

Thèse
présentée pour l'obtention du titre de
Docteur de l'Université de Lorraine
en Mathématiques
par
Khaled SALHI

Risques extrêmes en finance : analyse et modélisation

Soutenue publiquement le 5 décembre 2016

Membres du Jury :

Madalina DEACONU
Caroline HILLAIRET
Benjamin JOURDAIN
Antoine LEJAY
Anis MATOUSSI
Ashkan NIKEGHBALI
Nizar TOUZI

Chargée de recherche, Inria (Directrice de thèse)
Professeur, ENSAE - ParisTech, CREST
Professeur, Ecole des Ponts ParisTech (Président)
Directeur de recherche, Inria (Co-directeur de thèse)
Professeur, Université du Maine
Professeur, Université de Zürich (Rapporteur)
Professeur, Ecole Polytechnique (Rapporteur)



AVERTISSEMENT

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document.

D'autre part, toute contrefaçon, plagiat, reproduction illicite encourt une poursuite pénale.

Contact : ddoc-theses-contact@univ-lorraine.fr

LIENS

Code de la Propriété Intellectuelle. articles L 122. 4

Code de la Propriété Intellectuelle. articles L 335.2- L 335.10

http://www.cfcopies.com/V2/leg/leg_droi.php

<http://www.culture.gouv.fr/culture/infos-pratiques/droits/protection.htm>

À mon papa.

Remerciements

Je voudrais, en premier lieu, exprimer toute ma gratitude à mes directeurs de thèse, Madalina Deaconu et Antoine Lejay, non seulement pour leur disponibilité et leur soutien au cours de ces trois années, mais aussi pour avoir fait preuve, à plusieurs reprises, de leur confiance en moi, et pour m'avoir aidé, en m'envoyant à des nombreuses conférences et séminaires, à faire connaître mon travail de recherche.

Je remercie sincèrement Nicolas Champagnat pour son encadrement de mon stage de fin d'études d'ingénieur et de mon stage de master. Je le remercie également pour avoir resté toujours assez présent et pour toute l'aide et les conseils qu'il a continués à m'apporter durant ma thèse.

Je n'aurais jamais pensé à travailler dans le domaine de la finance mathématique si je n'avais pas été élève ingénieur à l'École Polytechnique de Tunisie dont je suis fier de compter parmi ses anciens élèves.

Je remercie les membres du jury : les rapporteurs Ashkan Nikeghbali et Nizar Touzi, les examinateurs Caroline Hillairet, Benjamin Jourdain et Anis Matoussi, pour avoir pris le temps d'examiner et d'évaluer mon travail et pour leurs remarques qui m'ont permis d'améliorer la qualité de ce manuscrit.

Je remercie également tous les anciens et actuels membres de l'équipe Tosca Inria, avec lesquels j'ai eu le plaisir d'avoir plusieurs interactions et réunions, notamment lors de nos journées Tosca annuelles. Un remerciement particulier à Denis Talay, notre chef d'équipe, pour les remarques et les questions pertinentes qui m'ont permis d'améliorer mon travail et d'envisager plusieurs perspectives.

Mes remerciements vont aussi à tous les enseignants avec qui j'ai assuré mon service de monitorat, en particulier Anne Gégout-Petit, Aurélie Gueudin, Sophie Mézières, Bruno Pinçon et Jean-François Scheid. C'était un plaisir de travailler avec eux.

Je voudrais également exprimer ma reconnaissance à tous les gens que j'ai eu la chance de côtoyer durant ces quelques dernières années et qui ont rendu mon séjour à Nancy aussi agréable, principalement Nadine et Mario, Gilles et Paco, Ayoub, Rim, Mohamed et Aymen, ainsi que Wahiba et Ines.

Enfin, je remercie chaleureusement ma chère femme Marwa pour tout l'amour et l'accompagnement qu'elle m'a procurés durant cette thèse. Je remercie également mon cher père Ezzeddine, à qui je dédie ce manuscrit, ma mère Zohra, mon frère Melek et mes deux sœurs Kaouther et Salma, pour TOUT.

Table des matières

Introduction et principaux résultats	9
I Mesure de risque et estimation de la Value-at-Risk	23
1 La Value-at-Risk : Estimation et validation	25
1.1 Value-at-Risk	26
1.2 Modèle gaussien et ses limites dans la mesure de risque	27
1.2.1 La loi normale	28
1.2.2 Estimation de la Value-at-Risk	29
1.3 Backtesting sur la VaR	31
1.3.1 Test du ratio de dépassement	32
1.3.2 Test de couverture non-conditionnelle	36
1.3.3 Test d'indépendance	38
1.3.4 Test de couverture conditionnelle	40
1.3.5 Résultats empiriques des différents backtests	40
1.4 Conclusion	42
2 Modèle des lois puissances	43
2.1 Les fonctions à variation lente	44
2.2 Estimateur de Hill	45
2.3 Estimation de la fonction à variation lente $\ell(x)$	48
2.4 Choix du seuil d'estimation r_n	50

2.5	Etude empirique : estimation des paramètres de la loi puissance, calcul de la VaR	52
2.6	Backtesting sur la VaR estimée sous un modèle des lois puissances . .	54
2.7	Conclusion	56
3	Détection des régimes : régime crise - régime hors crise	59
3.1	Problématique des modèles de Markov cachés (MMC)	61
3.2	Problème d'évaluation, problème de filtrage et problème de lissage . .	62
3.3	Problème d'apprentissage et algorithme de Baum-Welsh	65
3.3.1	Principe de l'algorithme	65
3.3.2	Application sur des données boursières	67
3.4	Problème de reconnaissance et algorithme de Viterbi	68
3.4.1	Principe de l'algorithme	68
3.4.2	Application sur des données boursières	70
3.5	Etude des crises	70
3.5.1	Temps de retour à la crise	70
3.5.2	Densité empirique	71
3.5.3	Classification par alignement séquentiel multiple	73
3.6	Conclusion	75
4	Modèle des lois puissances à changement de régimes	77
4.1	Procédure d'estimation	77
4.2	La probabilité d'être en crise	78
4.3	Estimation des paramètres des queues des distributions	80
4.4	Backtesting	81
4.5	Étude comparative : Benchmark	83
4.5.1	Base de données : Actions boursières de Euronext Paris	85
4.5.2	Discussion	85

4.6	Conclusion	88
II	Processus de Lévy en finance	91
5	Modèle exponentielle-Lévy et évaluation des options européennes	93
5.1	Processus de Lévy	93
5.1.1	Exponentielle de Doléans-Dade d'un processus de Lévy	96
5.1.2	Changement de mesure de probabilité	97
5.2	Modèle exponentielle-Lévy	98
5.2.1	Mesures risque-neutres	99
5.2.2	Évaluation d'options via la fonction caractéristique	104
5.2.3	Discrétisation et transformée de Fourier rapide (FFT)	106
5.3	Exemples des modèles exp-Lévy	108
5.3.1	Modèle de Black et Scholes	109
5.3.2	Modèle diffusion-sauts de Merton	110
5.3.3	Modèle variance gamma	114
5.4	Conclusion	116
6	La Value-at-Risk et la Value-at-Risk Conditionnelle sous le modèle exp-Lévy	123
6.1	La CVaR et l'algorithme de Rockafellar et Uryasev	124
6.2	VaR et CVaR monétaires	126
6.2.1	Méthode de calcul	126
6.2.2	Étude de l'erreur numérique dans le cadre du modèle de Black et Scholes	128
6.2.3	Application aux modèles exponentielle-Lévy	131
6.3	VaR et CVaR en termes de rendement	132
6.3.1	Méthode de calcul	132
6.3.2	Étude de l'erreur numérique dans le cadre du modèle de Black et Scholes	134

6.3.3	Application aux modèles exponentielle-Lévy	135
6.4	Conclusion	136
7	Minimisation de la CVaR pour le risque de couverture des options dans le cadre des modèles exp-Lévy	139
7.1	Couverture des options dans les marchés incomplets	139
7.2	Présentation du problème	142
7.3	Problème dual	144
7.4	Approximation du problème dual	147
7.5	Cas d'une option d'achat européenne	149
7.5.1	Résolution du problème approché	149
7.5.2	Illustration des résultats	152
7.6	Conclusion	154
	Conclusion et perspectives	155
	Références bibliographiques	157

Introduction et principaux résultats

Les événements extrêmes ont une importance majeure dans l'économie réelle et les marchés financiers ne font pas l'exception. À titre d'exemple, nous mentionnons la crise de 1929, le crash de 1987 et la crise de 2007. Ce sont des événements rares, qui ont une faible probabilité d'apparition comparée à la probabilité des événements qui se produisent le plus souvent. Ces événements ont un coût économique énorme. Du point de vue de la modélisation mathématique, ils s'interprètent par la queue de distribution gauche (pertes extrêmes) et la queue de distribution droite (gains extrêmes) des rendements des actifs financiers.

La première partie de la thèse s'inscrit dans le cadre de la gestion de risque avec application sur des données de la bourse de Paris. La gestion de risque a pour objectif de modéliser le comportement des queues de distributions et de construire des outils capables de prévoir et de limiter les grosses pertes. Dans cette partie de la thèse, nous construisons un modèle pour les queues de distribution lourdes qui prend en compte les changements de régime. En effet, le modèle de la distribution gaussienne et la loi des grands nombres, fondements de l'étude des grandeurs moyennes, privilégient les événements proches de la moyenne tout en négligeant les valeurs extrêmes. La modélisation par la loi normale conduit à une sous-estimation de l'occurrence des événements rares. Ainsi, la Value-at-Risk (VaR), le montant de perte qui n'est dépassée qu'avec une probabilité donnée sur un horizon temporel bien déterminé, calculée dans le cadre du modèle gaussien, est sous-estimée. Cette mesure a toujours été une norme en matière de mesure de risque au sein des établissements financiers et dans les accords de Bâle [156].

Une alternative au modèle gaussien, qui prend en considération des queues de distributions lourdes, est le modèle des lois stables, appelées aussi Lévy-stables. Ces lois s'obtiennent par la généralisation du théorème centrale limite à des sommes de variables aléatoires indépendantes de même loi et de variance infinie. Elles donnent plus d'importance aux valeurs extrêmes et proposent un équivalent aux outils de la gestion de portefeuilles utilisés dans le cadre du modèle gaussien. Il s'agit du MEDAF appliqué dans le cadre du marché stable [61]. Cependant, avec des tests numériques sur les quantiles de la queue de cette distribution, notamment sur la VaR, il s'avère que ces lois sur-estiment le plus souvent l'épaisseur de la queue et donnent trop d'ampleur aux événements rares.

Pour répondre aux défauts de sous-estimation et de sur-estimation des risques dans

ces deux modèles, nous proposons le modèle des lois puissances. Ce modèle est basé sur l'hypothèse qu'il existe un réel α tel que la fonction de répartition F , des pertes d'un portefeuille donné, s'écrit :

$$1 - F(x) = \mathbb{P}(L \geq x) = \ell(x) x^{-\alpha},$$

avec $\ell(x)$ une fonction dite à variation lente quand $x \rightarrow +\infty$. Une telle loi est appelée loi puissance d'indice de queue α , qui résume la lourdeur de la queue de la distribution et caractérise également l'existence des moments. Les lois stables ont une queue de distribution de loi puissance d'indice α inférieur à 2. Cette contrainte est imposée implicitement par la définition de ces lois. Cependant, l'étude des queues de distribution des données du marché modélisées par des lois puissances, sans contrainte sur α , conduit en pratique à des valeurs de α le plus souvent supérieures à 3. C'est cette différence entre l'indice de décroissance intrinsèque, obtenu dans la définition des lois stables, et le vrai comportement de la queue de la distribution des données financières réelles qui entraîne la sur-estimation de l'épaisseur de la queue de distribution par les lois stables. Pour cette raison, nous privilégions un modèle de loi puissance qui apporte plus de précisions sur les mesures des valeurs extrêmes.

Notre étude empirique utilise principalement un estimateur inspiré de celui de Hill pour approcher l'indice de décroissance de la queue de distribution des pertes financières, ainsi que l'estimateur de Weissman pour approcher la fonction à variation lente $\ell(x)$. Nous validons notre modèle par des tests portant sur la VaR mesurée à l'aide des données numériques, observées sur les marchés financiers.

Le modèle des lois puissances apporte une meilleure estimation de la VaR comparée aux estimations données par le modèle gaussien et le modèle des lois stables. Cependant, en regardant la VaR estimée à partir des données financières réelles sur des petites périodes et non pas sur l'ensemble de tout l'historique, nous constatons un regroupement des erreurs d'estimation. Autrement dit, nous remarquons des périodes où les estimations de la VaR ne détectent pas les événements rares et elles les sous-estiment, et des périodes où ces estimations donnent beaucoup d'ampleurs au risque réel et elles le sur-estiment. Cela s'explique par le fait que ces modèles reposent sur l'hypothèse que les pertes journalières sont indépendantes et identiquement distribuées. Or, il est plus légitime de supposer que le marché financier passe par des périodes de crise et des périodes de hors-crise et de tenir compte des changements de régimes qui en résultent.

Ainsi, nous construisons une variante du modèle des lois puissances pour la queue de distribution, qu'on appellera dans ce manuscrit le *modèle des lois puissances à changement de régimes*. Nous admettons qu'à un jour donné le portefeuille peut être soit dans un régime de crise soit dans un régime de hors-crise. Nous supposons par la suite que la queue de distribution des pertes dans un régime de crise suit une loi puissance d'indice α_c et la queue de distribution des pertes dans un régime de hors-crise suit une loi puissance d'indice α_h . Nous décrivons la loi de transition entre les deux régimes et nous construisons ainsi un modèle à changement de régimes pour la queue de distribution. Dans notre approche, nous cherchons à bénéficier de la stabilité des modèles des lois puissances dans l'estimation de la VaR et de la

détection des regroupements de la volatilité, donnée par les modèles conditionnels.

Dans la partie empirique, la théorie des modèles de Markov cachés (MMC) est appliquée pour la classification des données et la détermination des différentes périodes de crise et de hors-crise. Ensuite le quantile des lois puissances est estimé en se basant sur la connaissance de l'historique des régimes du marché. Pour estimer la VaR à un instant donné t , notre modèle donne plus de poids au régime présent à l'instant t , ce qui réduit les sous-estimations et les sur-estimations observées dans le modèle classique des lois puissances, à chaque changement de régime.

Nos modèles sont intensivement testés sur des données de marché historiques de la Bourse de Paris (Euronext Paris). Nous nous sommes principalement intéressés aux rendements journaliers des prix de clôtures des actions, ce qui représente en pratique la situation des petits investisseurs et pour lesquels le problème de calibration du modèle est le plus délicat à cause de faibles quantités de données dont ils disposent.

La deuxième partie de la thèse s'intéresse à une autre façon de considérer les événements extrêmes, celle d'intégrer les sauts dans la modélisation du prix d'actif. Parmi les classes les plus connues des modèles à temps continu et dont les trajectoires peuvent avoir des discontinuités, nous proposons les modèles exponentielle-Lévy où le logarithme du prix est décrit par un processus de Lévy. Un processus de Lévy est un processus stochastique à accroissements indépendants et stationnaires : si $(X_t)_{t \geq 0}$ est un processus de Lévy, $X_t - X_s$ avec $t > s$ est indépendant de l'histoire du processus avant le temps s et sa loi ne dépend pas de t ou s séparément mais seulement de $t - s$. Cette propriété des accroissements évoque une analogie avec les fonctions linéaires : on peut dire que les processus de Lévy sont, dans un certain sens, des « processus linéaires » ou additifs.

Malgré cette simplicité apparente, les processus de Lévy ont des nombreuses propriétés intéressantes et constituent un domaine d'étude en plein développement : plusieurs ouvrages ont été publiés récemment [5, 21, 43, 145] et une série de conférences internationales dédiées aux processus de Lévy et leurs applications a rencontré un grand succès.

Sur le plan de la modélisation financière, les processus de Lévy fournissent une classe de modèles avec sauts qui est suffisamment riche pour bien décrire les données empiriques et assez simple pour réaliser de nombreux calculs analytiquement. L'intérêt de tels modèles en finance est principalement dû aux trois facteurs suivants.

Premièrement, dans un modèle à trajectoires continues, comme un modèle de diffusion, le processus de prix se comporte localement comme un mouvement brownien et la probabilité que le prix bouge beaucoup, pendant un temps court, est très petite si la valeur de volatilité n'est pas déraisonnablement grande. Cela implique que, dans de tels modèles, les prix des options « hors de la monnaie » doivent être beaucoup plus petits que ce que l'on observe en réalité. Par contre, si le processus de prix peut sauter, même pour une maturité courte on ne peut pas négliger la probabilité d'un mouvement inattendu du prix qui déplacerait l'option dans la monnaie.

Deuxièmement, du point de vue de la couverture, les modèles à trajectoires continues correspondent en général aux marchés complets ou aux marchés qui peuvent être complétés en ajoutant un ou deux actifs supplémentaires comme dans les modèles à volatilité stochastique. Ceci implique que le flux terminal de n'importe quel actif contingent peut être parfaitement répliqué, et les options cotées deviennent donc redondantes. Cependant, en présence de sauts la couverture parfaite est impossible et les options permettent aux participants du marché de couvrir les risques qui ne peuvent pas être répliqués par une stratégie de trading ne faisant intervenir que le sous-jacent.

Le troisième argument en faveur de l'utilisation de modèles discontinus, qui est peut-être le plus important, est la présence-même de sauts dans les prix. Il est clair que des variations de prix, comme celles observées sur le marché, ne peuvent pas être expliquées par un modèle à trajectoires continues, mais elles doivent être prises en compte pour une gestion de risques fiable.

Dans cette partie, nous considérons deux problèmes liés respectivement à la gestion de risque et à la couverture des options dans le cadre des modèles exponentielle-Lévy. Premièrement, nous développons deux méthodes numériques pour la mesure de la Value-at-Risk (VaR) et la Value-at-Risk Conditionnelle (CVaR) qui est une mesure complémentaire à la VaR et qui correspond à la valeur moyenne de la perte sachant que cette dernière a dépassé la VaR. La première méthode permet d'évaluer la VaR et la CVaR monétaires, tandis que la deuxième s'intéresse à la VaR et la CVaR en termes de rendements. Le principe de ces méthodes est d'exprimer la VaR et la CVaR comme respectivement l'argument du minimum et le minimum d'une certaine fonction que l'on évalue numériquement par des techniques d'inversion de Fourier. Le premier avantage de notre méthodologie est qu'elle nécessite la seule connaissance de la fonction caractéristique du processus de Lévy. Le deuxième est qu'elle évalue simultanément la VaR et la CVaR. Ceci représente un atout par rapport à d'autres méthodes qui les calculent séparément en évaluant d'abord la fonction de densité par inversion de Fourier [28, 94, 95].

La deuxième problématique est la détermination de la stratégie autofinancée qui minimise la CVaR du risque de couverture. Sans contrainte sur le capital initial, la solution du problème est la stratégie de la sur-réplication et la CVaR minimale vaut zéro. Cependant, le capital initial requis pour construire une stratégie de sur-réplication est très élevé et égal au prix de l'actif sous-jacent dans certains cas. Cela rend ce choix non-adéquat en pratique et il faudrait ajouter une contrainte budgétaire sur le capital initial. Dans le cas du marché complet, où il existe une unique mesure risque-neutre, le problème est résolu dans [119] par l'application du lemme de Neyman-Pearson. Des généralisations de ce lemme ont été proposées dans [47, 138, 139, 141]. Néanmoins, elles ne peuvent pas être appliquées dans le cadre des modèles exponentielle-Lévy car ces généralisations reposent sur des hypothèses de convexité et de compacité dont nous montrerons la non-validité dans ce cas. Nous proposons comme alternative une approximation de la CVaR minimale par la résolution d'un problème approché. Cela est fait en relaxant la contrainte budgétaire et en résolvant le problème sous la mesure risque-neutre d'Esscher.

Finalement, nous mentionnons que cette thèse a été motivée initialement par une collaboration entre ALPHABILITY, une PME qui proposait des solutions informatiques pour les sociétés de gestion de portefeuilles, et notre équipe Tosca d’Inria, spécialiste des méthodes probabilistes et de la modélisation stochastique. La collaboration avait pour objectif de définir et d’étudier des mesures fiables du risque des marchés financiers pour proposer des stratégies de gestion de portefeuille prenant en compte de critères de risque.

Principaux résultats

La **première partie** de la thèse se divise en quatre chapitres.

Le **premier chapitre** de la thèse présente la méthodologie d’estimation de la Value-at-Risk à partir des données réelles et les techniques de validation du modèle choisi pour décrire les rendements de l’actif financier, dites les backtests. La présentation est faite dans le cadre simple du modèle gaussien.

Notons $R_t = \log(S_{t+1}/S_t)$ le rendement d’un actif financier à un instant t avec S_t son prix de clôture journalier. La perte est définie par $L_t = -R_t$ et la Value-at-Risk (quotidienne en termes de rendements), à l’instant t et à un niveau de confiance a , est la perte potentielle maximale que l’on peut subir le lendemain en dehors d’un événement de probabilité $1 - a$. Mathématiquement parlant, elle représente le $a^{\text{ème}}$ quantile supérieur de la distribution de la perte L_{t+1} ,

$$\text{VaR}_t^{(a)} = q^{L_{t+1}}(a) = \inf\{x \in \mathbb{R} | \mathbb{P}(L_{t+1} \leq x) > a\}.$$

Quand les variables aléatoires $(S_t)_{t \geq 0}$ sont continues, on peut écrire directement

$$\mathbb{P}(L_{t+1} \geq \text{VaR}_t^{(a)}) = \mathbb{P}(R_{t+1} \leq -\text{VaR}_t^{(a)}) = 1 - a.$$

La méthodologie de mesure de risque se décline en trois étapes :

- Choisir un modèle pour décrire la distribution (ou la queue de distribution) des rendements (ou pertes) de l’actif boursier ;
- Estimer la VaR sous ce modèle ;
- Valider la VaR estimée par des backtests pour vérifier l’adéquation de ce modèle dans la description de la queue de la distribution des données financières réelles.

Dans ce chapitre de présentation, les rendements sont modélisés par des variables aléatoires i.i.d. qui suivent une loi normale. L’estimation de la VaR se fait sur une fenêtre glissante de taille k pour capturer les caractéristiques dynamiques des données du marché, qui varient en fonction des périodes de temps. En d’autres termes, les paramètres de la loi de R_{t+1} sont estimés à partir des données antérieures $(R_{t-k+1}, \dots, R_t)$, celles de la loi de R_{t+2} à partir de $(R_{t-k+2}, \dots, R_{t+1})$ et ainsi de suite. Finalement, pour valider l’adéquation du modèle dans la mesure de

la VaR, nous avons eu recours aux procédures de backtesting. En effet, étant donné que la véritable valeur de la VaR n'est pas observable, les techniques usuelles de validation des estimations telles que l'erreur quadratique moyenne et l'erreur absolue moyenne en pourcentage ne peuvent pas être utilisées. Le backtesting consiste à tester un modèle de mesure de risque sur une grande base de données historiques pour évaluer son efficacité et valider son application dans la pratique. Le backtest le plus connu est celui du ratio de dépassement : pour une VaR à un niveau de confiance a , la probabilité que R_{t+1} soit inférieur à $-\text{VaR}_t^{(a)}$ est $1 - a$. Nous estimons la VaR sur une fenêtre glissante pour tout l'historique dont nous disposons et nous comptons le pourcentage de dépassement ($R_{t+1} < -\text{VaR}_t^{(a)}$) sur l'ensemble de toutes les estimations. La situation idéale est celle où ce ratio est égal à $1 - a$. S'il y est supérieur, nous disons que la VaR est sous-estimée sous le modèle choisi. S'il y est inférieur, nous disons que le modèle sur-estime la VaR. D'autres backtests plus sophistiqués, d'intervalle de confiance, de couverture non-conditionnelle et d'indépendance entre les dépassements, sont étudiés dans ce chapitre.

Les résultats de l'application sur les données de la bourse de Paris sont cohérents avec ce qui a été prouvé depuis longtemps sur d'autres données financières. En effet, nous remarquons que le modèle gaussien sous-estime la VaR et que la queue de distribution des données réelles est plus lourde que celle de la loi normale.

Dans le **chapitre 2**, nous étudions le modèle des lois puissances qui prend en considération les queues de distributions lourdes. Ce choix a été encouragé par les résultats obtenus par les graphes de Pareto tracés à partir des pertes journalières de nos données boursières. Sur ce graphe nous traçons, dans un certain sens, les quantiles des pertes contre ceux d'une loi uniforme. Si les données suivent une loi puissance, la fonction quantile doit avoir l'allure d'une droite sur la queue de distribution. Cette observation est vérifiée pour le graphe de Pareto obtenu à partir des pertes quotidiennes pour la plupart des actions dans notre base de données.

Sous le modèle des lois puissances, la loi de la perte L_t est donnée par

$$\mathbb{P}(L_t \geq x) = \ell(x) x^{-\alpha},$$

avec α l'indice de queue et $\ell(x)$ une fonction à variation lente que l'on suppose égale à une constante C à partir d'un certain rang (ou qui tend rapidement vers C) dans l'étude empirique. L'estimation des paramètres α et C est issue de la théorie des valeurs extrêmes et des statistiques d'ordre. Pour une série des pertes $(L_i)_{1 \leq i \leq n}$, nous définissons la statistique d'ordre associée $(L_{(i)})_{1 \leq i \leq n}$ telle que $L_{(i)} \leq L_{(i+1)}$ pour tout $1 \leq i \leq n - 1$. L'estimation de α est donnée par l'estimateur de Hill défini par

$$\hat{\alpha}_{r_n} = \left[\frac{1}{r_n} \sum_{i=1}^{r_n} \log(L_{(n-i+1)}) - \log(L_{(n-r_n)}) \right]^{-1},$$

pour un $0 < r_n < n$ fixé. Si $r_n \rightarrow +\infty$ et $\frac{r_n}{n} \rightarrow 0$ lorsque $n \rightarrow +\infty$, l'estimateur de Hill converge en probabilité vers α .

Il faut noter que les données du marché fournies par ALPHABILITY, bien que couvrant dix années, restent peu nombreuses puisque l'on n'étudie que les données

journalières. Ce qui induit un faible nombre de données disponibles pour les estimations. Il est donc inutile de chercher une méthode très fine, puisque les erreurs statistiques seront dans tous les cas importantes. Afin de réaliser un bon compromis entre précision et coût numérique, on va opter pour une méthode d'estimation simple, inspirée de la discussion ci-dessous.

L'estimateur de l'indice $\gamma = 1/\alpha$, que nous allons utiliser dans notre étude, est donné par le calcul de la pente sur le graphe de Pareto avec une technique de moindres carrés sur un intervalle qui se concentre sur la queue de distribution mais qui ne prend pas en compte les valeurs les plus grandes des pertes, qui sont aussi les plus entachées d'erreurs. Ce choix est une variante de l'estimateur de Hill, qui représente une version pondérée de l'estimateur des moindres carrés. Nous proposons aussi un estimateur pour la constante C .

Nous considérons dans ce chapitre l'estimation des paramètres de la loi puissance en prenant une fenêtre glissante de 252 jours. Sur une fenêtre de 252 jours, seulement environ la moitié des données correspond à des pertes. Donc entre 5 et 10 données seulement sont utilisées pour estimer la queue de distribution. Ainsi, notre choix de ne pas considérer les quantiles supérieurs à 99% s'avère pertinent pour éviter que les valeurs les plus grandes perturbent trop l'estimation de γ .

Les backtests de validation ont abouti à des résultats meilleurs que ceux du modèle gaussien et nous avons constaté que le modèle des lois puissances sous-estime moins la Value-at-Risk pour les faibles niveaux de risque. Cela revient au fait que ce modèle considère les queues de distribution lourdes et donne une plus grande importance aux pertes extrêmes.

Cependant, nous remarquons qu'à chaque passage d'une période, où les rendements subissent de grandes fluctuations, à une période de petites fluctuations, la VaR est sur-estimée. Inversement, à chaque passage d'une période de petites fluctuations à une période de grandes fluctuations, la VaR est sous-estimée. Nous expliquons cela par le fait que les données ne sont pas en réalité i.i.d. et que le marché passe par des périodes de grandes fluctuations (crise) et des périodes de petites fluctuations (hors-crise) et que la considération des modèles i.i.d. aboutit à des erreurs d'estimations à chaque changement de régimes. Cette problématique fera l'objet des deux chapitres suivants, où nous proposons un modèle des lois puissances à changement de régimes qui permettra de réduire ces erreurs d'estimations.

Dans le **chapitre 3**, nous introduisons la notion du *modèle de Markov caché* (MMC) que nous appliquons pour classer les données boursières et déterminer les différents régimes du marché. Nous supposons que le marché passe par deux régimes : un régime de *crise* (c), où les rendements de l'actif sont caractérisés par des grandes variations, et un régime dit *hors crise* (h), où le prix de l'actif est stable et ses rendements ne subissent par conséquent que des petites fluctuations. Un modèle à changement de régimes suppose que cette information est résumée dans un processus X d'états cachés, à valeurs dans $\{c, h\}$, qui est à l'origine des rendements observés. Les paramètres du modèle subissent des changements discrets et occasionnels décri-

vant la transition d'un régime à un autre. Ainsi, la loi des pertes est donnée par

$$\mathbb{P}(L_t \geq x) = \mathbb{P}(L_t \geq x | X_t = c) \mathbb{P}(X_t = c) + \mathbb{P}(L_t \geq x | X_t = h) \mathbb{P}(X_t = h).$$

Dans la littérature, il existe plusieurs techniques pour construire ce processus d'états. Les travaux [4, 86] utilisent des fonctions logistiques des variables endogènes retardées. L'article [147] applique des fonctions probablistes et l'article [24] se base sur le filtre de Hamilton dans son estimation.

Toutes ces méthodes reposent sur une connaissance a priori ou un choix de la forme des probabilités de transition. Afin d'utiliser le moins possible d'a priori, nous proposons une procédure d'apprentissage non-supervisé pour estimer ces probabilités à partir des données réelles. Pour cela, nous optons dans ce travail pour le modèle de Markov caché. Ce dernier considère que le processus X d'états cachés est une chaîne de Markov de loi initiale ρ et de matrice de transition Q . De plus, il fait l'hypothèse que, conditionnellement aux états X_t , les observations R_t sont mutuellement indépendantes et que chaque observation ne dépend que de l'état X_t et l'hypothèse que la loi conditionnelle de R_t sachant X_t ne dépend pas de t et admet une densité.

Sous ces hypothèses, un modèle de Markov caché est entièrement caractérisé par $M = (\rho, Q, \psi)$ où ρ est la loi initiale du processus d'état X , Q est sa matrice de transition et ψ est le noyau markovien du processus R sachant le processus X .

En partant d'une suite de rendements $(R_1, \dots, R_t)$, nous déterminons les paramètres $M = (\rho, Q, \psi)$ qui maximisent la vraisemblance d'apparition de cette suite d'observations. L'algorithme de Baum-Welsh a été appliqué pour résoudre ce problème, appelé souvent problème d'apprentissage. Ensuite, en ayant la suite d'observations $(R_1, \dots, R_t)$ et les paramètres du modèle $M = (\rho, Q, \psi)$, nous cherchons la suite d'états $(X_1, \dots, X_t)$ qui est le plus vraisemblablement à l'origine de ces observations $(R_1, \dots, R_t)$. Ce problème est appelé le problème de reconnaissance et est résolu par l'algorithme de Viterbi.

Dans le **chapitre 4**, nous proposons un modèle pour les queues de distribution, basé sur les lois puissances et prenant en compte ces changements de régimes. Nous admettons que la queue de distribution des pertes se comporte selon une loi puissance. Cependant nous pensons qu'il s'agit de deux lois puissances : une pour la queue de distribution des pertes en régime crise et une pour la queue de distribution des pertes en période de hors crise. Ainsi, la queue de distribution des pertes est décrite par la somme pondérée des deux lois :

$$\mathbb{P}(L_t \geq x) = C_c x^{-\alpha_c} \mathbb{P}(X_t = c) + C_h x^{-\alpha_h} \mathbb{P}(X_t = h), \quad \text{pour } x \geq x_0.$$

L'estimation des paramètres de crise se fait sur des fenêtres glissantes contenant uniquement des données du régime de crise. De la même façon, les paramètres du régime hors-crise sont estimés sur des fenêtres glissantes sur le régime hors-crise. Nous fixons $t_0 = 1500$, le nombre minimal de données pour avoir une estimation du modèle de Markov caché. Pour $t \geq t_0$, la $\text{VaR}_t^{(a)}$, à un niveau de confiance a , est donnée par le quantile $q^{L_{t+1}}(a)$. La procédure d'estimation de la VaR est la suivante :

- (a) On estime les paramètres $M = (\rho, Q, \psi)$ du MMC à partir de l'échantillon $(R_1, \dots, R_t)$ par l'algorithme de Baum-Welsh.
- (b) On estime la suite des états cachés $(X_1, \dots, X_t)$ par l'algorithme de Viterbi.
- (c) On estime les paramètres (C_c, α_c) de la loi Pareto du régime de crise à partir des dernières 252 données du régime crise dans $(R_1, \dots, R_t)$.
- (d) On estime les paramètres (C_h, α_h) de la loi Pareto du régime de hors-crise à partir des dernières 252 données du régime hors-crise dans $(R_1, \dots, R_t)$.
- (e) On calcule la probabilité d'être le lendemain en état de crise :

$$p_{t+1} = \mathbb{P}_M(X_{t+1} = c | R_{1:t}).$$

- (f) Le quantile $q^{L_{t+1}}(a)$, à un niveau a , de la perte à l'instant $t+1$ est ainsi solution de

$$1 - a = p_{t+1}C_c x^{-\alpha_c} + (1 - p_{t+1})C_h x^{-\alpha_h}.$$

Cette équation est résolue numériquement en utilisant la méthode de Newton-Raphson.

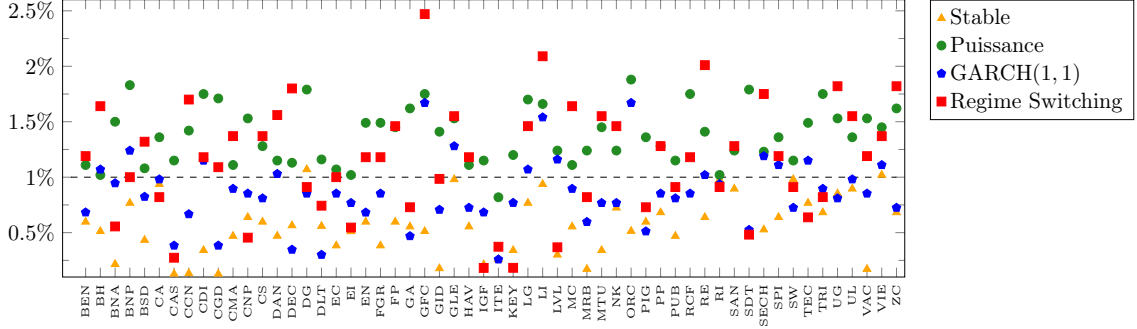
L'application des procédures de backtesting sur la VaR estimée sous le modèle des lois puissances à changement de régimes certifie ce constat d'existence de deux régimes du marché. En effet, les délais d'estimation et les sous/sur-estimations à chaque changement de régime, que l'on a constatés dans le modèle gaussien et le modèle des lois puissances, sont éliminés par l'introduction des notions crise et hors-crise. Nous concluons que prendre en compte le régime du marché dans le calcul de la VaR conduit à une adaptation rapide au changement de régime.

La performance de prédiction du modèle des lois puissances à changement de régimes a été comparée à celles du modèle simple des lois puissances, du modèle des lois stables et du modèle GARCH. Nous estimons la VaR à 99% sous chacun des quatre modèles pour un ensemble de 56 actifs sélectionnés dans notre base de données. Cet ensemble contient les actifs ayant des cotations durant toute la période allant de janvier 2001 à février 2011. Parmi ces actifs, nous trouvons certaines des actions les plus liquides qui composent l'indice CAC40. La performance de chaque modèle est résumée par la moyenne de ses résultats en termes du Backtesting sur la VaR, sur l'ensemble des actifs sélectionnés.

La figure suivante (copie de la figure 4.7 de la page 83) présente les ratios de dépassement de la VaR sous chacun des quatre modèles pour les 56 actifs sélectionnés. Clai-
rement, le modèle des lois stables sur-estime la VaR et le modèle des lois puissances la sous-estime. Les résultats du modèle GARCH et du modèle des lois puissances à changement de régimes fluctuent autour de la valeur cible 1%. En termes de ratio moyen de dépassement, le modèle des lois puissances à changement de régimes est le plus proche de 1%, puis le modèle GARCH, puis le modèle des lois puissances et au final le modèle des lois stables. La réalisation des autres backtests sur l'ensemble des 56 actifs indique que le modèle des lois puissances à changement de régimes donne les meilleurs résultats en termes de backtesting.

En guise de conclusion, l'introduction de la notion de crise nous a permis d'améliorer la performance de la prédiction de la Value-at-Risk en construisant un modèle qui

Comparaison des résultats du backtest des différents modèles sur 56 actifs.



prend en compte à la fois la lourdeur de la queue de la distribution et aussi les changements de régimes. Ce modèle a en plus éliminé le regroupement des dépassements de la VaR et les délais d'estimation constatés dans les modèles i.i.d.. Cependant, les backtests du ratio de dépassement n'avaient pas une tendance générale supérieure ou inférieure à la valeur cible 1%, comme c'est le cas des lois puissances ou des lois stables. Les résultats des backtests fluctuent autour de la valeur cible. Cette observation n'est pas spécifique au modèle des lois puissances à changement de régimes, mais caractérise l'ensemble des modèles conditionnels comme le montre aussi le modèle GARCH.

La **deuxième partie** de la thèse, qui s'intéresse aux processus à sauts et aux processus de Lévy en particulier, se compose de trois chapitres.

