



الجمهورية الجزائرية الديمقراطية الشعبية



République Algérienne Démocratique et Populaire

Ministère de L'Enseignement Supérieur et de la Recherche Scientifique

Université Saad Dahlab Blida1

Faculté des Sciences

Département de Mathématiques



Mémoire

En vue de l'obtention du Diplôme de **MASTER**

Spécialité : *Modélisation Stochastique et Statistique*

Thème

**Prévision de la valeur à risque par le modèle GARCH et la théorie
des valeurs extrêmes.**

Application aux données financières (SP500, INTEL, IBM).

Présenté par : - SADOUN WASSILA

Soutenu le : 02 Octobre 2021.

- FDAILAINE FOUZIA

Devant le jury composé de :

- | | | |
|-----------------------------------|-------|--------------------|
| * Président : Mr. TAMI OMAR | M.C.B | Université Blida 1 |
| * Examineur : Mr. LAIDI MOHAMED | M.C.B | ENST |
| * Promoteur : Mr. FRIHI REDHOUANE | M.A.A | Université Blida 1 |

Année Universitaire : 2020 - 2021

Remerciements

Avant tout, nous remercions Dieu de nous avoir données le courage, la patience et la capacité pour mener ce travail à terme.

Nous remercions nos parents, nos sœurs, nos frères qui se sont sacrifiés pour nos bien et qui nous ont encouragée et soutenue le long de notre cursus universitaire.
Nous remercions les personnes qui ont contribué au bon déroulement de notre travail.

Tout d'abord, et bien sûr, nos infinie gratitude et nos profonds remerciements sont destinés à notre promoteur Mr Le professeur Frihi Redhouane qui nous a accueillie au sein du bureau, de nous avoir bien conseillée et suivie pendant notre travail.

Nous remercions Mr Tami Omar et Mr Laidi Mohamed d'avoir accepter d'examiner notre travail et de participer au jury de notre soutenance.

Nous remercions vivement tous nos professeurs qui nous ont accompagnée durant notre cursus universitaire et pour leur aide et le savoir aussi précieux qu'ils nous ont apportée au sein de département de mathématiques a Saad Dahlab.

Ils savent combien ils comptent pour nous, nos familles et cousins, nous les remercions pour leurs soutient permanent.

Nous remercions en vrac : Maroua, Widad, Samia, Belkis, Ikram, Lamia, Salima, Sabah pour ses histoires drôles et ses blagues, Amina, Sabrina, Nadjat, Nacéra, Noudjoud, Mohamed, Ayoub, Razki, Riad de jouer Ludo King et des attitudes drôles que nous sortis .

Dédicaces

À nos chères familles

Nous dédions ce travail à nos chers parents et familles respectives, Merci à eux.

À nos cher(e)s ami(e)s À nos chers collègues

Nous ne pouvons trouver les mots justes et sincères pour vous exprimer nos affection et nos pensées, Vous êtes pour nous des frères, sœurs et des amis sur qui nous pouvons compter. En témoignage de l'amitié qui nous uni et des souvenirs de tous les moments que nous avons passé ensemble, Nous vous dédie ce travail et nous vous souhaitons une vie pleine de santé et de bonheur.

À notre cher encadreur Mr. Frihi Redhouane

Un remerciement particulier et sincère pour tous vos efforts fournis. Vous avez toujours été présente. Que ce travail soit un témoignage de nos gratitude et nos profond respect.

ملخص

بالنسبة الى اي مستثمر من الضروري القيام بتقدير قيمة الخطر وقيمة الخسارة التي يمكن التعرض لها من خلال وضعيات مختلفة .للاستثمار كثير من الطرق العملية متاحة له ولكل منها ايجابيات و سلبيات، لذا وجب عليه معرفة اي منها اكثر نجاعة . دراستنا تتمحور حول مقياسين هما قيمة المخاطرة وقيمة المخاطرة الشرطية ، التي تقدر قيمتهما من خلال طريقتين تجريبيتين GPD و GARCH(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. وقد قمنا بمعالجة معطيات حقيقية لسلاسل ايرادات لثلاثة مؤشرات في البورصة (IBM ، INTEL ، SP500).

كلمات مفتاحية: نظرية القيم القصوى ، تعميم القيم القصوى ، GARCH ، GPD ، قيمة المخاطرة.

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 faible, afin de réduire le risque de son placement. Dans ce travail, nos études s'articulent autour des mesures de risque : la valeur à risque (VaR) qui est estimé par deux méthodes : empiriques, GPD et GARCH 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. Ce modèle permet de représenter une chronique avec comme finalité de prévoir des valeurs futures de la VaR qu'est devenue un standard de la gestion du risque dans le monde financier à partir de traitement d'une application sur des données réelles pour les séries des rendements pour les trois indices boursiers (INTEL, IBM, SP500).

Mots clés: Théorie des valeurs extrêmes (TVE), GEV, GPD, GARCH, VaR.

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 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 model. 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 three stock market indices (INTEL, IBM, SP500).

Keywords: Extreme value theory (EVT), GEV, GPD, GARCH, VaR.

Table des matières

Introduction générale	5
1 Valeurs Extrêmes	9
1.1 Introduction	9
1.2 Définition de la théorie des valeurs extrêmes	9
1.3 Types des Valeurs Extrêmes (GEV et GPD)	11
1.3.1 Théorème de Fisher-Tippett	12
1.3.2 La loi des excès	15
1.4 Domaines d'attraction	17
1.4.1 Caractérisation générales	17
1.5 Statistique d'ordre	18
1.6 Estimation des Paramètres de la loi des Valeurs Extrêmes	18
1.6.1 Estimateur de Pickand	19
1.6.2 Estimateur de Hill	19
2 Modèle GARCH	21
2.1 Introduction	21
2.2 Modèle ARCH	21
2.3 Présentation du processus GARCH	22
2.4 Les modèles ARCH généralisées : GARCH	23
2.4.1 Modèle GARCH(p,q)	23
2.4.2 GARCH faible	23
2.4.3 GARCH semi-fort	23
2.4.4 GARCH fort	24
2.5 Propriétés des processus GARCH	25
2.5.1 Modèle GARCH(1,1)	26
2.6 Inférence statistique sur le Modèle GARCH	28
2.6.1 La méthode de maximum de vraisemblance	28
2.6.2 Estimation des paramètres du modèle GARCH	29
3 Théorie du Risque	32
3.1 Introduction	32
3.2 Définition de la VaR	32
3.3 Exemple de valeur à risque (VaR)	33
3.4 Aperçu général sur les mesures du risque	34
3.5 Mesure du risque Cohérente	34
3.5.1 Types du risque	36

3.5.2	Nature des risques :	36
3.5.3	Limites de la VaR	37
3.5.4	Mesures Alternatives à la VaR	38
3.6	Avantages et inconvénients de la valeur à risque (VaR)	39
3.6.1	– Avantages de la Value-at-Risk :	39
3.6.2	– Inconvénients de la Value-at-Risk :	40
4	Estimation des Valeurs à Risque	41
4.1	Introduction	41
4.2	Définition de la VaR	42
4.3	Estimation de la VaR par la théorie des valeurs extrêmes	42
4.3.1	VaR par la méthodes des blocs GEV	42
4.3.2	VaR par la méthode POT	43
4.4	Estimation de VaR par le modèle GARCH	43
5	Simulations et Applications	45
5.1	Introduction	45
5.2	Simulation	45
5.2.1	La VaR avec Lois Extrêmes (TVE)	45
5.3	Données réelles :	52
5.3.1	SP500 :	52
5.3.2	INTEL :	55
5.3.3	IBM :	57
	Conclusion générale	59
	Bibliographie	62
	Annexe	63

Liste des Tableaux

Tableau 1 : Tableau des paramètres estimés de loi GEV.

Tableau 2 : Tableau des paramètres estimés de loi GEV.

Tableau 3 : Tableau des paramètres estimés de loi GEV.

Tableau 4 : Tableau des paramètres estimés de loi GPD.

Tableau 5 : Tableau des paramètres estimés de loi GPD.

Tableau 6 : Tableau des paramètres estimés de loi GPD.

Tableau 7 : Tableau Estimation du paramètre ϵ par la méthode de Hill.

Tableau 8 : Tableau Valeur de VaR pour $\epsilon = 0.4$.

Tableau 9 : Tableau valeur de VaR pour $\epsilon = 0.5$.

Tableau 10 : Tableau valeur de VaR pour $\epsilon = 0.7$.

Tableau 11 : Estimation de la VaR dans le cas d'un modèle ARCH(2) de paramètre ($\alpha = 0.2, \beta = 0.4$).

Tableau 12 : Estimation de la VaR dans le cas d'un modèle GARCH(1,1) de paramètre ($\alpha = 0.2, \beta = 0.7$).

Tableau 13 : Estimation de la VaR dans le cas d'un modèle GARCH(1,2) de paramètre ($\alpha = 0.2, \beta_1 = 0.2, \beta_2 = 0.7$).

Tableau 14 : Estimation de la VaR dans le cas d'un modèle ARMA(1,1) de paramètre ($ar_1 = 0.5, ma_1 = 0.3$) + GARCH(1,1) de paramètre ($\alpha = 0.1, \beta = 0.7$).

Tableau 15 : Estimation de la VaR dans le cas d'un modèle AR(1) + GARCH(1,1) de paramètre ($\alpha = 0.1, \beta = 0.7$).

Tableau 16 : Estimation de la VaR dans le cas d'un modèle de GARCH(1,1) de paramètre ($\alpha = 0.1, \beta = 0.7$).

Tableau 17 : Tableau descriptive de la série de rendement du SP500.

Tableau 18 : Estimation d'indice SP500 par modèles GARCH (1.1).

Tableau 19 : Tableau des différent AIC pour différent modèles.

Tableau 20 : Tableau descriptive de la série de rendement du INTEL.

Tableau 21 : Estimation d'INTEL par modèle GARCH(1,1).

Tableau 22 : Tableau des différent AIC pour différent modèles.

Tableau 23 : Tableau descriptive de la série de rendement du IBM.

Tableau 24 : Estimation d'IBM par modèle GARCH(1,1).

Tableau 25 : Tableau des différent AIC pour différent modèles.

Introduction générale

La théorie des valeurs extrêmes est une branche probabiliste fascinante qu'on rencontre dans des nombreux domaines pratiques tels que l'assurance, la réassurance et la finance. Ce domaine de recherche a d'abord attiré aussi bien l'intérêt des probabilistes que celui des ingénieurs des risques financiers et assuranciers. Il est basé sur la connaissance de la loi asymptotique des extrêmes (maximum et minimum) d'un échantillon aléatoire. La distribution de type-Pareto (ou loi limite est de Fréchet) est a priori la plus appropriée dans le contexte financier et assurantiel car elle correspond à des observations individuelles dont la distribution possède des queues épaisses, et elle est cohérente avec les modèles GARCH des séries temporelles.

Le présent mémoire est alors une petite synthèse des travaux de recherches concernant la théorie des valeurs extrêmes et la théorie du risque ainsi que les modèles GARCH. Ce mémoire s'articule autour du plan du travail suivant :

Dans le chapitre 1, nous rappelons quelques éléments théoriques essentiels de la théorie des valeurs extrêmes (TVE). Il contient des rappels sur la statistique d'ordre, qui est très utile en théorie des valeurs extrêmes et on fait une introduction sur l'étude du comportement asymptotique du maximum d'un échantillon. Cette étude faisant appel à la notion de fonctions à variations régulières, on rappelle préalablement la définition de telles fonctions et on en donne quelques propriétés.

Dans le chapitre 2, nous donnons un aperçu sur quelques les modèles GARCH.

Dans le chapitre 3, nous avons présenté la théorie du risque, par contre le chapitre 4 est consacré pour l'estimation de la valeur à risque pour les valeurs extrêmes et le modèles ARMA à effet GARCH. Le dernier chapitre est réservé pour la simulation et l'applications aux données financières. Nous allons clôturé ce travail par une conclusion générale.

Chapitre 1

Valeurs Extrêmes

1.1 Introduction

La théorie des valeurs extrêmes (TVE) est une branche de la statistique mathématiques qui s'intéresse aux valeurs extrêmes des distribution de probabilité.

La théorie des valeurs extrêmes est appliquée en hydrologie pour prévoir les crues, en démographie pour prévoir la distribution de probabilité de l'âge maximum que l'être humain pourra atteindre, en assurance pour prévoir les grands sinistres, en finance pour prévoir les crises financières ou encore en météorologie les pics de précipitations.

1.2 Définition de la théorie des valeurs extrêmes

La théorie des valeurs extrêmes étudie les événements rares ayant une faible probabilité d'apparition, et a pour but d'étudier la loi du maximum (ou du minimum) d'une suite des variables aléatoires réelles dans le cas où la loi du phénomène n'est pas connue.

On considère (X_1, \dots, X_n) une suite de n variables aléatoires identiques indépendantes distribuées (i.i.d) de fonction de répartition F définie par :

$$F(x) = \mathbb{P}(X_i \leq x) \text{ pour } i = 1, \dots, n$$

Pour étudier le comportement extrême des événement, on considère la variable aléatoire :

$$M_n = \text{Max}(X_1, \dots, X_n)$$

que représente le maximum d'un échantillon de taille n .

Comme les variables aléatoires sont (i.i.d) alors la fonction de répartition de M_n peut être écrite de la façon suivante :

$$\begin{aligned}F_{M_n}(x) &= \mathbb{P}(M_n \leq x) \\ &= \mathbb{P}(X_1 \leq x, \dots, X_n \leq x) \\ &= \mathbb{P}(X_1 \leq x) \times \dots \times \mathbb{P}(X_n \leq x) \\ &= [F(x)]^n\end{aligned}\tag{1.1}$$

et sa densité est $f_n(x) = nf(x)F^{n-1}(x)$.

Remarque 1 *On obtient la correspondance entre minimum et maximum par la relation suivante :*

$$\text{Min}(X_1, \dots, X_n) = -\text{Max}(-X_1, \dots, -X_n)$$

Ainsi tous les résultats que nous allons présenter pour les maximas pourront être transposés pour les minimas.

Lemme 1 La suite $(\pi \frac{M_n}{n}, n \geq 1)$ converge en loi vers W de fonction de répartition définie par : $P(W \leq x) = e^{-\frac{1}{x}}, x > 0$.

La loi de W appartient à la famille des lois de Fréchet.

Démonstration, On note F_n la fonction de répartition de $\pi \frac{M_n}{n}$ on a

$$\begin{aligned} F_n(x) &= \mathbb{P}(M_n \leq \frac{nx}{\pi}) \\ &= \mathbb{P}(x_1 \leq \frac{nx}{\pi})^n \\ &= (1 - \int_{\frac{nx}{\pi}}^{+\infty} \frac{1}{\pi(1+y^2)} dy)^n \end{aligned} \tag{1.2}$$

Pour $x > 0$, on a

$$\int_{\frac{nx}{\pi}}^{+\infty} \frac{1}{1+y^2} dy = \int_{\frac{nx}{\pi}}^{+\infty} \frac{1}{\pi(y^2)} dy + \int_{\frac{nx}{\pi}}^{+\infty} [\frac{1}{\pi(1+y^2)} - \frac{1}{\pi(y^2)}] dy = \frac{1}{nx} + O((nx)^{-3})$$

On a alors pour $x \geq 0$,

$$F_n(x) = (1 - \frac{1}{nx} + O((nx)^{-3}))^n$$

On déduit que

$$\lim_{x \rightarrow +\infty} F_n(x) = e^{-\frac{1}{x}}, x > 0$$

Ainsi la suite $(\pi \frac{M_n}{n}, n \geq 1)$ converge en loi vers W de fonction de répartition

$$P(W \leq x) = e^{-\frac{1}{x}}, x > 0$$

1.3 Types des Valeurs Extrêmes (GEV et GPD)

Historiquement, la première approche développée dans l'analyse des valeurs extrêmes pour une population donnée est celle des blocs maxima connus par la distribution des Valeurs Extrêmes Généralisées (GEV)(Generalized Extreme Value). Cette approche est apportée aux données qui consistent en un ensemble maximums : annuels journaliers, semestriels journaliers, trimestriels journaliers... etc, et qui regroupent trois lois des valeurs extrêmes à savoir la loi Gumbel, loi de Weibull et loi de Fréchet.

Cependant, cette approche a été critiquée dans les autres grandes valeurs de l'échantillon. Pour pallier ce problème, Pickands (1975)([1]) a introduit une nouvelle approche dans l'analyse des valeurs extrêmes connue par la Distribution de Parto Généralisée(GPD)(Peaks Over Threshold), qui est apportée aux données qui dépassent un certain seuil élevé bien déterminé

1.3.1 Théorème de Fisher-Tippett

L'estimation de la loi des extremums à partir de la fonction parente F pose une difficulté majeure car cette dernière n'est pas connue. Même si la loi de variable parent X est connue avec exactitude la loi du terme maximum n'est pas toujours facilement calculable. Pour contourner cette limite, les chercheurs Fisher-Tippett(1928)([2]) se sont intéressés au comportement asymptotique des variables aléatoires M_n . Il s'agit de déterminer la loi vers laquelle M_n converge, quand n tend vers l'infini afin de remplacer F par cette dernière pour les plus grandes valeurs de n .

