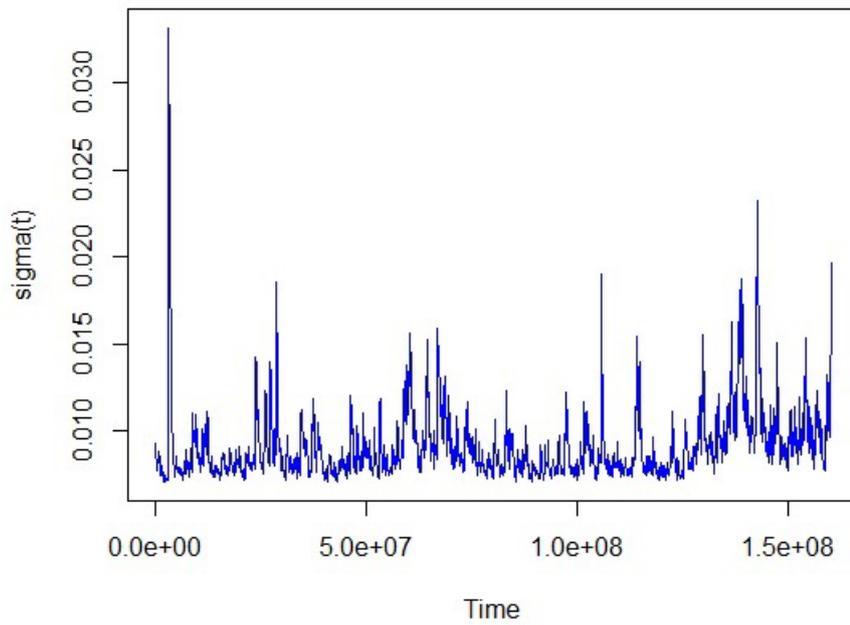


Figure 3.30 – Représentation graphique des valeurs de sigma pendant la période de prévision de l'indice SMI.

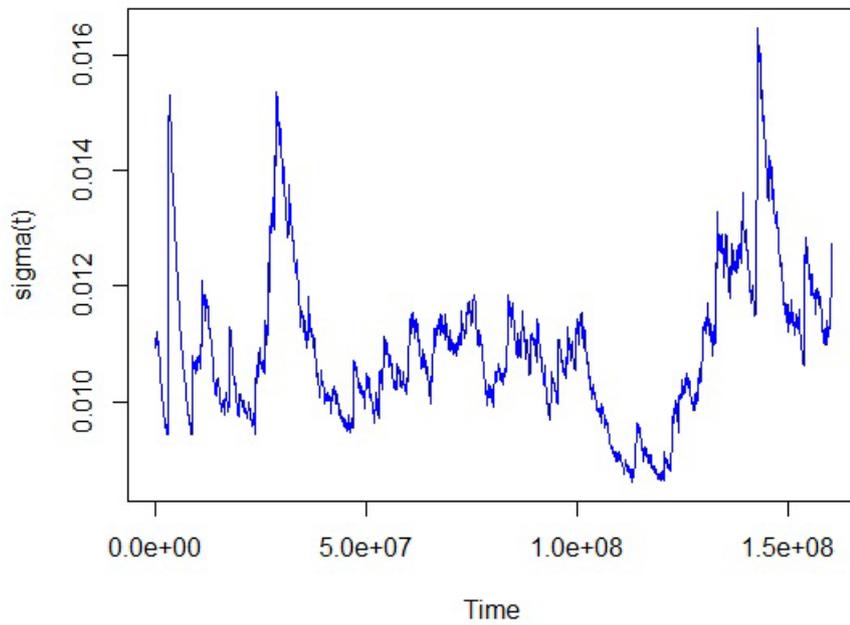


Figure 3.31 – Représentation graphique des valeurs de sigma pendant la période de prévision pour l'indice CAC40.

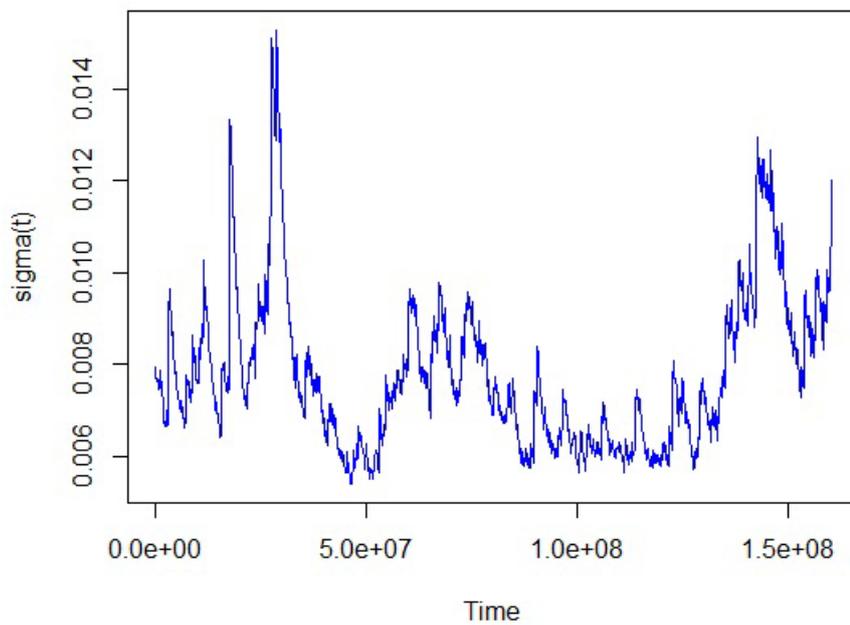


Figure 3.32 – Représentation graphique des valeurs de sigma pendant la période de prévision pour l'indice FTSE.