Le **chapitre 5** commence par une brève présentation des processus de Lévy, complétée d'un recueil de leurs principales propriétés, utilisées dans la suite de la thèse. Dans la section suivante, nous nous intéressons principalement aux modèles exp-Lévy, c'est-à-dire, les modèles où le prix d'un actif est décrit par l'exponentielle d'un processus de Lévy. Dans ce cadre, nous passons en revue les différentes méthodes de construction des mesures risque-neutre sous lesquelles le prix actualisé de l'actif est une martingale. Nous détaillons deux exemples construits via la transformée d'Esscher, la mesure risque-neutre d'Esscher et la mesure risque-neutre d'entropie minimale, pour lesquelles nous donnons les conditions d'existence et les paramètres des modèles risque-neutre. Ensuite, nous décrivons une méthode, due à Carr et Madan [32], permettant de valoriser les options européennes dans les modèles exp-Lévy à l'aide de la transformée de Fourier. La méthode est fondée sur l'observation suivante : si on multiplie le prix de l'option par un facteur d'amortissement :

$$c_\alpha(k) = e^{\alpha k} \times S_0 e^{-rT} \mathbb{E}^*[(e^{X_T} - e^k)_+],$$

alors la quantité obtenue est, sous certaines conditions, intégrable et on peut évaluer sa transformée de Fourier :

$$\hat{c}_\alpha(v) = \frac{S_0 e^{-rT} \Phi_T^*(v - (\alpha + 1)i)}{\alpha^2 + \alpha - v^2 + i(2\alpha + 1)v},$$

où Φ_T^* est la fonction caractéristique de X_T sous la mesure risque-neutre $\mathbb{P}^*$. Les prix d'options peuvent donc être évalués en calculant la transformée de Fourier inverse

de $\hat{c}_\alpha$. La dernière section de ce chapitre est consacrée à l'application des méthodes de construction des mesures risque-neutres et de la méthode d'évaluation de prix d'options dans le cadre de trois modèles exp-Lévy, à savoir le modèle continu de Black et Scholes, le modèle diffusion-saut de Merton et le modèle d'activité infinie variance gamma. Dans cette section, nous éclaircissons les ressemblances et les divergences entre la mesure risque-neutre d'Esscher et la mesure risque-neutre d'entropie minimale et nous étudions la sensibilité du prix de l'option par rapport aux paramètres des différents modèles.

Dans le **chapitre 6**, nous développons deux méthodes numériques pour calculer simultanément la Value-at-Risk cumulative et la Value-at-Risk conditionnelle (CVaR) cumulative dans le cadre des modèles exp-Lévy. La première méthode permet d'évaluer la VaR et la CVaR monétaires, tandis que la deuxième s'intéresse à la VaR et la CVaR en termes de rendements. Ces méthodes reposent sur l'inversion d'une transformée de Fourier exprimée en fonction de la fonction caractéristique du processus de Lévy. La CVaR à un niveau de confiance a , est une mesure complémentaire de la VaR définie par

$$\text{CVaR}_a(L) = \frac{1}{1-a} \int_a^1 \text{VaR}_x(L) dx,$$

où L est la fonction perte considérée.

Rockafellar et Uryasev [136] proposent un algorithme pour calculer simultanément la VaR et la CVaR par la résolution d'un problème unidimensionnel d'optimisation convexe présenté par le théorème ci-dessous.

En notant $a \in (0, 1)$ le niveau de risque et en introduisant la fonction

$$z \mapsto F_a(z) = z + \frac{1}{1-a} \mathbb{E}[(L - z)_+],$$

la VaR et la CVaR sont les solutions du problème suivant

$$\begin{aligned} \text{CVaR}_a(L) &= \min_{z \in \mathbb{R}} F_a(z), \\ \text{VaR}_a(L) &= \min\{y : y \in \underset{z \in \mathbb{R}}{\text{argmin}} F_a(z)\}. \end{aligned}$$

En particulier, on a toujours $\text{VaR}_a(L) \in \underset{z \in \mathbb{R}}{\text{argmin}} F_a(z)$ et $\text{CVaR}_a(L) = F_a(\text{VaR}_a(L))$.

Dans une première section, nous nous intéressons aux VaR et CVaR monétaires sur la période totale, c'est-à-dire nous considérons comme fonction perte $L = S_0 - S_T$. Donc, afin de procéder à la minimisation, l'enjeu est de calculer l'espérance $\mathbb{E}[(S_0 - S_T - z)_+]$ pour tout $z \in \mathbb{R}$. Pour $z \geq S_0$, cette espérance est nulle. Pour $z < S_0$, nous posons $w = \log(1 - z/S_0)$ et l'espérance se réécrit sous la forme $\mathbb{E}[(S_0 e^w - S_T)_+]$ pour $w \in \mathbb{R}$. En multipliant cette dernière par un facteur d'amortissement :

$$f_\alpha(w) = e^{-\alpha w} \times \mathbb{E}[(S_0 e^w - S_T)_+],$$

nous obtenons une quantité qui est, sous certaines conditions, intégrable et nous pouvons évaluer sa transformée de Fourier :

$$\hat{f}_\alpha(u) = \frac{S_0 \Phi_T(u + i(\alpha - 1))}{\alpha^2 - \alpha - u^2 - iu(2\alpha - 1)},$$

où Φ_T est la fonction caractéristique de X_T sous $\mathbb{P}$. Ainsi, l'espérance $\mathbb{E}[(S_0 - S_T - z)_+]$ peut être évaluée en calculant la transformée de Fourier inverse de $\hat{f}_\alpha$ et nous pouvons par la suite procéder à la minimisation et obtenir une évaluation numérique de la VaR et la CVaR monétaires. Nous appliquons cette méthode pour les trois modèles exp-Lévy mentionnés précédemment. Pour le modèle de Black et Scholes, nous pouvons obtenir des formules explicites pour la VaR et la CVaR et la comparaison nous donne des erreurs pour la méthode numérique de l'ordre de 10^{-2} (quelques centimes d'euro si les actions sont cotées en euro) pour la VaR monétaire et des erreurs inférieures à 10^{-4} pour la CVaR monétaire.

Dans un deuxième temps, nous nous intéressons aux VaR et CVaR en termes de rendements sur la période totale. La fonction perte considérée est $L = -\log(S_T/S_0) = -X_T$. Donc, la procédure de minimisation nécessite d'abord le calcul de $\mathbb{E}[(-X_T - z)_+]$ pour tout $z \in \mathbb{R}$. D'une façon analogue, notre méthode est fondée sur l'évaluation de cette espérance en calculant la transformée de Fourier inverse d'une certaine fonction $\hat{g}_\alpha$ et nous pouvons par la suite procéder à la minimisation. De même, nous appliquons cette méthode pour calculer la VaR et la CVaR en termes de rendements dans les cadres des modèles de Black et Scholes, de Merton et variance gamma. La comparaison avec les résultats des formules explicites obtenues sous le modèle de Black et Scholes nous donne des erreurs numériques de l'ordre de 10^{-3} pour la VaR et inférieures à 10^{-4} pour la CVaR.

Dans le **chapitre 7**, nous nous intéressons à la couverture des options européennes sous des modèles exp-Lévy. Notons $\tilde{H}$ le payoff actualisé d'un produit dérivé et $(\tilde{S}_t)_{t \in [0, T]}$ le prix actualisé de l'actif sous-jacent. Une stratégie autofinancée et admissible (V_0, ξ) est donnée par un capital initial $V_0 \geq 0$ et par un processus prévisible ξ tel que la valeur actualisée, à un instant $t \in [0, T]$, du portefeuille de couverture est positive et donnée par

$$\tilde{V}_t = V_0 + \int_0^t \xi_s d\tilde{S}_s.$$

La valeur $U_0 = \sup_{\mathbb{P}^* \in \mathcal{P}} \mathbb{E}^*[\tilde{H}]$, où $\mathcal{P}$ est l'ensemble des mesures risque-neutres, est le plus petit montant V_0 tel qu'il existe une stratégie de sur-réplication ($\mathbb{P}(\tilde{H} - \tilde{V}_T \leq 0) = 1$). Cependant, le montant U_0 est généralement très important et peut être égal à S_0 sous certains modèles exp-Lévy [19, 58]. En fixant une contrainte budgétaire $V_0 \leq V_0^m < U_0$, nous cherchons à déterminer la meilleure couverture en termes de la CVaR :

$$\min_{(V_0, \xi) \in \mathcal{A}_{V_0^m}} \text{CVaR}(\tilde{H} - \tilde{V}_T),$$

où $\mathcal{A}_{V_0^m}$ est l'ensemble des stratégies autofinancées et admissibles tel que $V_0 \leq V_0^m$.

En exploitant à nouveau l'algorithme de Rockafellar et Uryasev, nous réécrivons ce problème sous la forme $\min_{z \in \mathbb{R}} f(z)$, où

$$f(z) = z + \frac{1}{1-a} \min_{(V_0, \xi) \in \mathcal{A}_{V_0^m}} \mathbb{E}[(L(V_0, \xi) - z)_+].$$

L'enjeu majeur est alors d'évaluer $f(z)$ pour pouvoir procéder par la suite à une

optimisation unidimensionnelle et déterminer la CVaR minimale. La fonction $f(z)$ admet une forme équivalente qui ne peut être résolue que dans des marchés complets. Pour cela, nous proposons une approximation $\hat{f}(z)$, basée sur la relaxation de la contrainte du problème dual associé à la forme équivalente de $f(z)$. Nous donnons une expression de $\hat{f}(z)$ qui sera calculée numériquement par des techniques d'inversion de Fourier. Cela nous permet de proposer une approximation de la CVaR minimale. Nous illustrons ces résultats pour des exemples d'exponentielle-Lévy.

Nous finirons par un **chapitre de conclusion** présentant les questions ouvertes, les limitations et les perspectives de différents travaux réalisés dans cette thèse.

Première partie

Mesure de risque et estimation de la Value-at-Risk

Chapitre 1

La Value-at-Risk : Estimation et validation

La Value-at-Risk reste l'un des principaux indicateurs de risque pour la gestion des portefeuilles financiers. Elle peut être estimée par des méthodes paramétriques ou non-paramétriques [91]. Les méthodes non-paramétriques reposent seulement sur les distributions empiriques (historique, ré-échantillonnage) sans l'ajustement d'un modèle. Dans le cas de faible quantité des données disponibles, ces méthodes ne parviennent pas à fournir une estimation précise de la probabilité des événements extrêmes. L'approche paramétrique remplace en partie le problème induit par le manque de données en ajustant les paramètres d'un modèle sur l'historique des données et en calculant par la suite la VaR soit par des méthodes analytiques, soit par des méthodes numériques.

La méthodologie RiskMetrics [121] est largement utilisée pour évaluer le risque associé à un portefeuille en établissant des relations quantitatives entre les variations de l'indicateur de risque par rapport aux facteurs de risque (les actions et les prix des produits dérivés par exemple). Aujourd'hui, cette méthodologie intègre les lois à queue lourde, mais elle a été initialement développée dans le contexte gaussien, qui est encore assez répandu dans la gestion du risque et appliqué par les accords de Bâle [156].

Nous commençons par définir la Value-at-Risk dans la section 1.1. Nous nous plaçons ensuite dans le cadre du modèle gaussien. Dans la section 1.2, nous introduisons la méthodologie d'estimation de cet indicateur de risque. Les différents backtests de validation, qui permettent de vérifier l'adéquation de la VaR mesurée sous un certain modèle avec le vrai comportement des données financières réelles, sont détaillés dans la section 1.3. Cela est fait en appliquant les méthodologies d'estimation et de validation à un échantillon de notre base de données.

1.1 Value-at-Risk

La *Value-at-Risk* (VaR) est un indicateur utilisé par la majorité des institutions financières pour évaluer le risque du marché. Cette indicateur mesure, de manière quantitative, la perte potentielle maximale que nous pouvons subir en dehors d'un événement de probabilité p donnée et sur une période de durée donnée. Le principal attrait de cette mesure de risque réside dans le fait d'interpréter le risque par un chiffre simple et compréhensible pour tous.

Nous considérons un actif financier de cours $(S_t)_{t \geq 0}$, où le temps t est mesuré par unité de jours. Nous appelons rendements logarithmiques journaliers ou plus simplement *rendements journaliers* de l'actif $(S_t)_{t \geq 0}$ la suite $(R_t)_{t \geq 1}$, où

$$R_t = \ln(S_{t+1}) - \ln(S_t) = \ln \frac{S_{t+1}}{S_t}.$$

Nous appelons *pertes journalières* la suite $(L_t)_{t \geq 1}$ donnée par $L_t = -R_t$. Lorsque $L_t < 0$, nous parlons d'un gain.

Définition 1.1. Soit X une variable aléatoire réelle et $a \in [0, 1]$. Nous définissons $q^X(a-)$ et $q^X(a)$ les quantiles respectivement inférieur et supérieur de niveau a par :

$$q^X(a-) = \inf\{x \in \mathbb{R} | \mathbb{P}(X \leq x) \geq a\}, \quad (1.1)$$

$$q^X(a) = \inf\{x \in \mathbb{R} | \mathbb{P}(X \leq x) > a\}. \quad (1.2)$$

Avec cette notion, nous pouvons introduire la définition de la Value-at-Risk. Nous supposons dans la suite que a est un paramètre $a \in [0, 1]$ qui est proche de 1.

Définition 1.2. La Value-at-Risk monétaire, à un instant t , d'un titre financier de cours $(S_t)_{t \geq 0}$ sur un horizon de temps T (exprimé en jour), et à un seuil de confiance a est le $a^{\text{ème}}$ quantile supérieur de la distribution de pertes en unité monétaire sur une durée T , c'est-à-dire

$$\text{VaR}_{[t, t+T]}^{(a)}(\text{€}) = q^{S_t - S_{t+T}}(a). \quad (1.3)$$

Si les variables aléatoires $(S_t)_{t \geq 0}$ sont continues, nous pouvons écrire directement

$$\mathbb{P}\left(S_t - S_{t+T} > \text{VaR}_{[t, t+T]}^{(a)}(\text{€})\right) = 1 - a. \quad (1.4)$$

Lorsque l'horizon temporel se réduit à un jour, cette définition peut s'exprimer à l'aide de la perte journalière L_{t+1} et nous introduisons $\text{VaR}_{[t, t+1]}^{(a)}$ la Value-at-Risk journalière en termes de rendement.

Afin d'alléger l'écriture, nous désignerons par la notation $\text{VaR}_t^{(a)}$ la Value-at-Risk sur l'horizon temporel $[t, t+1]$.

Définition 1.3. La Value-at-Risk journalière en termes de rendement, à un instant t et à un niveau confiance a , est le $a^{\text{ème}}$ quantile supérieur de la distribution de la perte journalière L_{t+1} ,

$$\text{VaR}_t^{(a)} = q^{L_{t+1}}(a). \quad (1.5)$$

Si les variables aléatoires $(S_t)_{t \geq 0}$ sont continues, nous pouvons écrire directement

$$\mathbb{P}(L_{t+1} > \text{VaR}_t^{(a)}) = \mathbb{P}(R_{t+1} < -\text{VaR}_t^{(a)}) = 1 - a. \quad (1.6)$$

Nous nous intéressons principalement à la VaR journalière en termes de rendement, que nous appelons dans la suite la VaR, lorsqu'il n'y aura pas d'ambiguïté. Cela représente le cas d'un investisseur qui veut, à un instant donné et pour un niveau de confiance donné, estimer la perte potentielle maximale, en termes de rendement, qu'il pourrait subir le lendemain.

Remarque 1.1. Nous pouvons toujours convertir la VaR en termes de rendement en unité monétaire par le biais de la formule suivante :

$$\text{VaR}_t^{(a)}(\text{€}) = S_t \left(1 - \exp \left(-\text{VaR}_t^{(a)} \right) \right). \quad (1.7)$$

En pratique, nous cherchons le seuil tel que, en moyenne, 99 fois sur 100, les pertes lui sont inférieures. Nous parlons alors de la VaR à un niveau de confiance de 99%.

Nous pouvons calculer la VaR pour un seul actif ou pour un portefeuille d'actifs de composition fixée. Nous nous restreignons dans ce travail à un seul actif.

Dans ce chapitre, nous choisissons le modèle gaussien comme modèle de l'actif boursier, pour présenter la méthodologie d'estimation et de validation de la VaR. Dans ce modèle, le logarithme des prix de l'actif est décrit par un mouvement brownien. Ainsi, les rendements suivent une loi normale.

1.2 Modèle gaussien et ses limites dans la mesure de risque

Le mouvement brownien standard est très utilisé en finance pour étudier l'évolution des rendements des actifs, à titre d'exemple le modèle de Black-Scholes [25] n'est rien d'autre qu'un mouvement brownien géométrique. Ces modèles d'actifs financiers supposent que leurs cours sont continus en temps et suivent une distribution log-normale dont les accroissements logarithmiques, autrement dit les rendements, sont indépendants, stationnaires et de loi gaussienne. Dans cette section, nous allons montrer que la modélisation des rendements financiers par une loi normale pour mesurer la VaR, n'est pas conforme aux données du marché et que ce modèle sous-estime la probabilité des événements extrêmes.

1.2.1 La loi normale

Dans cette partie, nous comparons la densité empirique des rendements d'un actif financier et la densité de la loi normale. A titre d'exemple, nous représentons sur la figure 1.1 les distributions empiriques des rendements du cours journalier du titre BNP Paribas du 1^{er} janvier 2001 au 25 février 2011. Nous traçons également sur la même figure la densité normale $\mathcal{N}(m, \sigma^2)$, où m est la moyenne empirique des rendements observés et σ^2 la variance empirique de cet échantillon.

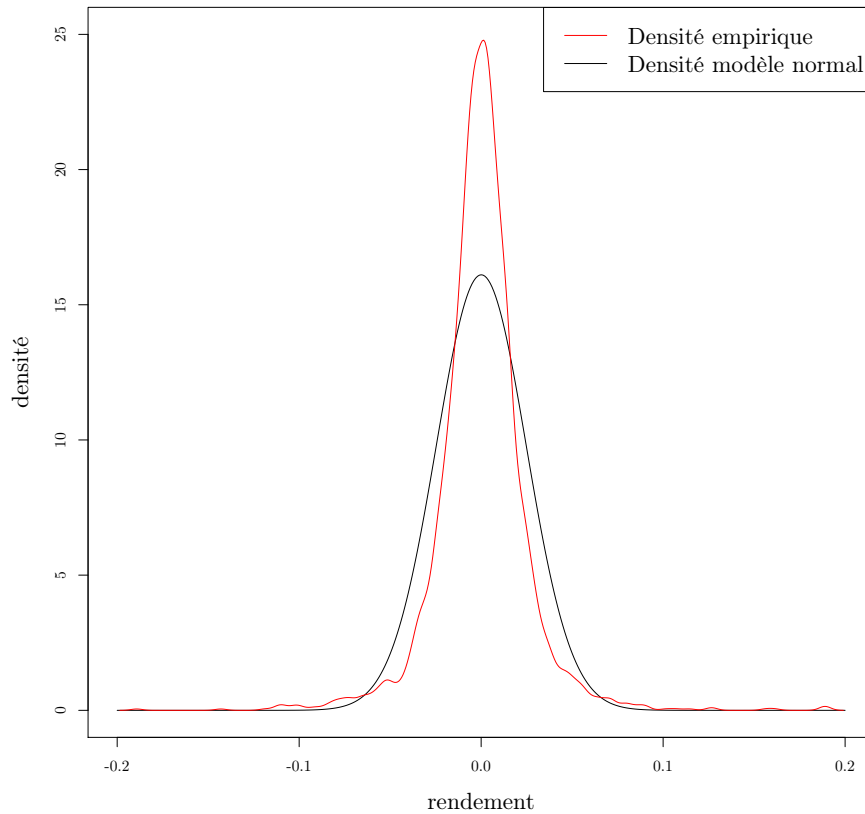


FIGURE 1.1 – Distribution empirique des rendements de l'action BNP Paribas du 1^{er} janvier 2001 au 25 février 2011, et distribution gaussienne estimée sur l'ensemble des données.

La densité empirique des rendements semble symétrique par rapport à m , mais elle est plus pointue que la distribution de la loi normale. La queue de la distribution empirique est plus épaisse que la queue de la distribution normale. Pour faire une comparaison dans ce sens entre la densité empirique et la densité gaussienne, nous considérons le kurtosis qui est un indicateur de l'aplatissement relatif d'une distribution donnée comparée à une distribution normale.

Définition 1.4. *Le kurtosis d'une variable aléatoire réelle X de moyenne μ et*

d'écart-type σ est le moment d'ordre quatre de la variable centrée réduite $\frac{X-\mu}{\sigma}$:

$$K_X = \mathbb{E} \left[\left(\frac{X - \mu}{\sigma} \right)^4 \right]. \quad (1.8)$$

On vérifie facilement que le kurtosis de la loi normale vaut 3. Un kurtosis supérieur à 3 indique une distribution relativement pointue, tandis qu'un kurtosis inférieur à 3 signale une distribution relativement aplatie.

Nous calculons dans le tableau 1.1 le kurtosis empirique pour plusieurs données de cours d'actions. Nous obtenons des valeurs variées, toujours supérieures à 3 et parfois largement plus grandes.

Actions	BNP	CA	CNP	GA	GLE	LG	MC	TEC	UL
Kurtosis	12.39	6.82	7.11	7.34	9.46	8.34	8.42	9.00	10.17

TABLE 1.1 – Kurtosis d'actifs financiers de la bourse de Paris sur la période 2001-2011

Ces résultats montrent que les distributions de ces séries sont plus pointues que la gaussienne. Ceci constitue une première indication que la loi normale ne modélise pas correctement les rendements des actifs financiers.

En outre, le fait que la densité gaussienne est caractérisée par une très forte décroissance, les modèles gaussiens excluent les grandes fluctuations. En particulier, ils ne rendent pas compte des événements extrêmes qui se produisent sur les marchés.

1.2.2 Estimation de la Value-at-Risk

Dans le modèle gaussien, le logarithme des prix de l'actif est décrit par un mouvement brownien. Ainsi, les rendements suivent une loi normale. L'expression de la VaR découle de cette hypothèse de normalité.

Proposition 1.1. *Dans le modèle gaussien, la VaR à un instant donné t et au seuil de confiance a , est donnée par*

$$\text{VaR}_t^{(a)} = -\mathbb{E}[R_t] - \Phi(1 - a)\sigma_t, \quad (1.9)$$

où R_t est le rendement de l'actif, σ_t est son écart-type et Φ est la fonction quantile de la loi normale centrée réduite.

Lorsque nous disposons d'un échantillon de n observations des rendements journaliers $R = (R_1, \dots, R_n)$, nous pouvons estimer la VaR d'horizon 1 jour à partir de la formule précédente comme suit : nous calculons les estimateurs sans biais de la moyenne et l'écart-type par

$$\hat{R} = \frac{1}{n} \sum_{i=1}^n R_i \quad \text{et} \quad \hat{\sigma} = \left(\frac{1}{n-1} \sum_{i=1}^n (R_i - \hat{R})^2 \right)^{1/2},$$

et nous obtenons une estimation de la VaR par

$$\widehat{\text{VaR}}_n^{(a)} = -\hat{R} - \Phi(1 - a)\hat{\sigma}. \quad (1.10)$$

Estimation sur une fenêtre glissante Le modèle gaussien suppose que les rendements sont stationnaires. Or il n'est pas légitime de supposer que les paramètres d'un tel modèle sont constants au fil du temps. Pour cela, notre méthode d'estimation repose sur l'approche de la fenêtre glissante illustrée par la figure 1.2. Nous considérons les rendements $(R_t)_{t \geq 1}$ de l'actif. Nous fixons une taille de fenêtre ω . Pour tout $t \geq \omega$, nous considérons la suite $(R_{t-\omega+1}, \dots, R_t)$ des rendements sur une fenêtre de ω jours antérieurs. Nous estimons ainsi la $\text{VaR}_t^{(a)}$ à partir de cette suite en appliquant la formule (1.10). La motivation derrière l'utilisation d'une fenêtre glissante est de capter les caractéristiques dynamiques des données qui varient en fonction des périodes du temps.

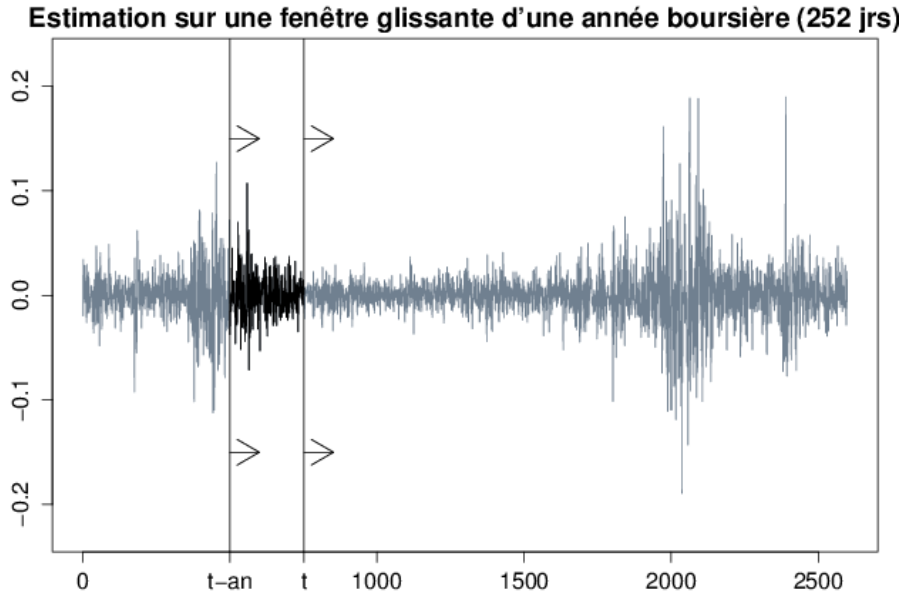


FIGURE 1.2 – Estimation sur une fenêtre glissante de 252 jours

La littérature propose généralement des fenêtres glissantes de taille une année boursière (252 jours) ou 4 années boursières (1000 jours). L'étude de l'influence de la taille de la fenêtre d'estimation sur la pertinence du modèle dans la mesure de la VaR est présentée dans la section 2.4.1. Cette étude nous amène à considérer des fenêtres de 252 jours. Dans la figure 1.3, nous représentons le prix et les estimations de la volatilité et la Value-at-Risk journalière monétaire $\widehat{\text{VaR}}_t^{(0.99)}$ (€) sur le cours de l'action BNP Paribas en utilisant le modèle normal. La volatilité et la VaR sont estimées sur une fenêtre glissante de 252 jours, l'équivalent d'une année boursière.

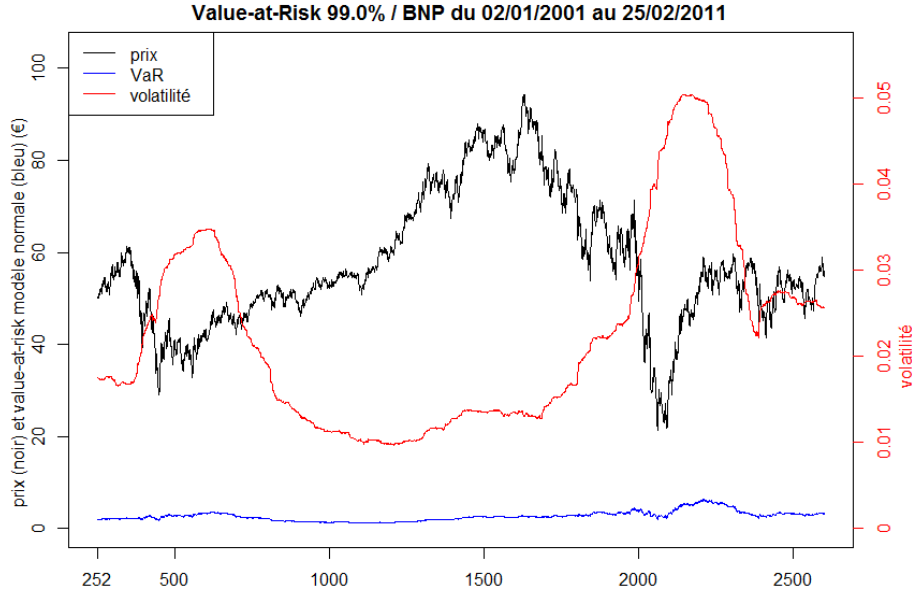


FIGURE 1.3 – Prix, volatilité et Value-at-Risk journalière monétaire à 99% de l’actif BNP Paribas en utilisant le modèle normal. La volatilité et la Value-at-Risk sont calculées sur une fenêtre glissante de 252 jours.

1.3 Backtesting sur la VaR

Notre objectif est de valider l’adéquation du choix du modèle gaussien avec la description des données du marché et de tester la précision du calcul de la VaR donnée par la proposition 1.1. Étant donné que la véritable valeur de la VaR n’est pas observable, les techniques usuelles de validation des prévisions telles que l’erreur quadratique moyenne et l’erreur absolue moyenne en pourcentage ne peuvent pas être utilisées. Dès lors, des procédures de *backtesting* ont été développées.

De manière générale, le backtesting consiste à tester un modèle de mesure de risque sur une grande base de données historiques pour évaluer son efficacité et valider son application dans la pratique. Il s’agit d’un ensemble de procédures statistiques, appelées parfois tests rétro-actifs de validité, dont le but est de vérifier que la fréquence de dépassements de la VaR estimée par les pertes réelles observées est conforme avec le niveau de confiance de l’estimation. Cela implique de comparer systématiquement l’historique des prévisions de la VaR aux rendements observés de l’actif [30].

Un dépassement se produit quand la perte à un instant $t + 1$ est supérieure à la VaR estimée à l’instant t , à partir des données précédentes. Nous fixons ω la taille de la fenêtre d’estimation et nous définissons la suite $(I_t(a))_{t \geq \omega}$ des dépassements de la VaR comme suit :

$$I_t(a) = \begin{cases} 1 & \text{si } R_{t+1} \leq -\widehat{\text{VaR}}_t^{(a)}, \\ 0 & \text{sinon.} \end{cases} \quad (1.11)$$

Christoffersen [36] réduit le problème de validation d’un modèle dans l’estimation de la VaR à celui de l’étude de sa suite de dépassements. Ainsi, pour qu’un modèle

estime la VaR avec précision, la suite de dépassements qui en résulte doit satisfaire les deux propriétés suivantes :

1. *La propriété de couverture non-conditionnelle* : Chaque jour la probabilité d'avoir un dépassement doit être exactement égale au niveau de risque $p = 1 - a$:

$$\mathbb{P}(I_t(a) = 1) = \mathbb{E}[I_t(a)] = p. \quad (1.12)$$

2. *La propriété d'indépendance* : Cette propriété impose une restriction sur la façon dont les dépassements sont distribués. La suite des dépassements issues d'un même niveau de risque doit être une suite de variables aléatoires indépendantes. Un dépassement aujourd'hui n'induit pas un dépassement le lendemain. Cette propriété assure que tout regroupement de dépassements est évité.

Les deux propriétés de la suite de dépassements sont souvent combinées dans l'hypothèse suivante : les variables aléatoires $(I_t(a))_{t \geq \omega}$, sont indépendamment et identiquement distribuées selon une loi de Bernoulli de paramètre $p = 1 - a$. Ces deux propriétés représentent une caractérisation complète de la précision de la mesure de la VaR, à un niveau de confiance donné a . En d'autres termes, toute suite d'estimations de VaR $(\widehat{\text{VaR}}_t^{(a)})_{t \geq \omega}$, qui coïncide avec les quantiles $(q^{L_{t+1}}(a))_{t \geq \omega}$ à un niveau a des pertes, doit produire une suite de dépassements $(I_t(a))_{t \geq \omega}$ vérifiant les propriétés de couverture non-conditionnelle et d'indépendance. Inversement, si une série donnée d'estimateurs de la VaR produit une suite de dépassements respectant ces deux propriétés, alors l'estimation de la VaR est précise au sens où l'estimateur $\widehat{\text{VaR}}_t^{(a)}$ coïncide avec le vrai quantile de perte $q^{L_{t+1}}(a)$ pour tout $t \geq \omega$.

Plusieurs backtests de la VaR ont été proposés dans ces dernières années. Ces tests examinent soit l'une des deux propriétés soit les deux ensemble. Nous discutons dans la suite les tests du ratio de dépassement et de couverture non-conditionnelle qui examinent la première propriété, ensuite nous présentons un test d'indépendance et finalement nous introduisons le test de couverture conditionnelle qui examine simultanément les deux propriétés.

1.3.1 Test du ratio de dépassement

Principe du test

Ce test permet de valider la première propriété de couverture non-conditionnelle.

Définition 1.5. Soient un échantillon de n mesures de la VaR à un niveau de confiance a et $(I_t(a))_{1 \leq t \leq n}$ la suite des dépassements associée. Nous définissons le ratio de dépassement comme étant la moyenne empirique des dépassements :

$$\hat{p} = \frac{1}{n} \sum_{j=1}^n I_j(a) = \frac{k}{n}, \quad (1.13)$$

où k est le nombre de dépassements.

Proposition 1.2. *Lorsque la VaR est estimée à un niveau de confiance a , l'espérance du ratio de dépassement doit être égale à $p = 1 - a$. Un ratio de dépassement supérieur à p implique une sous-estimation de la VaR. Contrairement, un ratio de dépassement inférieur à p implique une sous-estimation de la VaR sous le modèle choisi.*

Sous l'hypothèse que $\sum_{i=1}^n I_t(a)$ est la somme de n variables aléatoires indépendantes, cette somme suit une loi binomiale $\mathcal{B}(n, p)$ où n est la taille de l'échantillon des prévisions et p est le niveau de risque égal à $1 - a$. Un intervalle de confiance du ratio de dépassement à un niveau de confiance $(1 - \theta)$ est donné par la proposition suivante.

Proposition 1.3 ([2, 26]). *Soit un échantillon de n mesures de la VaR à un niveau de confiance a et soit k le nombre de dépassements dans cet échantillon. Un intervalle de confiance dit exact du ratio de dépassement $\hat{p} = \frac{k}{n}$, à un niveau de confiance $(1 - \theta)$ est donné par :*

$$p = 1 - a \in \left[\frac{1}{1 + \frac{n-k+1}{k} F_{2(n-k+1), 2k}(1 - \theta/2)}, \frac{\frac{k+1}{n-k} F_{2(k+1), 2(n-k)}(1 - \theta/2)}{1 + \frac{k+1}{n-k} F_{2(k+1), 2(n-k)}(1 - \theta/2)} \right],$$

où F_{ν_1, ν_2} est la fonction de répartition inverse d'une loi de Fischer de degrés de liberté ν_1 et ν_2 .

Étude empirique sur le ratio de dépassement

Dans cette section, nous calculons le ratio de dépassement pour des estimations de la VaR sur les cours d'un ensemble d'actions boursières entre 2001 et 2011. Nous étudions les résultats du backtest ainsi que leurs variations en fonction des paramètres d'estimation.

Dans la figure 1.4, nous réalisons ce backtest pour les rendements des actions BNP Paribas, CA, CNP et FGR entre 2001 et 2011. Nous disposons de 2598 données de rendements pour chaque actif. La VaR est calculée sur une fenêtre glissante de 252 jours et à un niveau de confiance de 99%, c'est-à-dire $a = 0.99$. La courbe de $(-\widehat{\text{VaR}}_t^{(a)})_{t=252, \dots, 2598}$ est représentée en bleu. Les dépassements de la VaR sont signalés en vert. Nous constatons que les ratios de dépassement pour ces actifs sont beaucoup plus grands que le niveau de risque $p = 1 - a = 1\%$. Par ailleurs, p n'appartient pas à l'intervalle de confiance à 99% du ratio de dépassement. Ainsi, le modèle gaussien sous-estime la VaR à un niveau a pour ces actifs.

Des résultats similaires ont été trouvés pour les autres données boursières dont nous disposons. La figure 1.5 montre le ratio de dépassement $\hat{p}$ ainsi que son intervalle de confiance à 99% pour un échantillon de 12 actifs. Pour chacun de ces actifs, la VaR est toujours estimée sur une fenêtre glissante de 252 jours et à un niveau $a = 99\%$ sur la période 2001-2011. La première constatation est que pour les 12 actifs le

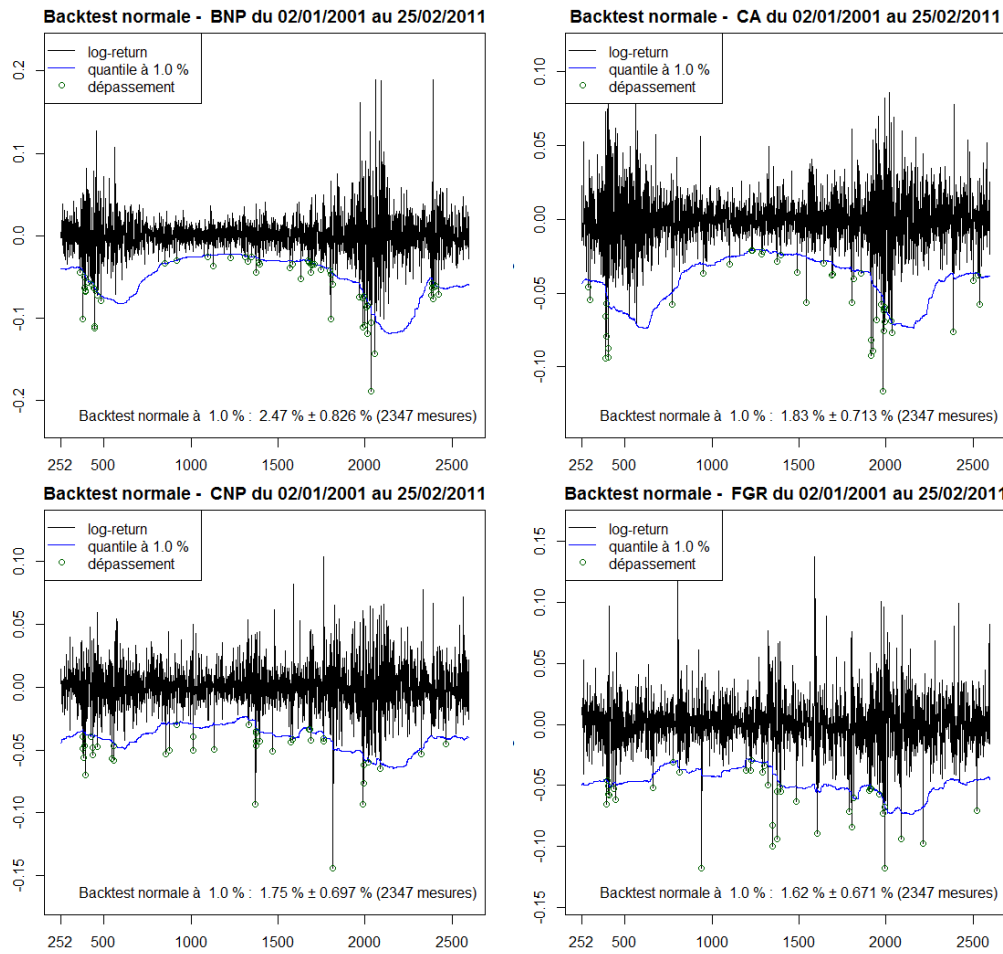


FIGURE 1.4 – Backtesting sur la Value-at-Risk à 99% estimée sur une fenêtre glissante de 252 jours sur le cours des actions BNP, CA, CNP et FGR.

ratio de dépassement est supérieur au niveau de risque $p = 1 - \alpha$ avec une valeur moyenne plus proche à 2%, d'où une sous-estimation du risque réel présent sur le marché. Nous remarquons aussi que, seulement pour 2 actifs parmi 12, l'hypothèse de l'appartenance de p à l'intervalle de confiance à 99% de $\hat{p}$ n'est pas rejetée. Ce qui confirme bien la non-adéquation du modèle gaussien dans l'estimation du risque avec des niveaux de confiance α très élevés.