Le théorème de Fisher-Tippett (1928)([2]) a permis de caractériser la loi de la distribution des valeurs extrêmes sans que celle-ci soit conditionnée par la loi de la variable parente.

Définition 1 (*Loi de même type*)

On dit que deux variables aléatoires réelles X et Y sont de même type s'il existe des constantes réelles $a \geq 0$ et $b \in \mathbb{R}$ tels que $Y = aX + b$

i.e si F et H sont des lois respectives des variables X et Y alors on a $F(ax + b) = H(x)$.

Autrement dit, les variables de même type ont la même loi à un facteur de localisation et d'échelle près.

De façon analogique au théorème central limite, peut-on trouver des constantes de normalisation : a_n et b_n , avec $a_n > 0$ et $b_n \in \mathbb{R}$ et une loi non-dégénérée de loi H telle que :

$$P\left(\frac{M_n - b_n}{a_n}\right) = \lim_{n \rightarrow +\infty} (F(a_n x + b_n))^n \rightarrow H(x) \quad (1.3)$$

Fisher et Tippett trouvent en 1928 une solution à ce problème au moyen d'un théorème qui port leur nom et qui est l'un des fondement de la théorie des valeurs extrêmes.

Théorème 1 (*Théorème de Fisher-Tippett*)

S'il existe deux suites des constantes de normalisation (a_n) et (b_n) et une loi non-dégénérée de loi H telle que :

$(\frac{M_n - b_n}{a_n}) \xrightarrow{\text{loi}} H$, alors H est l'une des trois lois limites :

$$\Lambda(x) = \exp\{-\exp[-x]\}, x \in R \text{ (distribution de Gumbel)} \quad (1.4)$$

$$\Phi_\alpha(x) = \begin{cases} \exp\{-x^{-\alpha}\} & , \text{si } x > 0 \text{ (distribution de Fréchet)} \\ 0 & , \text{si } x \leq 0 \end{cases} \quad (1.5)$$

$$\Psi_\alpha(x) = \begin{cases} \exp\{-(-x)^\alpha\} & , \text{si } x \leq 0 \text{ (distribution de Weibull)} \\ 1 & , \text{si } x > 0 \end{cases} \quad (1.6)$$

Ce théorème donne un résultat intéressant : quelle que soit la loi limite de la variable parente, la loi limite des extrêmes a toujours la même forme. Bien que le comportement de ces lois soit complètement différent, elles peuvent être combinées en un seule paramétrisation contenant un unique paramètre ξ qui contrôle lourdeur de la queue de loi appelé indice des valeurs extrêmes (ou indice de queue) :

$$H_\xi(x) = \begin{cases} \exp\left((-1 + \xi x)^{-\frac{1}{\xi}}\right) & , \xi \neq 0 \quad , 1 + \xi x > 0 \\ \exp(-\exp(-x)) & , \xi = 0 \quad , -\infty \leq x \leq +\infty \end{cases} \quad (1.7)$$

Où H est une fonction de répartition non-dégénérée. Cette loi est appelée loi des valeurs extrêmes généralisées (*Generalized Extreme Value*), que l'on notera *GEV*.

Le théorème de Fisher-Tippett fournit en quelque sorte la contrepartie du Théorème Central Limite (TCL) dans le cas d'événements extrêmes. Cependant, contrairement au TCL, où la loi limite est le seul loi limite possible dans le cas des extrêmes, trois types de loi limite sont possibles.

$$\text{Frechet} : \alpha = \frac{1}{\xi}, \xi > 0$$

$$\text{Weibull} : \alpha = \frac{-1}{\xi}, \xi < 0$$

$$\text{Gumbel} : \xi = 0$$

Le choix de l'une des trois distribution du domaine d'attraction de la théorie des valeurs extrêmes dépend de la valeur de l'indice de queue ξ .

1. Si $\xi > 0$, la GEV devient la distribution de Fréchet caractérisée par des queues de distribution épaisses et décroissantes à l'image d'une fonction puissance (loi de Student ou loi de Pareto).

2. Si $\xi = 0$, la GEV suit la loi de Gumbel avec des queues qui décroissent de manière exponentielle comme des lois normale et log-normale.

3. Si $\xi < 0$, nous avons la distribution de Weibull dont les queues de distribution sont bornées comme celles de la distribution uniforme et densité des lois des valeurs extrêmes, avec $\xi = 1$ pour la loi de Fréchet $\xi = 0$, pour la loi de Gumbel et $\xi = -1$ pour la loi de Weibull.

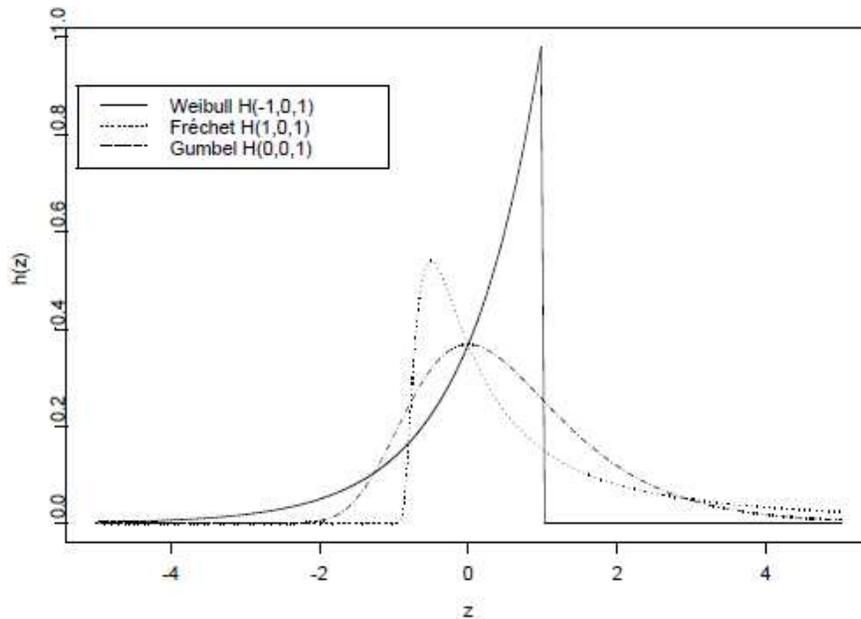


FIGURE 1.1 – Densités des lois des valeurs extrêmes

L'approche basée sur la GEV a été critiquée dans la mesure où l'utilisation d'un seul maxima conduit à une perte d'information contenue dans les autres grandes valeurs de l'échantillon. Pour pallier ce problème, la méthode POT (Peaks-over-Threshold) ou méthode des excès au-delà d'un seuil élevé a été introduite par Pickands[1975]([1]).

1.3.2 La loi des excès

L'approche basée sur les distributions GEV peut être réductrice du fait que l'utilisation d'un seul maxima conduit à une perte d'information continue dans les autres grandes valeurs de l'échantillon. La solution est de considérer plusieurs grandes valeurs au lieu de la plus grande. La nouvelle approche de la théorie des valeurs extrêmes appelée POT (ou Peak Over Threshold) consiste à utiliser les observation qui dépassent un certain seuil, plus particulièrement les différences entre ces observation et le seuil appelées excès. Il est clair que cette méthode nécessite la détermination d'un seuil ni trop faible pour ne pas prendre en considération des valeurs non extrêmes, ni trop élevé pour avoir suffisamment d'observation. Notons le seuil par u . Cette méthode initialement développée par Pickands (1975)([1]) et abondamment étudiée par divers auteurs tels que de Smith (1987)([3]), Davison et Smith (1990)([4]), ou Reiss et Thomas (2001)([5]).

Définition 2 Soit X une v.a de fonction de répartition F et de point terminal x_F tel que $x_F = \sup x \in R$, $F(x) < 1$. On cherche à partir de la loi F de X à définir une loi conditionnelle F_u par rapport au seuil u pour les variables aléatoires dépassant ce seuil. On définit alors la loi conditionnelle des excès F_u par :

$$F_u(y) = \mathbb{P}[X - u \leq y / X > u] \quad (1.8)$$

$$F_u(y) = \frac{F(u+y) - F(u)}{1 - F(u)}, y \geq 0 \quad (1.9)$$

Notons $Y = X - u$ pour $X > u$ et pour n v.a observées X_1, \dots, X_n nous pouvons écrire : $Y_j = X_i - u > 0$ pour $1 \leq j \leq N_u$ où $N_u = \text{card}[i : i = 1, \dots, n, X_i > u]$ est nombre des dépassement du seuil u par les excès correspondants.

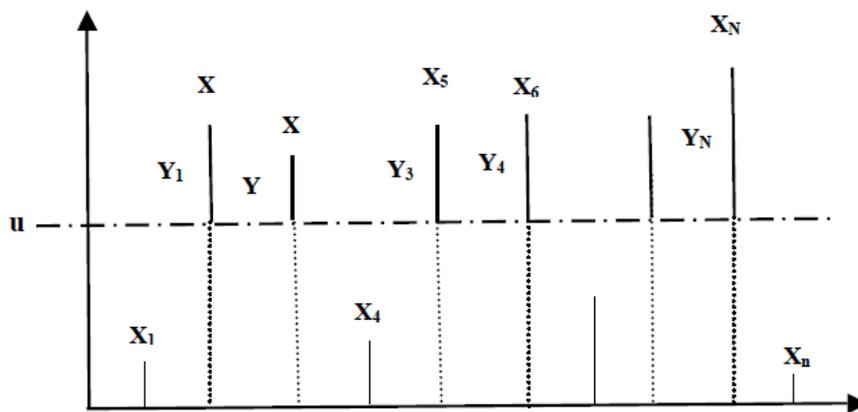


FIGURE 1.2 – Méthode des excès

Le théorème de Pickands-balkema de Haan ci-après donne la forme de la loi limite pour les valeurs extrêmes : sous certaines conditions de convergence, la loi limite est une loi de Pareto généralisée que l'on notera GPD.

Théorème 2 (de Pickands-Balkema-de Haan)

Une fonction de répartition F appartient au domaine d'attraction maximal de $G_{\xi,\beta}$, si et seulement si, il existe une fonction positive $\beta(u)$ telle que :

$$\lim_{u \rightarrow x_F} \sup_{0 \leq y \leq x_F - u} |F_u(y) - G_{\xi,\beta(u)}(y)| = 0 \quad (1.10)$$

Où $F_u(y)$ est la fonction de répartition conditionnelle des excès pour u élevé, x_F est le point terminal de F , et $G_{\xi,\beta}$ est la GPD pour un seuil assez élevé donnée par :

$$G_{\xi,\beta}(y) = \begin{cases} 1 - (1 - \xi \frac{y}{\beta})^{-\frac{1}{\xi}} & , \xi \neq 0 \\ 1 - \exp(-\frac{y}{\beta}) & , \xi = 0 \end{cases} \quad (1.11)$$

Où $y \geq 0$ pour $\xi \geq 0$ et $0 \leq y \leq -\frac{\xi}{\beta}$ pour $\xi < 0$.

Ce théorème montre l'existence d'une relation étroite entre la GPD et la GEV :

Pickands 1975 a montré que n'importe quelle loi F , l'approximation GPD définie par l'équation(1.11) n'est vérifiée que s'il existe des constantes de normalisation et une loi non-dégénérée telle que le résultat donné par l'équation(1.5) est vérifié.

Dans ce cas, si H est écrite sous la forme d'une GEV de l'équation(1.9) alors l'indice de queue ξ est le même que celui de la GPD donnée dans l'équation(1.12).

De même, la loi GPD regroupe les trois distribution suivantes :

- Lorsque $\xi > 0$, on obtient la loi de Pareto usuelle.
- Lorsque $\xi < 0$, on obtient la loi de Pareto du type II.
- Lorsque $\xi = 0$, on obtient la loi exponentielle de paramètre β .

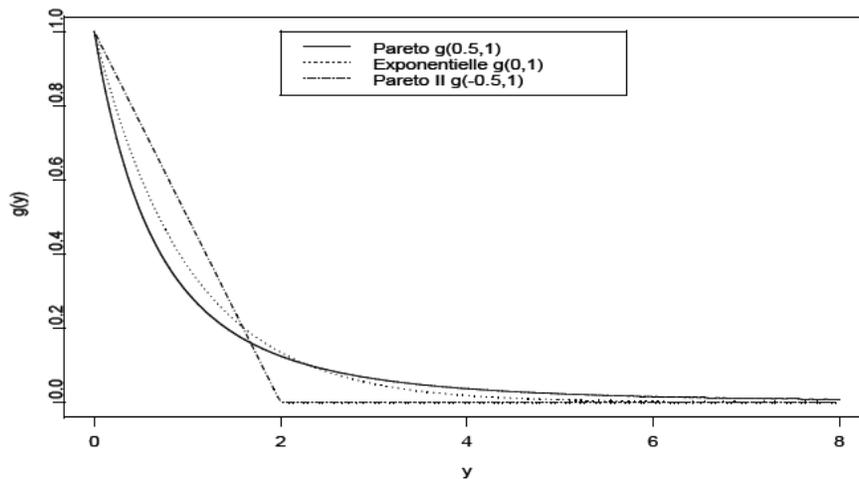


FIGURE 1.3 – Densités des lois des valeurs extrêmes

La dérivée de la distribution cumulative de GPD donne la fonction de densité de probabilité suivante :

$$G_{\xi,\beta}(y) = \begin{cases} \beta^{-1}(1 + \xi\frac{y}{\beta})^{-\frac{1}{\xi}-1} & , \xi \neq 0 \\ \beta^{-1}\exp(-\frac{y}{\beta}) & , \xi = 0 \end{cases} \quad (1.12)$$

- Densité des lois des valeurs extrêmes, avec $\xi = 0,5$ pour la loi de Pareto $\xi = 0$ pour la loi exponentielle et $\xi = -0,5$ pour la loi de Pareto II. La GPD a les propriétés suivantes :

$$E(Y) = \frac{\beta}{1 - \xi} , (\xi < 1)$$

et

$$V(Y) = \frac{\beta^2}{(1 - \xi)^2(1 - 2\xi)} , (\xi < \frac{1}{2})$$

1.4 Domaines d'attraction

Après avoir caractérisé les lois limites, il reste à déterminer les lois L pour lesquelles la loi du maximum renormalisé converge vers une loi maxstable donnée L' . On dit alors que la loi L (ou sa fonction de répartition F) appartient au bassin d'attraction de L' (ou de sa fonction de répartition F_0) pour la convergence du maximum renormalisé. On le notera $L \in D(L')$ (ou $F \in D(F_0)$).

1.4.1 Caratérisation générales

On définit la fonction U par :

$$U(t) = F^{-1}(1 - \frac{1}{t}), t > 1$$

Où F^{-1} est l'inverse généralisé de F . On sait que si $F \in D(H(\xi))$, alors on a un changement d'échelle près.

$$\lim_{s \rightarrow +\infty} \frac{U(sw) - U(s)}{a(s)} = \frac{w^\xi - 1}{\xi}$$

En particulier, si $x, y > 0$ et $y \neq 0$

$$\lim_{s \rightarrow +\infty} \frac{U(sx) - U(s)}{U(sy) - U(s)} = \lim_{s \rightarrow +\infty} \frac{U(sx) - U(s)}{a(s)} \frac{a(s)}{U(sy) - U(s)} = \begin{cases} \frac{x^\xi - 1}{y^\xi - 1}, \xi \neq 0 \\ \frac{\log x}{\log y}, \xi = 0 \end{cases} \quad (1.13)$$

Où $a(x) = a[x]$ pour $x \geq 1$, $x[x]$ désignant la partie entière de x

1.5 Statistique d'ordre

Soient X_1, \dots, X_n i.i.d de fonction de répartition F .

Définition 3 La statistique d'ordre de l'échantillon (X_1, \dots, X_n) est le rearrangement croissant de cette échantillon. On la note $(X_{(1,n)}, \dots, X_{(n,n)})$. On a

$$X_{(1,n)} \leq \dots \leq X_{(n,n)}$$

Et il existe une permutation aléatoire $\sigma_n \in S_n$ l'ensemble des permutations tel que :

$$(X_{(1,n)}, \dots, X_{(n,n)}) = (X_{\sigma_n(1)}, \dots, X_{\sigma_n(n)})$$

En particulier, on a :

$$X_{(1,n)} = \underset{1 \leq i \leq n}{\text{Min}}(X_i)$$

et

$$X_{(n,n)} = \underset{1 \leq i \leq n}{\text{Max}}(X_i)$$

on note F^{-1} l'inverse généralisé de F .