Ratio de dépassement en fonction de la taille de fenêtre d'estimation On considère de nouveau l'action BNP Paribas sur la période 2001-2011. Nous disposons de 2599 données de prix de clôture de cet actif. Nous calculons la VaR à un niveau $\alpha = 99\%$ sur une fenêtre glissante de taille ω qui varie de 252 jusqu'à 2100. Nous traçons ensuite les résultats des backtests correspondants en fonction de la longueur ω de la fenêtre d'estimation dans la figure 1.6.

La seule partie pertinente dans ce graphique correspond aux valeurs en abscisse inférieures à 1750. Au delà de cette valeur, la courbe du ratio de dépassement décroît. En effet, nous disposons de 2599 données de prix de clôture de l'actif BNP. Ainsi, pour

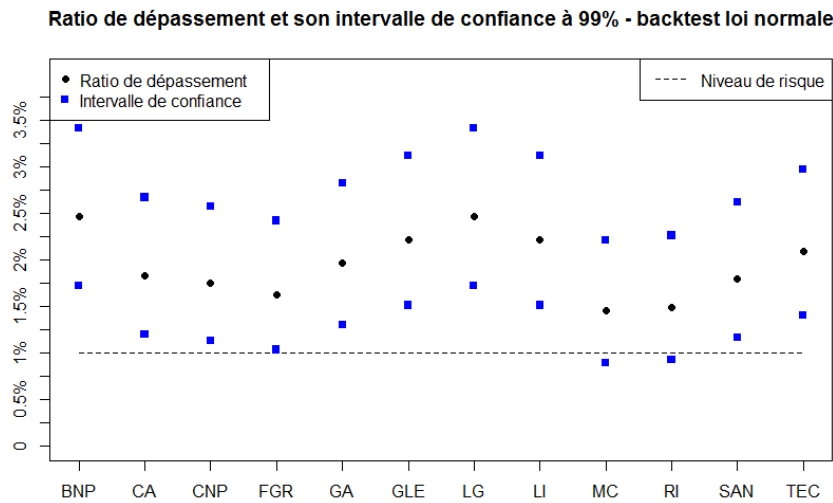


FIGURE 1.5 – Le ratio de dépassement et son intervalle de confiance à 99% pour un échantillon de 12 actifs. La VaR est toujours estimée sur une fenêtre glissante de 252 jours et à un niveau $\alpha = 99\%$ sur la période 2001-2011.

des valeurs plus grandes que 1750, le ratio en ordonnée est calculé sur un échantillon de dépassements de taille $2599 - 1750 = 849$, ce qui conduit à trop d'imprécision sur ce ratio. De plus, une fenêtre trop grande inclut des données recouvrant une ou deux crises financières sur laquelle un modèle gaussien stationnaire ne peut pas être ajusté.

La figure 1.6 montre que la valeur de la VaR et sa performance en termes de backtesting dépendent de la longueur de la fenêtre d'estimation. En effet, en élargissant la taille de l'échantillon de l'estimation, nous avons plus tendance à confondre les périodes de crises et les périodes de stabilité du marché ce qui revient à caractériser la distribution des rendements par une seule volatilité moyenne. Ceci entraîne une sous-estimation de la VaR surtout en période de crise. Toutefois, il faut garder une fenêtre d'estimation assez grande de telle sorte que les erreurs d'estimation des paramètres du modèle et de la probabilité de dépassement soient petites. Bien qu'il n'existe pas une référence unique pour le choix de la taille de la fenêtre, la littérature propose souvent des fenêtres de 252 jours (une année boursière) ou de 1000 jours. De plus, le backtesting sur nos données du marché montre que les estimations de la VaR les plus concluantes sont celles d'une fenêtre glissante d'une année boursière. Sauf contre-indication, nous choisissons cette durée pour estimer les paramètres des modèles et tester les méthodes proposées dans la suite.

Quelle que soit la longueur de la fenêtre d'estimation, nous observons sur la figure 1.6 que les pertes réelles se produisent au delà de la VaR mesurée à un pourcentage de plus de 2%. Ceci confirme que le modèle gaussien sous-estime l'épaisseur de la queue de la distribution et ne prend pas en considération les grandes fluctuations.

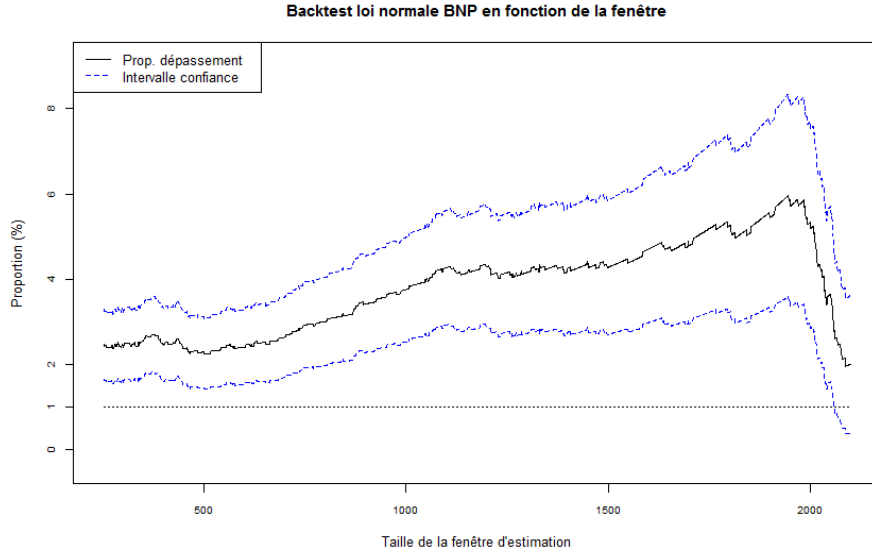


FIGURE 1.6 – La variation du ratio de dépassement et son intervalle de confiance à 99% en fonction de la taille de la fenêtre d’estimation.

Ratio de dépassement en fonction du niveau de confiance a de la VaR

Dans ce paragraphe, nous étudions l’effet de la variation du niveau de confiance a de la VaR sur sa performance en termes de backtesting. Nous considérons quatre actifs de notre base de données. Nous fixons la taille de la fenêtre d’estimation à 252 jours et nous faisons varier le niveau de confiance a dans $[90\%, 99.5\%]$ et donc le niveau de risque $p = 1 - a$ dans $[0.5\%, 10\%]$. Les résultats des backtests en fonction du niveau de risque p sont représentés sur la figure 1.7.

Dans cette figure, la partie qui correspond à des valeurs en abscisse inférieures à 3% est celle où le risque est sous-estimé sous le modèle gaussien. Pour des valeurs plus grandes, nous constatons que l’estimation devient plus consistante et que le ratio de dépassement se rapproche du niveau de risque p et même coïncide avec ce dernier pour certaines valeurs de p entre 3% et 7% selon les actifs. En effet, le modèle gaussien repose sur la loi normale pour décrire les rendements des actifs financiers. Dans la section 2.3.2, nous avons bien vérifié que cette loi sous-estime la queue de distribution empirique et que cette dernière est plus épaisse que la queue de distribution normale. Cela explique le fait que, pour des faibles niveaux de risque, le backtesting rejette le modèle gaussien. Cependant, pour des niveaux de risque plus élevés, nous ne sommes plus dans la queue de distribution et nous nous rapprochons plus de sa valeur centrale. Dans ce cas, le modèle gaussien devient adapté à l’estimation du risque.

1.3.2 Test de couverture non-conditionnelle

Dans le but de tester l’hypothèse que les variables aléatoires $(I_t(a))_{t \geq 0}$ suivent une loi de Bernoulli de paramètre $p = 1 - a$, Kupiec propose le test suivant [101].

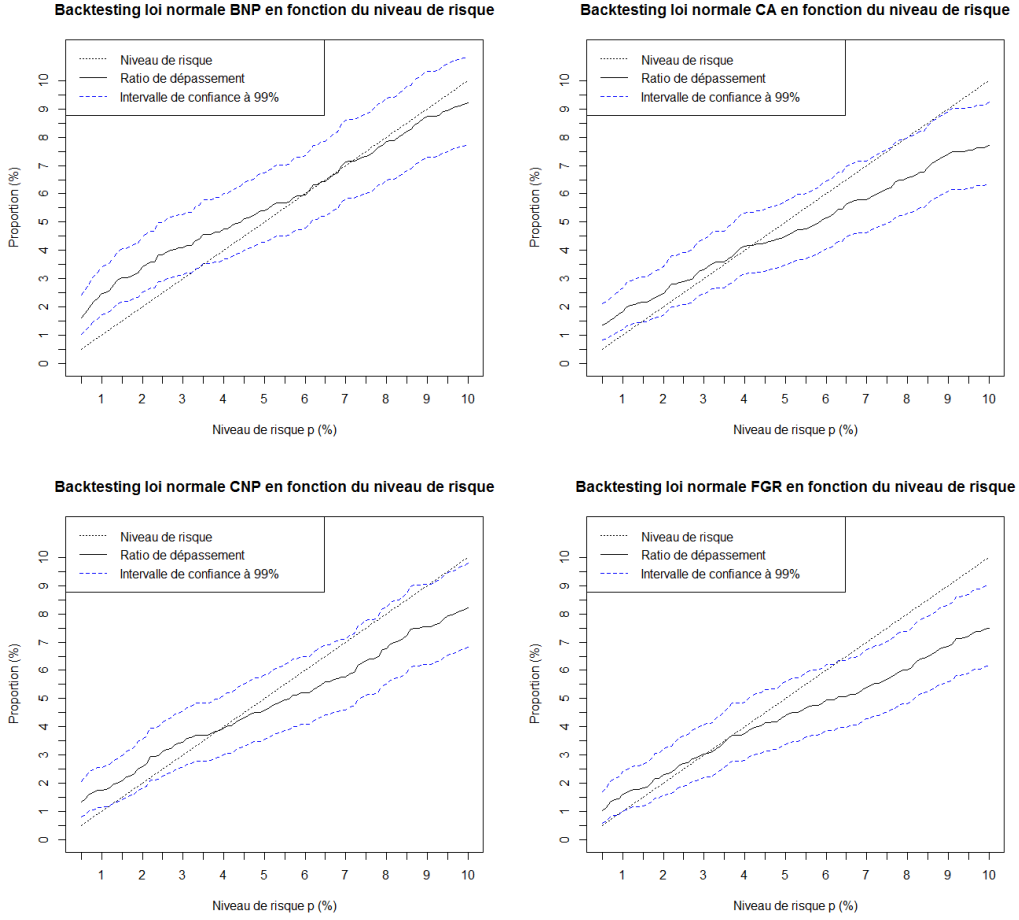


FIGURE 1.7 – Le ratio de dépassement et son intervalle de confiance à 99% en fonction du niveau de risque a de la VaR pour les actifs BNP, CA, CNP et FGR. La VaR est estimée sur les rendements de l'actif entre 2001-2011.

Si le modèle gaussien estime bien la VaR à un niveau a , alors les variables aléatoires $(I_t(a))_{t \geq \omega}$ sont indépendantes et distribuées selon une loi de Bernoulli de paramètre $p = 1 - a$. Le test de couverture non-conditionnelle, proposé par Kupiec [101], suppose l'indépendance des $(I_t(a))_{t \geq \omega}$ et teste l'hypothèse que ces variables suivent une loi de Bernoulli de paramètre p . Il est défini de la manière suivante.

Proposition 1.4 (Test de couverture non-conditionnelle). *Soit un échantillon de n mesures de VaR estimée à un niveau de confiance a . Soit k le nombre de dépassements égal à $\sum_{t=1}^n I_t(a)$ et $\hat{p} = \frac{k}{n}$ le ratio de dépassement correspondant. Nous testons l'hypothèse nulle suivante :*

$$H_0 : \hat{p} = p = 1 - a.$$

Sous l'hypothèse nulle H_0 , la statistique du test

$$LR_{uc} = 2 \log \left(\left(\frac{1 - \hat{p}}{1 - p} \right)^{n-k} \left(\frac{\hat{p}}{p} \right)^k \right) \quad (1.14)$$

suit une loi $\chi^2_{(1)}$ d'un seul degré de liberté.

Preuve. Pour chaque instant $1 \leq t \leq n$, si $I_t(a)$ suit une loi de Bernoulli de paramètre $p = 1 - a$ alors $\mathbb{E}[I_t(a)] = p$. Donc, le niveau de risque p peut être approché par le ratio de dépassement $\hat{p} = \frac{k}{n}$ qui n'est autre que la moyenne empirique de la suite des dépassements $(I_t(a))_{1 \leq t \leq n}$.

La fonction de vraisemblance d'un échantillon de n observations de $I_t(a)$, indépendantes et identiquement distribuées selon une loi de Bernoulli de paramètre $\hat{p}$ et contenant k dépassements ($I_t(a) = 1$), s'écrit :

$$L(\hat{p}) = \prod_{t=1}^n (1 - \hat{p})^{1-I_t} \hat{p}^{I_t} = (1 - \hat{p})^{n-k} \hat{p}^k. \quad (1.15)$$

Sous l'hypothèse de la couverture non-conditionnelle $H_0 : \hat{p} = p$, la fonction de vraisemblance devient

$$L(p) = (1 - p)^{n-k} p^k. \quad (1.16)$$

Nous savons que sous l'hypothèse nulle H_0 , la statistique du test du rapport de vraisemblance suit une loi $\chi^2_{(1)}$ avec un nombre de degrés de liberté égal au nombre de contraintes imposées par l'hypothèse H_0 . Ainsi, la statistique de notre test

$$LR_{uc} = -2 \log [L(p)/L(\hat{p})], \quad (1.17)$$

suit, asymptotiquement, une loi de χ^2 avec un degré de liberté. Ce qui prouve le résultat annoncé. $\square$

1.3.3 Test d'indépendance

La propriété de couverture non-conditionnelle toute seule n'est pas suffisante pour caractériser un bon modèle d'estimation de la Value-at-Risk. En effet, nous pouvons imaginer le cas où les dépassements de la VaR arrivent tous dans un petit intervalle de temps, tout en ayant un ratio de dépassement correct sur l'échantillon entier. Dans un tel cas, nous aimerions rejeter le modèle de calcul de la VaR. Autrement dit, nous aimerions éviter tout regroupement des dépassements.

Nous établissons alors un test d'indépendance de la manière suivante. Nous supposons que la suite des dépassements $(I_t(a))_{t \geq \omega}$ est dépendante au cours du temps et que nous pouvons modéliser cette dépendance par une chaîne de Markov de matrice de transition

$$\Pi_1 = \begin{bmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{bmatrix} \quad (1.18)$$

avec p_{ij} la probabilité de transition

$$p_{ij} = \mathbb{P}(I_t(a) = j | I_{t-1}(a) = i). \quad (1.19)$$

Notons que $p_{00} = 1 - p_{01}$, $p_{10} = 1 - p_{11}$ et que p_{01} par exemple représente la probabilité d'avoir un non-dépassement suivi par un dépassement et p_{11} la probabilité d'avoir deux dépassements successifs.

Ayant un échantillon de n observations de $I_t(a)$, la fonction de vraisemblance de Π_1 s'écrit :

$$L(\Pi_1) = p_{00}^{n_{00}} p_{01}^{n_{01}} p_{10}^{n_{10}} p_{11}^{n_{11}} = (1 - p_{01})^{n_{00}} p_{01}^{n_{01}} (1 - p_{11})^{n_{10}} p_{11}^{n_{11}} \quad (1.20)$$

avec n_{ij} le nombre d'observations de la valeur i suivie par une observation de la valeur j , $i, j = 0, 1$. Les probabilités p_{01} et p_{11} peuvent être estimées par

$$\hat{p}_{01} = \frac{n_{01}}{n_{00} + n_{01}} = \frac{n_{01}}{n - k}, \quad (1.21)$$

$$\hat{p}_{11} = \frac{n_{11}}{n_{10} + n_{11}} = \frac{n_{11}}{k}, \quad (1.22)$$

et la matrice de transition est estimée par

$$\hat{\Pi}_1 = \begin{bmatrix} 1 - \hat{p}_{01} & \hat{p}_{01} \\ 1 - \hat{p}_{11} & \hat{p}_{11} \end{bmatrix}. \quad (1.23)$$

Dans le cas où $n_{11} = 0$, nous obtenons $\hat{p}_{11} = 0$. La fonction de vraisemblance de $\hat{\Pi}_1$ se réduit alors à

$$L(\hat{\Pi}_1) = (1 - p_{01})^{n_{00}} p_{01}^{n_{01}}. \quad (1.24)$$

Notons que permettre la dépendance dans la suite de dépassements revient à permettre à la probabilité p_{01} d'être différente de p_{11} . Nous souhaitons éviter la dépendance positive qui accorde à la probabilité d'avoir deux dépassements successifs une valeur supérieure à celle d'avoir un dépassement après un non-dépassement ($p_{11} > p_{01}$), ce qui encourage par conséquent le clustering des dépassements. D'autre part, si les dépassements sont indépendants alors la probabilité d'avoir un dépassement à un instant t ne dépend pas de l'état de $I_{(t-1)}(a)$, et donc $p_{01} = p_{11}$. Nous pouvons ainsi approcher leur valeur par le ratio de dépassement $\hat{p} = \frac{k}{n}$.

Sous l'hypothèse d'indépendance, la matrice de transition est estimée par

$$\hat{\Pi} = \begin{bmatrix} 1 - \hat{p} & \hat{p} \\ 1 - \hat{p} & \hat{p} \end{bmatrix} \quad (1.25)$$

et sa fonction de vraisemblance est égale à $L(\hat{p})$ donnée par (1.15).

Finalement, pour tester l'hypothèse d'indépendance $p_{01} = p_{11}$, nous construisons le test de rapport de vraisemblance.

Proposition 1.5 (Test d'indépendance). *Nous testons l'hypothèse nulle suivante :*

$$H_0 : p_{01} = p_{11}.$$

Sous l'hypothèse nulle H_0 , la statistique du test

$$LR_{ind} = -2 \log \left[L(\hat{p}) / L(\hat{\Pi}_1) \right] \quad (1.26)$$

suit une loi de $\chi^2_{(1)}$ d'un seul degré de liberté.

1.3.4 Test de couverture conditionnelle

Afin de tester simultanément si les dépassements de la VaR sont indépendants et le ratio de dépassements $\hat{p}$ est égal à $p = 1 - a$, nous construisons le test de couverture conditionnelle qui correspond à tester l'hypothèse $p_{01} = p_{11} = p$.

Proposition 1.6 (Test de couverture conditionnelle). *Nous testons l'hypothèse nulle suivante :*

$$H_0 : p_{01} = p_{11} = p.$$

Sous l'hypothèse nulle H_0 , la statistique du test

$$LR_{CC} = LR_{UC} + LR_{ind} = -2 \log \left[L(p) / L(\hat{\Pi}_1) \right] \quad (1.27)$$

suit une loi de $\chi^2_{(2)}$ de deux degrés de liberté.

Preuve. Ayant un échantillon de n observations de $I_t(a)$ et en estimant la dépendance par la matrice $\hat{\Pi}_1$ donnée par (1.25), la fonction de vraisemblance de $L(\hat{\Pi}_1)$ est donnée par (1.20).

L'hypothèse nulle H_0 assure l'indépendance $p_{01} = p_{11}$ et que leur valeur est égale à p . Ainsi, sous cette hypothèse, la fonction de vraisemblance se réduit à $L(p)$, donnée par (1.16) et la statistique du test du rapport de vraisemblance suit une loi de $\chi^2_{(2)}$ de deux degrés de liberté car il y a deux contraintes imposées par l'hypothèse H_0 . $\square$

1.3.5 Résultats empiriques des différents backtests

Dans cette section, nous appliquons les trois tests d'hypothèse, à savoir l'appartenance du niveau de risque p à l'intervalle de confiance à 99%, la couverture non-conditionnelle et la couverture conditionnelle, sur un échantillon des données réelles issues de la bourse de Paris.

Dans un premier temps, nous considérons les rendements des actifs BNP Paribas, CA, CNP et FGR entre 2001 et 2011. Nous calculons la VaR sur une fenêtre glissante de 252 jours et à un niveau de confiance $a \in [90\%, 99.5\%]$ et nous appliquons les trois backtests sur la VaR mesurée. Nous traçons ensuite les résultats (rejet, non-rejet) en fonction du niveau de risque $p = 1 - a$ dans la figure 1.8. Dans un deuxième temps, nous élargissons l'échantillon d'actifs et nous appliquons les backtests sur la VaR mesurée sur une fenêtre glissante de 252 jours et à des niveaux de risque 1%, 5% et

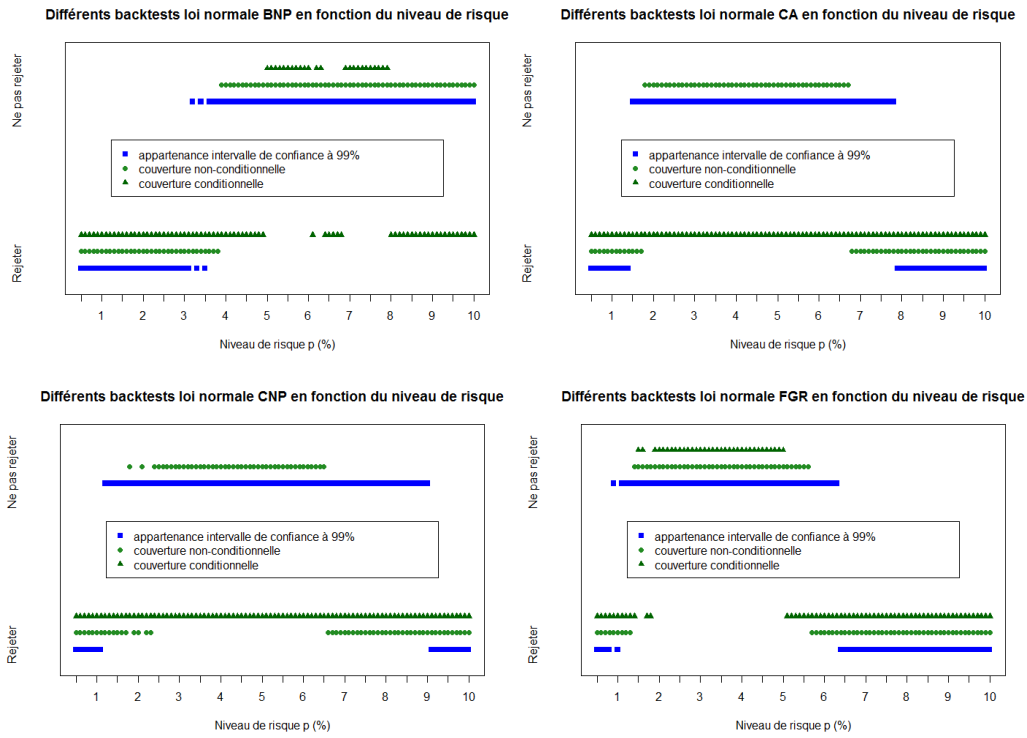


FIGURE 1.8 – Les résultats des backtests sur la VaR en fonction du niveau de risque p . La VaR est mesurée à un niveau de confiance $a = 1 - p = 0.99$ et sur une fenêtre glissante de 252 jours sur les rendements des actifs BNP, CA, CNP et FGR.

10% pour 12 actifs sélectionnés. La table 1.2 donne les résultats de ces backtests en termes de rejet et non-rejet.

La figure 1.8 montre que pour des niveaux de risque inférieurs à 3%, qui nécessitent une bonne modélisation de la queue de distribution, le modèle gaussien échoue et les trois backtests le rejettent. Pour des niveaux plus élevés, ce modèle peut être acceptable et permet d'obtenir des descriptions consistantes sur le risque du marché financiers. Cela peut être remarqué aussi dans la table 1.2 pour un niveau de risque 5%. D'autre part, à partir de 8% de risque, dans la figure 1.8 et pour un niveau de 10% de risque dans la table 1.2, l'hypothèse de couverture conditionnelle est systématiquement rejetée. L'explication que nous donnons est la suivante : pour des niveaux de risque élevés, 10% par exemple, nous ne sommes plus dans la queue de distribution et la VaR représente le quantile à 90% de la distribution de perte. Ce quantile va être dépassé plus souvent que dans le cas des faibles niveaux de risque, ce qui crée forcément des regroupements dans les dépassements de la VaR estimée par les pertes réelles. D'où, le rejet de l'hypothèse d'indépendance, et par conséquent l'hypothèse de la couverture conditionnelle.

Stock	$\alpha = 1\%$			$\alpha = 5\%$			$\alpha = 10\%$		
	CI	UC	CC	CI	UC	CC	CI	UC	CC
BNP	□	□	□	★	★	★	★	★	□
CA	□	□	□	★	★	□	□	□	□
CNP	□	□	□	★	★	□	□	□	□
FGR	□	□	□	★	★	★	□	□	□
GA	□	□	□	★	★	□	□	□	□
GLE	□	□	□	★	★	□	★	★	□
LG	□	□	□	★	★	□	★	★	□
LI	□	□	□	★	★	□	★	★	□
MC	★	□	★	★	★	★	□	□	□
RI	★	□	□	□	□	□	□	□	□
SAN	□	□	□	★	★	□	□	□	□
TEC	□	□	□	★	★	□	□	□	□
Total	2	0	1	11	11	3	4	4	0

★Ne pas rejeter, □Rejeter

TABLE 1.2 – Les résultats des backtests sur la VaR en fonction du niveau de risque p . La VaR est mesurée à des niveaux de risque 1%, 5% et 10% et sur une fenêtre glissante de 252 jours pour un échantillon de 12 actifs.

1.4 Conclusion

Dans ce chapitre, nous nous sommes servis du modèle gaussien pour expliquer la méthodologie d'estimation de la Value-at-Risk à partir des données réelles, ainsi que celle des backtests de validation qui mesurent la performance du modèle choisi dans l'estimation de risque. Dans le chapitre suivant nous appliquons cette méthodologie et nous estimons la VaR avec un modèle qui se focalise sur la queue de distribution des pertes. Ce modèle, appelé modèle des lois puissances, nous permet de considérer des queues plus épaisses.

Chapitre 2

Modèle des lois puissances

Ce chapitre et le chapitre 3 ont fait l’objet d’un rapport de recherche [33].

Le chapitre précédent montre que la modélisation par la loi normale permet de décrire le comportement de toute la distribution, mais manque de précision sur les évaluations des événements rares. De plus, il est largement connu que les prix et les rendements des actifs boursiers admettent certaines propriétés statistiques communes, appelés « faits stylisés » [42]. L’asymétrie et la queue épaisse de la distribution des rendements sont les deux propriétés principales des prix observés sur le marché, et qui ne sont pas vérifiées par le modèle gaussien.

Bien que *les lois stables* ont été proposées depuis les années 60 [62, 114], et une autre alternative au modèle gaussien, *le modèle log-normal subordonné à variance finie*, a été proposé en 1973 par P. Clark [38], l’utilisation systématique de la théorie des valeurs extrêmes (TVE) est tout de même récente. Cet intérêt est dû au krach de 1987, qui a mis en évidence l’importance d’une meilleure compréhension des grandes pertes. Pour les premières utilisations de la TVE pour la modélisation de la queue de la distribution, nous citons [22, 90, 109, 110]. Ensuite, plusieurs auteurs se sont intéressés à l’étude des valeurs extrêmes et des lois puissances dans les marchés développés, principalement par le biais des indices boursiers : nous citons pour les indices américains S&P 500, Dow Jones et NASDAQ [113], pour l’indice allemand DAX [55], pour l’indice australien ASX-All Ordinaries [152] et pour Nikkei 225 (Japon) et Euro Stoxx 50 (zone Euro) [76].

Le modèle des lois stables, qui se calibre sur toute la distribution, manifeste un comportement de type lois puissances sur la queue. Cependant, le choix de ce modèle impose un indice de queue $\alpha < 2$, impliquant que la variance des rendements est infinie. Ce paramètre est difficile à estimer, surtout lorsqu’il est proche de 2 [57]. De plus, de nombreuses études critiques (voir par exemple [57, 165]) montrent que l’indice de queue est supérieur à 2, et il est autour de 3 pour les rendements à courtes périodes. Nous confirmons, plus tard, ces observations dans l’étude comparative que nous allons développer dans le chapitre 4.

Pour cela, nous nous intéressons, dans ce chapitre, au *modèle des lois puissances*, qui se focalise sur la queue de la distribution. Il s'agit des lois de la forme

$$1 - F(x) = \mathbb{P}(L_t \geq x) = \ell_F(x) x^{-\alpha}, \quad (2.1)$$

où

- la fonction $F(x) = \mathbb{P}(L_t \leq x)$ est la fonction de répartition de la distribution des pertes des actifs financiers ;
- la fonction $\ell_F(x)$ est une fonction à variation lente ;
- le paramètre $\alpha > 0$, appelé l'indice de queue, est l'indice caractéristique de la décroissance de la fonction $1 - F(\cdot)$.

Nous commençons par donner la définition de la fonction à variation lente, ainsi que quelques unes de ses propriétés, dans la section 2.1. La section 2.2 est consacrée à l'estimation de l'indice de queue α . L'estimation de la fonction à variation lente fait l'objet de la section 2.3. Notons que l'estimation de α est faite sur un échantillon, de taille r_n , des plus grandes pertes. Le choix de ce paramètre r_n est discuté dans la section 2.4. Finalement, les sections 2.5 et 2.6 présentent l'étude empirique sur les données réelles : l'estimation des paramètres, le calcul de la VaR et les backtests de validation.

2.1 Les fonctions à variation lente

Nous introduisons dans cette section la notion de la fonction à variation lente, ainsi que quelques unes de ses propriétés.

Définition 2.1. *Une fonction mesurable ℓ de $\mathbb{R}_+$ dans $\mathbb{R}_+$ est dite à variation lente, au voisinage de l'infini, si pour tout $t > 0$,*

$$\lim_{x \rightarrow +\infty} \frac{\ell(xt)}{\ell(x)} = 1.$$

On note $\mathcal{R}_0$ l'ensemble des fonctions à variation lente.

Des exemples simples de fonctions à variation lente peuvent être les fonctions constantes ou celles qui convergent vers une constante.

Théorème 2.1.

(i) *Théorème de convergence uniforme. Si $\ell \in \mathcal{R}_0$, alors la convergence*

$$\lim_{x \rightarrow +\infty} \frac{\ell(xt)}{\ell(x)} = 1$$

est uniforme pour $t \in [a, b]$ avec $0 < a < b < \infty$.

(ii) **Théorème de représentation.** La fonction $\ell \in \mathcal{R}_0$ si et seulement si elle peut être représentée sous la forme

$$\ell(x) = c(x) \exp \left\{ \int_1^x \frac{\epsilon(u)}{u} du \right\}$$

avec $c(x) \rightarrow c \in (0, \infty)$ et $\epsilon(x) \rightarrow 0$ lorsque $x \rightarrow \infty$.

Des propriétés qui découlent facilement du théorème de représentation sont les suivantes.

Proposition 2.1. Soit ℓ une fonction à variation lente. Alors

(i) Nous avons,

$$\lim_{x \rightarrow +\infty} \frac{\log(\ell(x))}{\log(x)} = 0. \quad (2.2)$$

(ii) Pour tout $\delta > 0$, il existe un réel x_δ tel que pour toute constante $A > 0$ et $x > x_\delta$

$$Ax^{-\delta} < \ell(x) < Ax^\delta.$$

Proposition 2.2 ([17], p.79.). Soit ℓ une fonction à variation lente. Il existe alors une fonction à variation lente, ℓ^* , appelée le conjugué de Bruyn de ℓ , tel que

$$\ell(x)\ell^*(x\ell(x)) \xrightarrow{x \rightarrow +\infty} 1.$$

Nous pouvons par exemple vérifier que ℓ^* est à variation lente en utilisant la représentation de Karamata [134]. Le conjugué de Bruyn est asymptotiquement unique dans le sens où si $\tilde{\ell}$ est aussi à variation lente tel que $\ell(x)\tilde{\ell}(x\ell(x)) \rightarrow 1$, alors $\ell^* \sim \tilde{\ell}$ au voisinage de l'infini. Nous notons de plus que $(\ell^*)^* \sim \ell$.

Dans les parties qui suivent, nous nous intéressons à l'estimation de l'indice de queue α par l'estimateur de Hill [88]. Dans un premier temps, nous étudions son comportement asymptotique. Ensuite, nous proposons une approximation pour la fonction à variation lente $\ell(x)$. Enfin, nous définissons l'indicateur de risque VaR et nous l'évaluons à l'aide de cet estimateur.

2.2 Estimateur de Hill

Dans cette partie, nous cherchons à donner une estimation de l'indice de queue de la distribution des pertes sous le modèle des lois puissances

$$1 - F(x) = \mathbb{P}(L \geq x) = \ell_F(x)x^{-\alpha}.$$

Proposition 2.3. La fonction quantile vérifie

$$q^L(1-p) = p^{-\gamma} \ell_Q\left(\frac{1}{p}\right), \quad \text{pour } 0 < p < 1, \quad (2.3)$$

où $\gamma = \frac{1}{\alpha}$ et ℓ_Q est encore une fonction à variation lente définie sur $]1, \infty[$ et reliée à ℓ_F via la conjugaison de Bruyn.

Preuve. Posons $p = 1 - F(x) = (x\ell(x))^{-\alpha}$ où $\ell(x) = (\ell_F(x))^{-\frac{1}{\alpha}}$ est encore une fonction à variation lente. Posons $\gamma = \frac{1}{\alpha}$. En appliquant la proposition 2.2, nous pouvons résoudre l'équation $p^{-\gamma} = x\ell(x)$ en x . Pour p proche de 0, nous obtenons $x = p^{-\gamma}\ell^*(p^{-\gamma})$ avec ℓ^* le conjugué de Bruyn de ℓ . Comme $q^L(1-p) \sim x$, il suffit de définir la fonction $\ell_Q(y) = \ell^*(y^\gamma)$ pour $y \in]1, \infty[$ pour finir la preuve et avoir (2.3). $\square$

Nous appliquons la fonction $\log$ à (2.3)

$$\log(q^L(1-p)) = -\gamma \log(p) + \log\left(\ell_Q\left(\frac{1}{p}\right)\right), \quad \text{pour } 0 < p < 1. \quad (2.4)$$

Comme pour toute fonction à variation lente ℓ_Q , nous avons d'après la proposition 2.1,

$$\frac{\log\left(\ell_Q\left(\frac{1}{p}\right)\right)}{\log(p)} \xrightarrow[p \rightarrow 0]{} 0.$$

Nous obtenons alors

$$\frac{\log(q^L(1-p))}{-\log(p)} \xrightarrow[p \rightarrow 0]{} \gamma,$$

ce qui explique la linéarité asymptotique du graphe de Pareto dans le cas des lois puissances.

Nous considérons, par la suite, les données suivantes :

- La série des pertes journalières observées sur le marché est $(L_i)_{1 \leq i \leq n}$.
- La statistique d'ordre de cette série est $(L_{(i)})_{1 \leq i \leq n}$ telle que $L_{(i)} \leq L_{(i+1)}$ pour tout $1 \leq i \leq n-1$.

Proposition 2.4. *Pour toute suite i_n telle que $i_n/n \rightarrow a$, $L_{(i_n)}$ est un estimateur consistant pour $q^L(a)$ lorsque $n \rightarrow \infty$.*

En particulier, pour i et n fixés tel que $i \leq n$, nous pouvons voir $L_{(i)}$ comme une approximation de $q^L(\frac{i}{n+1})$.

Cette observation permet de vérifier graphiquement, à partir de (2.4), la linéarité de la fonction quantile. Nous considérons le graphe de Pareto (Pareto Quantile Plot) obtenu en traçant $(L_{(i)})_{1 \leq i \leq n}$, nous traçons $\log L_{(i)}$ en ordonnée contre $-\log(1 - i/(n+1))$ en abscisse. Si les données suivent une loi puissance, un tel graphe doit se comporter de façon asymptotiquement linéaire.

Le graphe de Pareto obtenu à partir des log-pertes journalières de l'action BNP Paribas sur la période 2001-2011 est donné par la figure 2.1. Nous observons que la fonction quantile a l'allure d'une droite sur la queue de distribution. Déterminer un estimateur de γ (et donc de α) revient à déterminer la pente de cette partie presque linéaire, correspondant aux valeurs extrêmes.