Lemme 2 Soit X_1, \dots, X_n des variables aléatoires indépendantes et la fonction de répartition F .

Soit U_1, \dots, U_n des variables aléatoires indépendantes de loi uniforme sur $[0, 1]$, Alors :

$(F_1(U(1, n)), \dots, F_1(U(1, n)))$ meme loi que $(X_{(1,n)}, \dots, X_{(n,n)})$.

1.6 Estimation des Paramètres de la loi des Valeurs Extrêmes

La vraisemblance d'un modèle aléatoire est la densité de sa loi de probabilité.

- Soit $(X_n, n \geq 1)$ une suite des variables aléatoires indépendantes la vraisemblance du modèle généralisé des extrêmes s'écrit :

$$l(x; \mu, \sigma, \xi) = \left(\frac{1}{\sigma}\right)^m \prod_{i=1}^m \exp\left(-\left(1 + \xi\left(\frac{x_i - \mu}{\sigma}\right)^{-\frac{1}{\xi}}\right)\left(1 + \xi\left(\frac{x_i - \mu}{\sigma}\right)^{-1-\frac{1}{\xi}}\right)\right)$$

$$\log l(\mu, \sigma, \xi) = -m \log(\sigma) - \left(\frac{1}{\sigma} + 1\right) \sum_{i=1}^m \log\left(1 + \xi \left(\frac{x_i - \mu}{\sigma}\right)\right) - \sum_{i=1}^m \left(1 + \xi \left(\frac{x_i - \mu}{\sigma}\right)\right)^{-\frac{1}{\xi}}$$

on pose

$$\rho = \frac{1}{\sigma}$$

et on introduit λ tel que :

$$\log(\lambda) = \frac{\mu}{\sigma}$$

alors la vraisemblance se présente comme suit avec $\xi = 0$

$$L(\mu, \sigma, \xi) = -m \log(\lambda \rho) - \rho \sum_{i=1}^m x_i - \lambda \sum_{i=1}^m \exp(-\rho x_i)$$

1.6.1 Estimateur de Pickand

Théorème 3 Soit $(X_n, n \geq 1)$ une suite des variables aléatoires indépendantes de même fonction de répartition $F \in D(H(\xi))$, où $\xi \in R$. Si

$$\lim_{n \rightarrow +\infty} k(n) = \infty$$

et

$$\lim_{n \rightarrow +\infty} \frac{k(n)}{n} = 0$$

alors l'estimateur de Pickand

$$\widehat{\xi}_{(k(n), n)}^p = \frac{1}{\log 2} \log\left(\frac{X_{(n-k(n)+1, n)} - X_{(n-2k(n)+1, n)}}{X_{(n-2k(n)+1, n)} - X_{(n-4k(n)+1, n)}}\right)$$

converge en probabilité vers ξ .

1.6.2 Estimateur de Hill

Lemme 3 soit L une fonction a variation lente. Alors on a pour tout

$$\rho \geq 0, L(x) = o(x^\rho) \text{ en } +\infty$$

$$\text{et } \int_x^\infty t^{-\rho-1} L(t) \sim \frac{1}{\rho} x^{-\rho} L(x) \text{ en } +\infty$$

Théorème 4 Soit $(X_n, n \geq 1)$ une suite des variables aléatoires indépendantes de même fonction de répartition $F \in D(H(\xi))$, où $\xi \geq 0$. Si

$$\lim_{n \rightarrow +\infty} k(n) = \infty$$

et

$$\lim_{n \rightarrow +\infty} \frac{k(n)}{n} = 0$$

alors l'estimateur de Hill définit par :

$$\widehat{\xi}_{(k(n),n)}^H = \frac{1}{k(n)} \sum_{i=(n-k(n)+1)}^n \log X_{(i,n)} - \log X_{(n-k(n)+1,n)}$$

converge en probabilité vers ξ .

Chapitre 2

Modèle GARCH

2.1 Introduction

Pour des nombreuses applications, l'introduction d'un grand nombre de retards p dans l'équation de la variance conditionnelle du modèle ARCH(p) est nécessaire pour tenir compte de la longue mémoire de la volatilité qui caractérise certaines séries monétaires et financières. Ce nombre important des paramètres peut conduire à la violation de la contrainte de non-négativité de la variance et poser des problèmes d'estimations. Dans cette perspective une extension importante, le modèle autorégressif conditionnellement hétéroscédastique généralisé (GARCH) est suggérée par Bollerslev (1986)([6]). Cette approche exige moins des paramètres à estimer que la formulation ARCH(p) pour modéliser les phénomènes de persistance des chocs. La variance conditionnelle de la variable étudiée est déterminée par le carré des p termes d'erreur passés et des q variances conditionnelles retardées.

2.2 Modèle ARCH

Dans le but de pallier aux insuffisances des représentations Autorégressive Moyenne Mobile (ARMA(p, q)) pour les problèmes monétaires et financiers, Engle (1982)([7]) propose une nouvelle classe des modèles autorégressifs conditionnellement hétéroscédastiques (ARCH) apte à capturer le comportement de la volatilité dans le temps. Le modèle est formé de deux équations : La première met en relation le rendement et certaines variables qui l'expliquent et la seconde modélise la variance conditionnelle des résidus. Le principe proposé par Engle consiste à introduire une dynamique dans la détermination de la volatilité en supposant que la variance est conditionnelle aux informations dont nous disposons. Il avance une spécification ARCH(p) où le carré des innovations, c'est-à-dire, la variance du terme d'erreur au temps t dépend de l'importance des termes d'erreur au carré des p périodes passées.

Le modèle ARCH(p) permet de générer des épisodes de volatilité importante suivis d'épisodes de volatilité plus faibles.

$$Y_t = \beta X_t + \varepsilon_t$$

où

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

où le terme y_t correspond aux variables expliquant les rendements, il peut être un modèle ARMA(p,q). L'expression $I_t = \sigma(X_{t-s})_{s \leq t}$ désigne la tribu engendrée par les X_{t-s} dans la modélisation ARCH, le processus ε_t peut s'écrire sous la forme :

$$\varepsilon_t = \eta_t h_t$$

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

$$\eta_t \sim N(0, 1)$$

ε_t un bruit blanc indépendant.

α_i les coefficients d'autorégressif.

Exemple 1 Un processus ε_t satisfait une représentation ARCH(1) si

$$\varepsilon_t = \eta_t h_t$$

avec

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

où η_t est un bruit blanc faible, tel que $E(\eta_t) = 0$ et $V(\eta_t) = 1$, $\alpha_0 > 0$ et $0 < \alpha_1 < 1$, la composante h_t désigne une variable qui conditionnellement à l'ensemble d'information des valeurs passées de ε_t , i.e. à $I_{t-1} = \sigma\{\varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-j}, \dots\}$, est déterministe et positive.

Dans ce cas, le processus ε_t est caractérisé par des autocorrélations nulles $E(\varepsilon_t \varepsilon_s) = 0$ pour $t \neq s$ ce qui signifie que les sont non corrélés dans le temps. en effet, ε_s reste un bruit blanc mais dit faible.

2.3 Présentation du processus GARCH

Le processus GARCH (Generalized Auto Regressive Conditional Hetermskedasticity) a été introduit en 1986 par Bollerslev. Le processus GARCH est une extension du processus ARCH, il présente les mêmes propriétés et les mêmes fondements que le processus ARCH. Le modèle GARCH a deux dimensions(p,q) alors que le modèle ARCH en a une dimension.

2.4 Les modèles ARCH généralisées : GARCH

2.4.1 Modèle GARCH(p,q)

Définition 4 On considère un modèle autorégressif exprimé sous la forme :

$Y_t = \beta X_t + \varepsilon_t$ avec $\beta > 0$ et $\forall t \in \mathbb{R}_+$, ε_t est un bruit blanc faible qui satisfaisant la propriété $E[\varepsilon_t/I_{t-1}] = 0$

$$\varepsilon_t = \eta_t h_t$$

l'équation de la variance conditionnelle d'un processus GARCH s'écrit :

$$h_t = \sqrt{\alpha_0 + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^q \beta_j h_{t-j}^2} \quad (2.1)$$

où $\eta_t \sim N(0, \sigma^2)$, η_t est bruit blanc faible tel que $\alpha_0 > 0$, $\alpha_i > 0$ pour $i = 1, 2, \dots, p$ et $\beta_j > 0$ pour $j = 1, 2, \dots, q$ satisfaisantes pour garantir la positivité de $\alpha_0 + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^q \beta_j h_{t-j}^2$

2.4.2 GARCH faible

Drost & Nijman (1993)([8]) ont convenu d'appeler GARCH faible "weak GARCH" tout bruit blanc faible $\varepsilon_t = \eta_t h_t$ tel que :

- $E[\varepsilon_t/I_{t-1}] = 0$, $\forall t \in \mathbb{Z}$, c'est la propriété différence de martingale.
- Il existe des constantes α_i , $i = 0, 2, \dots, p$ et β_j , $j = 1, 2, \dots, q$ tel que

$$h_t^2 = V(\varepsilon_t/I_{t-1}) = \alpha_0 + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^q \beta_j h_{t-j}^2, \forall t \in \mathbb{Z}$$

2.4.3 GARCH semi-fort

Lorsque le processus d'innovation v_t et ε_t^2 est lui même supposé être un bruit blanc faible, alors ils appellent GARCH semi-fort "semi-strong" GARCH le même processus ε_t lorsqu'il s'agit d'une différence de martingale avec un processus d'innovation v_t qui est lui même une différence de martingale. Les processus GARCH semi-forts ainsi définis coïncident bien avec l'idée initiale de Engle et Bollerslev puisqu'il est clair réciproquement que si l'on suppose que v_t est une différence de martingale, on en déduit que :

$$v_t = \varepsilon_t^2 - h_t^2$$

où h_t^2 est bien la variance de ε_t conditionnelle à l'information passée.

2.4.4 GARCH fort

On dit que le processus est GARCH(p,q) fort dans le cas d'un GARCH semi-fort tel que l'innovation standardisée v_t soit un bruit blanc fort (suite des variables indépendantes et de même loi) et $\eta_t \sim N(0, 1)$.

Pour motiver l'introduction des processus GARCH, on peut réécrire l'équation (2.1) à l'aide des opérateurs $\alpha(\cdot)$ et $\beta(\cdot)$. Dans ce nouveau contexte ces opérateurs sont définis par :

$$\alpha(L) = \alpha_1 L + \alpha_2 L^2 + \dots + \alpha_p L^p$$

Et

$$\beta(L) = \beta_1 L + \beta_2 L^2 + \dots + \beta_q L^q$$

On peut donc écrire

$$\varepsilon_t = \eta_t \sqrt{\alpha_0 + \alpha(L)\varepsilon_t^2 + \beta(L)h_t^2}$$

où L est l'opérateur de retard.

On a donc :

$$h_t^2 = \alpha_0 + \alpha(L)\varepsilon_t^2 + \beta(L)h_t^2$$

Si toutes les racines de $1 - \beta(L)$ sont en dehors du cercle unité, on a :

$$h_t^2 = (1 - \beta(L))^{-1} \alpha_0 + (1 - \beta(L))^{-1} \alpha(L) \varepsilon_t^2$$

Tout comme pour le modèle ARCH, on peut par inversion exprimer le processus ε_t^2 sous la forme d'un processus ARMA définie dans une innovation

$$v_t = \varepsilon_t^2 - h_t^2$$

En introduisant cette notation dans l'équation (2.1) il vient :

$$\varepsilon_t^2 - v_t = \alpha_0 + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^q \beta_j (\varepsilon_{t-j}^2 - v_{t-j})$$

D'où l'on tire que :

$$\varepsilon_t^2 = \alpha_0 + \sum_{i=1}^n (\alpha_i + \beta_i) \varepsilon_{t-i}^2 - \sum_{j=1}^q \beta_j v_{t-j} + v_t, t \in \mathbb{Z}$$

$n = \max(p, q)$.

avec la convention $\alpha_i = 0 \forall i > p$, $\beta_j = 0 \forall j > q$

2.5 Propriétés des processus GARCH

Les propriétés théoriques des processus GARCH se déduisent de la même façon que nous avons développé les propriétés des processus ARCH.

Proposition 1 *Le processus ε_t est un bruit blanc. si $E(\varepsilon_t^2) < \infty$.*

on a

$$E[\varepsilon_t] = E[E(\varepsilon_t/I_{t-1})]$$

et

$$\begin{aligned} Cov(\varepsilon_t, \varepsilon_{t-k}) &= E(\varepsilon_t \varepsilon_{t-k}) \\ &= E[\varepsilon_{t-k} E(\varepsilon_t/I_{t-1})] = 0, \forall k > 0 \end{aligned} \tag{2.2}$$

Proposition 2 *Une condition nécessaire de l'existence de la variance d'un processus GARCH(p,q) est*

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

Remarque 2 *Si cette condition est vérifiée avec les contraintes de non négativité donnée ci-dessus, elle est également suffisante. Donc le processus GARCH est faiblement stationnaire ou stationnaire au second ordre.*

Dans le cas où l'inégalité précédente est saturée, c'est à dire que

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

on dira alors que le processus GARCH est intégré, et on parlera de processus IGARCH.

Proposition 3 *Le processus ε_t^2 d'une représentation GARCH(p,q) peut être représenté sous la forme d'un processus ARMA(max(p,q),q) définie dans une innovation*

$$v_t = \varepsilon_t^2 - h_t^2 \text{ tel que :}$$

$$\varepsilon_t^2 = \alpha_0 + \sum_{i=1}^n (\alpha_i + \beta_i) \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j v_{t-j} + v_t$$

$$n = \max(p, q)$$

avec la convention $\alpha_i = 0$, si $i > p$ et $\beta_j = 0$ si $j > q$.

2.5.1 Modèle GARCH(1,1)

Les données empiriques des marchés financiers sont souvent modélisées par le modèle avec erreur GARCH (1,1). Il est donné par l'équation

$$Y_t = \beta X_t + \varepsilon_t$$

avec $\varepsilon_t = \eta_t h_t$

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

avec $\alpha_0 > 0$, $\alpha_1 > 0$ et $\beta_1 > 0$. Dans ce modèle, les carrés des résidus suivent un processus ARMA(1,1)

$$\varepsilon_t^2 = \alpha_0 + (\alpha_1 + \beta_1) \varepsilon_{t-1}^2 + \beta_1 v_{t-1} + v_t$$

Il est stationnaire pour $\alpha_0 > 0$, $\alpha_1 > 0$, et $\beta_1 > 0$ où $v_t = \varepsilon_t^2 - h_t^2$ est un processus d'innovation pour ε_t^2 . Sous la condition de stationnarité de second ordre, la variance inconditionnelle du processus ε_t existe et est constante au cours du temps.

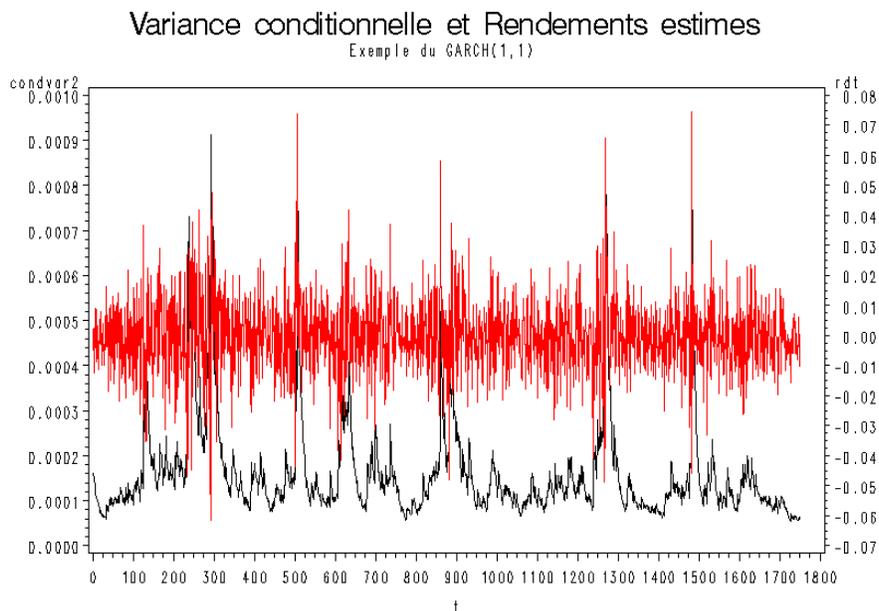


FIGURE 2.1 – variance conditionnelle et rendements estimés
exemple GARCH (1,1)

Sachant que $V(\varepsilon_t) = E(\varepsilon_t^2)$, il suffit à partir de la forme ARMA(1,1) sur ε_t^2 définir la variance du processus :

$$V(\varepsilon_t) = \frac{\alpha_0}{1 - (\alpha_1 + \beta_1)}$$

Selon Bollerslev T. (1986) ([6]), la kurtosis existe si

$$3\alpha_1^2 + 2\alpha_1\beta_1 + \beta_1^2 < 1$$

et est donnée par :

$$\begin{aligned} K_u &= \frac{E[\varepsilon_t^4]}{E^2[\varepsilon_t^2]} \\ &= 3 \frac{1 - (\alpha_1 + \beta_1)^2}{1 - (\alpha_1 + \beta_1)^2 - 2\alpha_1^2} \end{aligned} \tag{2.3}$$

Elle est toujours supérieure à trois. Ainsi, si α_1 tend vers zéro, l'hétéroscédasticité disparaît et la valeur de la kurtosis tend vers trois. Enfin, on peut montrer que pour un processus GARCH la kurtosis est directement liée à l'hétéroscédasticité conditionnelle.

Considérons le cas de la kurtosis associée à la loi non conditionnelle dans un processus GARCH conditionnellement gaussien tel que $\eta_t \sim N(0, 1)$.

Dans ce cas, les moments conditionnels d'ordre 2 et 4 du processus ε_t sont liés :

$$E[\varepsilon_t^4 / I_{t-1}] = 3[E(\varepsilon_t^2 / I_{t-1})]^2$$

En effet, on rappelle que si une variable centrée x suit une loi normale centrée, alors

$$E(x^4) = 3(V(x))^2 = 3(E(x^2))^2$$

Si on applique l'espérance sur les deux cotés de l'équation précédent, il devient

$$\begin{aligned} E(\varepsilon_t^4) &= E(E[\varepsilon_t^4 / I_{t-1}]) \\ &= 3E([E(\varepsilon_t^2 / I_{t-1})]^2) \\ &\geq 3E([E(\varepsilon_t^2 / I_{t-1})])^2 \\ &= 3E^2(\varepsilon_t^2) \end{aligned} \tag{2.4}$$

On peut calculer la kurtosis comme suit :

$$\begin{aligned}
 K_v &= \frac{E[\varepsilon_t^4]}{E^2[\varepsilon_t^2]} \\
 &= \frac{3E([E(\varepsilon_t^2/I_{t-1})]^2)}{E^2[\varepsilon_t^2]} \\
 &= 3\frac{E^2[\varepsilon_t^2]}{E^2[\varepsilon_t^2]} + \frac{3}{E^2[\varepsilon_t^2]}(E([E(\varepsilon_t^2/I_{t-1})]^2) - E^2[\varepsilon_t^2]) \\
 &= 3 + \frac{3}{E^2[\varepsilon_t^2]}(E([E(\varepsilon_t^2/I_{t-1})]^2) - E^2[E(\varepsilon_t^2/I_{t-1})]) \\
 &= 3 + 3\frac{V[E(\varepsilon_t^2/I_{t-1})]}{E^2[\varepsilon_t^2]} > 3
 \end{aligned} \tag{2.5}$$

La kurtosis est donc liée à une mesure de l'hétéroscédasticité conditionnelle.

2.6 Inférence statistique sur le Modèle GARCH

2.6.1 La méthode de maximum de vraisemblance

Pour comprendre cette approche, nous allons tout d'abord considérer le cas le plus simple d'un processus ARCH pour Y_t . Nous étudierons ensuite le cas des processus GARCH, et enfin des modèles de régression avec erreur GARCH.

L'estimateur des paramètres de modèle ARCH se base très souvent sur la maximisation de la fonction de vraisemblance. Nous supposons que le processus Y_t est conditionnellement gaussien. La vraisemblance associée à Y_t conditionnellement au passé I_{t-1} est donc

$$L(Y_t/I_{t-1}, \theta) = \frac{1}{h_t\sqrt{2\pi}} \exp\left(-\frac{(y_t - m_t(\theta))^2}{2h_t^2(\theta)}\right)$$

et dépend du vecteur $\theta = (\alpha_0, \dots, \alpha_p) \in R^p$

La fonction de vraisemblance de (y_1, y_2, \dots, y_T) conditionnelle est par conséquent

$$L(y_1, y_2, \dots, y_T, \theta) = \prod_{t=1}^T L(y_t/I_{t-1}, \theta)$$

L'estimateur est alors définie comme le vecteur $\widehat{\theta}_T = (\widehat{\alpha}_0, \dots, \widehat{\alpha}_p)$ qui maximise le logarithme de cette fonction vraisemblance :

$$\widehat{\theta}_t = \arg \max_{\theta} \log L(y_1, y_2, \dots, y_T, \theta)$$

2.6.2 Estimation des paramètres du modèle GARCH

La vraisemblance associée à Y_t conditionnellement au passé I_{t-1} s'écrit :

$$L(y_t/I_{t-1}, \theta) = \frac{1}{h_t \sqrt{2\pi}} \exp\left(-\frac{(y_t - m_t(\theta))^2}{2h_t^2(\theta)}\right)$$

mais cette fois, la variance h_t^2 suit un processus ARMA et dépend donc des valeurs passées de la variance conditionnelle h_1^2, \dots, h_t^2 . Ces valeurs n'étant pas observées en pratique, la maximisation en direct de la vraisemblance est rendue impossible. En pratique, on estime successivement les valeurs de h_1^2, \dots, h_t^2 avant de calculer la vraisemblance. Ainsi, pour un vecteur $\theta_0 = (\alpha_0^0, \dots, \alpha_p^0, \beta_0^0, \dots, \beta_q^0)$ fixé des paramètres, on calcule récursivement

$$\widehat{h}_s^2 = \alpha_0^0 + \sum_{i=1}^p \alpha_i^0 Y_{s-i}^2 + \sum_{j=1}^q \beta_j^0 \widehat{h}_{s-j}^2$$

avec la condition $Y_i = 0$, et $h_i^2 = 0$ si $i \leq 0$. On remplace donc la fonction de vraisemblance par :

$$L(y_t/I_{t-1}, \theta) = \frac{1}{\widehat{h}_t \sqrt{2\pi}} \exp\left(-\frac{(y_t - m_t(\theta))^2}{2\widehat{h}_t^2}\right)$$

et la fonction de vraisemblance total est :

$$L(y_1, y_2, \dots, y_t, \theta) = \prod_{t=1}^T L(y_t/I_{t-1}, \theta)$$

Cette fonction de vraisemblance peut être calculée pour différentes valeurs du vecteur θ_0 et sa maximisation livre l'estimateur de maximum de vraisemblance.

Définition 5 *Sous certaines conditions de régularité, l'estimateur du PMV est asymptotiquement convergent et normal.*

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

Où la matrice de variance covariance asymptotique de l'estimateur du PMV est calculée à partir de :

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

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

Où E_0 désigne l'espérance prise par rapport à la vraie loi.

Naturellement dans la pratique les matrices I et J sont directement estimées en remplaçant l'espérance E_0 par la moyenne empirique et le paramètre inconnu θ par son estimateur convergent $\hat{\theta}$. Ainsi, on utilise :

$$\hat{I} = \frac{1}{T} \sum_{i=1}^T \frac{\partial \log L(\theta)}{\partial \theta} \Big|_{\theta=\hat{\theta}} \frac{\partial \log L(\theta)}{\partial \theta'} \Big|_{\theta=\hat{\theta}}$$

$$\hat{J} = -\frac{1}{T} \sum_{i=1}^T \frac{\partial^2 \log(\theta)}{\partial \theta \partial \theta'} \Big|_{\theta=\hat{\theta}}$$

et la variance estimée de vérifie alors

$$Var[\sqrt{T}(\hat{\theta} - \theta)] = \hat{J}^{-1} \hat{I} \hat{J}^{-1}$$

Exemple 2 Appliquons cette formule au cas d'un modèle de régression linéaire avec erreur ARCH(p) :

$$\begin{cases} Y_t = \beta X_t' + \varepsilon_t \\ \varepsilon_t = \eta_t h_t(\theta) \end{cases}$$

avec $\eta \sim N(0, 1)$ i.i.d et

$$E(\varepsilon/I_{t-1}) = 0$$

$$V(\varepsilon/I_{t-1}) = \alpha_0 + \sum_{i=1}^p \alpha_i \varepsilon_{t-1}^2$$

Dans ce cas, on a donc la représentation de Gourieuroux (1992)([9])

$$\begin{cases} E(Y_t/I_{t-1}, X_t) = m_t(\theta) = X_t' \\ V(Y_t/I_{t-1}, X_t) = h_t^2(\theta) = \alpha_0 + \sum_{i=1}^p \bar{\alpha}_i (Y_{t-i} - cX_{t-i}')^2 \end{cases}$$

car $\varepsilon_{t-i} = Y_{t-i} - cX_{t-i}'$ par l'exemple Où $\theta = (c, \alpha_0, \alpha_1, \dots, \alpha_p) \in R^{P+2}$

La log-vraisemblance s'écrit :

$$\log L(\theta) = -\frac{T}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^T \log(\alpha_0 + \sum_{i=1}^p \alpha_i (Y_{t-i} - cX_{t-i}')^2)$$

$$- \frac{1}{2} \sum_{t=1}^T \frac{(Y_t - cX_t')^2}{\alpha_0 + \sum_{i=1}^p \alpha_i (Y_{t-i} - cX_{t-i}')^2}$$

Les estimateurs du MV sous l'hypothèse de normalité ou du PMV, notés $\widehat{\theta}_T$, des paramètres $\theta \in R^p$, satisfont un système non linéaire à p équations :

$$\frac{\partial \log L_T(\theta)}{\partial \theta} \Big|_{\theta=\widehat{\theta}} = 0$$

avec

$$\begin{aligned} \frac{\partial \log L_T(\theta)}{\partial \theta} \Big|_{\theta=\widehat{\theta}_T} = & - \sum_{t=1}^{mT} \frac{1}{h_t^2(\theta)} \frac{\partial h_t(\theta)}{\partial \theta} \Big|_{\theta=\widehat{\theta}_T} + \sum_{t=1}^T \frac{(Y_t - m_t(\theta))^2}{h_t^3(\theta)} \frac{\partial h_t(\theta)}{\partial \theta} \Big|_{\theta=\widehat{\theta}_T} \\ & + \sum_{t=1}^T \frac{Y_t - m_t(\theta)}{h_t^2(\theta)} \frac{\partial m_t(\theta)}{\partial \theta} \Big|_{\theta=\widehat{\theta}_T} \end{aligned}$$

Chapitre 3

Théorie du Risque

3.1 Introduction

Dans nos sociétés il y a beaucoup de risques, par ce que ce dernier occupe désormais la place centrale dans les politiques publique. On ne compte plus les activités traitées de risque pour la santé ou l'environnement. Aussi, compte tenu de l'évolution et changement dans le domaine financier.

La gestion de risque est une préoccupation majeure pour les gouvernements et les entreprises donnant lieu à des tentatives de mesure à tout niveau.

Cependant, les mesures de risque financier se manifestent explicitement dans beaucoup de différents types des problèmes d'assurances et de finance, elles se manifestent implicitement dans les problèmes impliquant le déficit et les probabilités de ruine. Dont le rôle important que peut jouer dans la prévision du comportement des marchés financiers. La mesure de risque est un problème actuariel important dans la politique de remboursement des sinistres.

Il existe plusieurs façons de mesurer le risque , la mesure la plus répandue est la Valeur-à-Risque (*VaR*) établie par JP Morgan. (1994)([10]) *VaR* comme mesure de risque est profondément.

3.2 Définition de la VaR

La *VaR* est un outil très répandu dans les marchés financiers dû à sa promesse implicite d'améliorer la gestion des risques en offrant une mesure complète des risques.

La Valeur-à-Risque a comme fonction de donner une information synthèse sur le risque d'un titre financier ou d'un portefeuille des titres financiers.

Elle se définit comme la perte maximale espérée à l'intérieur d'un horizon temporelle étant donné un niveau de confiance. L'horizon temporelle peut être journalier, hebdomadaire, semestriel, annuel... etc, et le niveau de confiance peut être 90%, 95%, 99%... etc. La dimension temporelle et le degré de confiance de la *VaR* sont généralement déterminés par le gestionnaire de portefeuille ou l'entreprise. Afin de donner un bon estimé de la *VaR* d'un point de vue méthodologique, on fait appel à des notions avancées de probabilité statistique et stochastique.

3.3 Exemple de valeur à risque (VaR)

La Valeur-à-Risque d'une position se calcule en estimant le montant de la perte potentielle, la probabilité de cette perte et la période à laquelle elle pourrait se produire.

Elle est généralement exprimée en pourcentage pour une période donnée. Par exemple, on pourrait dire qu'un actif a une valeur à risque de 1% à un horizon d'une semaine avec un intervalle de confiance de 2%. Cela signifie qu'il y a 2% de chances que l'actif perdra 1% de sa valeur au cours d'une semaine.

La Valeur-à-Risque peut également s'exprimer en valeur numérique. Par exemple, si la Valeur-à-Risque d'un portefeuille à un horizon d'un jour est de 1000 euro avec un intervalle de confiance de 5%, cela signifie qu'il y a 5% de chances que le portefeuille perde 1 000 euro de valeur en une journée.

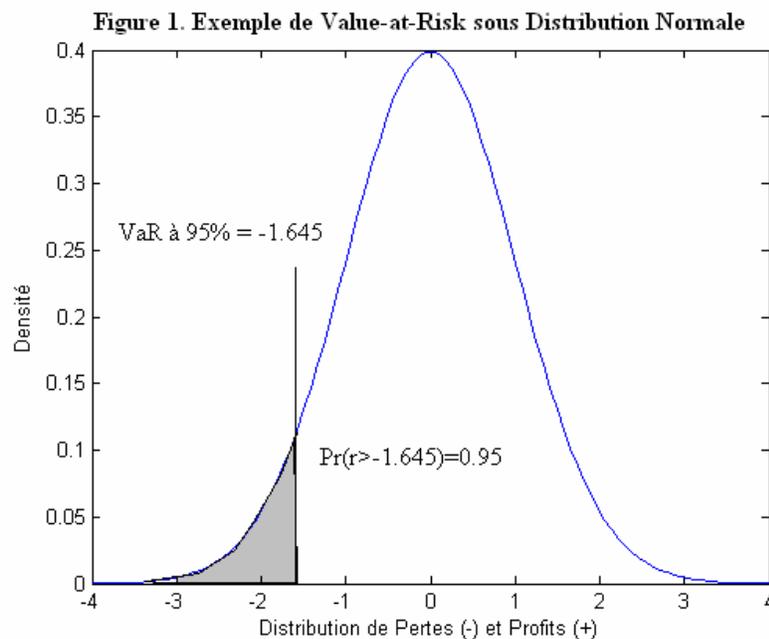


FIGURE 3.1 – *VaR* sous distribution normal

3.4 Aperçu général sur les mesures du risque

Il n'existe pas une définition bien précise du risque, bien que Mc Neil, Frey et Embrechts (2005)([11]) 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.

3.5 Mesure du risque Cohérente

Le risque d'un portefeuille mesure par VaR peut être supérieur à la somme des VaR de ses composantes.

Proposition 4 (*Caractérisations des mesures du risque cohérentes*)

Si ρ est une mesure du risque cohérente, alors il existe un ensemble des mesures de probabilité Q tel que :

$$\rho(X) = \sup_{Q \in \mathcal{Q}} \{\mathbb{E}_Q(X)\}$$

Selon Artzner et al. (1999)([12]) une mesure du risque est cohérente si elle adhère aux axiomes que nous énumérons maintenant .

Définition 6 (*Mesure Cohérente*)

Une mesure du risque $\rho : F \rightarrow R$ est cohérente si elle vérifiée ($X, Y \in F$) :

1-La sous additivité :

$$\rho(X + Y) \leq \rho(X) + \rho(Y)$$

plus généralement :

$$\rho(nX) = \rho(X + \dots + X) \leq n\rho(X), n = 1, 2, \dots$$

La fusion de deux centres de profit ne crée pas de risque supplémentaire. Au contraire, la diversification tend à réduire le risque global. Cette propriété permet ainsi une gestion décentralisée du besoin en capital dans les différents centres de profit sans courir le risque d'un besoin global supérieur à la somme des besoin individuels de chacun des certains niveaux requis de capital pourrait être incitée à se scinder artificiellement en deux entités afin de réduire son besoin en capital.

2. *L'homogénéité positive :*

$$\rho(\lambda X) \leq \lambda \rho(X) \text{ pour tout } \lambda \in \mathbb{R}_+.$$

plus généralement :

$$\rho(nX) = n\rho(X), n = 1, 2, \dots$$

De même qu'une fusion ne crée pas risque supplémentaire ($\rho(\lambda X) \leq \lambda \rho(X)$), une fusion sans diversification ne réduit pas le besoin global en capital.