Parmi les estimateurs les plus classiques de cette pente, nous trouvons celui de Hill introduit en 1975 [88], qui reste l'un des plus utilisés sur les marchés financiers vu son efficacité et sa simplicité.

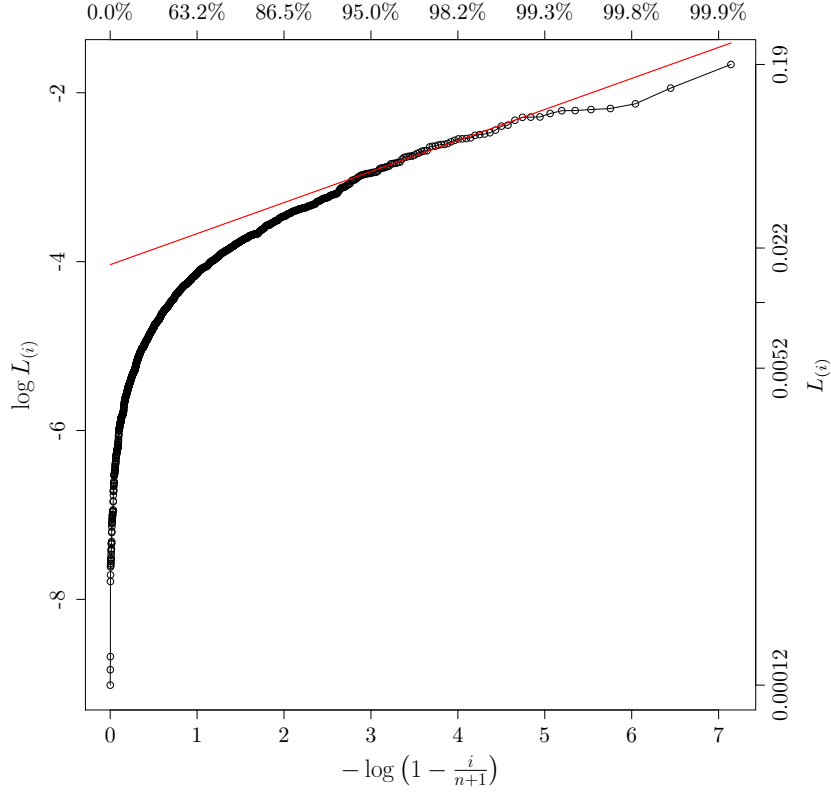


FIGURE 2.1 – Graphe de Pareto des log-pertes de l’actif BNP Paribas sur la période Jan. 2001-Fév. 2011.

Définition 2.2. Nous définissons l’estimateur de Hill par

$$\hat{\alpha}_{r_n} = \left[\frac{1}{r_n} \sum_{i=1}^{r_n} \log(L_{(n-i+1)}) - \log(L_{(n-r_n)}) \right]^{-1}, \quad (2.5)$$

pour un $0 < r_n < n$ fixé.

Afin d’alléger les notations, lorsqu’il n’y a pas d’ambiguïté sur le choix de r_n , nous utiliserons la notation $\hat{\alpha}$. La consistance de cet estimateur découle du résultat suivant.

Proposition 2.5 ([88]). Soit $(r_n)_{n \geq 1}$ une suite d’entiers positifs tels que $r_n \rightarrow +\infty$ et $\frac{r_n}{n} \rightarrow 0$ lorsque $n \rightarrow +\infty$. Nous avons la convergence en probabilité

$$\left(r_n^{-1} \sum_{i=1}^{r_n} \log(L_{(n-i+1)}) - \log(L_{(n-r_n)}) \right) \xrightarrow[n \rightarrow \infty]{\mathbb{P}} \gamma = \frac{1}{\alpha}. \quad (2.6)$$

Remarque 2.1. Nous constatons qu’une grande latitude est laissée pour le choix du paramètre r_n dans la proposition 2.5. En pratique, le choix de ce paramètre doit être un compromis entre une petite valeur, qui donnerait un nombre insuffisant de

données pour estimer α , et une valeur trop grande, qui inclurait dans l'estimateur des valeurs qui ne sont pas dans la queue de distribution. La littérature sur ce sujet, appelé le *tail index estimation problem*, sera discutée dans la section 2.4 de ce chapitre.

Afin de construire des intervalles de confiance sur l'estimation de α , par l'estimateur de Hill, le théorème de la limite centrale suivant est utilisé.

Théorème 2.2. *On considère $r_n \in \mathbb{N}$ tel que $r_n \rightarrow \infty$ et $r_n = o(n)$ lorsque $n \rightarrow \infty$. Nous posons :*

$$I(n, r_n, x) = \left(\frac{r_n}{n}\right) \left(1 + \frac{x}{r_n^{1/2}}\right),$$

et

$$J(F, n, r_n, x) = r_n^{1/2} \int_1^\infty \left(y^{-\alpha} - \frac{1 - F(F^{-1}(1 - I(n, r_n, x))y)}{I(n, r_n, x)} \right) \frac{dy}{y}.$$

Si $J(F, n, r_n, x) \xrightarrow[n \rightarrow \infty]{} A \in \mathbb{R}$ uniformément sur tout ensemble de compact, alors

$$r_n^{1/2}(\hat{\alpha}^{-1} - \alpha^{-1}) \xrightarrow[n \rightarrow \infty]{\mathcal{L}} \mathcal{N}(-A, \alpha^{-2}),$$

où $\hat{\alpha}$ est l'estimateur de Hill du paramètre α et $\mathcal{N}(m, \sigma^2)$ désigne la loi normale de moyenne m et de variance σ^2 .

La preuve de ce théorème se trouve dans [82].

Corollaire 2.1. *Un intervalle de confiance de l'estimateur de Hill à un niveau de confiance $(1 - \beta)$ est donné par :*

$$\alpha^{-1} \in \left[\hat{\alpha}^{-1} + \left(-A - \frac{\Phi(1 - \beta/2)}{\hat{\alpha}} \right) \frac{1}{r_n^{1/2}}, \hat{\alpha}^{-1} + \left(-A + \frac{\Phi(1 - \beta/2)}{\hat{\alpha}} \right) \frac{1}{r_n^{1/2}} \right],$$

où Φ est la fonction quantile de la loi normale centrée réduite.

Remarque 2.2. Lorsque la fonction à variation lente $\ell(x)$ est égale à une constante $C > 0$ pour x suffisamment grand (et également lorsque $\ell(x)$ converge vers $C > 0$ suffisamment vite), la constante A dans le théorème 2.2 est nulle [17]. Nous supposons implicitement dans la suite que les actifs du marché financier satisfont ces hypothèses.

2.3 Estimation de la fonction à variation lente $\ell(x)$

Nous nous placerons dans cette section sous l'hypothèse que la fonction à variation lente $\ell(x)$ converge vers une constante $C > 0$ lorsque $x \rightarrow +\infty$. L'estimateur de la constante C et ceux des quantiles ont été proposés initialement par Weissman [164].

Rappelons la relation (1.2), qui relie la fonction quantile à la fonction à variation lente : $1 - a = q^L(a)^{-\alpha} \ell(q^L(a))$.

En fixant a_0 proche de 1, le seuil tel que $\ell(q^L(a_0))$ est correctement approchée par sa limite C , nous choisissons un entier $0 \leq w_n < n$ tel que $(n - w_n)/(n + 1) \approx a_0$. Par la suite, en utilisant l'approximation de $q^L(a_0)$ par $L_{(n-w_n)}$ (voir la proposition 2.4) et celle de α par son estimateur de Hill, nous obtenons que la constante C est approchée par

$$\hat{C}_{w_n} \approx L_{(n-w_n)}^{\hat{\alpha}_{r_n}}(1 - a_0). \quad (2.7)$$

Par conséquent, pour tout $a \geq a_0$ le quantile de niveau a de la distribution des pertes est estimé par :

$$\hat{q}_{w_n, r_n}^L(a) = \left(\frac{\hat{C}_{w_n}}{1 - a} \right)^{\hat{\alpha}_{r_n}^{-1}} = L_{(n-w_n)} \left(\frac{1 - a_0}{1 - a} \right)^{\hat{\alpha}_{r_n}^{-1}}, \quad (2.8)$$

où $\hat{\alpha}_{r_n}$ désigne l'estimateur de Hill obtenu avec le choix r_n .

Afin de vérifier la consistance de cet estimateur du quantile et de construire un intervalle de confiance, nous allons utiliser l'écriture (2.4) de la fonction quantile. Remarquons que $q^L(1 - a) = a^{-\gamma} \ell^*(1/a)$. Ainsi, par utilisation du théorème des fonctions de répartition inverses : $L_{(n-w_n)}$ a la même loi que $q^L(1 - U_{(w_n+1)}) = U_{(w_n+1)}^{-\gamma} \ell^*(U_{(w_n+1)}^{-1})$, où $U_{(1)}, \dots, U_{(n)}$ sont les statistiques d'ordre d'une famille i.i.d de n v.a. uniformes sur $[0, 1]$. En utilisant ces deux relations,

$$\begin{aligned} \log \frac{\hat{q}_{w_n, r_n}^L(a)}{q^L(a)} &\stackrel{\mathcal{L}}{=} \gamma \left(-\log U_{(w_n+1)} - \log \frac{n+1}{w_n+1} \right) \\ &\quad + (\hat{\alpha}_{r_n}^{-1} - \gamma) \log \left(\frac{w_n+1}{(n+1)(1-a)} \right) + \log \frac{\ell^*(U_{(w_n+1)}^{-1})}{\ell^*(1/(1-a))}. \end{aligned} \quad (2.9)$$

Il est connu [17] que

$$\sqrt{w_n} \left(-\log U_{(w_n+1)} - \log \frac{n+1}{w_n+1} \right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1)$$

lorsque $n \rightarrow +\infty$, $w_n \rightarrow +\infty$ et $w_n/n \rightarrow 0$. Le théorème 2.2 donne la loi limite de $\sqrt{r_n}(\hat{\alpha}_{r_n}^{-1} - \gamma)$. Le seul terme dont nous ne pouvons pas estimer le comportement asymptotique en toute généralité est le dernier. Cependant, il est nul si $\ell^*(x) = C > 0$ pour x assez grand (ou si $\ell^*(x)$ converge vers $C > 0$ suffisamment rapidement).

Sous cette hypothèse, il est alors possible de construire un intervalle de confiance pour $q^L(a)$. Par exemple, si nous choisissons $w_n = r_n$, en nous souvenant que le biais asymptotique de l'estimateur de Hill (la constante A du théorème 2.2) est nul sous notre dernière hypothèse, nous obtenons

$$\sqrt{r_n} \left(1 + \log^2 \frac{r_n+1}{(n+1)(1-a)} \right)^{-1/2} \left(\frac{\hat{q}_{r_n, r_n}^L(a)}{q^L(a)} - 1 \right) \xrightarrow{\mathcal{L}} \mathcal{N}(0, \gamma^2),$$

quand $n \rightarrow +\infty$, $r_n \rightarrow +\infty$ et $r_n/n \rightarrow 0$. Nous en déduisons l'intervalle de confiance pour $q^L(a)$ de niveau asymptotique $1 - \beta$, qui est donné par

$$\left[\frac{\hat{q}_{r_n, r_n}^L(a)}{1 + \Phi^{-1}(1 - \beta/2) \frac{\hat{\alpha}_{r_n}^{-1}}{\sqrt{k}} \sqrt{1 + \log^2 \frac{r_n+1}{(n+1)(1-a)}}}, \frac{\hat{q}_{r_n, r_n}^L(a)}{1 - \Phi^{-1}(1 - \beta/2) \frac{\hat{\alpha}_{r_n}^{-1}}{\sqrt{k}} \sqrt{1 + \log^2 \frac{r_n+1}{(n+1)(1-a)}}} \right].$$

2.4 Choix du seuil d'estimation r_n

Afin d'appliquer les estimateurs précédents aux historiques du marché, il faut préciser comment choisir le paramètre r_n qui donne le nombre de données utilisées dans la statistique d'ordre pour estimer α et la fonction à variation lente. Il existe une vaste littérature sur ce sujet appelé le *tail index estimation problem*. Pour des exemples de quelques ouvrages et articles de synthèses, voir [3, 17, 49, 146].

Le choix d'une proportion fixe (par exemple un échantillon de 10% [57], de $r_n = \sqrt{n}$, ou d'autres choix [110]) est rarement satisfaisant en pratique (voir [146, §. 3] pour une discussion).

L'une des approches les plus simples consiste à tracer la courbe $(k, \hat{\alpha}_k)$, qui s'appelle un graphe de Hill (*Hill plot*, voir par exemple la figure 2.2). Cependant, les courbes sont rarement constantes, et cette approche est difficile à mettre en œuvre (voir [17, §. 4.2.2, p. 104] ou [56]). Un autre type de graphe, appelé *alternative Hill plot*, est de tracer $(\theta, \hat{\alpha}_{[n^\theta]})$, $0 \leq \theta \leq 1$ [133], et de construire un estimateur de α à partir des moyennes $\tilde{\alpha} = (u-1)^{-1}k^{-1} \sum_{i=k+1}^{uk} \hat{\alpha}_i$ pour $u > 1$.

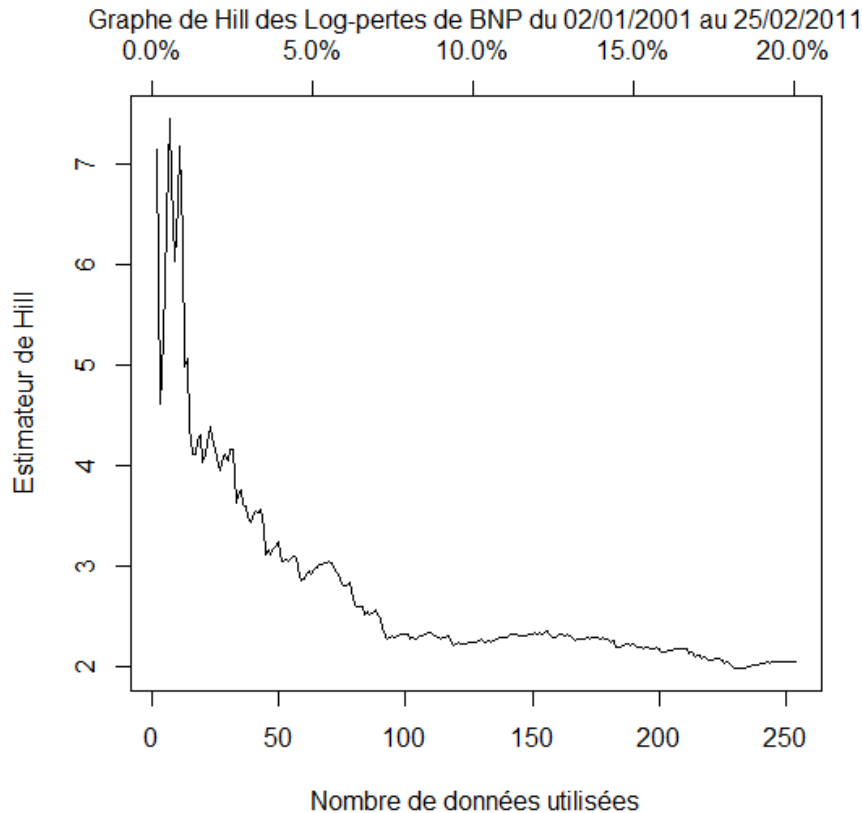


FIGURE 2.2 – Graphe de Hill pour les pertes de l'actif BNP Paribas.

Un autre diagnostic graphique est le *sum plot*, où le graphe $(k, \sum_{i=1}^k Y_i)$, avec Y_i définie par $Y_i = i(\log L_{(n-i+1)} - \log L_{(n-i)})$, est linéaire pour les petites valeurs de k . Une limite de r_n est donnée par l'indice où un tel graphe dévie d'une croissance

linéaire [50].

L'article [10] propose une estimation en utilisant comme estimateur celui qui est le plus proche d'estimateurs de α , construits par simulation Monte Carlo, en utilisant des variables aléatoires de Pareto pour de multiples valeurs de α .

Une autre série d'approches est d'utiliser un test d'adéquation (*goodness-of-fit*) sur les statistiques $(L_{(k)})_{k \geq r_n}$, après d'éventuelles transformations du lemme 1. Nous pouvons par exemple utiliser le fait que les variables $k \ln(L_{(n-k)}/L_{(n-k+1)})$ suivent des lois exponentielles indépendantes de paramètres α . Cette approche a été suggérée par B. Hill [88] (voir aussi [17, §. 4.2.1, p. 103] et [84]). Mais P. Hall et A. Welsh ont montré dans [84], que cette procédure tend à surestimer r_n pour les grands ensembles si la fonction à variation lente ℓ dans 2.1 n'est pas constante. S. Alfarano et T. Lux notent dans [3] qu'utiliser des tests d'adéquations sur les lois exponentielles (Kolmogorov-Smirnov, Anderson-Darling, χ^2) ne donne pas toujours des résultats probants. Dans [35], C.W. Cheong propose d'utiliser les statistiques de Cramer-von Mises et de Waston, et applique cela à des données financières de la bourse malaisienne.

Dans [77] et [18], il est montré que l'estimateur de Hill appartient à une classe plus large d'estimateurs, où le coefficient $\hat{\gamma} = 1/\hat{\alpha}$ est celui qui minimise la fonction

$$x \mapsto \sum_{j=1}^k w_{j,k} \left(\ln \left(\frac{L_{n-j+1}}{L_{n-k}} \right) - x \ln \left(\frac{k+1}{j} \right) \right)^2, \quad w_{j,k} = \frac{1}{(k+1)/j}.$$

Le choix optimal de k est cependant discuté lorsque la fonction à variation lente ℓ dans (2.1) n'est pas constante.

Enfin, une très large classe d'estimateurs de r_n repose sur un équilibre entre l'erreur Monte Carlo et le biais induit par l'estimateur (voir les références dans l'ouvrage [17] par exemple). Mais cette approche repose sur l'hypothèse que la fonction à variation lente n'est pas constante. Le modèle le plus traité est celui introduit par P. Hall dans [83] où la fonction $\ell(x)$ de (2.1) admet le développement proportionnel à $1 + Cx^{-\beta} + o(x^{-\beta})$ avec $C \neq 0$. Ces résultats montrent en particulier que le r_n optimal dépend fortement de la vitesse à laquelle la fonction $\ell(x)$ converge vers sa limite. Il est donc illusoire d'espérer estimer précisément le r_n optimal sans chercher à estimer cette vitesse de convergence. Étant donné le faible nombre de données dont nous disposons, il est peu probable que l'on puisse obtenir avec précision de telles informations.

Toutes les pistes mentionnées dans cette partie devront être explorées dans le futur, afin d'analyser la meilleure approche en termes de performance. Un point important vient du faible nombre de données disponibles pour faire nos estimations. Il est donc inutile de chercher une méthode très fine, puisque les erreurs statistiques seront dans tous les cas importantes. Du point de vue du problème soulevé par ALPHABILITY, il s'agit de trouver le bon compromis entre précision et coût numérique. C'est pourquoi nous nous concentrons maintenant sur une méthode très simple inspirée de la discussion ci-dessous.

2.5 Etude empirique : estimation des paramètres de la loi puissance, calcul de la VaR

L'étude numérique utilise des données du marché fournies par ALPHABILITY. Il est à mentionner que ces données, bien que couvrant dix années, restent peu nombreuses puisque l'on n'étudie que les données journalières (prix de clôture), d'où un faible nombre de données disponibles pour les estimations. Il est donc inutile de chercher une méthode très fine, puisque les erreurs statistiques seront dans tous les cas importantes. Afin de réaliser un bon compromis entre précision et coût numérique, nous allons opter pour une méthode d'estimation simple inspirée de la discussion ci-dessus.

Nous considérons ici l'estimation des paramètres de la loi puissance sur le cours de l'actif BNP Parisbas, en prenant des fenêtres de 252 jours.

L'estimateur de l'indice $\gamma = 1/\alpha$, que nous utilisons dans notre étude, est donné par le calcul de la pente sur le graphe de Pareto (figure 2.1) avec une technique de moindres carrés sur un intervalle, qui se concentre sur la queue de distribution mais qui ne prend pas en compte les valeurs les plus grandes des pertes, qui sont les plus entachées d'erreurs.

Ainsi, nous considérons que pour $i \in [i_0, i_1]$,

$$\log L_{(i)} = -\gamma \log \left(\frac{i}{n+1} \right) + K + \epsilon_i$$

où ϵ_i est le terme d'erreur, K une constante,

$$i_0 = \lfloor p_0 n \rfloor \text{ et } i_1 = \lfloor p_1 n \rfloor \text{ avec } p_0 = 0.95 \text{ et } p_1 = 0.99.$$

Comme nous l'avons déjà remarqué, $L_{(i_j)} \approx q^L(i_j/n) \approx q^L(p_j)$ pour $j = 0, 1$. Ainsi, γ et K sont estimés comme la pente $\hat{\gamma}^{LS}$ et la constante $\hat{K}^{LS}$ qui minimisent $\sum_{i=i_0}^{i_1} \epsilon_i^2$. Nous posons, par la suite, $\hat{\alpha}^{LS} = 1/\hat{\gamma}^{LS}$. Ce choix ne correspond pas à l'estimateur de Hill [88]. Cependant, ce dernier représente une version pondérée de l'estimateur moindre carrés, et nous utilisons simplement une version non pondérée. Cet estimateur est mieux adapté à notre cas vu le faible nombre de données.

En effet, sur une fenêtre de 252 jours, seulement environ la moitié des données correspond à des pertes, et donc entre 5 et 10 données seulement sont utilisées pour estimer la queue de distribution. Ainsi, notre choix de ne pas considérer les quantiles supérieurs à 99% s'avère pertinent pour éviter que les valeurs les plus grandes perturbent trop l'estimation de γ .

L'estimation de la constante limite de la fonction à variation lente est donnée par l'estimateur de Weissman, en prenant $\hat{C} = L_{(\lfloor na_0 \rfloor)}^{\hat{\alpha}^{LS}}(1 - a_0)$. Dans ce cas, $L_{(\lfloor na_0 \rfloor)} \approx q^L(a_0)$ et donc $F(x) \approx 1 - \hat{C}/x^{\hat{\alpha}^{LS}}$ pour $x \geq q^L(a_0)$. Nous prenons $a_0 = 90\%$.

Finalement, comme montré dans la section 2.2 du chapitre précédent, la Value-at-Risk, à un jour et à un niveau de confiance a , estimée à partir de la loi puissance est donnée par la formule (1.7) :

$$\widehat{\text{VaR}}_t^{(a)} = \left(\frac{\hat{C}}{1-a} \right)^{(\hat{\alpha}^{LS})^{-1}}.$$

La Value-at-Risk monétaire associée est donnée par :

$$\widehat{\text{VaR}}_t^{(a)}(\text{€}) = S_t \left(1 - \exp \left(-\frac{\hat{C}}{1-a} \right)^{(\hat{\alpha}^{LS})^{-1}} \right).$$

La figure 2.3 représente le prix, la volatilité et la VaR monétaire à un jour et un niveau de confiance de 99% sur le cours de la BNP Paribas. Une corrélation assez nette entre la VaR et la volatilité annuelle peut être observée sur cette figure.

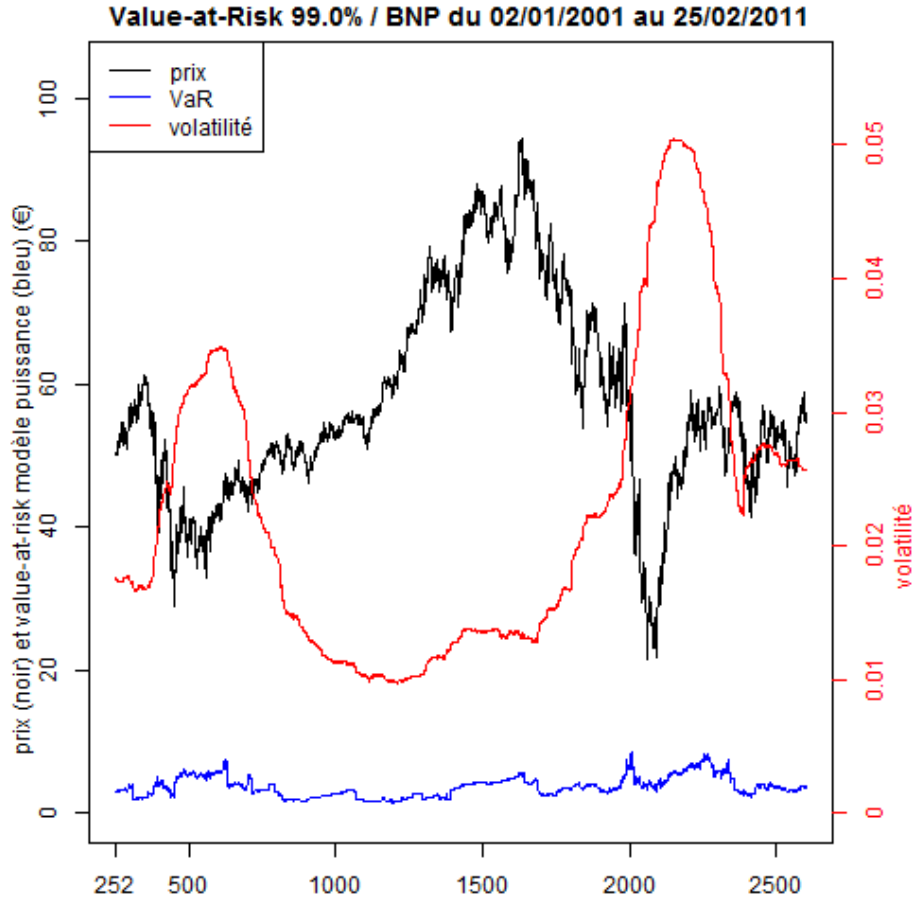


FIGURE 2.3 – VaR monétaire à 99% pour l'action BNP Paribas sur une fenêtre glissante de 252 jours avec les estimateurs $\hat{\alpha}^{LS}$ et $\hat{C}$.

Les paramètres de la loi puissance évalués sur une fenêtre glissante de 252 jours sont représentés dans la figure 2.4. Nous observons une très forte fluctuation des

valeurs estimées de α et C au cours du temps. La méthode est assez instable et imprécise, principalement à cause du faible nombre de données utilisées pour estimer ces paramètres. Nous observons néanmoins un phénomène de compensation entre les grandes valeurs de α et les petites valeurs de C (et inversement), qui conduit au final à une amplitude de fluctuations des valeurs de VaR plus réduite.

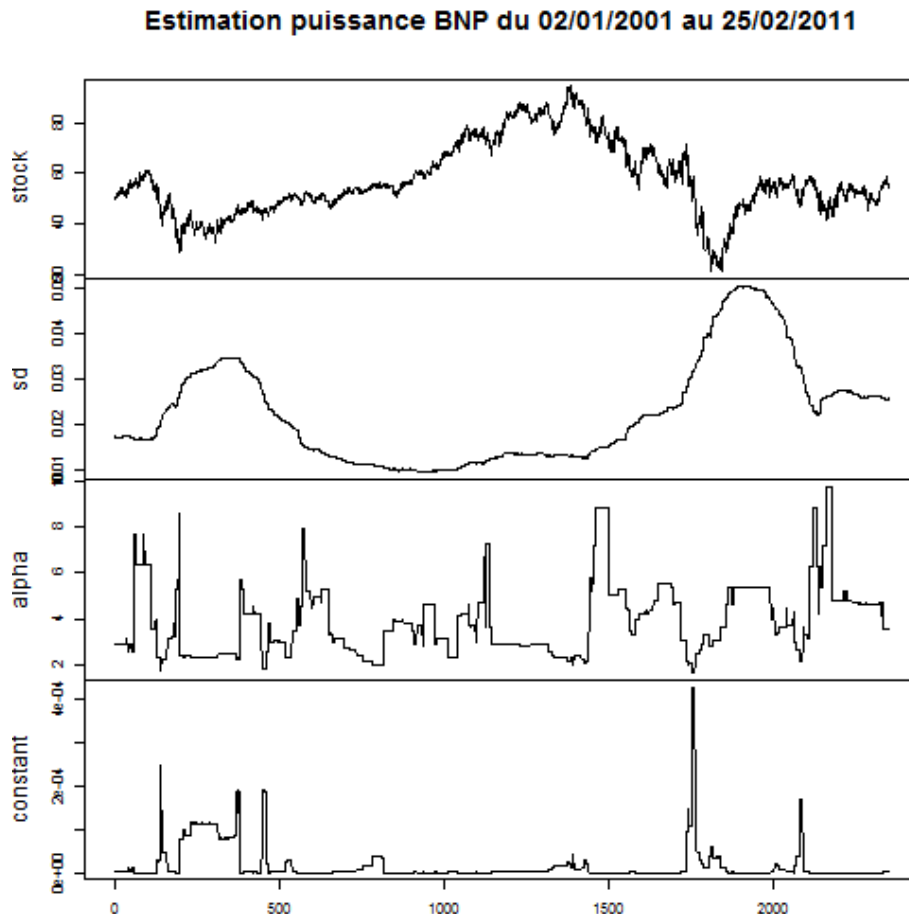


FIGURE 2.4 – Prix, écart-type, $\hat{\alpha}^{LS}$ et $\hat{C}$ pour l'action BNP Paribas pour une fenêtre glissante de 252 jours.

2.6 Backtesting sur la VaR estimée sous un modèle des lois puissances

Dans cette section, nous testons la performance du modèle des lois puissances dans l'estimation de la Value-at-Risk. Nous appliquons les différents backtests vu dans le chapitre précédent, sur la VaR estimée, et nous étudions leur variation en fonction des paramètres d'estimation.

La figure 2.5 montre le backtest du ratio de dépassement sur la VaR calculée à un niveau 99% et sur une fenêtre glissante de 252 jours sur les cours des actions BNP Paribas, CA, CNP et FGR sur la période de 2001 à 2011. La courbe bleue représente

la VaR estimée. Les dépassements par rapport à la perte réelle sont indiqués en vert. Nous constatons que les ratios de dépassement sont moindres que ceux qui sont donnés dans le cadre du modèle gaussien (figure 1.4 de la page 31) pour s'approcher du niveau de risque 1%. L'allure de la courbe de la VaR est semblable à celle du modèle gaussien avec un peu plus de fluctuations, mais surtout avec une prise en compte plus importante des grandes pertes.

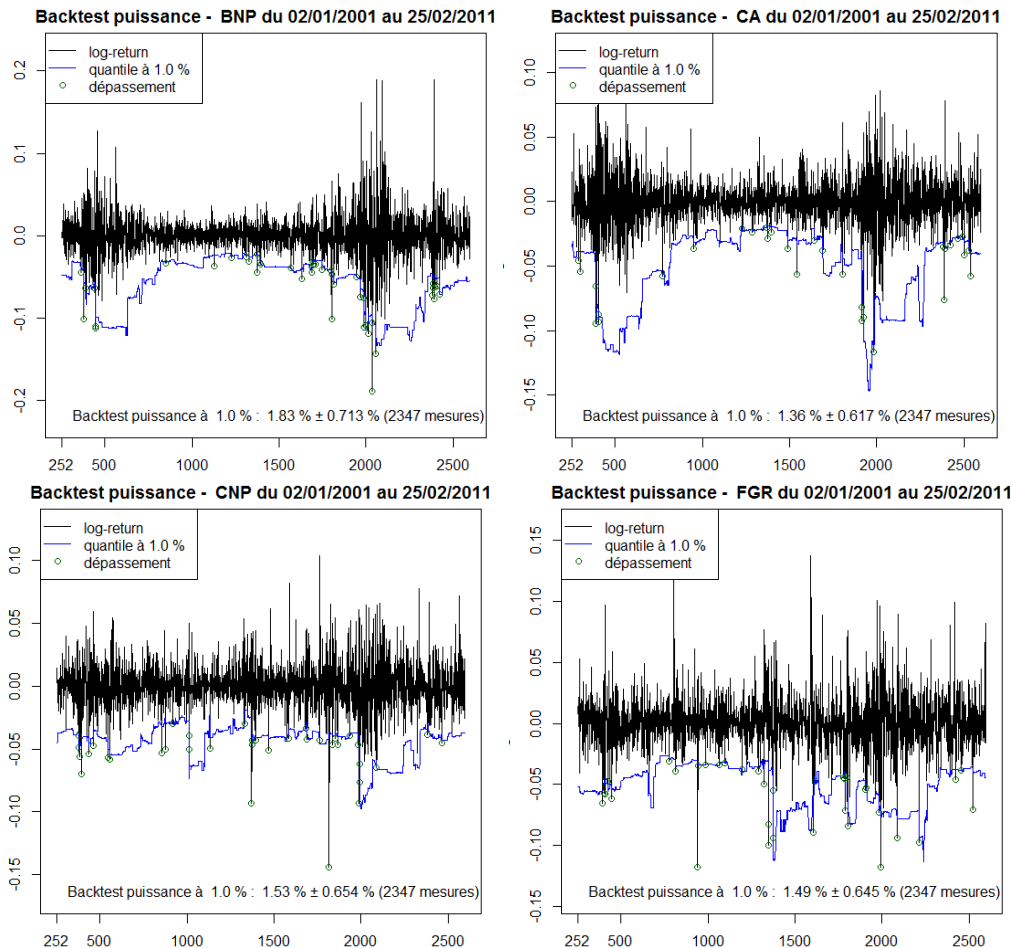


FIGURE 2.5 – Backtesting sur la VaR à 99% estimée sur une fenêtre glissante de 252 jours sur les cours des actions BNP, CA, CNP et FGR.

La comparaison sur un échantillon de 12 actifs des ratios de dépassement et leurs intervalles de confiance à 99% calculés dans le cas gaussien (figure 1.5) et le cas des lois puissances (figure 2.6), montre que le modèle des lois puissances permet de réduire la sous-estimation du risque. Nous remarquons aussi que le niveau de risque 1% appartient à l'intervalle de confiance exact du ratio de dépassement pour 8 actifs sur 12 dans le cas des lois puissances, contre 2 sur 12 pour la loi gaussienne. La VaR est estimée dans les mêmes conditions, à savoir à un niveau de risque 1% et sur une fenêtre glissante de 252 jours.

Résultats des backtests en fonction du niveau de risque. Dans ce paragraphe, nous varions le niveau de risque $p = 1 - \alpha$. Nous regardons l'effet de cette

Ratio de dépassement et son intervalle de confiance à 99% - backtest loi puissance

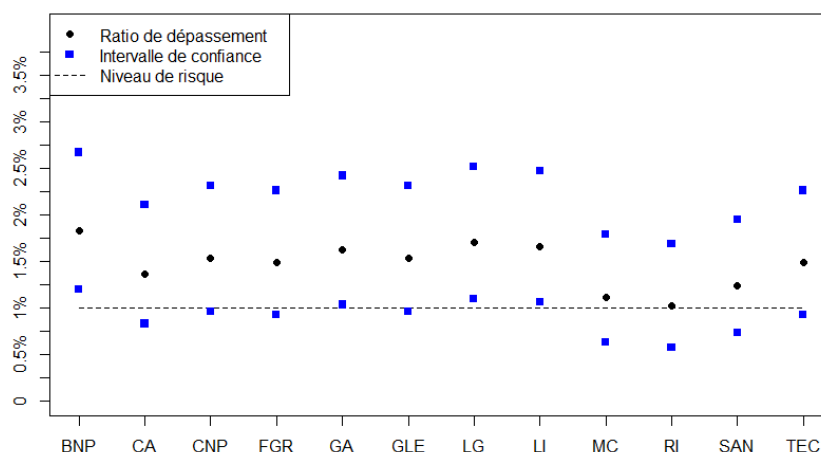


FIGURE 2.6 – Le ratio de dépassement et son intervalle de confiance à 99% pour un échantillon de 12 actifs. Dans le cadre du modèle des lois puissances, la VaR est estimée à un niveau $\alpha = 99\%$ et sur une fenêtre glissante de 252 jours sur la période 2001-2011.

variation sur les résultats des différents backtests. Ainsi, pour les niveaux de risque plus petits que 3%, nous constatons que les ratios de dépassement se rapprochent mieux des niveaux de risque associés, qu’avec le modèle gaussien. De plus, les niveaux de risque appartiennent à l’intervalle de confiance du ratio de dépassement comme le montre la figure 2.7 pour les actifs BNP, CA, CNP et FGR.

Un deuxième constat est que, pour la plupart des actifs, les backtests ne rejettent plus les différentes hypothèses d’appartenance à l’intervalle de confiance du ratio de dépassement, de couverture non-conditionnelle et de couverture conditionnelle pour les faibles niveaux de risque. Nous pouvons voir cela dans la figure 2.8 pour les actifs BNP, CA, CNP et FGR et dans la table 2.1 pour notre échantillon de 12 actifs. Sur 12 actifs, nous voyons dans la table 2.1, que l’hypothèse que le niveau de risque 1% appartient à l’intervalle de confiance à niveau 99% du ratio de dépassement, n’est pas rejetée pour 8 actifs.