3. *La monotonie :*

$$(X \geq Y) \Rightarrow \rho(X) \geq \rho(Y).$$

Si les pertes encourues avec le risque X sont toujours supérieures à celles obtenues avec Y , le besoin en capital pour X doit être supérieur à celui pour Y .

4. *L'invariance par translation :*

$$\rho(X + c) \leq \rho(X) - c$$

pour tout $c \in \mathbb{R}$. Spécialement, nous avons

$$\rho(X + \rho(X)) = \rho(X) - \rho(X) = 0$$

c'est-à-dire, en ajoutant $\rho(X)$ à la position initiale X , nous obtenons une position "neutre". On notera en particulier que $\rho(X) = -\alpha$ avec la convention que $\rho(0) = 0$.

5. *Convexe :*

La convexité implique que la diversification n'augmente pas le risque, car la valeur de risque de la portefeuille diversifiée

$$\lambda X + (1 - \lambda)Y$$

est inférieure ou égale à la moyenne pondérée des différentes valeurs de risque. ρ est dite convexe si pour tout $\lambda \in [0, 1]$

$$\rho(\lambda X + (1 - \lambda)Y) \leq \lambda \rho(X) + (1 - \lambda)\rho(Y)$$

Une mesure de risque ρ convexe et positivement homogène est dite cohérente.

3.5.1 Types du risque

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

1. **Risque de baisse** : Le risque de baisse de la valeur d'un titre se décompose en deux parties l'une dite systématique qui correspond au risque global du marché et l'autre spécifique au titre particulier dite non systématique. Nous verrons plus loin que le risque systématique est loin d'être négligeable puisqu'il est responsable de 73% à 90% du rendement moyen [à reformuler of p.105 de "Principes de Finance Moderne"].

La diversification au sein d'un portefeuille, si elle est bien menée permet d'éliminer le risque non systématique.

2. **Risque de liquidité** : Si un titre est peu liquide, il est fort probable qu'un décalage de cours important aura lieu au premier problème. Il faut toujours garder à l'esprit que pour qu'un échange ait lieu, il faut qu'il y ait accord entre deux parties. Autrement dit, si vous souhaitez vendre à un prix donné, il faut qu'une autre partie accepte d'acheter à ce prix. Si personne ne souhaite acheter, vous ne pouvez pas vendre ou vous devez accepter de vendre moins cher. Donc si des nombreuses personnes souhaitent vendre alors qu'aucun acheteur ne se présente ... le cours s'effondre. Le risque lié à liquidité est donc loin d'être négligeable.

3. **Risque de change** : Le risque de change est lié aux fluctuations des monnaies entre elles, donc aux investissements réalisés en monnaies étrangères. Ainsi, si vous achetez des valeurs sur le NYSE par exemple lorsque le dollar est en hausse, donc cher par rapport à l'euro, vous risquez de voir vos éventuelles plus-values partir en fumée, lorsque vous reconvertirez le fruit de votre vente de dollars en euro si le dollar venait de baisser.

3.5.2 Nature des risques :

1. **Risques de crédit** : Incapacité d'un débiteur (entreprise, institution financière, particulier...) à respecter ses engagements financiers vis-à-vis de sa banque.

2. **Risques Pays** : Défaillance d'un pays émergent ou incapacité d'une autre partie localisée dans un pays émergent de faire face à ses engagements en raison de facteurs (politique, économique, ...) échappant à son contrôle.

3. **Risques de marché** : Risque de perte suite à une évolution défavorable des paramètres de marché pouvant causer des impacts négatifs sur la position de la banque.

4. **Risques opérationnels** : Pertes résultant d'une inadéquation ou d'un échec au niveau des processus, des personnes, des systèmes (erreurs humaines, pannes systèmes, fraudes, litiges...) ou perte résultant d'événements externes (catastrophes naturelles, incendies, ...)

5. Risque de modèle : Risque de perte liée à l'utilisation d'un modèle inapproprié ou erroné pour la valorisation de positions ou le calcul d'indicateurs de risques.

3.5.3 Limites de la VaR

VaR en dépit de son universalité, plusieurs auteurs ont précisé ses déficiences. *VaR* comme mesure de risque est fortement critiquée :

1. *VaR* est difficile à optimiser. En outre, il est inadéquat d'employer la *VaR* dans la pratique en raison de sa non convexité elle possède beaucoup de minimum locaux. Ce qui peut avoir beaucoup d'extrémités locales, qui mènent au rang instable de risque.

2. *VaR* est un modèle de mesure de risque dépendante, parce que par définition elle dépend de la référence de probabilité initiale.

3. *VaR* n'est pas une mesure de risque cohérente, car elle n'obéit pas l'axiome de sous additivité proposée par Artzner et al (1999)([12]) et Acerbi et Tasche (2002)([13]) ce qui cause des contradictions, par conséquent il peut avoir un plus gros risque surgissant de la diversification. Ce résultat implique que l'agrégation des portefeuilles peut mener à une augmentation de risque.

4. La *VaR* ne mesure pas les pertes excédant le seuil de confiance, elle ignore toutes les informations concernant la queue de la distribution, et les propriétés statistiques de la perte significative au delà du seuil, par exemple elle ne s'inquiète de pas au risque de queue, ne nous indique rien au sujet de la taille potentielle de la perte qui l'excède, ne tient pas compte de la sévérité d'un événement encouru de dommages.

5. Pour les portefeuilles des titres, les modèles de *VaR* utilisés sont habituellement basés sur l'hypothèse (souvent implicite) que les rentabilités suivent une distribution normale.

Or, la normalité ne peut pas capturer l'épaisseur de la queue de distribution.

Sous des fluctuations extrêmes des prix des actifs ou une structure extrême de la dépendance des actifs, la *VaR* peut sous-estimer de manière significative le risque.

Yamai et Yoshihara (2002)([14]) prouvent que la *VaR* n'a pas de risque de queue quand les distributions sont de type elliptique.

La *VaR* est sous additive seulement si la distribution de la variable financière est normale (ou de manière générale elliptique), ce qui n'est pas souvent le cas même pour les rentabilités des actifs sur le marché.

3.5.4 Mesures Alternatives à la VaR

Comme alternatives à la VaR , il existe dans la littérature financière différentes mesures qui définissent le risque comme la moyenne de la queue de distribution des profits/pertes (TCE, ES, CVaR, ...). Ces différentes mesures fournissent le même niveau de risque dans le cas d'une distribution sous-jacente continue.

TCE

Artzner et al. (1999)([12]) ont proposé une mesure de risque appelée la Tail Conditional Expectation (CTE) qui est à la fois cohérent et mesurant la perte au-delà de la VaR :

$$TCE(X) = -E[X/X \leq -VaR(X)]$$

Dans le cas d'un saut en $x = -VaR(X)$, i.e $P(X = -VaR(X))$ positive, l'intervalle $]\infty, -VaR(X)]$ a une probabilité de γ^+ qui est supérieure à la probabilité $(1 - \alpha)$ sélectionnée pour la mesure de risque.

Cette mesure de la perte moyenne de la queue de distribution est parfois appelée Expected Shortfall (ES) ou Conditional VaR (CVaR). Cette multitude des noms donnés à cette mesure traduit le problème de la discontinuité de la distribution empirique. Dans ce cas, ces mesures aboutissent à des résultats différents.

ES

Acerbi et Tasche (2001)([13]) montrent que la CTE ne satisfait pas généralement la propriété de sous-additivité dans le cas d'une distribution discontinuité. Ils proposent l'ES (Expected Shortfall) comme mesure cohérente définie par :

$$ES(X) = -E[X/X < -VaR(X)]$$

Toutefois, cette mesure est définie sur l'intervalle $]\infty, -VaR(X)]$ qui a une probabilité égale à γ^- et qui est inférieure à $(1 - \alpha)$.

CVaR

Conditional VaR proposée par Rockafellar et Uryasev (2001)([15]) est définie en terme de la VaR et de l'ES pour donner une mesure définie exactement sur l'intervalle $(1 - \alpha)$ et qui satisfait l'ensemble des propriétés de cohérence définies par Artzner et al (1999)([12]) :

$$CVaR(X) = \lambda(x)VaR(x) + (1 - \lambda(x))ES(x)$$

et

$$\lambda(x) = \frac{\gamma - \gamma^-}{1 - \alpha}$$

TVaR

L'inconvénient de la VaR est qu'elle n'est pas une mesure cohérente du risque au sens de Artzner et al. (1999)([12]). En effet, elle n'est pas une mesure sous additive du risque. c'est à dire, si on considère deux risques X et Y alors pour tout $\alpha \in]0, 1[$ on peut avoir :

$$VaR_\alpha(X + Y) \geq VaR_\alpha(X) + VaR_\alpha(Y)$$

La $TVaR$ a l'avantage de prendre en compte les valeurs d'une variable aléatoire au-delà de la VaR . La Tail Valeur-à-Risque de niveau α associée au risque X est donnée par :

$$TVaR_\alpha = \frac{1}{1-\alpha} \int_\alpha^1 F^{-1}(\mu) d\mu$$

ou encore

$$TVaR_\alpha(X) = VaR_\alpha(x) + \frac{1}{1-\alpha} E[(X - VaR_\alpha(x))^+]$$

La quantité $E[(X - VaR_\alpha(X))^+]$ est l'expected shortfall de niveau de probabilité α est la perte moyenne au-delà de la VaR au niveau α . Si X représente la charge brute de réassurance alors $E[(X - VaR_\alpha(X))^+]$ correspond à la prime Stop-Loss pour un niveau de rétention égal à $VaR_\alpha(X)$.

3.6 Avantages et inconvénients de la valeur à risque (VaR)

3.6.1 – Avantages de la Value-at-Risk :

1. Elle tend à devenir un indicateur de risque largement utilisée par les établissements financiers car elle résume par une seule valeur tous les risques d'un portefeuille, quelle que soit leur nature (taux de change, actions, ...). Elle présente l'avantage d'être plus facile à comprendre par des investisseurs qui ne sont pas spécialistes en techniques de gestion de portefeuille ou de gestion de risque.

2. Les mesures traditionnelles du risque comme la déviation standard et le degré de sensibilité ne donnent pas une perception de l'ampleur des pertes possibles mais simplement une information sur le pourcentage de la déviation du prix ou du rendement de l'actif par rapport à sa moyenne. Par contre la VaR permet d'évaluer de manière quantitative la perte potentielle maximale qu'une entité financière peut subir à un niveau de probabilité donné et dans un temps donné.

3. Elle est probabiliste et fournit à un gestionnaire des risques l'information utile sur les probabilités associées avec montants spécifiques de perte.

3.6.2 – Inconvénients de la Value-at-Risk :

1. La VaR correspond à un quantile donné, elle ne prend pas en compte les risques au-delà de ce quantile.

2. Artzner et al. ([12]) ont montré que la VaR n'est pas une mesure de risque cohérente parceque elle ne satisfait pas la condition de sous-additivité : Si un portefeuille est composé de sous-portefeuilles A et B , alors la VaR du portefeuille total est inférieure à la somme des *VaR* des portefeuilles qui le composent.

Cette inégalité est notamment expliquée par la prise en compte de la diversification.

Chapitre 4

Estimation des Valeurs à Risque

4.1 Introduction

Le risque extrême de marché correspond à un risque de perte des portefeuilles des investisseurs dû aux variations des marchés financiers : marchés des instruments de base (actions, obligations, devises , matières premières), mais aussi marchés des produits dérivés (contrats à terme, options). Pour gérer le risque de marché, il faut donc mesurer de manière précise ce risque extrême. La mesure qui répond à celà est la Valeur-à-Risque (Value at Risk).

Mathématiquement, la notion de la Valeur-à-Risque se traduit ainsi :

$$Pr(\Delta V < VaR) = 1 - \alpha$$

Avec : ΔV = la variation de la valeur V du portefeuille sur la période de détention.

α = l'erreur

Plusieurs modèles ont été présentés pour l'estimation de la Valeur-à-Risque Manganeli & Engle (2001)([16]). L'élément clé qui distingue ces modèles est l'existence ou non d'une hypothèse de paramétrisation 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.

4.2 Définition de la VaR

Soit $0 < \alpha < 1$, F est la distribution de la perte X d'un investissement dans une période de temps donnée. Les valeurs typiques pour $1 - \alpha$ sont $1 - \alpha = 0,95$ et $1 - \alpha = 0,99$. La VaR est le quantile d'ordre $1 - \alpha$ définie par :

$$VaR_\alpha = F^{-1}(1 - \alpha)$$

où F^{-1} est la fonction inverse de F (dans le cas où F n'est pas inversible, on utilise l'inverse généralisé). Notons :

$$x_\alpha = VaR_\alpha = F^{-1}(1 - \alpha) = \inf\{x \in \mathbb{R} : F(x) \geq 1 - \alpha\}$$

4.3 Estimation de la VaR par 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 valeurs extrêmes (EVT) qui est différent de la théorie statistique habituelle fondée pour l'essentiel sur des raisonnements de type "tendance centrale". La théorie des valeurs extrêmes est utilisé pour estimer la valeur-à-risque.

4.3.1 VaR par la méthodes des blocs GEV

Supposons que l'échantillon des maximas suit exactement une loi GEV, Le quantile extrême x_p est la VaR .

En remplaçant le paramètre par son estimateur :

1. Approche par l'estimateur de Hill

L'estimation de la Valeur-à-Risque est donnée par la formule suivante :

$$\hat{x}_\alpha^H = \widehat{VaR}_\alpha = X_{n-k} + X_{n-k} \left[\left(\frac{k}{n}(1 - \alpha) \right)^{-\xi^H} - 1 \right]$$

Où ξ^H est l'estimateur de Hill, k nombre des points qui dépasses le seuil.

2. Approche par l'estimateur de Pickands

$$\widehat{VaR}_\alpha = x_\alpha^P = \frac{\left(\frac{k}{n(1-\alpha)}\right)^{\xi^P} - 1}{1 - 2^{-\xi^P}} (X_{(n-k+1,n)} - X_{(n-2k+1,n)}) + X_{(n-k+1,n)}$$

Où ξ^P est l'estimateur de Pickands, k nombre des points qui dépassent le seuil.

4.3.2 VaR par la méthode POT

La VaR n'est rien d'autre qu'un quantile extrême calculé à partir de la loi asymptotique des extrêmes (Distribution de Pareto Généralisée GPD), obtenue en modélisant les pertes (ou profits) extrêmes par la méthode des excès. Ainsi, la VaR_p correspondant au modèle GPD inconditionnel, pour un horizon donné T et une certaine probabilité p , ou de façon équivalente à un niveau de confiance typiquement $1 - \alpha = 0,95$ et $1 - \alpha = 0,99$ est :

$$\widehat{VaR}_\alpha = u + \frac{\widehat{\beta}}{\widehat{\xi}} \left[\left(\frac{k}{N_u} (1 - \alpha) \right)^{-\widehat{\xi}} - 1 \right]$$

Où $\widehat{\xi}$ et $\widehat{\beta}$ représentent les estimateurs des paramètres de la loi GPD, k nombre des points qui dépassent le seuil, n taille d'échantillon .

4.4 Estimation de VaR par le modèle GARCH

En vue de se rapprocher d'avantage de la réalité des marchés, i.e. la notion de persistance des marchés, on affine les modélisations par des approches de VaR dynamique, i.e. en faisant l'hypothèse que les rentabilités sont dynamiques, non pas constantes. Les modèles GARCH à volatilité conditionnelle fait partie de la classe des méthodes paramétriques de calcul de la VaR . Cette famille des modèles suppose que les variances conditionnelles des rendements sont auto-corrélées, ce qui permet de tenir compte du phénomène de persistance (ou clustering) des volatilités déjà mentionnées.