2.7 Conclusion

Dans ce chapitre, nous avons étudié dans un premier temps le modèle des lois puissances et la littérature de l’estimation de ses paramètres. Dans un deuxième temps, nous avons appliqué ce modèle à notre base d’actifs dans le but d’estimer la Value-at-Risk. Les backtests de validation ont abouti à de meilleurs résultats par rapport au modèle gaussien. Nous avons constaté que le modèle des lois puissances sous-estime moins la Value-at-Risk pour les faibles niveaux de risque. Ce modèle considère les

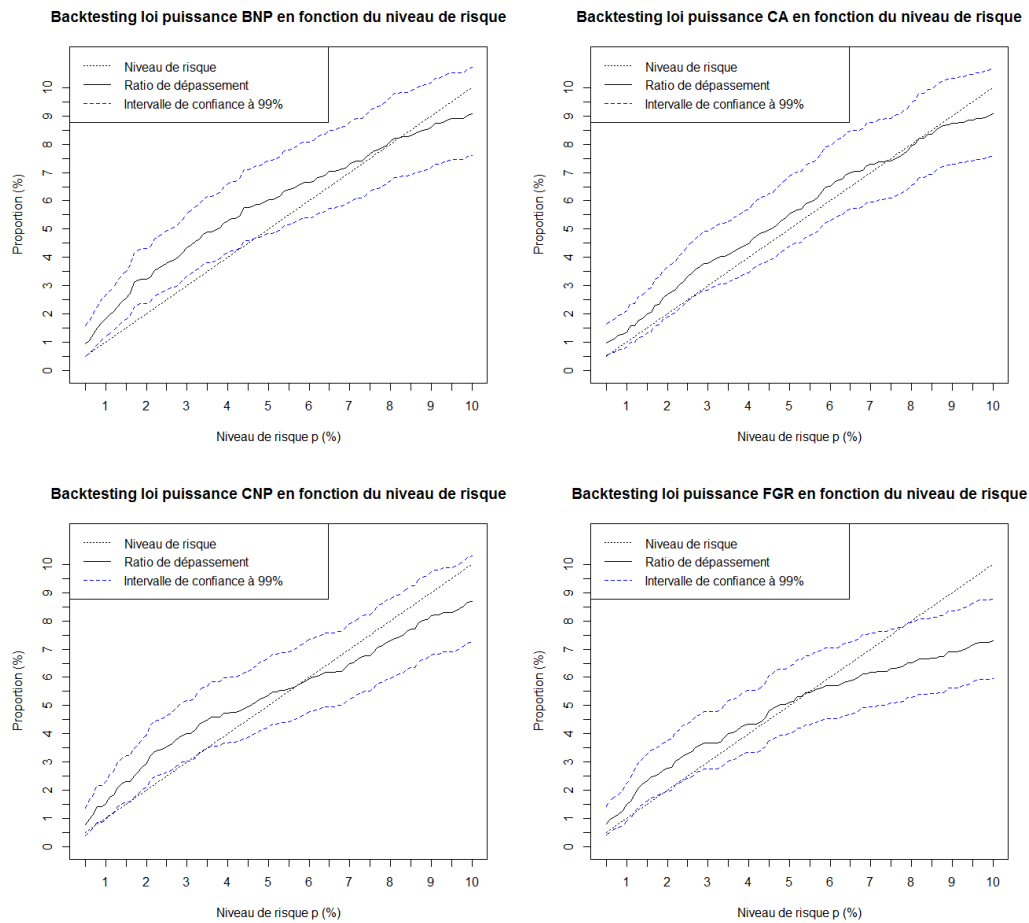


FIGURE 2.7 – Le ratio de dépassement et son intervalle de confiance à 99% en fonction du niveau de risque α de la VaR pour les actifs BNP, CA, CNP et FGR. La VaR est estimée sur les rendements de l'actif entre 2001-2011 sous le modèle des lois puissances.

queues de distribution lourdes et donne par conséquent une plus grande importance aux pertes extrêmes.

Cependant, nous remarquons, dans la figure 2.5, qu'à chaque passage d'une période de grandes fluctuations à une période de petites fluctuations, la VaR est sur-estimée. Inversement, à chaque passage d'une période de petites fluctuations à une période de grandes fluctuations, la VaR est sous-estimée. Nous expliquons cela par le fait que les données ne sont pas i.i.d. en réalité et que le marché passe par des périodes de grandes fluctuations (crise) et des périodes de petites fluctuations (hors-crise) et que la considération des modèles i.i.d. aboutit à des erreurs d'estimations à chaque changement de régimes. Cette problématique fera l'objet des deux chapitres suivants, où nous proposons un modèle des lois puissances à changement de régimes qui permettra de réduire ces erreurs d'estimations.

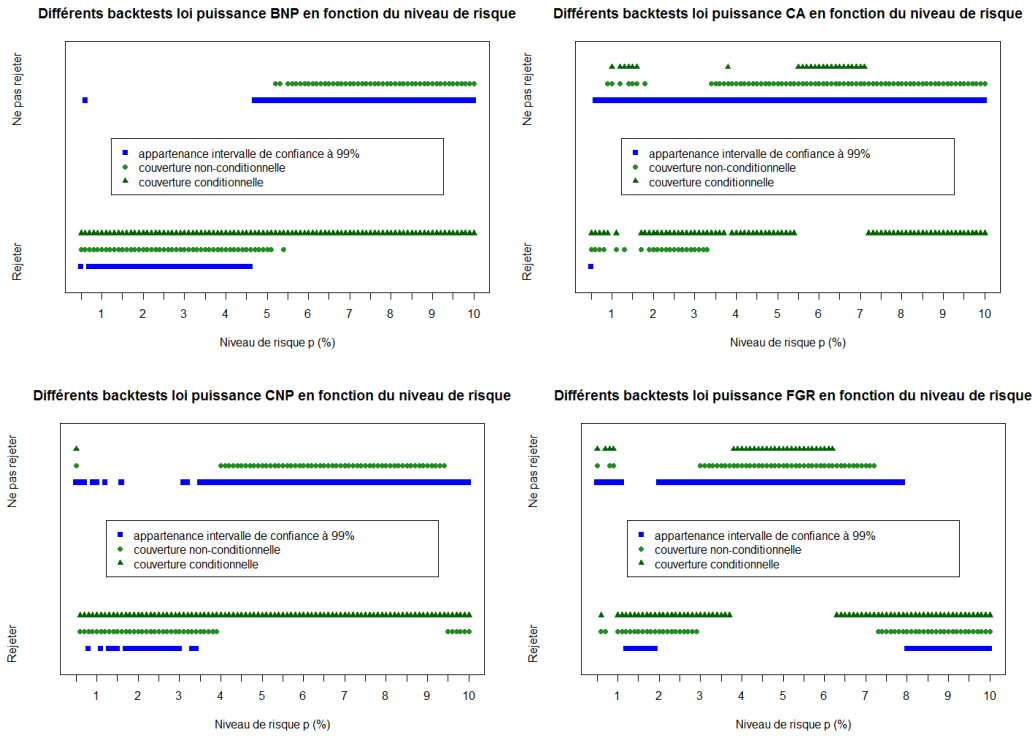


FIGURE 2.8 – Les résultats des backtests sur la VaR en fonction du niveau de risque p . La VaR est mesurée à un niveau de confiance $a = 1 - p = 0.99$ et sur une fenêtre glissante de 252 jours sur les rendements des actifs BNP, CA, CNP et FGR.

Stock	$a = 1\%$			$a = 5\%$			$a = 10\%$		
	CI	UC	CC	CI	UC	CC	CI	UC	CC
BNP	□	□	□	★	□	□	★	★	□
CA	★	★	★	★	★	□	★	★	□
CNP	★	□	□	★	★	□	★	□	□
FGR	★	□	□	★	★	★	□	□	□
GA	□	□	□	★	★	□	□	□	□
GLE	★	□	□	★	□	□	★	★	★
LG	□	□	□	★	★	□	★	★	★
LI	□	□	□	★	□	□	★	★	★
MC	★	★	★	★	★	★	★	★	□
RI	★	★	★	★	★	★	★	★	★
SAN	★	★	★	★	★	★	★	★	★
TEC	★	□	□	★	★	□	★	□	□
Total	8	4	4	12	9	4	10	8	5

★Ne pas rejeter, □Rejeter

TABLE 2.1 – Les résultats des backtests sur la VaR en fonction du niveau de risque p . La VaR est mesurée, dans le cadre du modèle des lois puissances, à des niveaux de risque 1%, 5% et 10% et sur une fenêtre glissante de 252 jours pour un échantillon de 12 actifs.

Chapitre 3

Détection des régimes : régime crise - régime hors crise

Le chapitre 2 et ce chapitre ont fait l'objet d'un rapport de recherche [33].

Bien que le modèle des lois puissances semble le mieux adapté pour estimer la Value-at-Risk parmi les modèles i.i.d., nous réalisons que cette indépendance a engendré un clustering de sous-estimations (respectivement de sur-estimations) au début de chaque période de larges (respectivement petites) fluctuations (figure 2.5). En effet, chaque passage à un régime de larges fluctuations pose un problème de sous-estimation du fait qu'on calcule le quantile d'une large fluctuation par des données de petite fluctuation. Inversement, le passage à un régime de petites fluctuations est caractérisé par des sur-estimations pour la même raison. Ainsi, nous construisons un modèle qui tient compte des changements de régimes, tout en restant focalisé sur la queue de distribution.

Les modèles à changement de régimes supposent l'existence d'un processus d'états X cachés à valeurs dans $x_1, \dots, x_K$ qui est à l'origine des rendements observés. Les paramètres du modèle subissent des changements discrets et occasionnels décrivant la transition d'un régime à un autre. Ainsi, la loi des pertes est donnée par :

$$\mathbb{P}(L_t \geq x) = \sum_{k=1}^K \mathbb{P}(L_t \geq x | X_t = x_k) \mathbb{P}(X_t = x_k). \quad (3.1)$$

Le processus d'états qui gouverne les transitions entre les différents régimes est supposé être une chaîne de Markov à K états :

$$\mathbb{P}(X_t = x_j | X_{t-1} = x_i) = p_{ij}, \quad \text{pour } i, j = 1, \dots, K.$$

Dans notre approche, nous supposons que le marché a deux régimes : un régime de *crise* (c) et un régime de *hors-crise* (h). Ensuite, nous construisons un modèle de queue de distribution pour les pertes en supposant que $L_t | X_t \sim \text{Pareto}(\alpha_{X_t}, C_{X_t})$:

$$\begin{aligned}\mathbb{P}(L_t \geq x) &= \mathbb{P}(L_t \geq x | X_t = c) \mathbb{P}(X_t = c) + \mathbb{P}(L_t \geq x | X_t = h) \mathbb{P}(X_t = h) \\ &= \mathbb{P}(X_t = c) C_c x^{-\alpha_c} + (1 - \mathbb{P}(X_t = c)) C_h x^{-\alpha_h}\end{aligned}$$

pour $x \geq x_0$.

Dans la littérature, il existe plusieurs techniques pour construire ce processus d'états. Les travaux [4, 86] utilisent des fonctions logistiques des variables endogènes retardées. L'article [147] applique des fonctions probabilistes, tandis que l'article [24] se base sur le filtre de Hamilton pour son estimation.

Toutes ces méthodes reposent sur une connaissance a priori ou un choix de la forme des probabilités de transition. Afin d'utiliser le moins possible d'a priori, nous optons pour une procédure d'apprentissage non-supervisé pour estimer ces probabilités à partir des données réelles. Pour cela, nous introduisons dans ce chapitre *le modèle de Markov caché* (MMC).

Les bases des modèles de Markov cachés ont été publiées dans une série d'articles à la fin des années 1960 et au début des années 1970, autour de Leonard E. Baum [11–15]. Ces modèles supposent, tout en disposant d'une série d'observations, l'existence d'une variable cachée qui est à l'origine de ces observations, et cherche à expliciter ce caractère caché. Depuis leur apparition, les modèles de Markov cachés ont eu diverses applications en finance. En effet, la variable cachée représente une information non-observable dans un signal donné. Elle est intéressante dans la mesure où elle fournit un résumé de l'information plus simple et plus fondamental que les données elles-mêmes. Ainsi la chaîne d'un MMC peut avoir une dynamique plus simple que le signal. Par exemple, l'introduction d'une chaîne de Markov cachée dans les modèles ARCH a permis de capter les retournements soudains du niveau de volatilité [85].

Comme dans toute approche de classification non supervisée, la question essentielle est de savoir quel genre d'information nous souhaitons extraire des données. Ainsi, parmi les études récentes appliquant ce concept, nous trouvons celle de Y. Zhang [169] où la prédiction des séries temporelles des rendements d'un portefeuille fait l'objet de l'application du MMC. Nous citons aussi la thèse de N. Mayo [115] sur la détection d'arbitrage et la prévision des volumes. Dans ses travaux, il a utilisé les modèles à variables cachés (MVC), une généralisation des MMC. L'hypothèse de markovianité dans le MMC sert à calculer une vraisemblance en un temps linéaire par rapport aux données, alors qu'il serait a priori exponentiel dans le modèle initial. Ce gain en temps de calcul est souvent vital en pratique.

Nous commençons ce chapitre par présenter le modèle de Markov caché, ainsi que les problèmes canoniques qui lui sont liés. Dans la section 2.2, nous étudions les trois premiers de ces problèmes et introduit les algorithmes forward et backward de Baum. Ensuite, nous résolvons, dans la section 3.3, le problème d'estimer les paramètres du modèle de Markov caché à partir d'une série d'observations, en appliquant l'algorithme de Baum-Welsh [14]. La section 3.4 est consacrée à l'algorithme de Viterbi [160] et à la reproduction de la chaîne d'états cachés. Dans la dernière section, nous faisons une étude sur la répartition des périodes de crises dans notre base de

données. En étudiant le temps de retour en crise, nous obtenons trois classes d'actifs boursiers. Nous retrouvons les mêmes trois classes en appliquant un algorithme d'alignement séquentiel multiple [108] sur l'ensemble des actifs étudiés.

3.1 Problématique des modèles de Markov cachés (MMC)

Considérons une chaîne de Markov $X = (X_t)_{t \geq 0}$ à valeurs dans un espace d'état fini $\mathcal{X} = \{c, h\}$. Notons ρ sa loi initiale et $Q = [Q(x, x')]_{x, x' \in \mathcal{X}}$ sa matrice de transition. Nous supposons que la chaîne X n'est pas directement observable, elle est dite *cachée*, et nous disposons d'une suite d'observations $R = (R_t)_{t \geq 0}$ à valeurs dans $\mathbb{R}$.

Notation : On désigne par $X_{k:\ell}$ la séquence $(X_k, \dots, X_\ell)$ si $k \leq \ell$ et l'ensemble vide si $k > \ell$.

Hypothèse 3.1 (Canal d'observation sans mémoire). *Soit $n \geq 1$, pour tout $x_{0:n} \in \mathcal{X}^{n+1}$ et tout $r_{0:n}$ dans $\mathbb{R}^{n+1}$:*

$$\mathbb{P}(R_{0:n} \in dr_{0:n} \mid X_{0:n} = x_{0:n}) = \prod_{t=0}^n \mathbb{P}(R_t \in dr_t \mid X_t = x_t).$$

Cette hypothèse caractérise la façon dont la loi des observations dépend de celle du processus d'état. En d'autres termes, conditionnellement aux états X_t , les observations R_t sont mutuellement indépendantes et chaque observation ne dépend que de l'état X_t .

Définition 3.1. *Un noyau markovien sur $\mathcal{X} \times \mathbb{R}$ est une famille $\psi = \{\psi(x, \cdot); x \in \mathcal{X}\}$ de fonctions de $\mathbb{R}$ dans $\mathbb{R}$ telle que :*

- (i) $\psi(x, r) \geq 0$ pour tout $x \in \mathcal{X}$, $r \in \mathbb{R}$,
- (ii) $\int_{\mathbb{R}} \psi(x, r) dr = 1$ pour tout $x \in \mathcal{X}$.

Nous supposons que la loi conditionnelle de R_t sachant X_t ne dépend pas de t et admet une densité,

$$\mathbb{P}(R_t \in dr \mid X_t = x) = \psi(x, r) dr.$$

Nous vérifions aisément que $\psi(x, r)$ est bien un noyau Markovien au sens précédent. Nous en déduisons le résultat suivant.

Proposition 3.1. *La loi jointe du couple (X, R) est entièrement caractérisée par la loi initiale ρ , la matrice de transition Q de la chaîne de Markov X et le noyau markovien d'émission ψ . En effet :*

$$\mathbb{P}(X_{0:n} = x_{0:n}, R_{0:n} \in dr_{0:n}) = \rho(x_0) \psi(x_0, r_0) \prod_{t=1}^n [Q(x_{t-1}, x_t) \psi(x_t, r_t)] dr_{0:n}, \quad (3.2)$$

pour tout $n \in \mathbb{N}$, $x_{0:n} \in \mathcal{X}^{n+1}$ et $r_{0:n} \in \mathbb{R}^{n+1}$. En particulier le couple (X, Y) est markovien.

À partir de la proposition précédente, nous pouvons définir une caractérisation du modèle de Markov caché (X, R) .

Définition 3.2 (Modèle de Markov caché). *Un modèle de Markov caché (X, R) , noté généralement (MMC), à espace d'observations continu est entièrement caractérisé par :*

(i) *la loi initiale ρ du processus d'état X :*

$$\rho(x) = \mathbb{P}(X_0 = x), x \in \mathcal{X};$$

(ii) *la matrice de transition Q du processus d'état X :*

$$Q(x, x') = \mathbb{P}(X_{t+1} = x' | X_t = x), x, x' \in \mathcal{X};$$

(iii) *le noyau markovien ψ du processus R sachant le processus X :*

$$\mathbb{P}(R_t \in dr | X_t = x) = \psi(x, r) dr, x \in \mathcal{X}, r \in \mathbb{R}.$$

On désigne par $M = (\rho, Q, \psi)$ les paramètres du modèle de Markov caché.

Les trois problèmes canoniques des modèles de Markov cachés sont :

P1 *Problème d'évaluation* : étant donné les paramètres $M = (\rho, Q, \psi)$ du modèle, donner la loi de probabilité de la suite $R_{0:n}$ en fonction des paramètres du modèle. Ce point est résolu par *l'algorithme forward de Baum*.

P2 Étant donné les paramètres $M = (\rho, Q, \psi)$ du modèle et une suite d'observations $R_{0:n}$, déterminer la suite d'états $X_{0:n}$ qui est le plus vraisemblablement à l'origine de ces observations $R_{0:n}$. On peut considérer trois cas :

- (i) *Problème de filtrage* : l'estimation de l'état courant X_n . Ceci est résolu par *l'algorithme forward de Baum* ;
- (ii) *Problème de lissage* : l'estimation d'un état intermédiaire X_k (pour $k < n$). Ceci est résolu par *l'algorithme backward de Baum* ;
- (iii) *Problème de reconnaissance* : l'estimation de la trajectoire complète $X_{0:n}$. *L'algorithme de Viterbi* permet de résoudre ce point.

P3 *Problème d'apprentissage* : étant donné une suite d'observations $R_{0:n}$, déterminer les paramètres $M = (\rho, Q, \psi)$ qui maximisent la vraisemblance d'apparition de cette suite d'observations. Ceci est résolu par *l'algorithme de Baum-Welsh*.

3.2 Problème d'évaluation, problème de filtrage et problème de lissage

Nous considérons un modèle de Markov caché $M = (\rho, Q, \psi)$ à espace d'états $\mathcal{X} = \{c, h\}$ et à espace d'observations égal à $\mathbb{R}$.

Premièrement, le problème d'évaluation **P1** vise à calculer *efficacement* la distribution de probabilité de la suite d'observations $R_{0:n}$ en fonction des paramètres du modèles M , c'est-à-dire la *fonction de vraisemblance*.

D'après (3.2), la loi des observations $R_{0:n}$ est donnée par

$$\begin{aligned}\mathbb{P}(R_{0:n} \in dr_{0:n}) &= \sum_{x_{0:n}} \mathbb{P}(X_{0:n} = x_{0:n}, R_{0:n} \in dr_{0:n}) \\ &= \sum_{x_{0:n}} \rho(x_0) \psi(x_0, r_0) \prod_{t=1}^n [Q(x_{t-1}, x_t) \psi(x_t, r_t)] dr_{0:n}\end{aligned}\quad (3.3)$$

pour $n \in \mathbb{N}$. Comme nous le verrons dans la remarque 3.1., cette formulation de la loi de $R_{0:n}$ est trop coûteuse en termes de complexité algorithmique. Pour cette raison, nous introduisons l'algorithme forward de Baum.

Considérons la loi jointe de $R_{0:n}$ et X_n :

$$\mathbb{P}(R_{0:n} \in dr_{0:n}, X_n = x) = \alpha_x^n(r_{0:n}) dr_{0:n} \quad (3.4)$$

pour $n \in \mathbb{N}$, $r_0, \dots, r_n \in \mathbb{R}$, et la variable *forward* :

$$p_n(x) = \alpha_x^n(R_{0:n}) \quad (3.5)$$

pour $n \in \mathbb{N}$, $x \in \mathcal{X}$.

Théorème 3.1 (Algorithme forward de Baum). *La suite $(p_n)_{n \in \mathbb{N}}$ des variables forward est solution de l'équation récurrente suivante :*

$$p_{n+1}(x) \leftarrow \psi(x, R_{n+1}) \sum_{x' \in \mathcal{X}} Q(x', x) p_n(x'), \quad n \geq 0, \quad x \in \mathcal{X}, \quad (3.6)$$

$$p_0(x) \leftarrow \rho(x) \psi(x, R_0), \quad x \in \mathcal{X}. \quad (3.7)$$

La loi des observations $R_{0:n}$ admet la densité suivante :

$$\mathbb{P}(R_{0:n} \in dr_{0:n}) = \left[\sum_{x \in \mathcal{X}} \alpha_x^n(r_{0:n}) \right] dr_{0:n}. \quad (3.8)$$

Remarque 3.1. La formulation (3.3) de la loi des observations $R_{0:n}$ est trop coûteuse. Pour chacune des $|\mathcal{X}|^{n+1} = 2^{k+1}$ trajectoires $x_{0:n}$ possibles de l'état, il faut en effet effectuer $2n + 1$ produits. La complexité de cet algorithme est donc en $O(n2^n)$. Elle est exponentielle en fonction du nombre n d'observations et donc impossible à utiliser en pratique. Cependant, chaque itération en temps de l'algorithme (3.7) fait appel à 10 opérations élémentaires ($N(2N + 1)$ dans le cas d'une chaîne de Markov X à N états). Ainsi le calcul (3.8) de la loi des observations $R_{0:n}$ est en $O(n)$. Il est donc linéaire en fonction du nombre d'observations.

La formulation (3.8) répond ainsi à la question du problème **P1**, en effet la fonction de vraisemblance associée au modèle M est :

$$\mathcal{L}_n(M) = \sum_{x \in \mathcal{X}} p_n(x). \quad (3.9)$$

Deuxièmement, le problème de filtrage **P2**-(i) vise à estimer au mieux l'état présent X_n . Or la loi de X_n sachant les observations $R_{0:n}$ est donnée par :

$$\mathbb{P}(X_n = x | R_{0:n}) = \frac{p_n(x)}{\sum_{x' \in \mathcal{X}} p_n(x')}, \quad x \in \mathcal{X}.$$

Ainsi, le problème de filtrage **P2**-(i) est résolu en prenant l'état dont la probabilité d'apparition est la plus grande. Nous remarquons, par ailleurs, que p_n apparaît comme une loi de probabilité conditionnelle *non-normalisée*. La constante de normalisation est la vraisemblance (3.9).

Troisièmement, le problème de lissage **P2**-(ii) vise à estimer au mieux l'état intermédiaire X_k . Pour cela nous introduisons l'algorithme backward de Baum.

Soit $n \geq 1$ un instant fixé. Pour tout instant $k < n$, considérons la loi conditionnelle des observations $R_{k+1:n}$ sachant que $X_k = x$:

$$\mathbb{P}(R_{k+1:n} \in dr_{k+1:n} | X_k = x) = \beta_x^k(r_{k+1:n}) dy_{k+1:n}$$

pour $x \in \mathcal{X}, r_{k+1}, \dots, r_n \in \mathbb{R}$, et la variable *backward* :

$$v_k(x) = \beta_x^k(R_{k+1:n})$$

pour $0 \leq k < n, x \in \mathcal{X}$.

Théorème 3.2 (Algorithme backward de Baum). *La suite $(v_k)_{k=0:n}$ des variables backward est solution de l'équation récurrente rétrograde suivante :*

$$v_k(x) \leftarrow \sum_{x' \in \mathcal{X}} Q(x, x') \psi(x', R_{k+1}) v_{k+1}(x'), \quad 0 \leq k < n, \quad x \in \mathcal{X}, \quad (3.10)$$

$$v_n(x) \leftarrow 1, \quad x \in \mathcal{X}. \quad (3.11)$$

Proposition 3.2. *Pour tout $k < n$, la loi jointe de X_k et $R_{0:n}$ est donnée par :*

$$\mathbb{P}(X_k = x, R_{0:n} \in dr_{0:n}) = \alpha_x^k(r_{0:k}) \beta_x^k(r_{k+1:n}) dr_{0:n}$$

pour tout $x \in \mathcal{X}$.

Pour tout $k < n$, la loi conditionnelle de l'état X_k sachant les observations $R_{0:n}$ est donnée par :

$$\mathbb{P}(X_k = x | R_{0:n}) = \frac{q_k(x)}{\sum_{x' \in \mathcal{X}} q_k(x')} \quad \text{où} \quad q_k(x) = p_k(x) v_k(x), \quad x \in \mathcal{X}.$$

Ainsi, pour répondre à **P2**-(ii), il suffit de prendre l'état qui a la probabilité d'apparition la plus grande.

Nous mentionnons aussi que les deux algorithmes « forward/backward » de Baum permettent de calculer les quantités $\alpha_x^n(r_{0:n})$ et $\beta_x^k(r_{k+1:n})$ d'une manière récurrente non-coûteuse, facile à implémenter.

3.3 Problème d'apprentissage et algorithme de Baum-Welsh

Dans cette section, les paramètres du modèle sont inconnus. Nous disposons uniquement d'une suite d'observations $R_{0:n}$ que nous utilisons pour estimer les paramètres $M = (\rho, Q, \psi)$. Un algorithme particulièrement efficace dans cette estimation est celui de Baum-Welsh [14]. Cet algorithme représente une procédure de maximisation locale. Il s'agit d'un dérivé de l'algorithme *Expectation-Maximization (EM)* [53, 117] où on cherche à ajuster les paramètres M qui maximisent localement la fonction de vraisemblance :

$$M \mapsto \mathcal{L}(M) = \mathbb{P}_M(R_{0:n} \in dr_{0:n})$$

où $\mathbb{P}_M$ est la probabilité sous laquelle le modèle M est valide et $r_{0:n}$ est un vecteur de $\mathbb{R}$.

3.3.1 Principe de l'algorithme

Partant d'une suite d'observation $R_{0:n}$ et d'un modèle initial M , la méthode de ré-estimation de Baum-Welsh consiste à ré-estimer d'une manière itérative les paramètres du modèle M jusqu'à arriver à un maximum local.

Considérons la probabilité de transition $x \rightarrow x'$ sachant le modèle M et les observations $R_{0:n}$:

$$\xi_{x,x'}^k(r_{0:n}) = \mathbb{P}_M(X_k = x, X_{k+1} = x' | R_{0:n} \in dr_{0:n}).$$

Ce dernier terme peut s'écrire en fonction de $\alpha_x^k(r_{0:k})$ et $\beta_x^k(r_{k+1:n})$:

$$\begin{aligned} \xi_{x,x'}^k(r_{0:n}) &= \frac{\mathbb{P}_M(X_k = x, X_{k+1} = x', R_{0:n} \in r_{0:n})}{\mathbb{P}_M(R_{0:n} \in r_{0:n})} \\ &= \frac{\mathbb{P}_M(X_k = x, X_{k+1} = x', R_{0:n} \in dr_{0:n})}{\sum_{\bar{x}, \bar{x}' \in \mathcal{X}} \mathbb{P}_M(X_k = \bar{x}, X_{k+1} = \bar{x}', R_{0:n} \in dr_{0:n})} \\ &= \frac{\alpha_x^k(r_{0:k})Q(x, x')\psi(x', r_{k+1})\beta_{x'}^{k+1}(r_{k+2:n})}{\sum_{\bar{x}, \bar{x}' \in \mathcal{X}} \alpha_{\bar{x}}^k(r_{0:k})Q(\bar{x}, \bar{x}')\psi(\bar{x}', r_{k+1})\beta_{\bar{x}'}^{k+1}(r_{k+2:n})}. \end{aligned}$$

Posons

$$\gamma_x^k(r_{0:n}) = \sum_{x' \in \mathcal{X}} \xi_{x,x'}^k(r_{0:n}).$$

Remarquons que $\gamma_x^k(r_{0:n}) = \mathbb{P}_M(X_k = x | R_{0:n} \in dr_{0:n})$. Ainsi, la proposition 3.2. nous donne :

$$\gamma_x^k(r_{0:n}) = \frac{\alpha_x^k(r_{0:k})\beta_x^k(r_{k+1:n})}{\sum_{x' \in \mathcal{X}} \alpha_{x'}^k(r_{0:k})\beta_{x'}^k(r_{k+1:n})}.$$

Une sommation en k de 0 à $n - 1$ de cette expression conduit à :

$$\begin{aligned}
\sum_{k=0}^{n-1} \gamma_x^k(r_{0:n}) &= \sum_{k=0}^{n-1} \mathbb{P}_M(X_k = x | R_{0:n} \in dr_{0:n}) \\
&= \sum_{k=0}^{n-1} \mathbb{E}_M[\mathbf{1}_{X_k=x} | R_{0:n} \in dr_{0:n}] \\
&= \mathbb{E}_M\left[\sum_{k=0}^{n-1} \mathbf{1}_{X_k=x} | R_{0:n} \in dr_{0:n}\right] \\
&= \mathbb{E}_M[\#\{0 \leq k < n : X_k = x\} | R_{0:n} \in dr_{0:n}].
\end{aligned}$$

Cette somme peut être interprétée comme l'estimateur du nombre de transitions issues de x de la suite $X_{0:n}$ au vu des observations $R_{0:n}$.

De même :

$$\begin{aligned}
\sum_{k=0}^{n-1} \xi_{x,x'}^k(r_{0:n}) &= \sum_{k=0}^{n-1} \mathbb{P}_M(X_k = x, X_{k+1} = x' | R_{0:n} \in dr_{0:n}) \\
&= \sum_{k=0}^{n-1} \mathbb{E}_M[\mathbf{1}_{\{X_k=x, X_{k+1}=x'\}} | R_{0:n} \in dr_{0:n}] \\
&= \mathbb{E}_M\left[\sum_{k=0}^{n-1} \mathbf{1}_{\{X_k=x, X_{k+1}=x'\}} | R_{0:n} \in dr_{0:n}\right] \\
&= \mathbb{E}_M[\#\{0 \leq k < n : X_k = x, X_{k+1} = x'\} | R_{0:n} \in dr_{0:n}]
\end{aligned}$$

représente l'estimateur du nombre de transitions $x \rightarrow x'$ de la suite $X_{0:n}$ au vu des observations $R_{0:n}$.

Ainsi :

$$\sum_{k=0}^{n-1} \xi_{x,x'}^k(r_{0:n}) \quad \text{et} \quad \sum_{k=0}^{n-1} \gamma_x^k(r_{0:n})$$

sont des estimateurs respectivement du nombre de transitions $x \rightarrow x'$ et du nombre de transitions issues de x de la suite $X_{0:n}$ au vu des observations $R_{0:n}$.

Nous proposons alors, partant d'un modèle initial $M = (\rho, Q, \psi)$, les « ré-estimateurs » suivants

$$\bar{\rho}(x) = \gamma_x^0(r_{0:n}), \tag{3.12}$$

$$\bar{Q}(x, x') = \frac{\sum_{k=0}^{n-1} \xi_{x,x'}^k(r_{0:n})}{\sum_{k=0}^{n-1} \gamma_x^k(r_{0:n})}, \tag{3.13}$$

$$\bar{\psi}(x, r) = \frac{\sum_{k=0}^n \mathbf{1}_{r_n \in dr} \gamma_x^k(r_{0:n})}{\sum_{k=0}^n \gamma_x^k(r_{0:n})}. \tag{3.14}$$

En effet, il est clair que :

$$\begin{aligned}\bar{\rho}(x) &= \text{estimateur du nombre d'état } x \text{ à l'instant initial } k = 0, \\ \bar{Q}(x, x') &= \frac{\text{estimateur du nombre de transitions } x \rightarrow x'}{\text{estimateur du nombre de transitions issues de } x}, \\ \bar{\psi}(x, r) &= \frac{\text{estimateur du nombre de transitions issues de } x \text{ et émettant } dr}{\text{estimateur du nombre de transitions issues de } x}.\end{aligned}$$

Ainsi, étant donné le modèle courant $M = (\rho, Q, \psi)$, l'utilisation des formules (3.12), (3.13) et (3.14) permet de ré-estimer le modèle et obtenir $\bar{M} = (\bar{\rho}, \bar{Q}, \bar{\psi})$. Baum a montré [14] que :

1. Soit le modèle initial M est un point critique de la fonction de vraisemblance :

$$\nabla_M \mathcal{L}_n(M) = \nabla_M \mathbb{P}_M(R_{0:n} \in dr_{0:n}) = 0$$

et alors $\bar{M} = M$;

2. Soit le modèle $\bar{M}$ est strictement plus vraisemblable au sens où :

$$\mathcal{L}_n(\bar{M}) = P_{\bar{M}}(R_{0:n} \in dr_{0:n}) > \mathcal{L}_n(M) = P_M(R_{0:n} \in dr_{0:n}).$$

L'itération de cette procédure donne un algorithme EM qui permet de déterminer un maximum local de la fonction de vraisemblance.

3.3.2 Application sur des données boursières

Nous disposons d'une série de log-rendements $R_{1:n}$. Nous appliquons l'algorithme de Baum-Welsh pour déterminer les paramètres $M = (\rho, Q, \psi)$ du modèle de Markov caché. Les résultats sont les suivants pour le cas de l'actif BNP :

- La loi initiale ρ de X :

$$\rho(c) = 4.97 \cdot 10^{-05}, \quad \rho(h) = 0.9999503.$$

Il est tout à fait logique d'avoir $\rho(h)$ presque égale à 1. En effet, $\rho(h) = \mathbb{P}(X_1 = h)$ est la probabilité que le premier jour de la suite soit en période de hors-crise.

- La matrice de transition Q de X :

$$\begin{pmatrix} Q(c, c) & Q(c, h) \\ Q(h, c) & Q(h, h) \end{pmatrix} = \begin{pmatrix} 0.979 & 0.021 \\ 0.007 & 0.993 \end{pmatrix}.$$

Nous voyons bien que la probabilité de passer d'une période de crise à une période hors-crise est 3 fois plus élevée que celle de passer d'une période hors-crise à une période crise. En effet, la première quantité s'interprète par la probabilité $Q(c, h)$ d'être aujourd'hui en période hors-crise sachant que la veille on était en crise, alors que la deuxième quantité est définie par la probabilité $Q(h, c)$ d'être aujourd'hui en crise sachant que la veille on était en période hors-crise. De l'autre côté, la probabilité de rester en état hors-crise est supérieure à celle de rester en état crise. Ces résultats, en termes de matrice de transition de X , sont acceptables et conformes avec la réalité du marché financier.

— Les moyennes et les variances par ligne du noyau d'émission ψ sont :

$$\begin{pmatrix} \overline{\psi(c, R)} & \text{Var}(\psi(c, R)) \\ \overline{\psi(h, R)} & \text{Var}(\psi(h, R)) \end{pmatrix} = \begin{pmatrix} -0.12 & 0.18 \\ 0.05 & 0.02 \end{pmatrix} \times 10^{-2}.$$

La variance de la série des rendements ayant l'état crise est beaucoup plus élevée que celle des log-rendements ayant l'état hors-crise. Cela s'explique par la grande variance qui caractérise les périodes de crise dans les marchés financiers.

En termes de performance informatique, le nombre moyen d'itérations, effectuées par l'algorithme, pour aboutir au résultat (les paramètres du modèle qui présentent un maximum local en termes de mieux correspondre aux rendements) est 15 itérations. Cette moyenne est calculée en lançant l'algorithme 10 fois sur les log-rendements de l'actif BNP.

En conclusion, les paramètres de notre modèle sont cohérents avec la réalité du fonctionnement du marché financier. De plus, des résultats similaires ont été trouvés lors de l'application du modèle de Markov caché sur d'autres actifs boursiers. On peut dire alors que la résolution du problème de ré-estimation a donné des résultats conformes à ce que nous attendons. L'étape suivante est la résolution du problème de reconnaissance et la détermination de la série des états, qui fait l'objet de la prochaine section.

3.4 Problème de reconnaissance et algorithme de Viterbi

3.4.1 Principe de l'algorithme

Étant donnés les paramètres $M = (\rho, Q, \psi)$ du modèle et une séquence d'observations $R_{0:n}$, la réponse au problème de reconnaissance **P2**-(iii) est fournie par un algorithme de programmation dynamique, *l'algorithme de Viterbi* [160], qui permet de maximiser la loi conditionnelle de la suite des états $X_{0:n}$ par rapport aux observations $R_{0:n}$.

Lorsqu'il s'agit de calculer la loi conditionnelle de l'état à un instant donné, les algorithmes « forward » et « backward » de Baum sont satisfaisants. En effet,

$$\mathbb{P}(X_k = x | R_{0:n}) = \frac{q_k(x)}{\sum_{x' \in \mathcal{X}} q_k(x')}, \quad k \leq n,$$

pour tout $x \in \mathcal{X}$, où $q_k(x) = p_k(x)v_k(x)$.

En revanche, lorsqu'il s'agit de calculer des statistiques comme la moyenne conditionnelle, l'utilisation des lois conditionnelles précédentes s'avère sans sens vu que

les états $\{c, h\}$ de la chaîne X sont des étiquettes et qu'on ne peut pas effectuer des opérations algébriques sur ces valeurs. Les statisticiens ont proposé de calculer des estimateurs comme *l'estimateur local du maximum à posteriori* tout en supposant que le maximum est atteint en un point unique :

$$X_k^{LMAP} = \arg \max_{x \in \mathcal{X}} \mathbb{P}(X_k = x | R_{0:n}) = \arg \max_{x \in \mathcal{X}} q_k(x).$$

Cependant, un autre obstacle apparaît. En effet, il arrive que la suite $X_{0:n}^{LMAP}$ engendrée ne soit pas cohérente avec le modèle. Autrement dit, il est possible que $X_k^{LMAP} = x$, $X_{k+1}^{LMAP} = x'$ alors que la probabilité de transition $Q(x, x') = 0$. Pour cette raison, nous optons finalement pour *l'estimateur global du maximum à posteriori*, défini par :

$$X_{0:n}^{MAP} = \arg \max_{x \in \mathcal{X}} \mathbb{P}(X_{0:n} = x_{0:n} | R_{0:n}). \quad (3.15)$$

Cet estimateur est calculé à l'aide d'un algorithme de programmation dynamique, appelé *algorithme de Viterbi* [160].

Tout d'abord, remarquons qu'il est équivalent dans (3.15) de chercher l'argument du maximum de la loi conditionnelle ou de la loi jointe vu que la maximisation se fait sur $x_{0:n}$. Nous introduisons alors la quantité :

$$\delta_x^n(R_{0:n}) = \max_{x_{0:n-1}} \mathbb{P}(X_{0:n-1} = x_{0:n-1}, X_n = x, R_{0:n} \in r_{0:n})$$

pour tout $n \geq 0$, $x_0, \dots, x_{n-1}, x \in \mathcal{X}$, $r_{0:n} \in \mathbb{R}^{n+1}$. Nous définissons, par la suite, la fonction valeur :

$$w_n(x) = \delta_x^n(R_{0:n}), \quad n \geq 0, \quad x \in \mathcal{X}.$$

Théorème 3.3 (Programmation dynamique). *La suite $(w_n)_{n \in \mathbb{N}}$ des fonctions valeur est solution de l'équation non-linéaire récurrente suivante :*

$$w_{n+1}(x) \leftarrow \psi(x, R_{n+1}) \max_{x' \in \mathcal{X}} [Q(x', x) w_n(x')], \quad n \geq 0, \quad x \in \mathcal{X}, \quad (3.16)$$

$$w_0(x) \leftarrow \rho(x) \psi(x, R_0), \quad x \in \mathcal{X}. \quad (3.17)$$

Corollaire 3.1 (Algorithme de Viterbi). *Soit n un instant fixé, nous commençons par calculer la séquence $w_{0:n}$. Nous obtenons, par la suite, l'estimateur $X_{0:n}^{MAP}$ par récurrence rétrograde :*

$$X_n^{MAP} \leftarrow \arg \max_{x \in \mathcal{X}} w_n(x) \quad (3.18)$$

puis, pour $k = n - 1, \dots, 0$:

$$X_k^{MAP} \leftarrow \arg \max_{x \in \mathcal{X}} [Q(x, X_{k+1}^{MAP}) w_k(x)]. \quad (3.19)$$

Pour des détails de la preuve, voir l'article de base [160].

3.4.2 Application sur des données boursières

Ayant maintenant une série d'observation $R_{1:n}$ représentant les log-rendements d'un actif boursier et les paramètres $M = (\rho, Q, \psi)$ calculés par l'algorithme de Baum-Welsh, nous cherchons à déterminer la série des états $X_{1:n}$ et à attribuer ainsi à chaque jour une étiquette crise (c) ou hors-crise (h). Ce problème est connu sous le nom du problème de reconnaissance et il est résolu par l'algorithme de Viterbi.

La figure 3.1 montre le résultat de l'application de cet algorithme sur les log-rendements du BNP. Nous voyons en rouge les périodes de crise et en bleu les périodes de hors-crise.

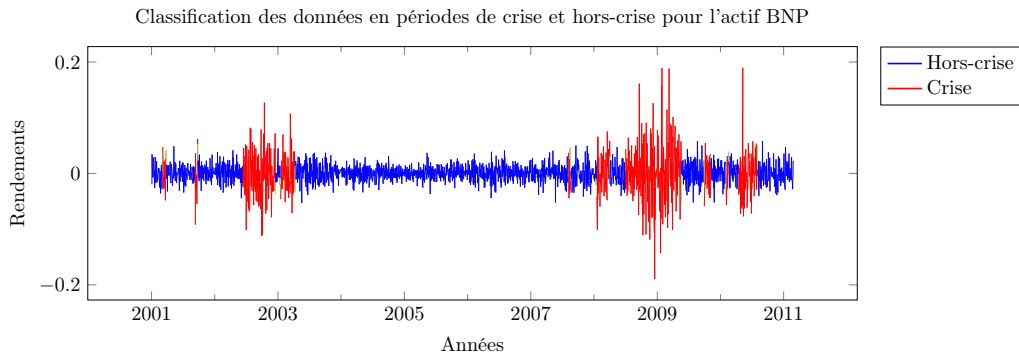


FIGURE 3.1 – Classification et détermination des périodes crise et des périodes hors-crise pour l'actif BNP.

3.5 Etude des crises

Le modèle de Markov caché nous permet d'attribuer à un jour donné un état crise ou hors-crise. Dans cette section, nous nous intéressons à cette série d'états et nous étudions quelques caractéristiques telles que la durée moyenne d'une crise et la distribution du temps de retour à la crise. Le but est de définir des classes d'actifs boursiers selon la répartition temporelle des périodes de crise et de hors-crise.

3.5.1 Temps de retour à la crise

L'étude des courbes de temps de retour à la crise pour différents actifs boursiers est une manière de voir la répartition temporelle des périodes de crise et hors-crise. En effet, une telle courbe est généralement à valeurs entre 0 et 100 jours avec un ou plusieurs pics qui peuvent aller jusqu'à 4 ans. Ainsi, la transition entre période de crise et période de hors-crise se fait d'une façon cyclique, sauf une ou plusieurs périodes de stabilité en hors-crise, caractérisée par des rendements à petite variation. En classant les actifs par rapport à ces pics, on distingue trois grandes classes.

- Une classe d’actifs ayant le pic au début de la courbe. Ces actifs sont caractérisés donc par une longue période de hors-crise suivie par un cycle d’oscillations entre crise et hors-crise. Nous trouvons dans cette classe les actifs BN, BELI, BIO, BNP, CA, CAPLI, CMA, FP, GFC, GLE, HAV, IPN, ITT, KSA, LI, MMT, RAL, SBTBR, TEC, TFI et UL.
- Une classe ayant le pic à la fin de la courbe. Ces actifs sont caractérisés par quelques petites périodes de crise et hors-crise suivies d’une longue période de hors-crise. Ainsi, nous constatons un phénomène d’influence d’un groupe d’actifs sur d’autres exprimé par l’agitation, que cette classe a subi au début des données et a transmis ensuite à la première classe. D’où, des courbes de temps de retour à la crise inversées. Cette deuxième catégorie inclut les actifs BDAL, COX, CS, CSAR, EAD, EN, EUR, EXE, GOE, RGR, RI, RMS, SAN, SELER et UBI.
- Une troisième classe caractérisée par plusieurs petits pics éparpillés sur l’ensemble de données. Ainsi, pour cette dernière, la transition entre crise et hors-crise se fait d’une manière continue avec quelques périodes de hors-crise qui dépassent les 200 jours. Ces actifs peuvent être considérés comme des actifs instables et très sensibles aux perturbations du marché financier. Comme actifs dans cette classe, on cite CNP, DG, DNX, FGR, FPG, ORG, SAF et SOP.

La figure 3.2 montre les allures de graphes pour ces trois classes de comportement des actifs boursiers.

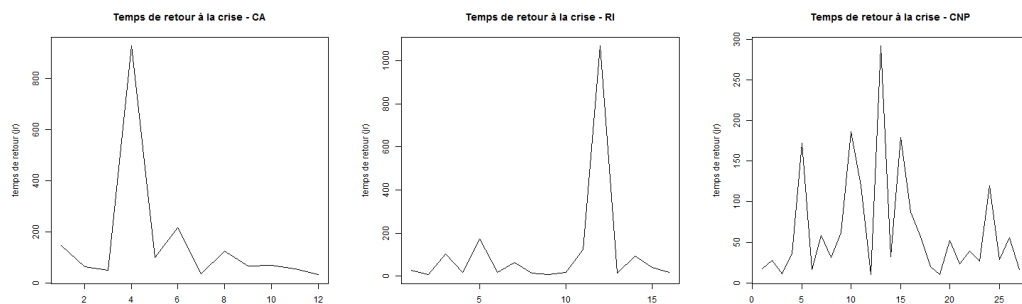


FIGURE 3.2 – Les allures des courbes des temps de retour à la crise, respectivement, pour les trois classes d’actifs.

3.5.2 Densité empirique

L’étude de la densité empirique de la distribution de temps de retour rencontre certaines difficultés. En effet, nous disposons généralement d’une dizaine de points pour estimer la densité. Pour cela, nous utilisons la méthode d’estimation par noyau (appelée aussi la méthode de Parzen-Rozenblatt) [129], une méthode non-paramétrique qui représente une généralisation de la méthode d’estimation par histogramme. Nous remarquons que les deux premières classes d’actifs ont la même allure de courbe avec une accumulation de la densité autour du point 100 jours et une petite accumulation

entre les points 800 et 1200 jours. En ce qui concerne la troisième classe la courbe est plus homogène. Les figures 3.3, 3.4 et 3.5 montrent des exemples de densité empirique pour chaque classe d'actifs. Une certaine discontinuité dans les courbes de densité est observée, mais cela peut s'expliquer par le faible nombre de données utilisées dans l'estimation.

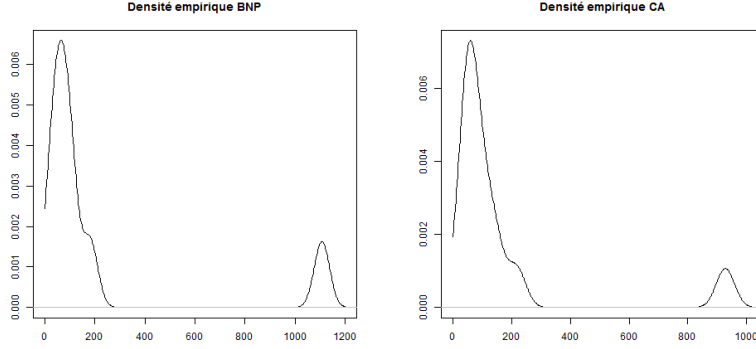


FIGURE 3.3 – Densité empirique d'actifs de la première classe.

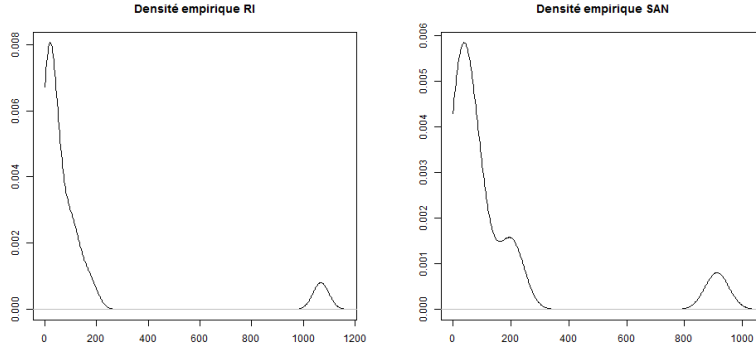


FIGURE 3.4 – Densité empirique d'actifs de la deuxième classe.

Étant dans une période de hors-crise, le modèle le plus intuitif pour le temps de retour à la crise τ est celui d'une loi géométrique de probabilité de succès

$$p = \mathbb{P}(X_{i+1} = c/X_i = h) = Q(h, c).$$

Ainsi, la probabilité que le retour en crise soit dans k jours est :

$$\mathbb{P}(\tau = k) = (1 - p)^{k-1}p = Q(h, h)^{k-1}Q(h, c).$$

La densité théorique de la distribution de temps de retour à la crise pour l'actif BNP Paribas est illustrée par la figure 3.6. Le résultat n'est pas conforme à ce que nous avons obtenu numériquement. En effet, la densité théorique ne prend pas en compte la longue période de hors-crise représentée par la pic autour du point 1000 jours dans la courbe empirique. Cela met en cause l'hypothèse de la loi géométrique du temps de retour à la crise. Ainsi, des tests plus sophistiqués, qui prennent en compte la limitation de la taille de l'échantillon (une dizaine de points), sont à étudier comme suite de ce travail.

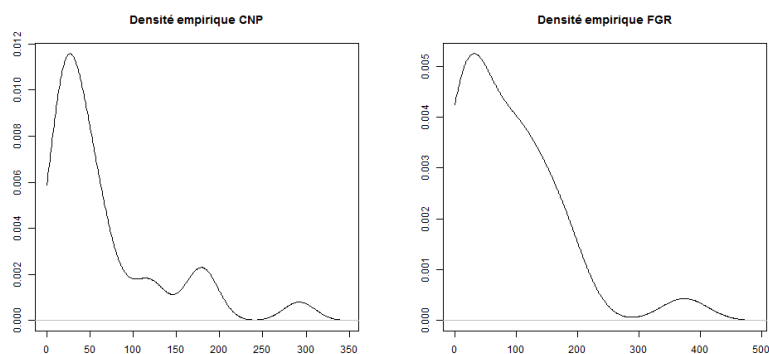


FIGURE 3.5 – Densité empirique d’actifs de la troisième classe.

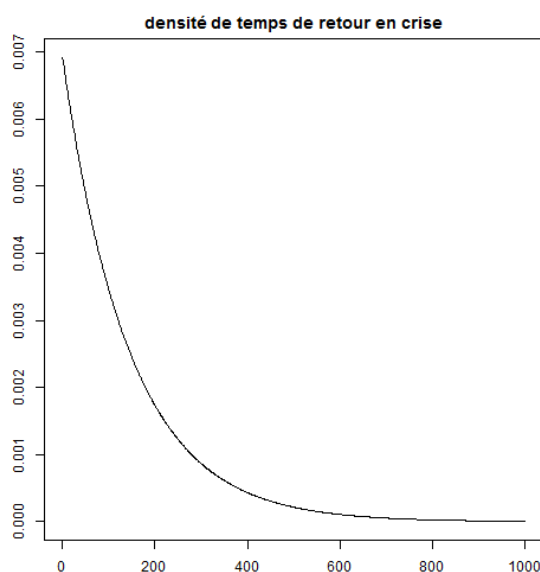


FIGURE 3.6 – Densité de temps de retour à la crise pour l’actif BNP

3.5.3 Classification par alignement séquentiel multiple

L’alignement séquentiel consiste à représenter deux ou plusieurs chaînes de caractères les unes sous les autres, de manière à en faire ressortir les régions homologues ou similaires. L’objectif de l’alignement est de disposer les caractères pour identifier des zones de concordance. Les premières utilisations de l’alignement séquentiel sont apparues en bio-informatique sur des séquences macromoléculaires biologiques (ADN, ARN ou protéines), ainsi que dans les correcteurs d’orthographe pour mesurer la distance entre un mot correct et un mot comportant une faute d’orthographe humaine. Récemment, certaines études s’intéressent à appliquer ce principe en finance. S. Peng et al. se sont basés sur l’alignement pour estimer la relation entre les valeurs mobilières et le taux de change [130]. En supposant que les fluctuations des taux de change sont à l’origine d’une fraction relativement importante de la variation du marché boursier de Taiwan, ils appliquent l’alignement séquentiel multiple sur des séquences de données de marché. Une autre étude de G. Wang [162] utilise

l'alignement pour mesurer la similarité et l'évolution de la topologie des marchés de change.

L'alignement séquentiel consiste à calculer la distance de Levenshtein [108] qui permet de quantifier l'écart entre deux chaînes de caractères. C'est une mesure de la similarité entre deux chaînes de caractères, égale au nombre minimal de caractères qu'il faut supprimer, insérer ou remplacer pour passer d'une chaîne à l'autre. Le résultat final est représenté dans une matrice appelée matrice de score. A partir de cette dernière, nous traçons un dendrogramme (arbre) illustrant l'arrangement des différents groupes similaires de séquences.

Nous appliquons ce concept sur les séquences d'états crise/hors-crise afin de classer les actifs par rapport au critère de crise. Nous cherchons à identifier la concordance des périodes de crise entre différents actifs. Le calcul des matrices de score donne les arbres suivants dans les figures 3.7 et 3.8.

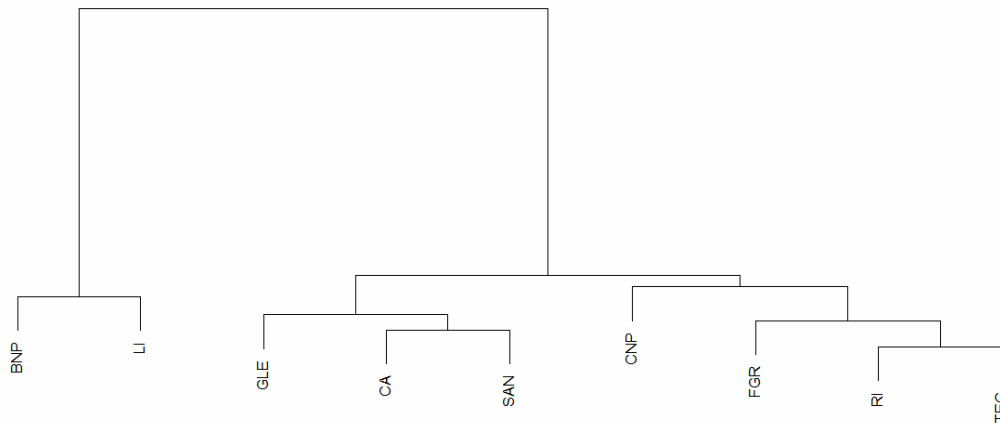


FIGURE 3.7 – Arbre de classification des actifs BNP, CA, CNP, FGR, SAN, LI, RI, GLE, TEC par rapport au critère de la crise.

L'alignement fait ressortir trois classes d'actifs selon la concordance en termes de répartition des périodes de crise. De plus, nous remarquons que, sur ces trois classes, il y en a deux qui sont plus proches. Cela nous confirme les résultats de l'étude du temps de retour à la crise. En effet, l'alignement séquentiel regroupe les actifs ayant des séquences d'états crise/hors-crise similaires. Cette similarité peut être simultanée ou avec un décalage temporel. Ainsi, il est clair que la classe ayant la longue période de hors-crise au début et celle ayant cette longue période de hors-crise à la fin vont être proches dans la classification de l'alignement séquentiel. Quelques actifs changent de classes selon que nous utilisons la première méthode ou la deuxième. Nous pensons que nous devons combiner les deux méthodes pour définir un classement final. Cette piste sera explorée dans le futur dans le but d'améliorer le temps de calcul sur une base de données d'actifs en déterminant la chaîne d'états pour un seul actif de chaque classe et l'utiliser pour tous les actifs de sa classe.

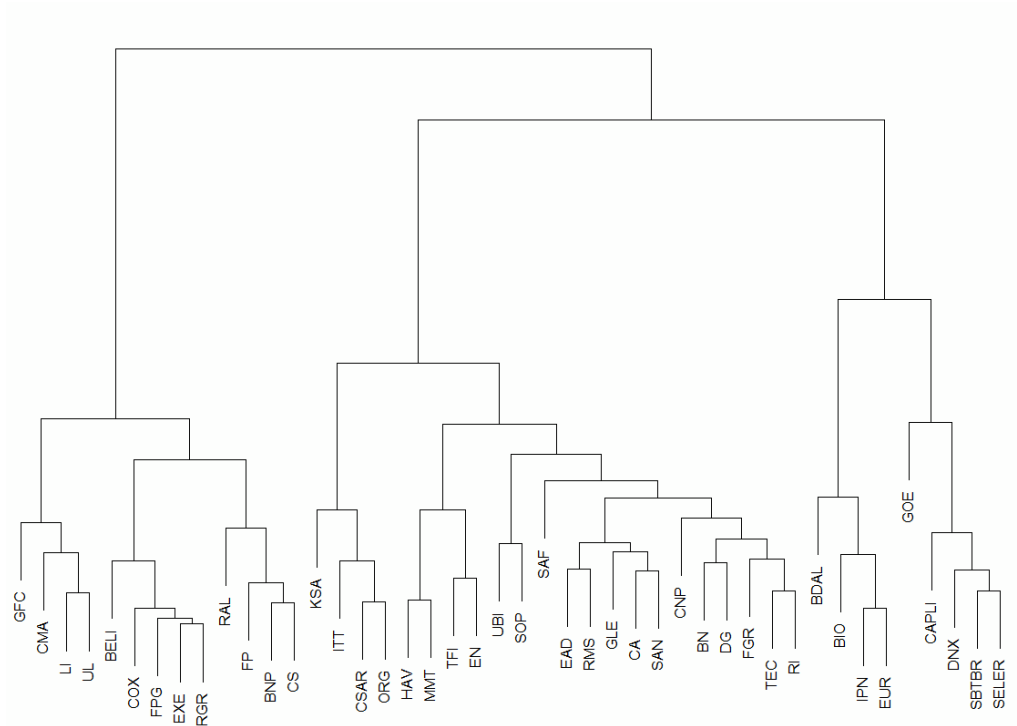


FIGURE 3.8 – Arbre de classification d’un ensemble de 44 actifs de la bourse de Paris par rapport au critère de la crise.

3.6 Conclusion

Ce chapitre a permis d’introduire le modèle de Markov caché, ainsi que ses applications sur les séries de rendements des actifs boursiers, à savoir la détection des régimes du marché, la loi de transition entre ces derniers et la classification des données en périodes de crise et périodes hors-crise. Par ailleurs, une partie a été consacrée à l’étude des crises tout en essayant d’identifier des classes d’actifs ayant des comportements similaires face aux périodes de crise. Cela pourra raffiner les modèles de mesure de risque en l’adaptant à chaque classe d’actifs boursiers. Le chapitre suivant se sert de ces informations pour construire un modèle des lois puissances à changement de régimes pour le calcul de la Value-at-Risk.

Chapitre 4

Modèle des lois puissances à changement de régimes

Ce chapitre a fait l'objet d'un article [144] publié dans *Physica A*.

Dans ce chapitre, nous construisons le modèle des lois puissances à changement de régimes. À chaque instant t , nous estimons le modèle de Markov caché sur les rendements antérieurs, nous calculons les paramètres des lois puissances pour chaque régime et nous évaluons la VaR. La procédure d'estimation est détaillée dans la section 4.1. Ensuite, dans la section 4.2, nous calculons la probabilité d'être en crise à l'instant $t + 1$. La section 4.3 illustre les résultats de l'estimation des paramètres des queues de distributions. Dans la section 4.4, nous validons le modèle par des procédures de backtesting. Dans la dernière section, nous effectuons une étude comparative entre le modèle des lois puissances à changement de régimes, le modèle simple des lois puissances, le modèle des lois stables et le modèle GARCH, sur une base de 56 actifs boursiers.

4.1 Procédure d'estimation

Comme nous avons mentionné dans l'introduction du chapitre précédent, nous supposons que le marché a deux régimes : un régime de *crise* (c) et un régime de *hors-crise* (h). Ensuite, nous construisons un modèle de queue de distribution pour les pertes en supposant que $L_t|X_t$ suit une loi de Pareto(α_{X_t}, C_{X_t}) :

$$\begin{aligned}\mathbb{P}(L_t \geq x) &= \mathbb{P}(L_t \geq x|X_t = c)\mathbb{P}(X_t = c) + \mathbb{P}(L_t \geq x|X_t = h)\mathbb{P}(X_t = h) \\ &= \mathbb{P}(X_t = c) C_c x^{-\alpha_c} + (1 - \mathbb{P}(X_t = c)) C_h x^{-\alpha_h}\end{aligned}$$

pour $x \geq x_0$.

L'estimation des paramètres de crise se fait sur des fenêtres glissantes contenant uniquement des données du régime de crise. De la même façon, les paramètres du régime hors-crise sont estimés sur des fenêtres glissantes sur le régime hors-crise.

Nous fixons $t_0 = 1500$ le nombre minimal de données pour avoir une estimation du modèle de Markov caché. Pour $t \geq t_0$, la $\text{VaR}_t^{(a)}$ à un niveau de confiance a est donnée par le quantile $q^{L_{t+1}}(a)$. La procédure d'estimation de la VaR est la suivante :

- (a) On estime les paramètres $M = (\rho, Q, \psi)$ du MMC à partir de l'échantillon $(R_1, \dots, R_t)$ par l'algorithme de Baum-Welsh.
- (b) On estime la suite des états cachés $(X_1, \dots, X_t)$ par l'algorithme de Viterbi.
- (c) On estime les paramètres (C_c, α_c) de la loi Pareto du régime de crise à partir des dernières 252 données crise dans $(R_1, \dots, R_t)$.
- (d) On estime les paramètres (C_h, α_h) de la loi Pareto du régime de hors-crise à partir des dernières 252 données hors-crise dans $(R_1, \dots, R_t)$.
- (e) On calcule la probabilité d'être le lendemain en état de crise :

$$p_{t+1} = \mathbb{P}_M(X_{t+1} = c | R_{1:t}).$$

- (f) Le quantile $q^{L_{t+1}}(a)$, à un niveau a , de la perte à l'instant $t+1$ est ainsi solution de

$$1 - a = p_{t+1} C_c x^{-\alpha_c} + (1 - p_{t+1}) C_h x^{-\alpha_h}.$$

Cette équation est résolue numériquement en utilisant la méthode de Newton-Raphson.

4.2 La probabilité d'être en crise

La Value-at-Risk à un instant t calcule la perte maximale potentielle qu'on peut subir à l'instant $t+1$. L'état du jour $t+1$ n'est pas connu. Ainsi nous devons l'estimer à partir des informations disponibles à l'instant t . Comme expliqué dans le chapitre précédent, l'algorithme de Viterbi détermine la suite des états $(X_1, \dots, X_n)$ en maximisant la probabilité d'être dans ces états. Ainsi, pour tout $1 \leq s \leq t$, la probabilité que $X_s = c$ n'est pas forcément égale à 0 ou 1, mais égale à

$$\delta_s(c) = \mathbb{P}(X_s = c | R_{1:t}).$$

Ainsi, la probabilité d'être en crise à l'instant $t+1$ peut être calculée en fonction de δ_t et de la matrice de transition Q en appliquant la formule suivante :

$$\begin{aligned} p_{t+1} &= \mathbb{P}_M(X_{t+1} = c | R_{1:t}) \\ &= \mathbb{P}(X_{t+1} = c | X_t = c) \mathbb{P}(X_t = c) + \mathbb{P}(X_{t+1} = c | X_t = h) \mathbb{P}(X_t = h) \\ &= Q(c, c) \delta_{t-1}(c) + Q(h, c) (1 - \delta_{t-1}(c)). \end{aligned}$$

La figure 4.1 illustre les probabilités conditionnelles $Q(c, c)$ et $Q(h, c)$ estimées sur une fenêtre étendue (une fenêtre qui augmente de 1 à chaque itération : $R_{1:t}$ puis $R_{1:t+1}$ puis $R_{1:t+2}$, ...). Nous constatons que l'estimation des paramètres du MMC sur une fenêtre étendue est stable et ne présente que des légères fluctuations. La

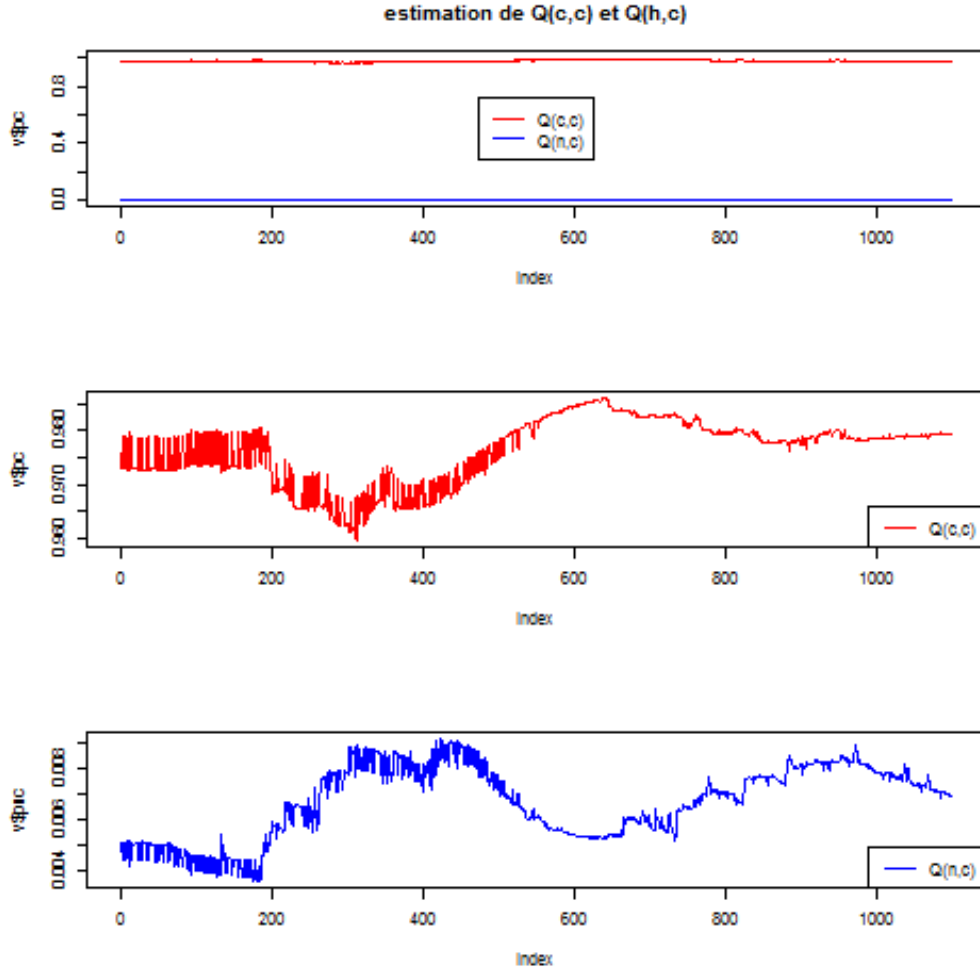


FIGURE 4.1 – Estimation du paramètre Q du modèle de Markov caché sur une fenêtre étendue pour l'actif BNP Paribas sur la période de novembre 2006 à février 2011. La taille initiale de la fenêtre d'estimation est fixée à 1500 pour avoir suffisamment de données.

probabilité de rester dans un état de crise $Q(c, c)$ est proche de 1 et la probabilité d'entrer dans un régime de crise $Q(h, c)$ est proche de 0.

Dans la figure 4.2, nous présentons la probabilité p_{t+1} d'être en crise le lendemain dans le cas de l'actif BNP. Du haut vers le bas, sont données pour chaque abscisse t : les rendements, la probabilité conditionnelle $Q(c, c)$ estimée sur la fenêtre $R_{1:t}$, la probabilité conditionnelle $Q(h, c)$ estimée sur la fenêtre $R_{1:t}$ et finalement en bas la probabilité $p_{t+1} = \mathbb{P}_M(X_{t+1} = c | R_{1:t})$.

Nous constatons sur la figure 4.2 que la probabilité de rester en crise $Q(c, c)$ augmente légèrement dans la période à grandes fluctuations au milieu de l'axe des abscisses. Nous voyons aussi que la probabilité estimée $Q(h, c)$ respecte bien la propriété de décroissance en fonction de $Q(c, c)$, donnée par la formule de Bayes. La probabilité p_{t+1} détecte semblablement les changements de régimes. Nous remarquons que p_{t+1}

accorde majoritairement des valeurs proches de 1 aux périodes de grandes fluctuations et des valeurs proche de 0 aux périodes de petites fluctuations avec quelques fausses alertes où elle ne donne pas la bonne valeur. Ceci est lié au fait que, parfois, un changement de régimes n'est pas détecté rapidement par l'algorithme de Viterbi. Cependant, les estimations des paramètres du MMC et des états sont faites sur des fenêtres de 1500 à 2599 données. Avec un échantillon plus large, nous pensons que les états étiquetés par les algorithmes de MMC vont être plus stables vis-à-vis de l'ajout ou du retrait de quelques données.

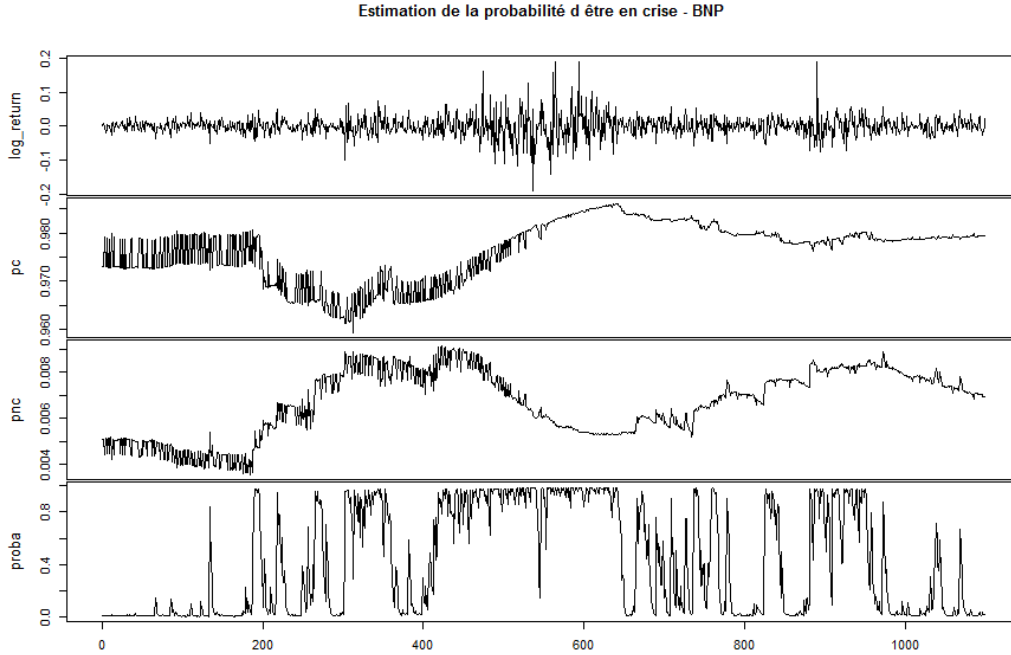


FIGURE 4.2 – Estimation des probabilités conditionnelles $Q(c, c)$ et $Q(h, c)$ sur une fenêtre étendue et calcul de la probabilité d'être le lendemain en crise pour l'actif BNP Paribas sur la période de novembre 2006 à février 2011. La taille initiale de la fenêtre d'estimation est fixée à 1500 pour avoir suffisamment de données.

4.3 Estimation des paramètres des queues des distributions

Nous estimons les paramètres du modèle des lois puissances à changement de régimes tel que décrit ci-dessus. À un instant $t \geq t_0$, nous utilisons les 252 derniers jours du sous-échantillon des données crise pour estimer les paramètres (C_c, α_c) de la queue de distribution du régime crise. Respectivement, nous utilisons les 252 derniers jours du sous-échantillon des données hors-crise pour estimer les paramètres (C_h, α_h) de la queue de distribution du régime hors-crise.

Les résultats de ces estimations sont illustrés dans la figure 4.3. La courbe du paramètre α_c présente une moyenne de 3.64, tandis que celle de α_h a une moyenne

de 4.30. Cela s'explique par le fait que les grandes pertes font généralement parties du régime crise, et par conséquent la queue de distribution du régime crise est plus lourde que celle du régime hors-crise.

Les paramètres estimés C_c , α_c , C_h et α_h représentent toujours des fluctuations mais moins importantes que celles de C et α dans le cas du modèle des lois puissances (figure 2.4). Néanmoins, nous constatons toujours le même phénomène de compensation dans les couples (C_c, α_c) et (C_h, α_h) .

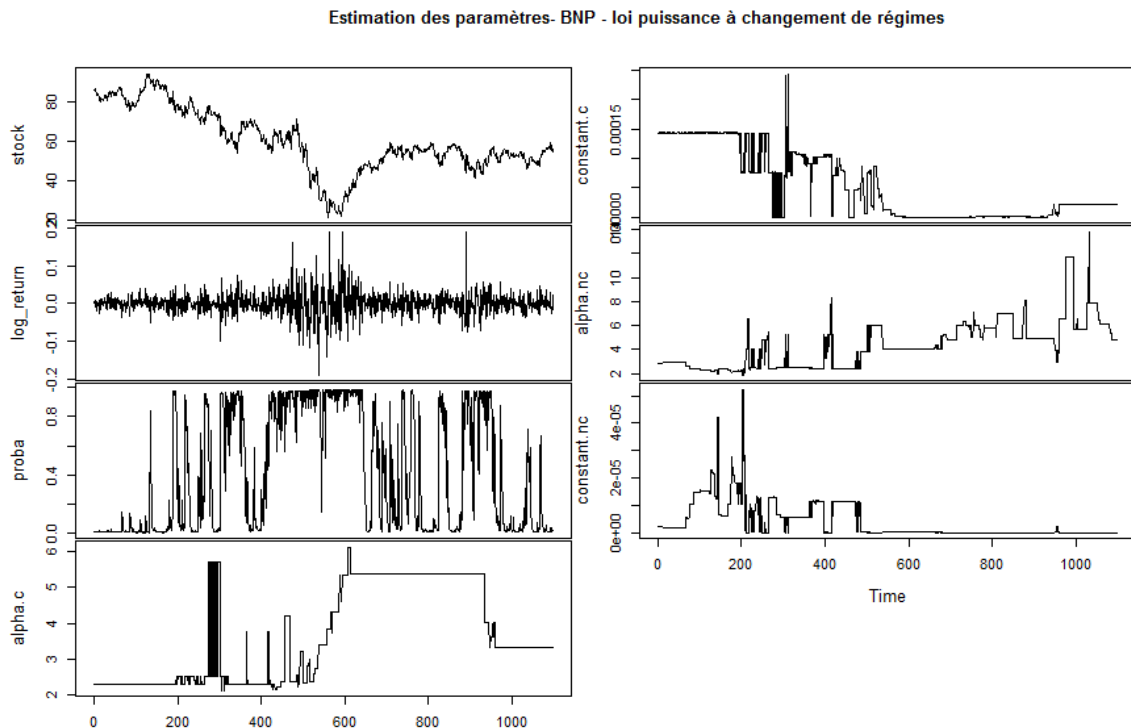


FIGURE 4.3 – Prix de clôture, rendements, probabilité d'être en crise, α_c , C_c , α_h et C_h pour l'actif BNP Paribas sur la période novembre 2006 à février 2011.

La VaR monétaire évaluée à partir de ces paramètres est représentée dans la figure 4.4. En comparant cette figure à sa semblable dans le cadre du modèle des lois puissance (figure 2.3) sur la même période du 11/2006 à 02/2011, nous remarquons qu'il y a plus de petites variations mais moins de grandes variations. L'évaluation de la VaR monétaire sous le modèle des lois puissances à changement de régimes est donc sujette à des petites écarts journaliers mais elle s'adapte mieux et quasi-instantanément aux variations du prix d'actif. Ceci conduit à de moins grandes variations dans la courbe de la VaR monétaire.