Dans ce qui suit, on finance il est suffisant de modéliser les volatilités par un modèle GARCH(1,1), l'espérance conditionnelle du rendement L_t et sa variance conditionnelle sont modélisés comme :

$$L_t = \mu + \rho \cdot L_{t-1} + \varepsilon_t$$

avec

$$\begin{aligned} \varepsilon_t &= \eta_t \times \sqrt{h_t} \\ \varepsilon_t^2 &= \alpha_0 + \alpha_1 \cdot \varepsilon_{t-1}^2 + \beta_1 \cdot h_{t-1} + N_t \dots \dots \dots (**) \end{aligned}$$

et donc :

$$h_t = E(\varepsilon_t^2 | I_{t-1}) = \alpha_0 + \alpha_1 \cdot \varepsilon_{t-1}^2 + \beta_1 \cdot h_{t-1} + N_t$$

N_t est réputé être un pur bruit (généralement gaussien), η_t est supposé être distribué selon $N(0,1)$, ε_t est alors un variable aléatoire d'espérance nulle, de variance non conditionnelle stable mais de variance conditionnelle instable (équation (**)) et L_t représente l'actif à l'instant t . On peut écrire que conditionnellement à l'information I_{t-1} disponible jusqu'à $t-1$:

$$\varepsilon_t | I_{t-1} \sim N(0, h_t).$$

L'estimation de la $VaR_{t|t-1}$ à partir d'un modèle GARCH(1,1) passe par l'estimation des paramètres du modèle sur la fenêtre $[t-T+1, t]$. Sur la base des valeurs estimées des coefficients, il est possible de construire à la date t , une prévision sur ce que devrait être la variance du rendement en $t-1$. C'est cette prévision sur la variance future du rendement qui intervient dans le calcul de la $VaR_{\alpha,t}$, Celle-ci est calculée comme :

$$\widehat{VaR}_{\alpha,t} = \widehat{L}_{t|t-1} + F^{-1}(\alpha, \widehat{\eta}) \times \sqrt{\widehat{h}_t}$$

Où $F^{-1}(\cdot)$ représente la fonction de répartition inverse de la loi choisie pour Z_t et des paramètres estimés $\widehat{\eta}$, α est l'ordre du quantile choisi (par exemple $1 - \alpha = 0.99$) et $\widehat{L}_{t|t-1}$ est la prévision en t sur le rendement en $t-1$, tel que :

$$\widehat{L}_{t|t-1} = \widehat{\mu} + \widehat{\rho} L_t$$

Chapitre 5

Simulations et Applications

5.1 Introduction

La *VaR* (Value-at-Risk) est très utilisée dans les milieux financiers. Elle permet de mesurer la perte maximale liée à une position de marché, sur une fenêtre de temps particulière et pour un niveau de confiance $(1-\alpha)$ donné en supposant que nous connaissons la loi du rendement du portefeuille. Dans le cadre d'une probabilité égale à $(1-\alpha)=95\%$, cela signifie simplement que dans 95% des cas, la perte ne devrait pas dépasser cette mesure de risque.

De plus la *VaR* repose sur l'idée de coupler le montant des pertes (quantification brute) et la fréquence des pertes (prise en compte du risque).

5.2 Simulation

Nous utilisons dans ce chapitre le logiciel R et les packages `evir`, `fGARCH`, `actuar` et `qrmdata`.

5.2.1 La VaR avec Lois Extrêmes (TVE)

La théorie des valeurs extrêmes donne deux méthodes principales, la première est la méthode Blocs Maximas (BM) et la deuxième est la méthode de Peaks Over Threshold (POT).

1-Estimation des paramètres de la loi GEV et GPD

Nous avons utilisé les packages `evir` et `actuar` pour estimer les paramètres de la loi GEV et la loi GPD par la méthode du maximum de vraisemblance.

Les résultats sont résumés dans les tableaux suivants :

Tableau 4.1 Tableau des paramètres estimés de la loi GEV.

Paramètre	$\mu=0$	$\xi=0.4$	$\beta=1$
Paramètre estimés	-0.03643969	0.35977865	0.98948516

Tableau 4.2 Tableau des paramètres estimés de la loi GEV.

Paramètre	$\mu=0$	$\xi=0.5$	$\beta=1$
Paramètre estimé	0.01364849	0.51515408	1.00651427

Tableau 4.3 Tableau des paramètres estimés de la loi GEV.

Paramètre	$\mu=0$	$\xi=0.7$	$\beta=1$
Paramètre estimé	0.06002841	0.69565317	1.04786965

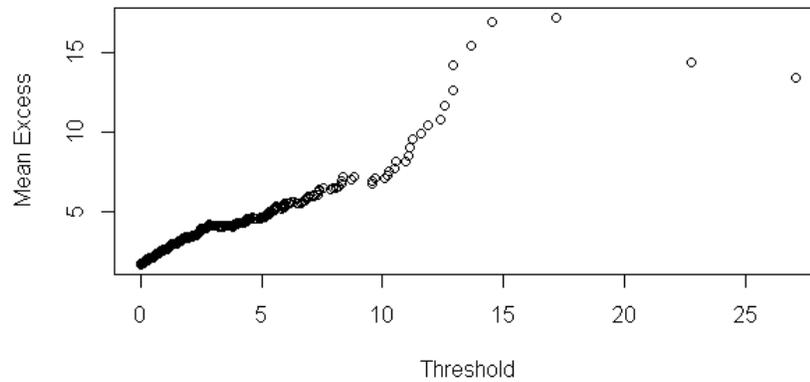


FIGURE 5.1 – La distribution moyenne des excès.

Tableau 4.4 Tableau des paramètres estimés de la loi GPD.

paramètre	$\xi=0.4$	$\beta=1$
Paramètre estimé	0.4027475	0.9719541

Tableau 4.5 Tableau des paramètres estimés de la loi GPD.

paramètre	$\xi=0.5$	$\beta=1$
paramètre estimé	0.5289681	0.9959586

Tableau 4.6 Tableau des paramètres estimés de la loi GPD.

paramètre	$\xi=0.7$	$\beta=1$
Paramètre estimé	0.6941242	0.9894211

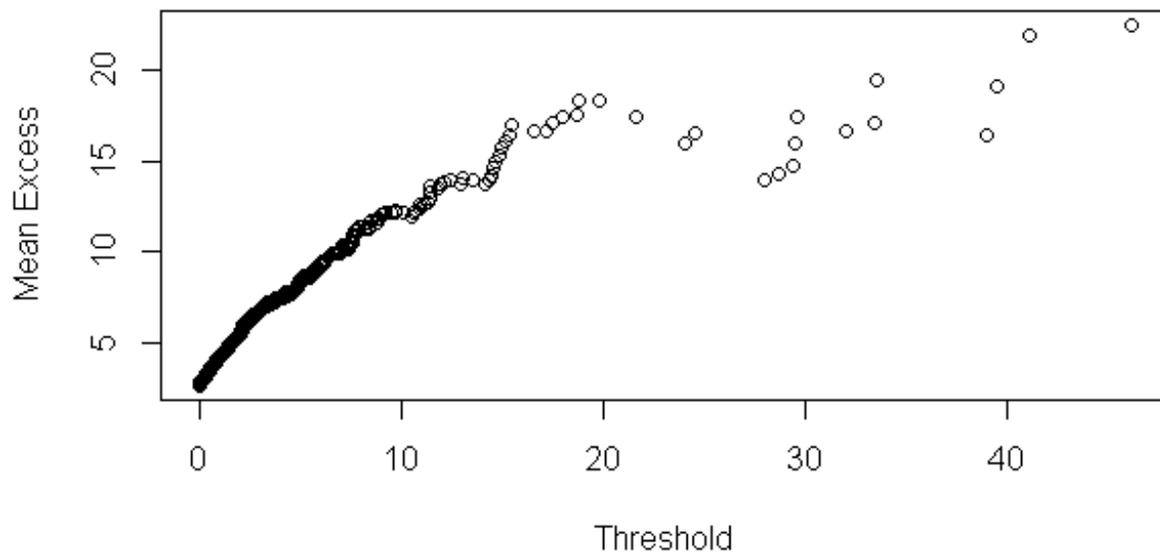


FIGURE 5.2 – La distribution moyenne des excès.

2-Estimation du paramètre ξ par la méthode de Hill

Nous avons étudié la performance de la méthode de Hill pour l'estimation du paramètre de forme de la queue des distribution. nous avons simulé des échantillons des lois GEV, GPD et loi de Pareto. Les résultats sont résumés dans le tableau ci-dessous. Nous remarquons que la méthode Hill estime bien le paramètre ξ .

Tableau 4.7 Tableau Estimation du paramètre ξ par la méthode de Hill.

Paramètre ξ	Distribution	0.4	0.5	0.7
Paramètre estimé	GEV	0.4174648	0.46	0.9020835
	GPD	0.4277611	0.4460523	0.821
	Pareto	0.4930951	0.4699288	0.9277265

3-Estimation de la VaR des échantillons simulés

Nous avons simulé des échantillons des lois GEV, GPD pour différentes valeurs des paramètres μ , ξ et β pour des échantillon de taille 1000 répétés 100 fois.

Nous avons utilisé l'expression de la VaR par la loi GPD pour l'estimer.

$$\widehat{VaR}_\alpha = \mu + \frac{\widehat{\beta}}{\widehat{\xi}} \left[\left(\frac{k}{n}(1 - \alpha) \right)^{-\widehat{\xi}} - 1 \right]$$

Les résultats obtenu sont résumé dans les tableaux ci-dessous.

Nous remarquons que les valeurs de la VaR pour les différents lois sont très proches pour les mêmes valeurs du paramètre de l'indice du queue de loi ξ .

Tableau 4.8 Valeur de la VaR pour $\xi = 0.4$.

VaR	$\mu=0$	$\mu=3$	$\mu=7$
GEV	-2.395803	0.4805403	4.501695
GPD	-2.509887	0.4798388	4.521018

Tableau 4.9 Valeur de la VaR pour $\xi = 0.5$.

VaR	$\mu=0$	$\mu=3$	$\mu=7$
GEV	-2.072345	1.06026	4.991657
GPD	-2.096483	0.9741262	5.00613

Tableau 4.10 Valeur de la VaR pour $\xi = 0.7$.

VaR	$\mu = 0$	$\mu = 3$	$\mu = 7$
GEV	-1.433524	1.558484	5.594281
GPD	-1.438656	1.541656	5.561875

4-Estimation de la VaR des modèles GARCH

Nous avons simulé un modèle ARCH(2) (GARCH(2,0)), un modèle GARCH(1,1), GARCH(1,2), et ARMA (1,1)+GARCH(1,1). Ensuite, nous avons calculé la VaR dans les différents modèles cités ci-dessus au niveau significatif de 95%. Nous avons simulé des échantillons de taille 10000 avec 100 répliques. La méthode du maximum de vraisemblance est utilisée pour estimer les paramètres des modèles. Nous avons pris 500 plus grandes valeurs maximales des données (5%). Les résultats obtenues sont résumés dans les tableaux suivants :

Tableau 4.11 Estimation de la VaR dans le cas d'un modèle ARCH(2) de paramètre $(\alpha = 0.2, \beta = 0.4)$.

Paramètre de la loi GPD des résidus	Pour 50 valeurs	Pour 500 valeurs
μ	$1.96e - 6$	$1.96e - 6$
ξ	- 3.92	- 2.9
β	0.05	0.04
β_0	$1.03e - 6$	$1.03e - 6$
α_1	0.214	- 0.214
α_2	0.417	0.417
VaR Résidu	122.44	$-1.96e - 6$
VaR (GARCH)	436.03	$-8.97e - 6$

Tableau 4.12 Estimation de la VaR dans le cas d'un modèle GARCH(1,1) de paramètre $(\alpha = 0.2, \beta = 0.7)$.

Paramètre de la loi GPD des résidus	Pour 50 valeurs	Pour 500 valeurs
μ	$-4.46e - 5$	$-4.46e - 5$
ξ	- 8.9	- 5.18
β	0.16	0.09
β_0	$1.28e - 6$	$1.28e - 6$
α_1	0.214	- 0.216
β_1	0.58	0.58
VaR Résidu	14958659	$-433.46e - 5$
VaR (GARCH)	54829734	0.000019

Tableau 4.13 Estimation de la VaR dans le cas d'un modèle GARCH(1,2) de paramètre $(\alpha = 0.2, \beta_1 = 0.2, \beta_2 = 0.7)$.

Paramètre de la loi GPD des résidus	Pour 50 valeurs	Pour 500 valeurs
μ	$4.46e - 6$	$4.46e - 6$
ξ	- 8.9	- 5.18
β	0.16	0.097
β_0	$1.28e - 6$	$1.28e - 6$
α_1	0.218	- 0.218
β_1	0.58	0.58
β_2	0.08	0.08
VaR Résidu	14958659	$-4.46e - 5$
VaR (GARCH)	49379470	0.00019

Tableau 4.14 Estimation de la VaR dans le cas d'un modèle ARMA(1,1) de paramètre $(ar_1 = 0.5, ma_1 = 0.3) + \text{GARCH}(1,1)$ de paramètre $(\alpha = 0.1, \beta = 0.7)$.

Paramètre de la loi GPD des résidus	Pour 50 valeurs	Pour 500 valeurs
μ	$-8.73e - 6$	$-8.73e - 6$
ar_1	0.502	0.502
ma_1	0.286	0.286
β_0	$1.15e - 6$	$1.15 * 10^{-6}$
α_1	0.11	0.11
β_1	0.77	0.77
seuil	0.0052	0.005
ξ	-0.05	-0.039
β	0.0019	0.001
VaR Résidu	0.00045	0.005
VaR (GARCH)	INF	INF

Tableau 4.15 Estimation de la VaR dans le cas d'un modèle AR(1) + GARCH(1,1) de paramètre ($\alpha = 0.1, \beta = 0.7$).

Paramètre de la loi GPD des résidus	Pour 50 valeurs	Pour 500 valeurs
μ	$2.13e - 6$	$2.13e - 6$
ar_1	0.66	0.66
β_0	$1.15e - 6$	$1.15*^{-6}$
α_1	0.0105	0.0105
β_1	0.782	0.782
seuil	0.008	0.0052
ξ	-0.005	-0.25
β	0.0019	0.0024
VaR Résidu	0.00045	0.0052
VaR (GARCH)	INF	INF

Tableau 4.16 Estimation de la VaR dans le cas d'un modèle de GARCH(1,1) de paramètre ($\alpha = 0.1, \beta = 0.7$).

Paramètre de la loi GPD des résidus	Pour 50 valeurs
μ	$1.15e - 4$
β_0	$8.78e - 6$
α_1	0.484
β_1	0.03
seuil	0.0068
ξ	-0.42
β	0.002
VaR Résidu	0.00068
VaR (GARCH)	0.015

5.3 Données réelles :

5.3.1 SP500 :

SP500 est un indice boursier basé sur 500 grandes sociétés cotées sur les bourses aux États-Unis. L'indice est possédé et géré par Standard & Poor's, l'une des trois principales sociétés de notation financière. Il couvre environ 80% du marché boursier américain par sa capitalisation. C'est un indice sans dividendes.

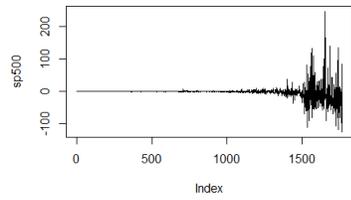
L'indice SP500 a été créé le 4 mars 1957. Il a détrôné le Dow Jones Industrial Average comme indice le plus représentatif du marché boursier américain parce qu'il est composé d'un plus grand nombre de compagnies et que sa valeur tient compte de la capitalisation boursière des compagnies contenues dans l'indice. De son côté, le Dow Jones Industrial Average est basé sur seulement 30 compagnies. La pondération des valeurs au sein du Dow ne s'effectue ni en fonction des capitalisations boursières, ni du flottant (comme pour les indices français), mais en fonction des cours de bourse. Une variation d'un dollar dans la valeur de la plus petite compagnie de l'indice a le même impact sur l'indice qu'une variation d'un dollar dans la valeur de la plus grosse compagnie.

En 1976, les entreprises industrielles représentaient 86% de la capitalisation boursière totale du S&P 500. En 2016 après quatre décennies de ralentissement de l'activité manufacturière et la montée de la finance comme secteur économique majeur, elles ne représentent que 10%.

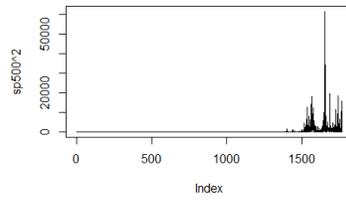
Le calcul du S&P 500 se fait en pondérant la valeur de la société sur le marché sur la valeur totale de toutes les sociétés du marché.

Tableau 4.17 Tableau descriptive de la série de rendement du SP500.