4.4 Backtesting

Dans cette section, nous appliquons les procédures de backtesting, présentées dans la section 1.3 du chapitre 1, sur la VaR mesurée sous le modèle des lois puissances

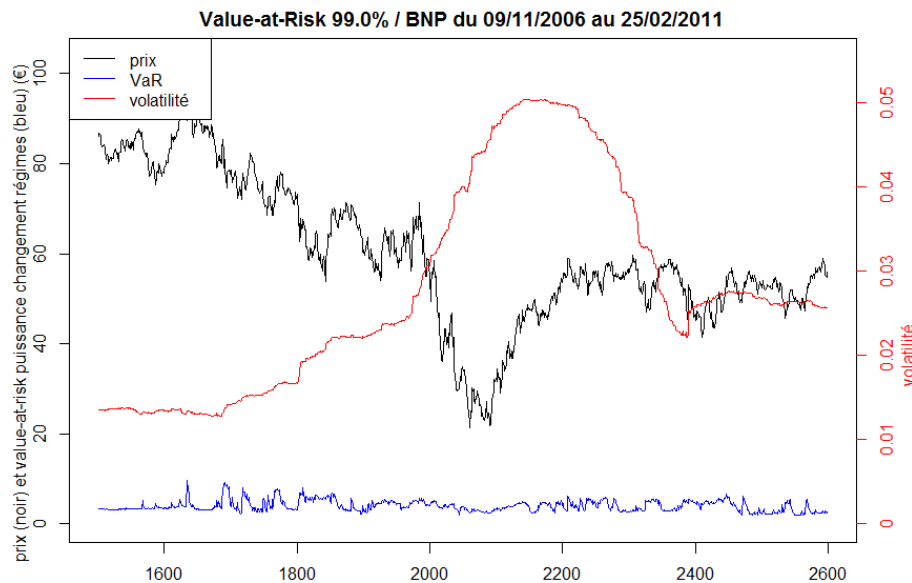


FIGURE 4.4 – VaR monétaire à 99% estimée sous le modèle des lois puissances à changement de régimes pour l'action BNP Paribas sur la période de novembre 2006 à février 2011.

à changement de régimes. Nous nous concentrons sur la performance du modèle dans le calcul de la VaR à un niveau de confiance 99%, c'est-à-dire un niveau de risque $p = 1\%$. En effet, nous avons vu dans les chapitres 2 et 3 que, pour des niveaux de risque élevés, le modèle des lois puissances et même le modèle gaussien fournissent une bonne estimation de la VaR. Cependant, le vrai défi est l'estimation des événements rares qui représentent les grandes pertes. Cela nécessite un calcul de la VaR avec des niveaux de risques très petits. Dans la suite du chapitre, nous considérons un niveau de confiance $a = 99\%$.

La figure 4.5 montre le backtest du ratio de dépassement sur la VaR estimée à un niveau 99% sur les cours des actions BNP, CA, CNP et FGR sur la période de novembre 2006 à février 2011. Les délais d'estimation et les sous/sur-estimations à chaque changement de régime, que nous avons constatés dans le modèle gaussien (figure 1.4) et le modèle des lois puissances (figure 2.5), sont éliminés par l'introduction des notions crise et hors-crise. Nous concluons que le fait de tenir compte du régime du marché dans le calcul de la VaR conduit à une adaptation rapide au changement de régime.

Nous illustrons dans la figure 4.6 les backtests du ratio de dépassement, ainsi que leurs intervalles de confiance à 99%, appliqués sur les VaR estimées à un niveau 99% de 12 actifs. Cet échantillon d'actifs contient des représentants des trois classes d'actifs, dégagées dans la section 3.5.1 du chapitre précédent. Nous voyons que le modèle des lois puissances à changement de régimes arrive à estimer la VaR correctement, même pour des actifs de la troisième classe, qui sont instables et très sensibles aux perturbations du marché. En effet, pour ces derniers, il y aura une succession des petites périodes de crise et de hors-crise et les probabilités de transition seront proches de $1/2$. Ainsi, le modèle ne va pas chercher très loin dans le passé et va calibrer les

paramètres des lois puissances, à peu près, sur les 504 données antérieures. Ce qui donnera tout de même une bonne estimation. Nous pouvons voir cela plus en détail dans la figure 4.7, qui est faite sur la base de 56 actifs.

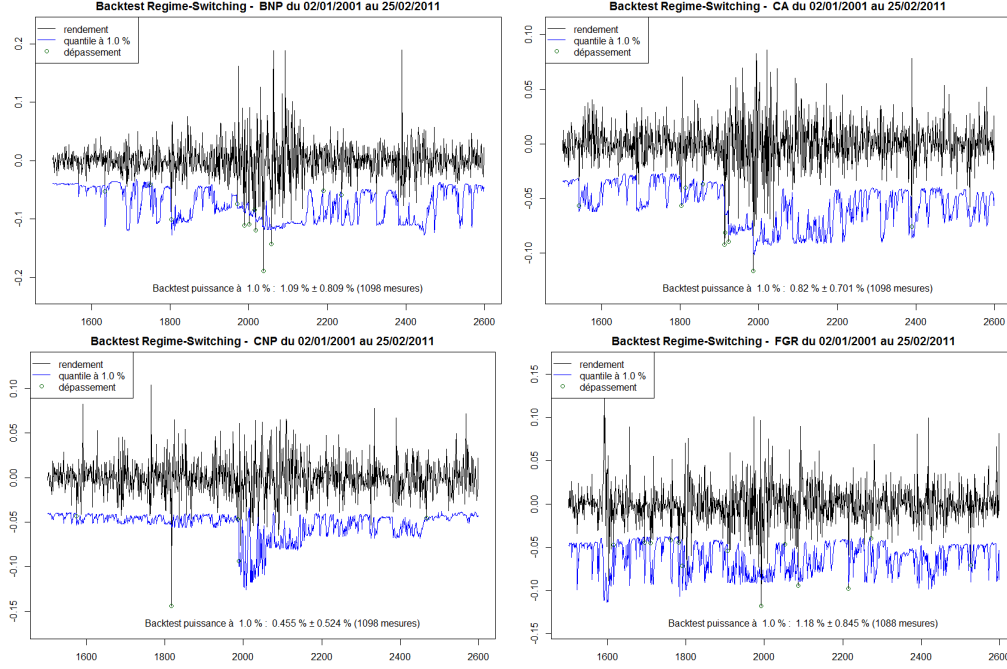


FIGURE 4.5 – Backtesting sur la VaR à 99% estimée dans le cadre du modèle des lois puissances à changement de régimes sur les cours des actions BNP, CA, CNP et FGR entre novembre 2006 et février 2011.

4.5 Étude comparative : Benchmark

Afin d’opposer les performances du modèle de lois puissances à changement de régimes à celles d’autres modèles de benchmark, nous faisons dans cette section une étude comparative entre quatre modèles du prix d’actif : le modèle des lois puissances et le modèles des lois puissances à changement de régimes pour les queues de distribution des pertes, le modèle des lois stables et le modèle de t-GARCH pour la distribution des rendements. L’étude est faite sur une base de 56 actions de la bourse de Paris

(I) Le **Modèle des lois puissances**. Ce modèle est décrit dans le chapitre 3. Il consiste à modéliser la queue de distribution de la perte L_t par

$$\mathbb{P}(L_t \geq x) = C x^{-\alpha}, \quad \text{pour } x \geq x_0.$$

(II) Le **Modèle des lois puissances à changement régimes**. C’est le modèle décrit dans ce chapitre. La queue de distribution de la perte L_t suit la loi suivante :

$$\mathbb{P}(L_t \geq x) = C_c x^{-\alpha_c} \mathbb{P}(X_t = c) + C_h x^{-\alpha_h} (1 - \mathbb{P}(X_t = c)), \quad \text{pour } x \geq x_0.$$

Ratio de dépassement et son intervalle de confiance à 99% - backtest loi normale

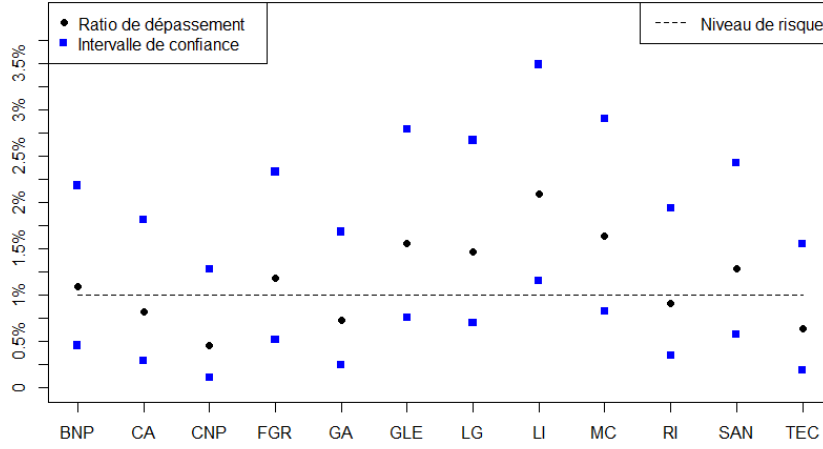


FIGURE 4.6 – Le ratio de dépassement et son intervalle de confiance à 99% pour un échantillon de 12 actifs. Dans le cadre du modèle des lois puissances à changement de régimes, la VaR est estimée à un niveau $a = 99\%$ sur la période novembre 2006-février 2011.

(III) Le **Modèle des lois stables**. Ce modèle est défini par la fonction caractéristique du rendement R_t ,

$$\phi_{R_t} = \begin{cases} \exp[i\mu t - \sigma^\alpha |t|^\alpha (1 - i\beta \operatorname{sign}(t) \tan(\frac{\pi\alpha}{2}))] & \text{si } \alpha \neq 1, \\ \exp[i\mu t - \sigma |t|(1 + \frac{2}{\pi} i\beta \operatorname{sign}(t) \ln |t|)] & \text{si } \alpha = 1, \end{cases} \quad (4.1)$$

avec $\alpha \in (0, 2)$ l'indice de queue, $\beta \in (-1, 1)$ le paramètre de dissymétrie, $\sigma \geq 0$ le paramètre d'échelle et $\mu \in \mathbb{R}$ le paramètre de position (pour plus de détails, voir [63, 128]). Notons que les lois stables apparaissent naturellement comme des classes universelles lorsque nous regardons le théorème central limite. En effet, une généralisation du théorème central limite dit que la somme de variables aléatoires ayant des queues de distribution décroissantes selon $1/|x|^{\alpha+1}$ avec $0 < \alpha < 2$ (ayant donc une variance infinie) tend vers une loi stable de paramètre α . D'autre part, notons que la loi normale n'est qu'une loi stable de paramètre $\alpha = 2$.

Pour l'estimation des paramètres α , β , σ et μ , nous appliquons l'algorithme de McCulloch [116] tel que implémenté dans la librairie `fbs` de R sur une fenêtre glissante de 252 jours. Nous calculons par la suite le quantile à travers une estimation directe de la densité de probabilités.

(IV) Le **Modèle t-GARCH**. Nous considérons le modèle t-GARCH(1,1) [27] pour décrire la série temporelle des rendements financiers. Plus précisément, nous supposons que (R_t) satisfait le modèle suivant :

$$R_t = \mu + \varepsilon_t = \mu + \sigma_t Z_t, \quad (4.2)$$

$$\sigma_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2, \quad (4.3)$$

où $(Z_t)_{t \geq 0}$ est une suite d'innovations i.i.d. qui suivent une loi de Student. Notre choix des innovations de type Student est basé sur plusieurs études qui montrent

que ce choix donne la meilleure performance pour le modèle de GARCH en termes d'ajustement aux données réelles [45].

Ce modèle est souvent proposé comme alternative aux lois puissances pour la modélisation des queues de distribution épaisses. Ceci s'explique par sa capacité à capturer les regroupements de la volatilité. Cependant, nous mentionnons que d'autres études, sur des différents marchés financiers, ont conclu à la supériorité des lois puissances dans la prédiction du quantile [37, 73, 74, 99].

Nous estimons les paramètres du modèle t-GARCH sur les rendements de nos actions boursières en utilisant la librairie `rugarch` de R sur une fenêtre élargie, c'est-à-dire toutes les données disponibles jusqu'à l'instant $t : (R_1, \dots, R_t)$. Cette approche est préférée car la détection des regroupements de volatilité nécessite plus de données. Nous calculons par la suite le quantile à travers une estimation directe de la fonction de densité.

4.5.1 Base de données : Actions boursières de Euronext Paris

Les instruments financiers considérés dans l'étude empirique sont des actions échangées sur NYSE Euronext Paris. Les données sur le marché sont fournis par eSignal (Données Interactives) et dans ce qui suit, les stocks sont identifiés par leur symbole eSignal. Sur l'ensemble des valeurs cotées à la bourse Euronext Paris en Février 2011 (plus de 600), nous avons sélectionné ceux ayant des cotations, durant toute la période allant de Janvier 2001 à Février 2011 (plus de 11 ans). Ce qui nous a conduit à un sous-ensemble de 56 actions parmi eux certaines des actions les plus liquides qui composent l'indice CAC40. Dans la présente étude, les séries temporelles considérées sont les rendements des prix de clôture des actions sélectionnées. Les paramètres des différents modèles sont estimés sur un échantillon, comme expliqué précédemment. Le choix d'une longueur de fenêtre définie sur un an, dans des modèles I et III, est un compromis entre la nécessité d'avoir suffisamment de données pour faire un bon étalonnage et inclure des crises récentes et l'augmentation du risque de ne pas respecter l'hypothèse de stationnarité avec une fenêtre de données plus large. En ce qui concerne la stationnarité, il est difficile d'établir une conclusion claire à partir des tests de Dickey-Fuller [54] et de Kwiatkowski-Phillips-Schmidt-Shin [103] qui testent l'hypothèse de stationnarité contre celle de racine unitaire. En outre, nous ne sommes pas en mesure d'identifier d'autres tailles de fenêtre qui dépasseront toujours la performance d'une fenêtre de taille 252 jours, en ce qui concerne le backtesting de la VaR ou les mesures de stationnarité.

4.5.2 Discussion

Dans cette section, la performance de prédiction du modèle des lois puissances à changement de régimes et celles des modèles de comparaison sont illustrées. Nous appliquons les méthodes d'estimation de la VaR et des backtests décrits précédemment dans la bases de données réelles. Notre comparaison est basée sur le backtest sur la VaR à un seul niveau de confiance 99%. La performance de chaque modèle

est résumée par la moyenne de ses résultats sur l'ensemble des actifs sélectionnés.

La figure 4.7 présente les ratios de dépassement de la VaR sous chacun des quatre modèles pour les 56 actifs sélectionnés. Clairement, le modèle des lois stables sur-estime la VaR et le modèle des lois puissances la sous-estime. Les résultats du modèle GARCH et du modèle des lois puissances à changement de régimes fluctuent autour de la valeur cible 1%. En termes de ratio moyen de dépassement, le modèle des lois puissances à changement de régimes est le plus proche de 1% avec une valeur de 1.12%, suivi par 0.86% pour le modèle GARCH, puis le modèle des lois puissances avec 1.37% et le modèle des lois stables avec 0.55%. Le nombre d'actifs dont leurs résultats du backtest appartiennent à $[0.90\%, 1.10\%]$, est 8 pour le modèle des lois puissances à changement de régimes, 8 pour le modèle GARCH, 5 pour le modèle des lois puissances et 7 pour le modèle des lois stables.

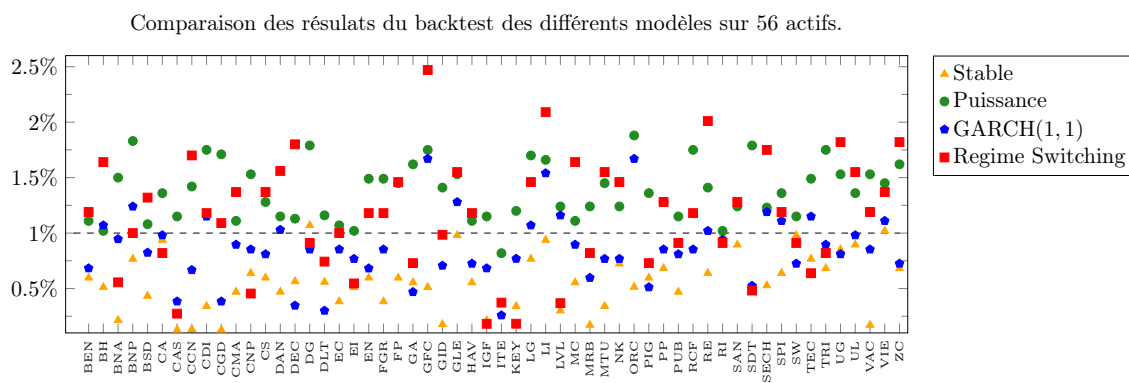


FIGURE 4.7 – Comparaison des résultats du backtest du ratio de dépassement sous les quatre différents modèles et pour un échantillon de 56 actifs. La VaR est estimée à un niveau de confiance $\alpha = 1 - p = 0.99$.

Vu le nombre limité disponible pour le backtest, les intervalles de confiance sont relativement larges. La taille moyenne de l'intervalle de confiance est de 0.39% pour les lois stables, 0.60% pour les lois puissances, 0.98% pour le modèle GARCH et de 0.80% pour le modèle des lois puissances à changement de régimes. Ce ne sera pas possible de donner une conclusion pertinente pour un seul actif. Cependant l'étude sur l'ensemble des 56 actifs, illustrés dans la table 4.1, indique que le modèle des lois puissances à changement de régimes donne les meilleurs résultats en termes de backtests. Sous le modèle des lois puissances à changement de régimes, 49 parmi les 56 intervalles de confiance exacts contiennent la valeur cible 1%, à l'opposé des modèles GARCH, lois puissances et lois stables, sous lesquels le taux d'appartenance du 1% à l'intervalle de confiance est respectivement 46/56, 43/56 et 36/56. Les tests de couverture non-conditionnelle et de couverture conditionnelle donnent le meilleur résultat pour le modèle GARCH suivi par le modèle des lois puissances à changement de régimes. Ces résultats sont aussi dans la table 4.1.

Nous interprétons ces résultats comme suit : la queue de distribution des lois stables est beaucoup plus lourde vu que son indice de queue α est toujours inférieur à 2 par construction. Cela est confirmée par le fait que l'indice de queue estimé est très proche de 2 pour tous les actifs de notre base de données. Nous expliquons ce

constat par le fait que les indices de queue des données réelles sont supérieurs à 2, mais les estimations ont été contraintes de s'arrêter à 2 pour répondre aux exigences du modèle des lois stables [116, 165]. D'autre part, la queue de distribution des lois puissances n'est pas assez lourde pour décrire le marché Euronext Paris durant les périodes de crise. Le modèle des lois puissances à changement de régimes repose sur une combinaison des queues de distribution de type lois puissances pour les périodes de crise et de stabilité de marché, ce qui produit une queue de distribution plus lourde.

La figure 4.8 illustre les estimations de quantile pour notre exemple typique BNP Paribas. Comme les modèles des lois stables et des lois puissances sont des modèles i.i.d., qui ne prennent pas en compte les changements de régimes, ils souffrent d'un retard d'adaptation à ces changements en termes d'estimation du quantile. Le début des périodes à petites variations des rendements est caractérisé par des sur-estimations de la VaR. Contrairement, la VaR est souvent sous-estimée en début des périodes à larges variations des rendements. Cependant, en ce qui concerne le modèle GARCH, par sa détection des regroupements de volatilité, il arrive à prendre en compte des changements de régimes et surmonte ces sous-estimations et sur-estimations. Notre approche de classification par MMC a pour but de fournir un modèle simple qui tire les avantages des modèles des lois puissances et ceux des modèles conditionnels tels que le modèle GARCH. Par la considération des périodes de crise et de hors-crise et la combinaison de leurs queues de distribution de type lois puissances, nous avons créé un ré-équilibre des estimations de quantile aux changements de régimes. Au début des périodes de larges fluctuations, le seuil de la VaR est augmenté grâce au terme additionnel calculé à partir des précédentes périodes de crises. Ce qui a permis de réduire les sous-estimations. D'une manière opposée, au début des périodes de petites fluctuations, l'addition d'un terme calculé à partir des précédentes périodes de hors-crisis a conduit à une réduction du seuil de la VaR, d'où la suppression de la sur-estimation. Ainsi, la considération de deux états cachés a permis de réduire le retard de détection des changements de régimes dans l'estimation du quantile tout en gardant le modèle simple.

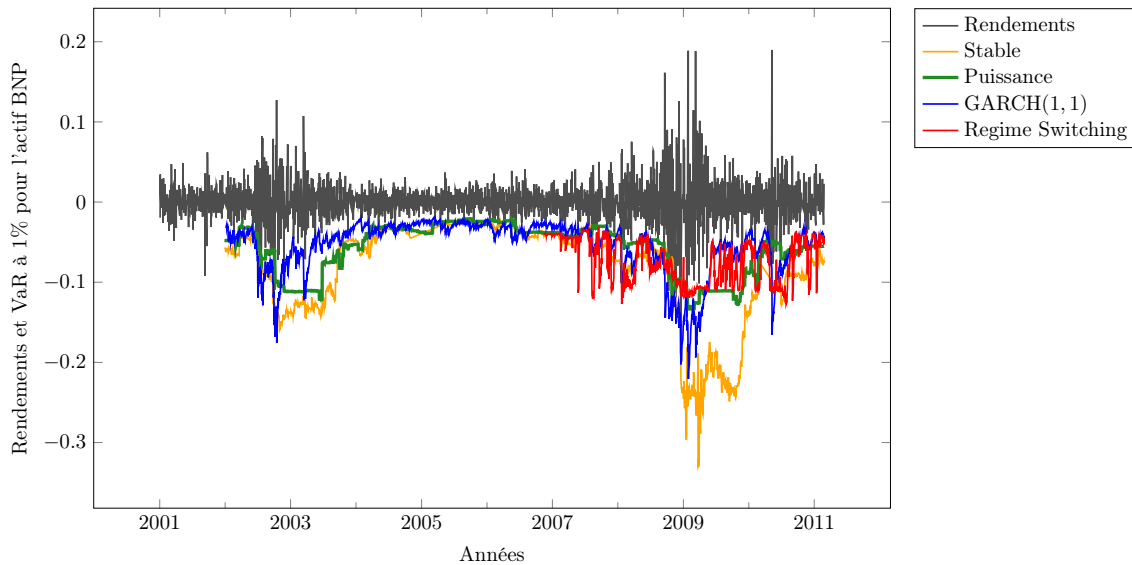


FIGURE 4.8 – Comparaison des estimations de la VaR à 99% sous les quatre modèles pour l’actif BNP Paribas. L’estimation sous les modèles des lois stables (jaune), des lois puissances (vert) et GARCH (bleu) est faite sur la période janvier 2001-février 2011, tandis que celle sous le modèle des lois puissances à changement de régimes (rouge) utilise uniquement la période novembre 2006-février 2011.

4.6 Conclusion

Dans ce chapitre, nous avons proposé un modèle pour la queue de distribution des pertes. Ce modèle, que nous avons appelé le modèle des lois puissances à changement de régimes, fournit une meilleure estimation des pertes extrêmes. L’introduction de la notion de crise nous a permis d’améliorer la performance de la prédiction de la Value-at-Risk en construisant un modèle qui prend en compte à la fois la lourdeur de la queue de la distribution et les changements de régimes. Ce modèle a en plus éliminé le regroupement des dépassements de la VaR et les délais d’estimation constatés dans les modèles i.i.d..

Stock	Stable			Puissance			GARCH(1, 1)			Regime Switch.		
	CI	UC	CC	CI	UC	CC	CI	UC	CC	CI	UC	CC
BEN	★	□	★	★	★	★	★	★	★	★	★	★
BH	★	□	□	★	★	★	★	★	★	★	★	★
BNA	□	□	□	★	□	□	★	★	★	★	★	★
BNP	★	★	★	□	□	□	★	★	★	★	★	★
BSD	□	□	□	★	★	★	★	★	★	★	★	★
CA	★	★	□	★	★	★	★	★	★	★	★	★
CAS	□	□	□	★	★	★	□	□	□	★	□	□
CCN	□	□	□	★	★	★	★	★	★	★	□	★
CDI	□	□	□	□	□	□	★	★	□	★	★	★
CGD	□	□	□	□	□	□	□	□	□	★	★	□
CMA	□	□	□	★	★	★	★	★	★	★	★	□
CNP	★	★	□	★	□	□	★	★	★	★	★	★
CS	★	★	★	★	★	★	★	★	★	★	★	★
DAN	□	□	□	★	★	★	★	★	★	★	★	□
DEC	★	□	★	★	★	★	□	□	□	★	□	□
DG	★	★	★	□	□	□	★	★	★	★	★	★
DLT	★	□	★	★	★	★	□	□	□	★	★	★
EC	□	□	□	★	★	★	★	★	★	★	★	★
EI	★	□	□	★	★	★	★	★	★	★	★	★
EN	★	★	★	★	□	★	★	★	★	★	★	★
FGR	□	□	□	★	□	□	★	★	★	★	★	★
FP	★	★	★	★	□	★	—	—	—	★	★	★
GA	★	□	□	□	□	□	□	□	★	★	★	★
GFC	★	□	□	□	□	□	□	□	□	□	□	□
GID	□	□	□	★	★	□	★	★	★	★	★	★
GLE	★	★	□	★	□	□	★	★	★	★	★	□
HAV	★	□	★	★	★	★	★	★	★	★	★	★
IGF	□	□	□	★	★	★	★	★	★	★	□	□
ITE	□	□	□	★	★	★	□	□	□	★	□	★
KEY	□	□	□	★	★	★	★	★	★	□	□	□
LG	★	★	★	□	□	□	★	★	★	★	★	□
LI	★	★	★	□	□	□	★	★	★	□	□	□
LVL	□	□	□	★	★	□	★	★	★	★	□	★
MC	★	★	★	★	★	★	★	★	★	★	★	★
MRB	□	□	□	★	★	□	★	★	★	★	★	□
MTU	□	□	□	★	□	★	★	★	★	★	★	★
NK	★	★	★	★	★	★	★	★	★	★	★	★
ORC	★	□	□	□	□	□	□	□	□	□	□	□
PIG	★	□	□	★	★	□	★	□	★	★	★	★
PP	★	★	★	★	★	★	★	★	★	★	★	★
PUB	★	□	□	★	★	★	★	★	★	★	★	★
RCF	□	□	□	□	□	□	★	★	★	★	★	★
RE	★	★	★	★	★	★	★	★	★	★	□	□
RI	★	★	★	★	★	★	★	★	★	★	★	★
SAN	★	★	★	★	★	★	—	—	—	★	★	★
SDT	□	□	□	□	□	□	★	★	★	★	★	★
SECH	★	□	□	★	★	★	★	★	★	★	□	□
SPI	★	★	□	★	★	□	★	★	★	★	★	★
SW	★	★	★	★	★	★	★	★	★	★	★	★
TEC	★	★	★	★	□	□	★	★	★	★	★	★
TRI	★	★	★	□	□	□	★	★	★	★	★	★
UG	★	★	★	★	□	□	★	★	★	★	□	□
UL	★	★	★	★	★	★	★	★	★	★	★	★
VAC	□	□	□	★	□	□	★	★	★	★	★	★
VIE	★	★	□	★	□	□	★	★	★	★	★	★
ZC	★	★	★	□	□	□	★	★	★	★	□	□
Total	36	24	23	43	32	30	46	45	46	49	41	39

★Ne pas rejeter, □Rejeter, — Non-convergence

TABLE 4.1 – Comparaison des résultats des backtests sur les hypothèses d'appartenance à l'intervalle de confiance du ratio de dépassement (CI), de couverture non-conditionnelle (UC) et de couverture conditionnelle (CC) sur la VaR à 99%.

Deuxième partie

Processus de Lévy en finance

Chapitre 5

Modèle exponentielle-Lévy et évaluation des options européennes

Ce chapitre a fait l'objet d'un survey soumis [143].

Plusieurs études ont cherché des alternatives aux modèles de type lois puissances, pour modéliser les distributions à queues épaisses. Il existe une vaste littérature sur ce sujet [131]. Du côté des processus stochastiques à temps continu, l'une des façons de considérer les événements extrêmes est d'intégrer des sauts dans la modélisations du prix de l'actif. Parmi les classes les plus connues des modèles à temps continu et dont les trajectoires peuvent avoir des discontinuités, nous proposons les modèles exponentielle-Lévy où le logarithme du prix est décrit par un processus de Lévy [158].

Nous commençons par donner la définition du processus de Lévy et rappeler ses principales propriétés dans la section 5.1. Puis, nous présentons le modèle exponentielle-Lévy dans la section 5.2, ainsi que la construction des mesures risque-neutres et l'évaluation numérique du prix de l'option par l'application de la transformée de Fourier rapide. Nous finissons le chapitre par la section 5.3, où nous discutons trois modèles exponentielle-Lévy : le modèle de Black et Scholes, le modèle diffusion-sauts de Merton et le modèle variance gamma.

5.1 Processus de Lévy

Un processus de Lévy est un processus stochastique à temps continu, continu à droite limité à gauche (càdlàg), partant de 0, dont les accroissements sont stationnaires et indépendants. Les exemples les plus connus sont le processus de Wiener et le processus de Poisson.

Définition 5.1. Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité. Un processus stochastique $(X_t)_{t \geq 0}$ càdlàg tel que $X_0 = 0$ est dit un processus de Lévy s'il possède les propriétés suivantes :

1. *Accroissements indépendants* : pour toute suite croissante $t_0 < \dots < t_n$, les variables aléatoires $X_{t_0}, X_{t_1} - X_{t_0}, \dots, X_{t_n} - X_{t_{n-1}}$ sont indépendantes.
2. *Accroissements stationnaires* : pour tout $h > 0$, la loi de $X_{t+h} - X_t$ ne dépend pas de t .
3. *Continuité stochastique* : pour tout $\varepsilon > 0$, $\lim_{h \rightarrow 0} \mathbb{P}(|X_{t+h} - X_t| \geq \varepsilon) = 0$.

La loi d'un processus de Lévy est complètement identifiée par son triplet caractéristique (b, σ^2, ν) , où $b \in \mathbb{R}$, $\sigma > 0$ et ν est une mesure de Lévy définie comme étant une mesure positive sur $\mathbb{R} \setminus \{0\}$ qui satisfait

$$\int_{\mathbb{R} \setminus \{0\}} 1 \wedge |x|^2 \nu(dx) < \infty.$$

En particulier, la fonction caractéristique de X_t peut être calculée à partir du triplet caractéristique comme suit.

Théorème 5.1 (Représentation de Lévy-Khintchine). *Soit $(X_t)_{t \geq 0}$ un processus de Lévy en $\mathbb{R}$ de triplet caractéristique (b, σ^2, ν) . Alors la fonction caractéristique de X_t s'écrit*

$$\Phi_t(u) := \mathbb{E}[e^{iuX_t}] = e^{t\Psi(u)}, \quad u \in \mathbb{R} \quad (5.1)$$

où Ψ , appelé l'exposant caractéristique, est donné par

$$\Psi(u) = ibu - \frac{1}{2}\sigma^2 u^2 + \int_{\mathbb{R}} (e^{iux} - 1 - iux \mathbf{1}_{|x| \leq 1}) \nu(dx). \quad (5.2)$$

La mesure de Lévy ν détermine l'intensité des sauts de différentes tailles. Ainsi, $\nu([a_1, a_2])$ est le nombre moyen de sauts, dans l'intervalle de temps $[0, 1]$, de taille appartenant à $[a_1, a_2]$. Cette quantité est définie par

$$\nu([a_1, a_2]) = \mathbb{E} \left[\sum_{t \in (0, 1]} \mathbf{1}_{\Delta X_t \in [a_1, a_2]} \right],$$

où $\Delta X_t = X_t - X_{t-}$.

Notons que $\nu([a_1, a_2])$ est finie pour tout compact $[a_1, a_2]$ qui ne contient pas 0. Si ce n'est pas le cas, le processus $(X_t)_{t \geq 0}$ doit avoir une infinité de sauts de taille finie dans tout intervalle de temps $[0, t]$, ce qui contredit la propriété de càdlàg de $(X_t)_{t \geq 0}$. Ainsi ν définit une mesure de Radon sur $\mathbb{R} \setminus \{0\}$. Cependant, ν n'est pas nécessairement une mesure finie. En effet, la restriction précédente permet toujours une explosion en 0 et $(X_t)_{t \geq 0}$ peut avoir une infinité de petits sauts en $[0, t]$. Dans un tel cas, la somme des sauts devient une série infinie et sa convergence nécessite des conditions supplémentaires sur ν .

Si la mesure de Lévy satisfait la condition $\int_{|x| \leq 1} |x| \nu(dx) < \infty$, le processus des sauts, défini par

$$X_t^J = \sum_{\substack{s \in (0, t] \\ \Delta X_s \neq 0}} \Delta X_s,$$

est un processus à variation finie. Dans un tel cas, le processus $(X_t)_{t \geq 0}$ admet la décomposition suivante

$$X_t = \gamma t + \sigma B_t + X_t^J,$$

où $\gamma = b - \int_{|x| \leq 1} x \nu(dx)$ et $(B_t)_{t \geq 0}$ est un mouvement brownien standard. L'exposant caractéristique est donné alors par

$$\Psi(u) = i\gamma u - \frac{1}{2}\sigma^2 u^2 + \int_{\mathbb{R}} (e^{iux} - 1) \nu(dx).$$

On note que le triplet de $(X_t)_{t \geq 0}$ n'est pas donné par (γ, σ^2, ν) , mais plutôt par (b, σ^2, ν) . En effet, b n'est pas une quantité intrinsèque et dépend de la fonction de troncature $\mathbf{1}_{|x| \leq 1}$ appliquée dans la représentation de Lévy-Khintchine. Cependant, γ admet une interprétation intrinsèque comme étant la pente de l'espérance de la partie continue de $(X_t)_{t \geq 0}$. Dans ce cas, l'espérance $\mathbb{E}[X_t]$ est donnée par $(\gamma + \int_{\mathbb{R}} x \nu(dx))t$.

Proposition 5.1 (Moments et cumulants d'un processus de Lévy).

1. Soit $(X_t)_{t \geq 0}$ un processus de Lévy dans $\mathbb{R}$ de triplet caractéristique (b, σ^2, ν) et soit $n \geq 1$. Alors $\mathbb{E}[|X_t|^n] < \infty$ si et seulement si $\int_{|x| \geq 1} |x| \nu(dx) < \infty$. Dans ce cas, $\Phi_t(u)$, la fonction caractéristique de X_t , est de classe C^n et les n premiers moments de X_t peuvent être calculés par différentiation :

$$\mathbb{E}[X_t^k] = \frac{1}{i^k} \frac{\partial^k}{\partial u^k} \Phi_t(u)|_{u=0}, \quad k = 1, \dots, n.$$

Les cumulants de X_t , définis par

$$c_k(X_t) := \frac{1}{i^k} \frac{\partial^k}{\partial u^k} \log \Phi_t(u)|_{u=0},$$

admettent en particulier des expressions simples :

$$c_1(X_t) \equiv \mathbb{E}[X_t] = t \left(b + \int_{|x| \geq 1} x \nu(dx) \right),$$

$$c_2(X_t) \equiv \text{Var } X_t = t \left(\sigma^2 + \int_{\mathbb{R}} x^2 \nu(dx) \right),$$

$$c_k(X_t) = t \int_{\mathbb{R}} x^k \nu(dx), \quad \text{pour } 3 \leq k \leq n.$$

2. Soit $(X_t)_{t \geq 0}$ un processus de Lévy sur $\mathbb{R}$ de triplet caractéristique (b, σ^2, ν) et soit $u \in \mathbb{R}$. $\mathbb{E}[e^{u X_t}] < \infty$ pour un certain $t > 0$, ou de manière équivalente pour tout $t > 0$, si et seulement si $\int_{|x| \geq 1} e^{ux} \nu(dx) < \infty$. Dans ce cas,

$$\mathbb{E}[e^{u X_t}] = e^{t\Psi(-iu)},$$

où Ψ est l'exposant caractéristique du processus de Lévy défini par (5.2).

Corollaire 5.1. Soit $(X_t)_{t \geq 0}$ un processus de Lévy dans $\mathbb{R}$ de triplet caractéristique (b, σ^2, ν) . Alors,

1. $(X_t)_{t \geq 0}$ est une martingale si et seulement si $\int_{|x| \geq 1} |x| \nu(dx) < \infty$ et

$$b + \int_{|x| \geq 1} x \nu(dx) = 0. \quad (5.3)$$

2. $(\exp(X_t))_{t \geq 0}$ est une martingale si et seulement si $\int_{|x| \geq 1} e^x \nu(dx) < \infty$ et

$$b + \frac{1}{2} \sigma^2 + \int_{\mathbb{R}} (e^x - 1 - x \mathbf{1}_{|x| \leq 1}) \nu(dx) = 0. \quad (5.4)$$

Pour les détails de la preuve, voir [145, théorèmes 25.3 et 25.17].

5.1.1 Exponentielle de Doléans-Dade d'un processus de Lévy

L'exponentielle de Doléans-Dade, connue aussi sous le nom d'exponentielle stochastique, d'une semimartingale $(Y_t)_{t \geq 0}$ est définie comme étant l'unique processus càdlàg $(\mathcal{E}(Y)_t)_{t \geq 0}$ tel que

$$d\mathcal{E}(Y)_t = \mathcal{E}(Y)_{t-} dY_t, \quad \mathcal{E}(Y)_0 = 1. \quad (5.5)$$

Le processus $\mathcal{E}(Y)$ est donné par :

$$\mathcal{E}(Y)_t = e^{Y_t - Y_0 - \frac{1}{2} \langle Y^c \rangle_t} \prod_{s \in (0, t]} (1 + \Delta Y_s) e^{-\Delta Y_s}, \quad (5.6)$$

où $(Y_t^c)_{t \geq 0}$ désigne la partie continue de Y . Le résultat suivant de Goll et Kallsen [78] décrit la relation entre l'exponentielle de Doléans-Dade et l'exponentielle ordinaire lorsque Y est un processus de Lévy.