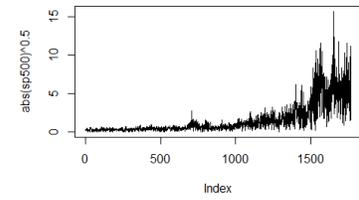
Min	1st Qu	Median	Mean	3rd Qu	Max	kurtosis	skewness
-125.460	-0.860	-0.080	-1.493	0.250	248.150	33.71843	1.711954



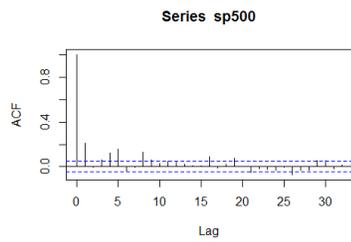
(a) La série de rendement du SP500



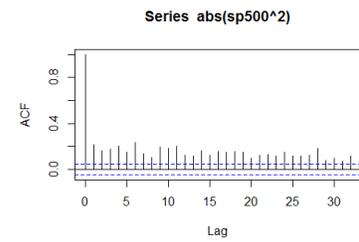
(b) Le carré de rendement du SP500



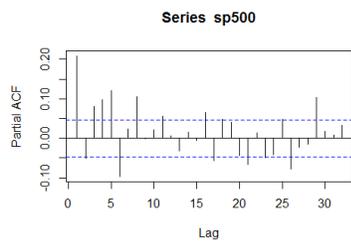
(c) La racine de rendement du SP500



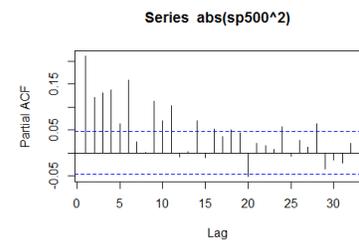
(a)



(b)



(c)



(d)

FIGURE 5.4 – Autocorrélation et Autocorrélation partielle de la série de rendement du SP500

Nous remarquons des graphes de la série n'est pas stationnaire et elle est confirmée par l'analyse des autocorrélation et autocorrélation partielle.

Le p-value du test Box-Ljung ≤ 0.05 , on accepte que les résidus sont autocorrélés (Voir Annexe 1).

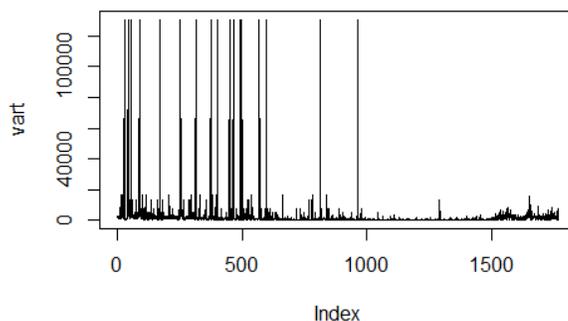


FIGURE 5.5 – La VaR de rendement du SP500

Tableau 4.18 Estimation d'indice SP500 par modèles GARCH (1.1).

Paramètre de la loi GPD des résidus	Pour 50 valeurs
μ	$2.25e - 13$
ω	8.27
α_1	1
β_1	10^{-8}
seuil	-156.27
ξ	-0.2947
β	988.80
VaR residu	113.4459
VaR (GARCH)	2917.62
Max(VaR)	130802
Min(VaR)	3.303515

Tableau 4.19 Tableau des différent AIC pour différent modèles.

AIC	$ma(3)$	$ar(3)$	$GARCH(1,1)$	$ARMA(3,3)$
	15230.95	26237.92	3.580722	15161.79

Les résultats obtenus montrent que le modèle GARCH(1,1) est meilleur que les deux autres modèles car son AIC est le plus faible.

Nous remarquons que le VaR de modèle GARCH est plus grand que le VaR calculé par le GPD.

5.3.2 INTEL :

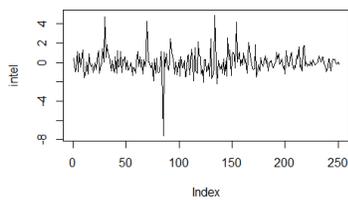
INTEL corporation est une entreprise américaine fondée le 18 juillet 1968 par Gordon Moore, Robert Noyce et Andrew Grove. Elle est le second fabricant mondial de semi-conducteurs après Samsung si on se fonde sur le chiffre d'affaires. Elle fabrique des microprocesseurs, c'est d'ailleurs elle qui a créé le premier microprocesseur X86, des cartes mères, des mémoires flash et des processeurs graphiques notamment.

Intel est cotée au Nasdaq sous le sigle INTC. Sa capitalisation boursière s'élève à 254,88 milliards de dollars (en date d'octobre 2018) avec pour principal concurrent Advanced Micro Devices (AMD).

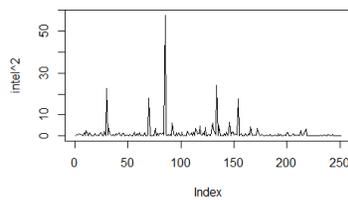
À compter de février 2021, l'entreprise est dirigée par Pat Gelsinger, jusqu'alors PDG de VMWare.

Tableau 4.20 Tableau descriptive de la série de rendement du INTEL.

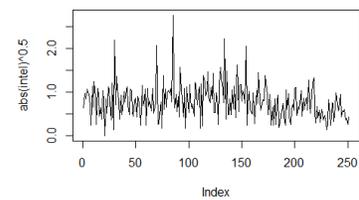
Min	1st Qu	Median	Mean	3rd Qu	Max	kurtosis	skewness
-7.60000	-0.62500	-0.07000	-0.01829	0.43500	4.92000	10.35384	-0.0501078



(a) La série de rendement du INTEL



(b) Le carré de rendement du INTEL



(c) La racine de rendement du INTEL

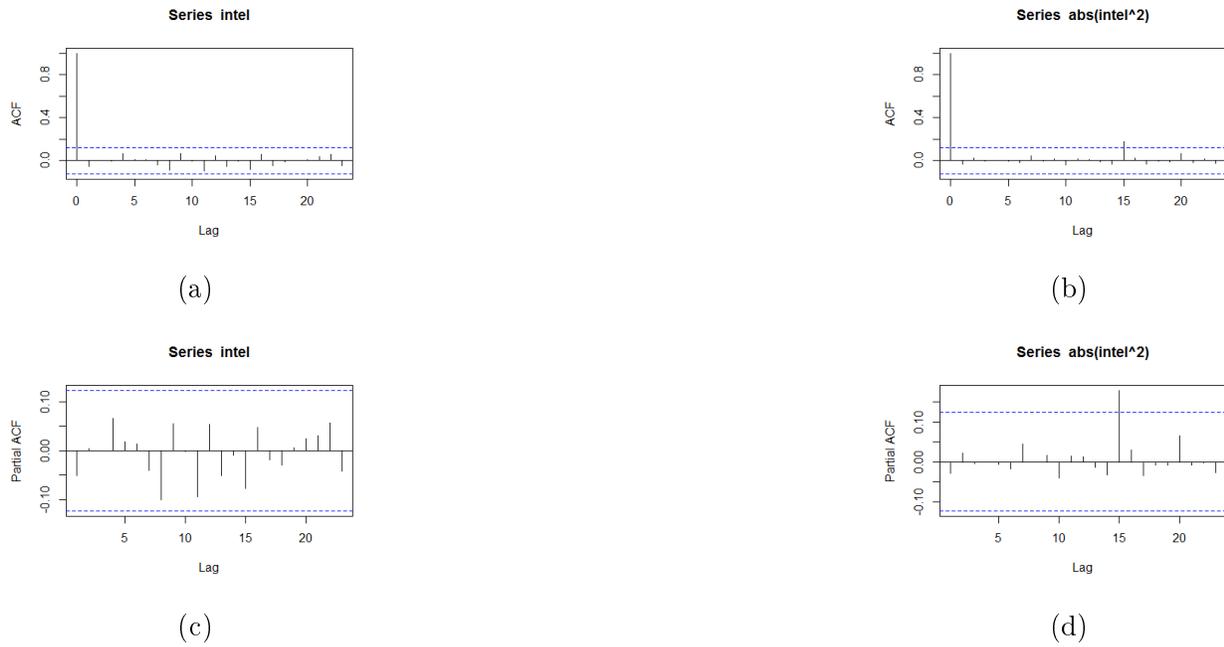


FIGURE 5.7 – Autocorrélation et Autocorrélation partielle de la série de rendement INTEL

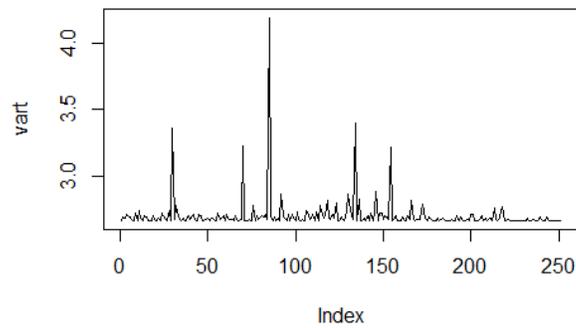


FIGURE 5.8 – La VaR de rendement du INTEL

Nous remarquons des graphes de la série n'est pas stationnaire et elle est confirmée par l'analyse des autocorrélation et autocorrélation partielle.

Le p-value du test Box-Ljung ≤ 0.05 , on accepte que les résidus sont autocorrélés (Voir Annexe 2).

Tableau 4.21 Estimation d'INTEL par modèle GARCH(1,1).

Paramètre de la loi GPD des résidus	Pour 50 valeurs
μ	55.03
β_0	0.68
α_1	1
β_1	10^{-8}
seuil	1.36
ξ	-0.489
β	6.13
VaR residu	2.706983
VaR (GARCH)	2.670068
Max (VaR)	4.19414
Min (GARCH)	2.663789

Tableau 4.22 Tableau des différent AIC pour différent modèles.

AIC	<i>ma</i> (3)	<i>ar</i> (3)	<i>GARCH</i> (1, 1)	<i>ARMA</i> (3, 3)
	21159.69	785.93	2.771714	789.39

Les résultats obtenus montrent que le modèle GARCH(1,1) est meilleur que les deux autres modèles car son AIC est le plus faible.

Nous remarquons que le VaR de modèle GARCH est plus grand que le VaR calculé par le GPD.

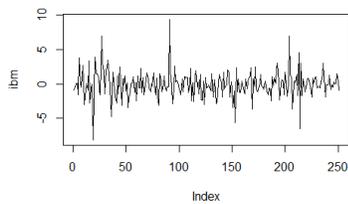
5.3.3 IBM :

IBM international Business Machines Corporation, connue sous le sigle IBM, est une société multinationale américaine présente dans les domaines du matériel informatique, du logiciel et des services informatiques.

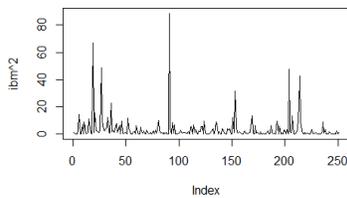
La société est née le 16 juin 1911 de la fusion de la Computing Scale Company et de la Tabulating Machine Company sous le nom de Computing Tabulating Recording Company (CTR). Celle-ci a changé de nom pour devenir International Business Machines Corporation le 14 février 1924. On lui prête le surnom de Big Blue en référence au bleu sombre, couleur longtemps associée à l'entreprise. Dans les années 1970 et les années 1980, IBM était la première capitalisation boursière au monde.

Tableau 4.23 Tableau descriptive de la série de rendement du IBM.

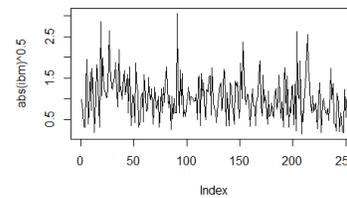
Min	1st Qu	Median	Mean	3rd Qu	Max	kurtosis	skewness
-8.19000	-1.08500	-0.09000	-0.06944	0.87000	9.42000	4.635905	0.34125



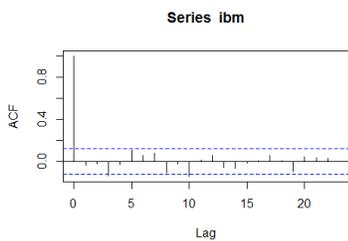
(a) La série de rendement du IBM



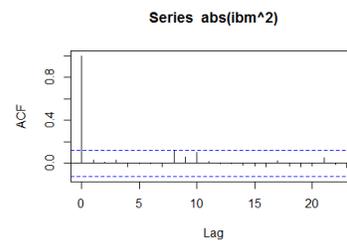
(b) Le carré de rendement du IBM



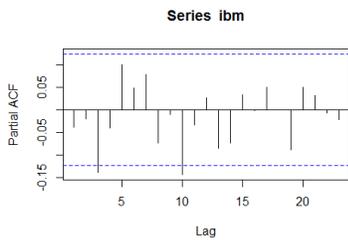
(c) La racine de rendement du IBM



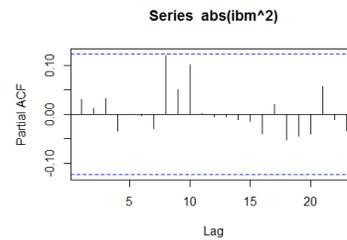
(a)



(b)



(c)



(d)

FIGURE 5.10 – Autocorrélation et Autocorrélation partielle de la série de rendement du IBM

Nous remarquons des graphes de la série n'est pas stationnaire et elle est confirmée par l'analyse des autocorrélation et autocorrélation partielle.

Le p-value du test Box-Ljung ≤ 0.05 , on accepte que les résidus sont autocorrélés (Voir Annexe 3).

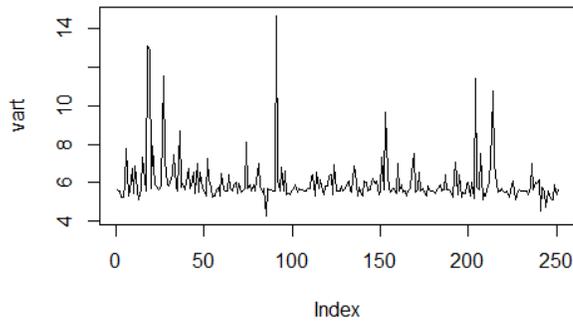


FIGURE 5.11 – La VaR de rendement du IBM

Tableau 4.24 Estimation d'IBM par modèle GARCH(1,1).

Paramètre de la loi GPD des résidus	Pour 50 valeurs
μ	$1.109e - 16$
β_0	$1.062e - 1$
α_1	1
β_1	10^{-8}
seuil	0.283
ξ	0.04
β	1.28
VaR residu	4.412589
VaR (GARCH)	5.642781
Max (VaR)	14.69751
Min (GARCH)	4.26218

Tableau 4.25 Tableau des différent AIC pour différent modèles.

AIC	$ma(3)$	$ar(3)$	$GARCH(1, 1)$	$ARMA(3, 3)$
	1035.82	4.138867	3.936395	1036.84

Les résultats obtenus montrent que le modèle GARCH(1,1) est meilleur que les deux autres modèles car son AIC est le plus faible.

Nous remarquons que le VaR de modèle GARCH est plus grand que le VaR calculé par le GPD.

Conclusion générale

Les sociétés financières, les banques et les compagnies d'assurances s'exposent à des risques importants de diverses natures. La gestion de risque est une préoccupation majeure pour les gouvernements et les entreprises donnant lieu à des tentatives de mesure du risque à tout niveau.

Le risque du marché extrême est un type important de risque financier, qui est généralement causé par des mouvements de prix extrêmes sur le marché financier. Bien que ce risque se produise avec des faibles probabilités, il peut entraîner des conséquences désastreuses sur le marché en engendrant des pertes financières importantes. Les catastrophes financières graves qui se sont produites dans le passé sont des preuves criantes : Les graves catastrophes financières qui se sont produites dans le passé, comme le fameux krach boursier, communément appelé lundi noir, qui a eu lieu aux États-Unis en octobre 1987, la crise asiatique de 1997-1998 et la crise des sous primes de 2007-2009.

Il existe plusieurs façons de mesurer le risque, la mesure la plus répandue est la Valeur à-Risque (VaR) établie par JP Morgan (1994)([10]). Elle permet de mesurer la perte maximale liée à une position de marché, sur une fenêtre de temps particulière et pour un niveau de confiance $(1-\alpha)$ donné en supposant que nous connaissons la loi du rendement du portefeuille.

De plus la VaR repose sur l'idée de coupler le montant des pertes (quantification brute) et la fréquence des pertes (prise en compte du risque).

Par ailleurs, la modélisation ARMA se révèle insuffisante pour les séries à cause de non normalité des résidus, cette série risque d'être caractériser par une dynamique non linéaire et une volatilité variable au cours du temps. Ce conditionnement nous a amené à entamer une modélisation non linéaire (les modèles Autorégressif Conditionnellement Hétéroscidastique) qui nous a donné un modèle ARMA (p,q) avec erreur GARCH(1,1).

Enfin, une approche innovante est utilisée, dans laquelle GARCH est combiné à EVT-POT en utilisant la procédure en deux étapes de McNeil (1998).

Des méthodes statistiques sont utilisées pour évaluer la performance de prévision de la Var des modèles à effet GARCH. Dans cette étude, il s'avère que les modèles GARCH sont assez performants avec toutes les résidus. Les techniques EVT (tant conditionnelles ou inconditionnelles) sont plus performantes que les approches GARCH. l'EVT inconditionnelle est la plus performante. Les applications faites sur les données financières (SP500, INTEL et IBM), ont montré que le modèle à effet GARCH(1,1) ajuste mieux ces dernières et que la prévision de la Var est plus efficace que la Var estimée par la loi GPD des résidus car elle prend en considération les données historiques des séries financières. Parmi les perspectives de ce travail sont la prédiction de la Var par l'approche bayésienne et dans le cas multivarié.

Bibliographie

- [1] Pickands (1975), Statistical Inference Using Extreme Order Statistics, *The Annals of Statistics* 3, p 119-131.
- [2] Fisher-tippet (1928), Limiting Forms of The Frequency Law of The Largest or Smallest Member of a Sample, *Proceedings of The Cambridge Philosophical Society* 24, p 180-190.
- [3] Smith (1987), Estimating Tail of Probability, *The Annals of Statistics*, p 1174-1207.
- [4] Davison & Smith (1990), Models for Exceedances Over High Thresholds, *Journal of Royal Statistical Society Ser B* 52, p 393-442.
- [5] Reiss & Thomas (2001), *Statistics Analysis of Extreme Value With Application to Assurance, Finance, Hydrology and other Fields*, Basel, Birkhauser, Verlag.
- [6] Bollerslev (1986), Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics* 31, p 307-327.
- [7] Engle (1982), Autoregressive Conditional Heteroscedasticity With Estimates of The Variance of UK Inflation *Econometrica* 50, p 987-1008.
- [8] Drost & Nijman (1993), Temporal Aggregation of GARCH Processes *Econometrica* 61, p 909-927.
- [9] Gouriéroux (1992), *Modèles ARCH et Application Financières* *économica* 35.
- [10] JP Morgan (1994), *Risk Metrics Technical Document* New York, Morgan Guaranty Trust Company October.
- [11] Mc Neil, Frey & Embrechts (2005), *Quantitative Risk Management*, édition Springer.

- [12] Artzner & al (1999), Coherent Measures of Risk, Journal of Mathematical Finance.
- [13] Acerbi & Tasche (2001), On The Coherence of Expected Shortfall, Journal of Banking and Finance Vol 26, p 1487-1503.
- [14] Yamai & Yoshiba (2002), Comparative Analysis of Expected Shortfall and Value at Risk Their Estimateion Error, Decomposition and Optimization , Monetary and Economic Studies 20, p 57-86.
- [15] Rockafellar & Uryasev (2001), Conditional Value-at-Risk for General loss distributions, Journal of Banking and Finance 26, P 1443–1471.
- [16] Manganelli & Engle (2001), Value-at-Risk Models in Finance, European Central Bank Working Paper Series 75.

Annexe

Application 1 *summary(sp500)*

Min 1st Qu Median Mean 3rd Qu Max
-125.460 -0.860 -0.080 -1.493 0.250 248.150

skewness(sp500)

1.711954

attr(,method)

moment

> *kurtosis(sp500)*

33.71843

attr(,method)

excess

One Sample t-test

data : sp500

t = 21.132, df = 1767, p-value < 2.2e-16

alternative hypothesis : true mean is not equal to 0

95 percent confidence interval :

234.3942 282.3549

sample estimates :

mean of x

258.3746

Box-Ljung test

data : sp500

X-squared = 208.74, df = 12, p-value < 2.2e-16

arima(x = at, order = c(0, 0, 3))

Call :

arma(x = at, order = c(0, 0, 3))

Coefficients :

ma1 ma2 ma3 intercept

0.2217 -0.0309 0.0534 0.0038

s.e. 0.0243 0.0234 0.0248 0.5314

sigma² estimated as 322.5 : log likelihood = -7610.47, aic = 15230.95

Call :

arma(x = at, order = c(3, 0, 3))

Coefficients :

ar1 ar2 ar3 ma1 ma2 ma3 intercept

0.6719 -0.6943 0.8192 -0.4887 0.6283 -0.7506 -0.0462

s.e. 0.0437 0.0263 0.0379 0.0471 0.0375 0.0480 0.7986

sigma² estimated as 309 : log likelihood = -7572.89, aic = 15161.79

Title :

GARCH Modelling

Title :

GARCH Modelling

Call :

garchFit(formula = sp500 arma(1, 1) + garch(1, 1), data = sp500, trace = F)

Mean and Variance Equation :

data arma(1, 1) + garch(1, 1)

<environment : 0x00000000085de898>

[data = sp500]

Conditional Distribution :

norm

*Coefficient(s) :**mu ar1 ma1 omega alpha1 beta1**-0.00780814 0.04614599 0.25105653 0.00075616 0.29526658 0.77108666**Std. Errors :**based on Hessian**Error Analysis :**Estimate Std. Error t value Pr(>|t|)**mu -0.0078081 0.0079998 -0.976 0.32904**ar1 0.0461460 0.0819461 0.563 0.57335**ma1 0.2510565 0.0819236 3.065 0.00218 ****omega 0.0007562 NA NA NA**alpha1 0.2952666 0.0096313 30.657 < 2e-16 *****beta1 0.7710867 0.0071758 107.456 < 2e-16 ****

*Signif. codes : 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1**Log Likelihood :**-3099.63 normalized : -1.754177**Description :**Mon Sep 20 11 :16 :36 2021 by user : dell**Standardised Residuals Tests :**Statistic p-Value**Jarque-Bera Test R Chi² 184.1238 0**Shapiro-Wilk Test R W 0.9837306 2.810839e-13**Ljung-Box Test R Q(10) 25.29016 0.004821825**Ljung-Box Test R Q(15) 34.61062 0.002792915**Ljung-Box Test R Q(20) 38.37075 0.007975751**Ljung-Box Test R² Q(10) 24.09723 0.007346442**Ljung-Box Test R² Q(15) 30.19106 0.01125089**Ljung-Box Test R² Q(20) 33.74266 0.02791909**LM Arch Test R TR² 30.90266 0.002038458**Information Criterion Statistics :**AIC BIC SIC HQIC**3.515145 3.533742 3.515122 3.522016*

Title :

GARCH Modelling

Call :

garchFit(formula = sp500 garch(1, 1), data = sp500, trace = F)

Mean and Variance Equation :

data garch(1, 1)

<environment : 0x0000000014f45360>

[data = sp500]

Conditional Distribution :

norm

Coefficient(s) :

mu omega alpha1 beta1

-0.01112766 0.00084867 0.29493656 0.76886473

Std. Errors :

based on Hessian

Error Analysis :

Estimate Std. Error t value Pr(>|t|)

mu -0.0111277 0.0069954 -1.591 0.112

omega 0.0008487 NA NA NA

*alpha1 0.2949366 0.0007570 389.593 <2e-16 ****

*beta1 0.7688647 0.0051837 148.325 <2e-16 ****

*— Signif. codes : 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1*

Log Likelihood :

-3159.568 normalized : -1.788097

Description :

Mon Sep 20 11 :19 :27 2021 by user : dell

*Standardised Residuals Tests :**Statistic p-Value**Jarque-Bera Test R Chi² 192.173 0**Shapiro-Wilk Test R W 0.9832182 1.57429e-13**Ljung-Box Test R Q(10) 132.8752 0**Ljung-Box Test R Q(15) 142.0001 0**Ljung-Box Test R Q(20) 145.9983 0**Ljung-Box Test R² Q(10) 21.5445 0.0176009**Ljung-Box Test R² Q(15) 25.98156 0.03821692**Ljung-Box Test R² Q(20) 28.16744 0.1054946**LM Arch Test R TR² 27.87794 0.005764005**Information Criterion Statistics :**AIC BIC SIC HQIC**3.580722 3.593121 3.580712 3.585303**threshold**7.091128**p.less.thresh**0**n.exceed**100**method**ml**par.ests**xi beta**0.1865105 26.5027771**par.ses**xi beta**0.1400775 4.5343201**varcov**[, 1] [, 2]**[1,] 0.01962171 -0.4614114**[2,] -0.46141141 20.5600588*

information
observed

converged
0

nllh.final
446.3967

attr(,class) gpd

var1 :113.4459
vart : 2917.62
max(vart) : 130802
min(vart) : 3.303515

Application 2 *Min 1st Qu Median Mean 3rd Qu Max*

-7.60000 -0.62500 -0.07000 -0.01829 0.43500 4.92000

> skewness(intel)

-0.0501078

attr(,method)

moment

> kurtosis(intel)

10.35384

attr(,method)

excess

One Sample t-test

data : intel

t = 158.18, df = 251, p-value < 2.2e-16

alternative hypothesis : true mean is not equal to 0

95 percent confidence interval :

54.35370 55.72424

sample estimates :

mean of x

55.03897

*Box-Ljung test**data : abs(at)**X-squared = 1623, df = 12, p-value < 2.2e-16**Call :**arima(x = at, order = c(0, 0, 3))**Coefficients :**ma1 ma2 ma3 intercept**2.1224 2.0054 0.8344 1.8383**s.e. 0.0139 0.0176 0.0102 13.5483**sigma² estimated as 9144 : log likelihood = -10574.85, aic = 21159.69**Call :**arima(x = at, order = c(3, 0, 0))**Coefficients :**ar1 ar2 ar3 intercept**0.9378 0.0539 -0.0136 -0.9527**s.e. 0.0628 0.0863 0.0629 2.7959**sigma² estimated as 1.257 : log likelihood = -387.97, aic = 785.93**Title :**GARCH Modelling**Call :**garchFit(formula = at garch(1, 1), data = at, cond.dist = std,
trace = F)**Mean and Variance Equation :**data garch(1, 1)**<environment : 0x000000001376d018>**[data = at]*

Conditional Distribution :

std

Coefficient(s) :

mu omega alpha1 beta1 shape

-1.0397e-16 1.2801e-06 2.4916e-02 9.6834e-01 3.9251e+00

Std. Errors :

based on Hessian

Error Analysis :

Estimate Std. Error t value Pr(>|t|)

mu -1.040e-16 5.180e-02 0.000 1.0000

omega 1.280e-06 1.970e-02 0.000 0.9999

*alpha1 2.492e-02 1.188e-02 2.098 0.0359 **

*beta1 9.683e-01 2.173e-02 44.561 < 2e-16 ****

*shape 3.925e+00 8.435e-01 4.653 3.27e-06 ****

—
*Signif. codes : 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1*

Log Likelihood :

-342.8501 normalized : -1.365937

Description :

Mon Sep 20 11 :52 :30 2021 by user : dell

Standardised Residuals Tests :

Statistic p-Value

Jarque-Bera Test R Chi² 1779.373 0

Shapiro-Wilk Test R W 0.8437366 3.335772e-15

Ljung-Box Test R Q(10) 5.086926 0.8852961

Ljung-Box Test R Q(15) 11.78751 0.6950362

Ljung-Box Test R Q(20) 12.84558 0.8839107

Ljung-Box Test R² Q(10) 1.083879 0.9997513

Ljung-Box Test R² Q(15) 18.54687 0.2350066

Ljung-Box Test R² Q(20) 19.03445 0.5195875

LM Arch Test R TR² 1.03353 0.999983

Information Criterion Statistics :

AIC BIC SIC HQIC

2.771714 2.841942 2.770941 2.799976

threshold

0.1182899

p.less.threshold

0

n.exceed

100

method

ml

par.ests

xi beta

0.1071124 0.7328883

par.ses

xi beta

0.1038945 0.1054004

varcov

[,1] [,2] [1,] 0.010794075 -0.007044488

[2,]-0.007044488 0.011109250

information

observed

converged

0

llh.final

79.65262

```
attr(,"class")
gpd
var1 : 2.706983
var2 : 2.670068
max(var) : 4.19414
> min(var) : 2.663789
```

Application 3 *summary(ibm)*

```
Min. 1st Qu. Median Mean 3rd Qu. Max.
-8.19000 -1.08500 -0.09000 -0.06944 0.87000 9.42000
skewness(ibm) :0.34125
kurtosis(ibm) :4.635905
One Sample t-test
```

```
data : ibm
t = -0.58234, df = 250, p-value = 0.5609
alternative hypothesis : true mean is not equal to 0
95 percent confidence interval :
-0.3042973 0.1654128
sample estimates :
mean of x
-0.06944226
Box-Ljung test
```

```
data : ibm
X-squared = 20.422, df = 12, p-value = 0.05951
Call :
arima(x = at, order = c(0, 0, 3))
```

```
Coefficients :
ma1 ma2 ma3 intercept
-0.0451 -0.0027 -0.1252 0.0001
s.e. 0.0620 0.0640 0.0604 0.0977
```

```
sigma2 estimated as 3.486 : log likelihood = -512.91, aic = 1035.82
Call :
arima(x = at, order = c(3, 0, 3))
```

Coefficients :

ar1 ar2 ar3 ma1 ma2 ma3 intercept
0.4555 -0.5340 -0.2960 -0.5109 0.5485 0.1534 0.0010
s.e. 0.4306 0.4666 0.3784 0.4396 0.4932 0.3599 0.1012

*sigma*² estimated as 3.416 : log likelihood = -510.42, aic = 1036.84

Title :

GARCH Modelling

Call :

garchFit(formula = ibm arma(1, 1) + garch(1, 1), data = ibm, trace = F)

Mean and Variance Equation :

data arma(1, 1) + garch(1, 1)
<environment : 0x00000000080464e0>
[data = ibm]

Conditional Distribution :

norm

Coefficient(s) :

mu ar1 ma1 omega alpha1 beta1
-0.040555 0.477938 -0.580558 2.794721 0.065398 0.144355

Std. Errors :

based on Hessian

Error Analysis :

Estimate Std. Error t value Pr(>|t|)
mu -0.04055 0.05432 -0.747 0.4553
ar1 0.47794 0.27447 1.741 0.0816 .
ma1 -0.58056 0.25193 -2.304 0.0212 *
omega 2.79472 2.67757 1.044 0.2966
alpha1 0.06540 0.06093 1.073 0.2831
beta1 0.14436 0.77637 0.186 0.8525
— Signif. codes : 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Log Likelihood :

-513.4278 normalized : -2.045529

Description :

Mon Sep 20 09 :53 :04 2021 by user : dell

Standardised Residuals Tests :

Statistic p-Value

Jarque-Bera Test R Chi² 288.1269 0

Shapiro-Wilk Test R W 0.9320928 2.440528e-09

Ljung-Box Test R Q(10) 17.22692 0.06949114

Ljung-Box Test R Q(15) 20.43485 0.1558781

Ljung-Box Test R Q(20) 24.32237 0.2285938

Ljung-Box Test R² Q(10) 6.705465 0.7529275

Ljung-Box Test R² Q(15) 6.919379 0.9598484

Ljung-Box Test R² Q(20) 8.071543 0.9913838

LM Arch Test R TR² 6.259796 0.9024235

Information Criterion Statistics :

AIC BIC SIC HQIC

4.138867 4.223141 4.137759 4.172781

Title :

GARCH Modelling

Call :

garchFit(formula = at garch(1, 1), data = at, cond.dist = std, trace = F)

Mean and Variance Equation :

data garch(1, 1)

<environment : 0x000000001402d208>

[data = at]

Conditional Distribution :

std

Coefficient(s) :

mu omega alpha1 beta1 shape

-1.1092e-16 8.2322e-01 1.0628e-01 7.0278e-01 3.3184e+00

Std. Errors :

based on Hessian

*Error Analysis :**Estimate Std. Error t value Pr(>|t|)**mu -1.109e-16 9.514e-02 0.000 1.000**omega 8.232e-01 2.333e+00 0.353 0.724**alpha1 1.063e-01 2.015e-01 0.527 0.598**beta1 7.028e-01 7.353e-01 0.956 0.339**shape 3.318e+00 7.714e-01 4.302 1.7e-05 *****— Signif. codes : 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1**Log Likelihood :**-489.0176 normalized : -1.948277**Description :**Mon Sep 20 09 :53 :57 2021 by user : dell**Standardised Residuals Tests :**Statistic p-Value**Jarque-Bera Test R Chi² 427.3813 0**Shapiro-Wilk Test R W 0.9220553 3.372012e-10**Ljung-Box Test R Q(10) 15.48441 0.1153737**Ljung-Box Test R Q(15) 18.5418 0.235253**Ljung-Box Test R Q(20) 22.72036 0.3027331**Ljung-Box Test R² Q(10) 4.569808 0.9180042**Ljung-Box Test R² Q(15) 4.960392 0.9924565**Ljung-Box Test R² Q(20) 5.724735 0.9992214**LM Arch Test R TR² 4.296576 0.9774707**Information Criterion Statistics :**AIC BIC SIC HQIC**3.936395 4.006624 3.935622 3.964657**threshold**0.2838917**p.less.thresh**0**n.exceed**100*

method

ml

par.ests

xi beta

0.04811669 1.28124898

par.ses

xi beta

0.08976778 0.17207911

varcov

[, 1] [, 2]

[1,] 0.008058255 -0.009681351

[2,] -0.009681351 0.029611219

information

observed

converged

0

llh.final

129.6024

attr(, class)

gpd

var1 : 4.412589

var2 : 5.642781

max(var) = 14.69751

min(var) = 4.26218