Proposition 5.2 (Relation entre l'exponentielle de Doléans-Dade et l'exponentielle ordinaire).

1. Soit $(X_t)_{t \geq 0}$ un processus de Lévy dans $\mathbb{R}$ de triplet caractéristique (b, σ^2, ν) et $S_t = e^{X_t}$ son exponentielle. Il existe alors un processus de Lévy $(\hat{X}_t)_{t \geq 0}$ tel que S_t est l'exponentielle de Doléans-Dade de $\hat{X}$: $S = \mathcal{E}(\hat{X})$ avec

$$\hat{X}_t := X_t + \frac{\sigma^2 t}{2} + \sum_{s \in (0, t]} \{e^{\Delta X_s} - 1 - \Delta X_s\}. \quad (5.7)$$

Le triplet caractéristique $(\hat{b}, \hat{\sigma}^2, \hat{\nu})$ de $\hat{X}$ est donné par :

$$\hat{b} = b + \frac{1}{2} \sigma^2 + \int_{|x| \leq 1} x \hat{\nu}(dx) - \int_{|x| \leq 1} x \nu(dx),$$

$$\hat{\sigma}^2 = \sigma^2,$$

$$\hat{\nu}(dx) = \nu \circ J^{-1}(dx) \text{ où } J(x) := e^x - 1 \text{ pour } x \in \mathbb{R}.$$

2. Soit $(\hat{X}_t)_{t \geq 0}$ un processus de Lévy dans $\mathbb{R}$ de triplet caractéristique $(\hat{b}, \hat{\sigma}^2, \hat{\nu})$ et $S = \mathcal{E}(\hat{X})$ son exponentielle de Doléans-Dade. Si $S > 0$ p.s. alors il existe un autre processus de Lévy $(X_t)_{t \geq 0}$ tel que pour tout t , $S_t = e^{X_t}$, avec

$$X_t = \log S_t = \hat{X}_t - \frac{\hat{\sigma}^2 t}{2} + \sum_{s \in (0, t]} \{\log(1 + \Delta \hat{X}_s) - \Delta \hat{X}_s\}. \quad (5.8)$$

Le triplet caractéristique (b, σ^2, ν) de X est donné par :

$$b = \hat{b} - \frac{1}{2}\sigma^2 + \int_{|x| \leq 1} x \nu(dx) - \int_{|x| \leq 1} x \hat{\nu}(dx),$$

$$\sigma^2 = \hat{\sigma}^2,$$

$$\nu(dx) = \hat{\nu} \circ H^{-1}(dx) \text{ où } H(x) := \log(1 + x) \text{ pour } x > -1.$$

Remarque 6.1.

- (i) D'après (5.6), $S = \mathcal{E}(\hat{X}) > 0$ p.s. si et seulement si $\text{supp}\{\hat{\nu}\} \subset (-1, \infty)$.
- (ii) ν admet une densité $\nu(x)$ si et seulement si $\hat{\nu}$ admet une densité $\hat{\nu}(x)$. Dans ce cas, nous avons

$$\hat{\nu}(x) = \frac{1}{1+x} \nu(\log(1+x)) \quad \text{pour } x > -1$$

et

$$\nu(x) = e^x \hat{\nu}(e^x - 1) \quad \text{pour } x \in \mathbb{R}.$$

5.1.2 Changement de mesure de probabilité

Le changement de probabilité joue un rôle important dans la construction d'un modèle d'évaluation des options, sans aucune opportunité d'arbitrage [39]. Ce modèle est dit risque-neutre et il en existe au moins un dès l'existence d'une mesure de probabilité équivalente sous laquelle le prix actualisé de l'actif sous-jacent est une martingale [52]. Dans cette section, nous présentons le résultat général d'équivalence des mesures de probabilité lorsque l'aléa est donné par un processus de Lévy.

Proposition 5.3 (voir [145], théorème 33.1 et théorème 33.2). *Soient $(X_t, \mathbb{P})$ et $(X_t, \mathbb{P}^*)$ deux processus de Lévy dans $\mathbb{R}$ de triplets caractéristiques respectifs (b, σ^2, ν) et $(b^*, \sigma^{*2}, \nu^*)$. Les probabilités $\mathbb{P}|_{\mathcal{F}_t}$ et $\mathbb{P}^*|_{\mathcal{F}_t}$ sont équivalentes si et seulement si les trois conditions suivantes sont satisfaites.*

- 1. $\sigma = \sigma^*$.
- 2. Les mesures de Lévy sont équivalentes avec

$$\int_{\mathbb{R}} (e^{\phi(x)/2} - 1)^2 \nu(dx) < \infty,$$

$$\text{où } \phi(x) = \log\left(\frac{d\nu^*}{d\nu}\right).$$

3. Si $\sigma = 0$ alors nous devons en plus avoir

$$b - \int_{|x| \leq 1} x\nu(dx) = b^* - \int_{|x| \leq 1} x\nu^*(dx).$$

De plus, lorsque $\mathbb{P}|_{\mathcal{F}_t}$ et $\mathbb{P}^*|_{\mathcal{F}_t}$ sont équivalentes, la dérivée de Radon-Nikodym est donnée par

$$\left. \frac{d\mathbb{P}^*}{d\mathbb{P}} \right|_{\mathcal{F}_t} = e^{U_t} \quad (5.9)$$

avec

$$U_t = \theta X_t^c - \frac{\theta^2 \sigma^2 t}{2} - \theta b t + \lim_{\varepsilon \rightarrow 0^+} \left(\sum_{\substack{s \in (0, t] \\ |\Delta X_s| > \varepsilon}} \phi(\Delta X_s) - t \int_{|x| > \varepsilon} (e^{\phi(x)} - 1) \nu(dx) \right). \quad (5.10)$$

Ici (X_t^c) note la partie continue de (X_t) et θ est une constante telle que

$$\left(b^* - \int_{|x| \leq 1} x\nu^*(dx) \right) - \left(b - \int_{|x| \leq 1} x\nu(dx) \right) = \sigma^2 \theta$$

si $\sigma > 0$ et égale à 0 si $\sigma = 0$.

Le processus $(U_t)_{t \geq 0}$ est un processus de Lévy de triplet caractéristique (b_U, σ_U^2, ν_U) donné par

$$\begin{aligned} b_U &= -\frac{\sigma^2 \theta^2}{2} - \int_{\mathbb{R}} (e^y - 1 - y \mathbf{1}_{|y| \leq 1}) (\nu \circ \phi^{-1})(dy), \\ \sigma_U^2 &= \sigma^2 \theta^2, \\ \nu_U &= \nu \circ \phi^{-1}. \end{aligned}$$

Le résultat ci-dessus montre une caractéristique intéressante des modèles à sauts par rapport aux modèles de diffusion : nous avons une grande liberté dans le changement de la mesure de Lévy, tout en conservant l'équivalence des mesures de probabilité.

5.2 Modèle exponentielle-Lévy

Les modèles exponentielle-Lévy sont obtenus en remplaçant le mouvement brownien avec une dérive dans le modèle classique de Black et Scholes, par un processus de Lévy.

Définition 5.2. Soit $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0, T]}, \mathbb{P})$ un espace de probabilité filtré qui satisfait les conditions usuelles. Sous le modèle exponentielle-Lévy (exp-Lévy), le prix de l'actif est décrit par un processus de la forme :

$$S_t = S_0 e^{X_t}, \quad (5.11)$$

où $S_0 > 0$ et $(X_t)_{t \in [0, T]}$ est un processus de Lévy dans $\mathbb{R}$ de triplet caractéristique (b, σ^2, ν) . Le prix actualisé est donné par $\tilde{S}_t = e^{-rt} S_t$ où r est le taux d'intérêt sans risque.

Dans ce qui suit, les notations X , S et $\tilde{S}$ font toujours référence au modèle exp-Lévy introduit dans la définition précédente.

Nous considérons une option d'achat européenne de maturité T et de strike K . Le prix d'expiration de cette option est donné par une variable aléatoire $H = (S_T - K)_+$. Soit $\mathcal{P}$ l'ensemble de toutes les mesures risque-neutres

$$\mathcal{P} = \left\{ \mathbb{P}^* \sim \mathbb{P}, \quad (\tilde{S}_t)_{t \in [0, T]} \text{ est une martingale sous } \mathbb{P}^* \right\}.$$

Dans un marché complet, il existe une unique mesure risque-neutre $\mathbb{P}^*$. Ainsi, le prix risque-neutre de l'option est donné par

$$C(K) = e^{-rT} \mathbb{E}^*[(S_T - K)_+], \quad (5.12)$$

où $\mathbb{E}^*$ est l'espérance sous $\mathbb{P}^*$. Avec les modèles exp-Lévy, nous sommes généralement dans le cas du marché incomplet. Par conséquent, plusieurs mesures risque-neutres peuvent être utilisées pour évaluer l'option. L'intervalle de prix de l'option est donné par

$$\left[\inf_{\mathbb{P}^* \in \mathcal{P}} e^{-rT} \mathbb{E}^*[(S_T - K)_+], \sup_{\mathbb{P}^* \in \mathcal{P}} e^{-rT} \mathbb{E}^*[(S_T - K)_+] \right].$$

Dans la pratique, nous choisissons une mesure $\mathbb{P}^* \in \mathcal{P}$ selon un critère de choix et nous évaluons l'option en utilisant la formule (5.12).

Une autre difficulté qui se manifeste pour les modèles exp-Lévy est l'absence d'une expression explicite pour la densité du processus de Lévy $(X_t)_{t \in [0, T]}$, dans la plupart des cas. Ainsi, trouver une formule fermée pour $C(K)$ s'avère difficile, voire impossible pour certaines mesures risque-neutres et certains processus de Lévy. Cependant, la fonction caractéristique de $(X_t)_{t \in [0, T]}$ admet souvent une expression explicite. Dans un tel cas, nous pouvons appliquer la transformée de Fourier rapide (FFT) pour évaluer numériquement le prix de l'option [32].

Supposons que la fonction caractéristique Φ_t de $(X_t)_{t \in [0, T]}$ sous $\mathbb{P}$ est analytiquement connue. L'évaluation du prix de l'option se fait en deux étapes :

1. Choisir une mesure risque-neutre $\mathbb{P}^* \in \mathcal{P}$ sous laquelle nous avons une expression explicite de la fonction caractéristique, notée Φ_t^* .
2. Appliquer la transformée de Fourier rapide à Φ_T^* pour évaluer le prix de l'option.

5.2.1 Mesures risque-neutres

La mesure risque-neutre représente un outil puissant pour l'évaluation des options. Plusieurs candidats pour une mesure risque-neutre sont proposés dans la littérature. Pour les construire, deux approches différentes sont appliquées :

- **Méthode de transformée d'Esscher** : La méthode de transformée d'Esscher est largement utilisée dans la théorie de risque. Elle consiste à appliquer cette transformée par rapport à un processus de risque. Le processus de risque peut être les rendements logarithmique $(X_t)_{t \in [0, T]}$ dans le cas de la *mesure risque-neutre d'Esscher* [29, 75], les rendements arithmétiques $(\hat{X}_t)_{t \in [0, T]}$ dans le cas de la *mesure risque-neutre d'entropie minimale* [70, 71, 122] ou la partie continue $(X_t^c)_{t \in [0, T]}$ du processus de Lévy $(X_t)_{t \in [0, T]}$ dans le cas de la *mesure risque-neutre de moyenne corrigée* [167].
- **Méthode de distance minimale** : Cette méthode est plus reliée au contexte de la maximisation de l'utilité et de la couverture d'options. Elle consiste à trouver la mesure $\mathbb{P}^* \in \mathcal{P}$ qui minimise une certaine distance par rapport à la mesure de probabilité historique $\mathbb{P}$. Cela inclut la *mesure risque-neutre minimale* [68] et la *mesure risque-neutre de variance optimale* [149]. La *mesure risque-neutre d'entropie minimale* peut aussi être retrouvée en utilisant cette méthode.

Mesure risque-neutre d'Esscher

La mesure risque-neutre d'Esscher est construite en appliquant la transformée d'Esscher par rapport au processus $(X_t)_{t \in [0, T]}$.

Définition 5.3. Soit $(X_t)_{t \in [0, T]}$ un processus de Lévy dans $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0, T]}, \mathbb{P})$. Nous appelons transformée d'Esscher, par rapport à $(X_t)_{t \in [0, T]}$, tout changement de $\mathbb{P}$ à une mesure équivalente $\mathbb{P}^*$, par un processus de densité Z_t , de la forme :

$$Z_t \equiv \frac{d\mathbb{P}^*}{d\mathbb{P}} \Big|_{\mathcal{F}_t} := \frac{e^{\theta X_t}}{\mathbb{E}[e^{\theta X_t}]}, \quad (5.13)$$

où $\theta \in \mathbb{R}$.

L'un des avantages majeurs d'adopter $(X_t)_{t \in [0, T]}$ comme processus de risque est le fait que le processus de densité $(Z_t)_{t \in [0, T]}$ dépend uniquement du prix actuel de l'actif. D'autre part, $(X_t)_{t \in [0, T]}$ est encore un processus de Lévy sous $\mathbb{P}^*$. Dans la suite, nous présentons la condition d'existence du processus de densité $Z_t = \frac{d\mathbb{P}^*}{d\mathbb{P}} \Big|_{\mathcal{F}_t}$ et nous donnons le triplet caractéristique de $(X_t)_{t \in [0, T]}$ sous $\mathbb{P}^*$, dans un tel cas.

Proposition 5.4. Soit $(X_t)_{t \in [0, T]}$ un processus de Lévy, à valeurs dans $\mathbb{R}$, de triplet caractéristique (b, σ^2, ν) . Pour tout $\theta \in \mathbb{R}$ tel que $\mathbb{E}[e^{\theta X_1}] < \infty$:

- (i) Le processus $(Z_t)_{t \in [0, T]}$, donné par (5.13), définit un processus de densité.
- (ii) Le processus $(X_t)_{t \in [0, T]}$ est un processus de Lévy sous $\mathbb{P}^*$, de triplet (b^*, σ^2, ν^*) où

$$\nu^*(dx) = e^{\theta x} \nu(dx), \quad \text{pour } x \in \mathbb{R},$$

et

$$b^* = b + \sigma^2 \theta + \int_{|x| \leq 1} x \nu^*(dx) - \int_{|x| \leq 1} x \nu(dx).$$

Preuve. (i) Rappelons que $\mathbb{E}[e^{\theta X_t}] = e^{t\Psi(-i\theta)} = (\mathbb{E}[e^{\theta X_1}])^t$. Ainsi, Z_t est intégrable pour tout t . En utilisant l'indépendance et la stationnarité des accroissements du processus de Lévy $(X_t)_{t \in [0, T]}$, nous obtenons pour tout $s < t$,

$$\begin{aligned}\mathbb{E}[Z_t | \mathcal{F}_s] &= \frac{1}{\mathbb{E}[e^{\theta X_t}]} \mathbb{E}[e^{\theta(X_t - X_s + X_s)} | \mathcal{F}_s] = \frac{1}{\mathbb{E}[e^{\theta X_t}]} \mathbb{E}[e^{\theta(X_t - X_s)} | \mathcal{F}_s] \mathbb{E}[e^{\theta X_s} | \mathcal{F}_s] \\ &= \frac{1}{\mathbb{E}[e^{\theta X_t}]} \mathbb{E}[e^{\theta(X_t - X_s)}] e^{\theta X_s} = \frac{e^{\theta X_s}}{\mathbb{E}[e^{\theta X_s}]} = Z_s.\end{aligned}$$

Ainsi, $(Z_t)_{t \in [0, T]}$ est une $\mathbb{P}$ -martingale.

(ii) Nous montrons que $(X_t)_{t \in [0, T]}$ est un processus de Lévy sous $\mathbb{P}^*$ en calculant sa fonction caractéristique sous $\mathbb{P}^*$:

$$\begin{aligned}\Phi^*(u) &= \mathbb{E}^*[e^{iuX_t}] = \int e^{iuX_t} d\mathbb{P}^* = \int e^{iuX_t} \frac{e^{\theta X_t}}{\mathbb{E}[e^{\theta X_t}]} d\mathbb{P} \\ &= \frac{\mathbb{E}[e^{(\theta + iu)X_t}]}{\mathbb{E}[e^{\theta X_t}]} = \exp(t(\Psi(-i(\theta + iu)) - \Psi(-i\theta))).\end{aligned}$$

Nous définissons $\Psi^*(u) \equiv \Psi(-i(\theta + iu)) - \Psi(-i\theta)$ pour $u \in \mathbb{R}$. Alors,

$$\begin{aligned}\Psi^*(u) &= \left(b(\theta + iu) + \frac{\sigma^2}{2}(\theta + iu)^2 + \int_{\mathbb{R}} (e^{(\theta + iu)x} - 1 - (\theta + iu)x \mathbf{1}_{|x| \leq 1}) \nu(dx) \right) \\ &\quad - \left(b\theta + \frac{\sigma^2}{2}\theta^2 + \int_{\mathbb{R}} (e^{\theta x} - 1 - \theta x \mathbf{1}_{|x| \leq 1}) \nu(dx) \right) \\ &= i(b + \sigma^2\theta)u - \frac{\sigma^2}{2}u^2 + \int_{\mathbb{R}} (e^{\theta x}(e^{iux} - 1) - iux \mathbf{1}_{|x| \leq 1}) \nu(dx) \\ &= i \left(b + \sigma^2\theta + \int_{|x| \leq 1} (e^{\theta x} - 1)x \nu(dx) \right) u - \frac{\sigma^2}{2}u^2 \\ &\quad + \int_{\mathbb{R}} (e^{\theta x}(e^{iux} - 1 - iux \mathbf{1}_{|x| \leq 1})) \nu(dx).\end{aligned}$$

En définissant

$$\nu^*(x) = e^{\theta x} \nu(x), \quad \text{pour } x \in \mathbb{R}$$

et

$$b^* = b + \sigma^2\theta + \int_{|x| \leq 1} x \nu^*(dx) - \int_{|x| \leq 1} x \nu(dx),$$

nous obtenons la représentation de Lévy-Khintchine pour Φ^* . Ainsi, $(X_t)_{t \in [0, T]}$ est un processus de Lévy sous $\mathbb{P}^*$, de triplet (b^*, σ^2, ν^*) . $\square$

Afin d'interpréter l'expression de b^* , rappelons que la mesure de Lévy ν peut avoir une singularité en 0. Ainsi, nous pouvons avoir une infinité de petits sauts, et la fonction caractéristique de leur somme $\int_{|x| \leq 1} (e^{iux} - 1) \nu(dx)$ ne converge pas nécessairement. Pour obtenir la convergence, cette intégrale de sauts doit être centrée et remplacée dans la représentation de Lévy-Khintchine par sa version compensée. Ce compensateur $\int_{|x| \leq 1} x \nu(dx)$ est par conséquent intégré dans la dérive. Pour passer de la mesure de probabilité $\mathbb{P}$ à $\mathbb{P}^*$, nous devons naturellement soustraire le compensateur associé à ν de la dérive et le remplacer par celui de ν^* .

Théorème 5.2 ([132]). Soit $(X_t)_{t \in [0, T]}$ un processus de Lévy sous $\mathbb{P}$ de triplet (b, σ^2, ν) . Supposons que X_1 est non-dégénéré et admet une fonction génératrice des moments, $u \mapsto \mathbb{E}[\exp(uX_1)]$, sur un intervalle ouvert (a_1, a_2) , avec $a_2 - a_1 > 1$. Supposons qu'il existe un réel $\theta \in (a_1, a_2 - 1)$ tel que

$$b + \sigma^2\theta + \frac{\sigma^2}{2} + \int_{\mathbb{R}} (e^{\theta x}(e^x - 1) - x\mathbf{1}_{|x| \leq 1}) \nu(dx) = r, \quad (5.14)$$

ou, d'une manière équivalente

$$\mathbb{E}[e^{(\theta+1)X_t}] = e^{rt}\mathbb{E}[e^{\theta X_t}], \quad (5.15)$$

où r est le taux d'intérêt sans risque. Alors θ est unique.

De plus, la mesure équivalente $\mathbb{P}^*$ donnée par la transformée d'Esscher avec $(X_t)_{t \in [0, T]}$

$$\left. \frac{d\mathbb{P}^*}{d\mathbb{P}} \right|_{\mathcal{F}_t} = \frac{e^{\theta X_t}}{\mathbb{E}[e^{\theta X_t}]},$$

est une mesure risque-neutre. Cette mesure est appelée la mesure risque-neutre d'Esscher.

Preuve. La proposition 5.4 garantit que $(X_t)_{t \in [0, T]}$ est un processus de Lévy sous toute mesure $\mathbb{P}^*$ construite via une transformée d'Esscher. Par indépendance et stationnarité des accroissements de $(X_t)_{t \in [0, T]}$, la propriété de martingale de $(S_t)_{t \in [0, T]}$ sous $\mathbb{P}^*$ est équivalente à $\mathbb{E}^*[\tilde{S}_t] = S_0$ pour $t > 0$. Par construction de la mesure via une transformée d'Esscher, nous avons

$$\begin{aligned} \mathbb{E}^*[\tilde{S}_t] &= \mathbb{E}^*[S_0 e^{X_t - rt}] = \mathbb{E}[S_0 e^{X_t - rt} e^{\theta X_t} (\mathbb{E}[e^{\theta X_t}])^{-1}] \\ &= S_0 e^{-rt} \frac{\mathbb{E}[e^{(\theta+1)X_t}]}{\mathbb{E}[e^{\theta X_t}]}. \end{aligned}$$

Ainsi, $\mathbb{E}^*[\tilde{S}_t] = S_0$ si et seulement s'il existe un réel θ tel que (5.15) est satisfaite. Ce réel θ doit être dans $(a_1, a_2 - 1)$ pour assurer l'existence de la fonction génératrice des moments de θ et $\theta + 1$.

Nous appliquons la proposition 5.1-2 pour réécrire (5.15) en termes de l'exposant caractéristique sous $\mathbb{P}$

$$\Psi(-i(\theta + 1)) - \Psi(-i\theta) = r.$$

Ensuite, en développant l'expression de Ψ (équation (5.2)), nous obtenons la condition (5.14). Ainsi, s'il existe θ qui vérifie cette condition, le prix actualisé $(\tilde{S}_t)_{t \in [0, T]}$ est une martingale sous la mesure construite via une transformée d'Esscher avec $(X_t)_{t \in [0, T]}$ et θ .

D'autre part, l'unicité de θ découle du fait que le logarithme de la fonction génératrice des moments $u \mapsto \ln \mathbb{E}[e^{uX_1}]$ est strictement convexe pour toute variable aléatoire X_1 non-dégénérée [132]. Ceci peut être démontré par un raffinement de l'argument de [23, section 9], où il montre seulement la convexité. $\square$

Mesure risque-neutre d'entropie minimale (MEMM)

La mesure risque-neutre d'entropie minimale a été étudiée dans différents contextes par plusieurs auteurs [60, 68, 70, 122, 149]. En particulier, la MEMM pour les modèles exp-Lévy a été discutée dans [34, 71, 79, 123]. Il se trouve que cette mesure peut être obtenue par une transformée d'Esscher en utilisant les rendements arithmétiques $(\hat{X}_t)_{t \in [0, T]}$ comme processus de risque. De plus, $(X_t)_{t \in [0, T]}$ est encore un processus de Lévy sous cette mesure. Dans cette section, nous rappelons la définition de l'entropie relative et nous donnons la condition que le paramètre d'Esscher doit satisfaire pour l'existence de la MEMM, ainsi que le triplet caractéristique de $(X_t)_{t \in [0, T]}$ sous cette mesure.

Définition 5.4. Soit $\mathcal{G}$ une sous-tribu de $\mathcal{F}$ et $\mathbb{Q}$ une mesure de probabilité sous $\mathcal{G}$. L'entropie relative sur $\mathcal{G}$ de $\mathbb{Q}$ par rapport à $\mathbb{P}$ est définie par

$$\mathbb{H}_{\mathcal{G}}(\mathbb{Q}|\mathbb{P}) := \begin{cases} \int \log \left(\frac{d\mathbb{Q}}{d\mathbb{P}} \Big|_{\mathcal{G}} \right) d\mathbb{Q}, & \text{si } \mathbb{Q} \ll \mathbb{P} \text{ sur } \mathcal{G}, \\ +\infty, & \text{sinon,} \end{cases} \quad (5.16)$$

où $\frac{d\mathbb{Q}}{d\mathbb{P}} \Big|_{\mathcal{G}}$ est la dérivée de Radon-Nikodym de $\mathbb{Q}|_{\mathcal{G}}$ par rapport à $\mathbb{P}|_{\mathcal{G}}$.

Théorème 5.3 ([71]). Soit $(X_t)_{t \in [0, T]}$ un processus de Lévy sous $\mathbb{P}$ de triplet (b, σ^2, ν) . Supposons qu'il existe $\beta \in \mathbb{R}$ tel que

$$\int_{x>1} e^x e^{\beta(e^x-1)} \nu(dx) < \infty \quad (5.17)$$

et

$$b + \sigma^2 \beta + \frac{\sigma^2}{2} + \int_{\mathbb{R}} ((e^x - 1)e^{\beta(e^x-1)} - x \mathbf{1}_{|x| \leq 1}) \nu(dx) = r, \quad (5.18)$$

où r est le taux d'intérêt sans risque. Alors,

1. Le réel β est unique et la mesure équivalente $\mathbb{P}^\dagger$ construite par la transformée d'Esscher par rapport à $(\hat{X}_t)_{t \in [0, T]}$,

$$\frac{d\mathbb{P}^\dagger}{d\mathbb{P}} \Big|_{\mathcal{F}_t} = \frac{e^{\beta \hat{X}_t}}{\mathbb{E} \left[e^{\beta \hat{X}_t} \right]}$$

est une mesure risque-neutre, où $(\hat{X}_t)_{t \in [0, T]}$ est donné par (5.7).

2. Le processus $(X_t)_{t \in [0, T]}$ est encore un processus de Lévy sous $\mathbb{P}^\dagger$. Son triplet caractéristique sous $\mathbb{P}^\dagger$ est $(b^\dagger, \sigma^2, \nu^\dagger)$ avec

$$\nu^\dagger(dx) = e^{\beta(e^x-1)} \nu(dx)$$

et

$$b^\dagger = b + \sigma^2 \beta + \int_{|x| \leq 1} x \nu^\dagger(dx) - \int_{|x| \leq 1} x \nu(dx).$$

3. La mesure de probabilité $\mathbb{P}^\dagger$ atteint l'entropie minimale dans $\mathcal{P}$

$$\mathbb{H}_{\mathcal{F}_T}(\mathbb{P}^\dagger|\mathbb{P}) = \min_{\mathbb{Q} \in \mathcal{P}} \mathbb{H}_{\mathcal{F}_T}(\mathbb{Q}|\mathbb{P}).$$

Nous appelons $\mathbb{P}^\dagger$ la mesure risque-neutre d'entropie minimale.

Ainsi, la mesure risque-neutre d'entropie minimale peut être simplement construite à partir d'une transformée d'Esscher par rapport aux rendements arithmétiques $(\hat{X}_t)_{t \in [0, T]}$. Bien que nous disposons du triplet caractéristique de $(X_t)_{t \in [0, T]}$ sous $\mathbb{P}^\dagger$, l'expression analytique de la fonction caractéristique sous cette mesure risque-neutre reste difficile à expliciter, même pour des modèles exponentielle-Lévy simples. Ainsi, l'évaluation d'options avec la transformée de Fourier rapide n'est pas possible sous cette mesure.

5.2.2 Évaluation d'options via la fonction caractéristique

Nous considérons le problème d'évaluation d'une option d'achat de maturité T et de strike K . Plusieurs techniques ont été appliquées pour répondre à cette question.

Soit $S_T = S_0 \exp(X_T)$ le prix de l'actif sous-jacent à la maturité, avec $(X_t)_{t \in [0, T]}$ un processus de Lévy sous $\mathbb{P}$ de triplet (b, σ^2, ν) . Notons $\mathbb{P}^*$ la mesure risque-neutre choisie et ρ_T^* (son expression analytique peut être inconnue) la densité risque-neutre de X_T . La fonction caractéristique de X_T sous $\mathbb{P}^*$ est exprimée par

$$\Phi_T^*(u) = \int_{\mathbb{R}} e^{iux} \rho_T^*(x) dx. \quad (5.19)$$

Soit $k = \log(K/S_0)$ le logarithme du strike normalisé. L'évaluation risque-neutre sous $\mathbb{P}^*$ nous donne

$$\begin{aligned} C(K) &= e^{-rT} \mathbb{E}^*[(S_T - K)_+] \\ &= S_0 e^{-rT} \mathbb{E}^*[(e^{X_T} - e^k)_+] \\ &= S_0 e^{-rT} \int_k^\infty (e^x - e^k) \rho_T^*(x) dx. \end{aligned}$$

Nous introduisons la fonction $c_0 : k \mapsto C(S_0 e^k)$. La technique d'inversion de Fourier consiste à calculer le prix de l'option $C(K) \equiv c_0(k)$ de la façon suivante : $c_0 = \text{FT}^{-1} \circ \text{FT}(c_0)$ avec FT est la transformée de Fourier. Comme

$$\lim_{k \rightarrow -\infty} c_0(k) = \lim_{K \rightarrow 0} C(K) = S_0,$$

nous remarquons que c_0 n'est pas intégrable vu que la limite de $c_0(k)$ ne tend pas vers 0 lorsque $k \rightarrow -\infty$. Pour cela, nous ne pouvons pas appliquer directement la technique d'inversion de Fourier car la transformée de Fourier de $c_0(k)$ ne converge pas.

Pour contourner ce problème de convergence, nous considérons le prix de l'option d'achat modifié

$$c_\alpha(k) = e^{\alpha k} c_0(k)$$

avec $\alpha > 0$.

Dans les deux propositions suivantes, nous présentons la formule fermée de la transformée de Fourier de $c_\alpha(k)$, à laquelle nous allons appliquer la transformée de Fourier inverse pour obtenir le prix de l'option $C(K)$.

Proposition 5.5. *Soit $\alpha > 0$ tel que $\mathbb{E}^*[e^{(\alpha+1)X_T}] < \infty$. La transformée de Fourier de $c_\alpha(k)$ est bien définie et elle est donnée par*

$$\hat{c}_\alpha(v) = \frac{S_0 e^{-rT} \Phi_T^*(v - (\alpha + 1)i)}{\alpha^2 + \alpha - v^2 + i(2\alpha + 1)v}, \quad \forall v \in \mathbb{R}, \quad (5.20)$$

avec Φ_T^* la fonction caractéristique de X_T sous $\mathbb{P}^*$.

Preuve. Supposons pour l'instant que $\hat{c}_\alpha(v)$ est bien définie. Nous avons

$$\begin{aligned} \hat{c}_\alpha(v) &= \int_{-\infty}^{\infty} e^{ivk} c_\alpha(k) dk \\ &= \int_{-\infty}^{\infty} e^{ivk} e^{\alpha k} C(S_0 e^k) dk \\ &= \int_{-\infty}^{\infty} e^{ivk} e^{\alpha k} \left(S_0 e^{-rT} \int_k^{\infty} (e^x - e^k) \rho_T^*(x) dx \right) dk \\ &= S_0 e^{-rT} \int_{-\infty}^{\infty} \rho_T^*(x) \left(\int_{-\infty}^x e^{(\alpha+iv)k} (e^x - e^k) dk \right) dx \\ &= S_0 e^{-rT} \int_{-\infty}^{\infty} \rho_T^*(x) \left(e^x \int_{-\infty}^x e^{(\alpha+iv)k} dk - \int_{-\infty}^x e^{(\alpha+1+iv)k} dk \right) dx \\ &= S_0 e^{-rT} \int_{-\infty}^{\infty} \rho_T^*(x) \left(\frac{e^{(\alpha+1+iv)x}}{\alpha + iv} - \frac{e^{(\alpha+1+iv)x}}{\alpha + 1 + iv} \right) dx. \end{aligned}$$

En appliquant (5.19), nous remplaçons l'intégrale par $\Phi_T^*(v - (\alpha + 1)i)$ et nous obtenons l'expression (5.20).

Nous prouvons maintenant l'existence de $\hat{c}_\alpha(v)$. Tout d'abord notons que $\mathbb{E}^*[e^{(\alpha+1)X_T}] < \infty$ implique que

$$\hat{c}_\alpha(0) < \infty, \quad (5.21)$$

car

$$\hat{c}_\alpha(0) = \frac{S_0 e^{-rT} \Phi_T^*(-(\alpha + 1)i)}{\alpha^2 + \alpha} = \frac{S_0 e^{-rT} \mathbb{E}^*[e^{(\alpha+1)X_T}]}{\alpha^2 + \alpha}.$$

D'autre part, comme $c_\alpha(0)$ est positive, nous avons

$$|\hat{c}_\alpha(v)| = \left| \int_{-\infty}^{\infty} e^{ivk} c_\alpha(k) dk \right| \leq \int_{-\infty}^{\infty} c_\alpha(k) dk = \hat{c}_\alpha(0).$$

En combinant cela avec (5.21), l'existence de $\hat{c}_\alpha(v)$ est prouvée. $\square$

Proposition 5.6. Soit $S_T = S_0 \exp(X_T)$ le prix de l'actif sous-jacent à la maturité de l'option et soit $\mathbb{P}^*$ la mesure risque-neutre choisie. Notons Φ^* la fonction caractéristique du processus $(X_t)_{t \in [0, T]}$ sous $\mathbb{P}^*$. Le prix de l'option est donné par

$$C(K) = \frac{e^{-\alpha \log(K/S_0)}}{\pi} \operatorname{Re} \left\{ \int_0^\infty e^{-iv \log(K/S_0)} \hat{c}_\alpha(v) dv \right\}, \quad (5.22)$$

avec $\hat{c}_\alpha$ donnée par (5.20).

Preuve. La transformée de Fourier inverse nous donne

$$c_\alpha(k) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-ivk} \hat{c}_\alpha(v) dv. \quad (5.23)$$

Alors,

$$\begin{aligned} C(K) &= \frac{e^{-\alpha \log(K/S_0)}}{2\pi} \int_{\mathbb{R}} e^{-iv \log(K/S_0)} \hat{c}_\alpha(v) dv \\ &= \frac{e^{-\alpha \log(K/S_0)}}{\pi} \operatorname{Re} \left\{ \int_0^\infty e^{-iv \log(K/S_0)} \hat{c}_\alpha(v) dv \right\}. \end{aligned}$$

La dernière égalité découle de l'observation suivante :

$$\int_{\mathbb{R}} e^{-iv \log(K/S_0)} \hat{c}_\alpha(v) dv = \int_0^\infty e^{-iv \log(K/S_0)} \hat{c}_\alpha(v) dv + \int_{-\infty}^0 e^{-iv \log(K/S_0)} \hat{c}_\alpha(v) dv,$$

et le second terme de la partie droite de l'égalité peut s'écrire comme étant le conjugué du premier terme,

$$\begin{aligned} \int_{-\infty}^0 e^{-iv \log(K/S_0)} \hat{c}_\alpha(v) dv &= \int_0^\infty e^{iu \log(K/S_0)} \hat{c}_\alpha(-u) du = \int_0^\infty \overline{e^{-iu \log(K/S_0)} \hat{c}_\alpha(u) du} \\ &= \overline{\int_0^\infty e^{-iu \log(K/S_0)} \hat{c}_\alpha(u) du}. \end{aligned}$$

Ainsi la preuve est faite. □

Nous avons considéré que l'évaluation du prix d'option d'achat. Évidemment, nous pouvons obtenir le prix d'une option de vente en utilisant la parité put-call. Cependant, le prix $P(K)$ d'une option de vente peut être alternativement obtenu par la méthode d'inversion de Carr-Madan en choisissant une valeur négative de α [107].

5.2.3 Discrétisation et transformée de Fourier rapide (FFT)

Calculer le prix d'une option d'achat $C(K) = e^{-rT} \mathbb{E}^*[(S_T - K)_+]$, sous la mesure risque-neutre $\mathbb{P}^*$, nécessite l'inversion de la transformée de Fourier dans (5.23). Dans le cas général, le problème ne peut pas être résolu analytiquement. Une approche numérique s'avère nécessaire. Pour avoir un meilleur temps de calcul, nous adopterons

une formulation à laquelle nous pouvons appliquer l'algorithme de transformée de Fourier rapide (FFT) [44, 161]. Ici nous définissons la transformée de Fourier discrète comme étant

$$F_u = \sum_{n=1}^N f_n \omega_N^{(n-1)(u-1)}, \quad u = 1, \dots, N \quad (5.24)$$

avec $\omega_N = e^{-\frac{2\pi i}{N}}$. Le langage R fournit un algorithme de FFT efficace pour cette formulation [16, 153].

Notons $g_k(v) \equiv e^{-ivk} \hat{c}_\alpha(v)$, avec $\hat{c}_\alpha(v)$ donnée par (5.20). Nous sommes alors intéressés par le calcul de l'intégrale $\int_0^\infty g_k(v) dv$, pour donner le prix de l'option (5.22).

En appliquant la méthode des trapèzes, nous avons

$$\int_0^\infty g_k(v) dv = \int_0^A g_k(v) dv + \varepsilon_T \quad (5.25)$$

$$= \frac{\Delta v}{2} \left[g_k(v_1) + 2 \sum_{n=2}^{N-1} g_k(v_n) + g_k(v_N) \right] + \varepsilon_D + \varepsilon_T \quad (5.26)$$

$$= \Delta v \left[\sum_{n=1}^N g_k(v_n) - \frac{1}{2} [g_k(v_1) + g_k(v_N)] \right] + \varepsilon_D + \varepsilon_T, \quad (5.27)$$

avec $A = (N-1)\Delta v$, ε_T est l'erreur de troncature et ε_D est l'erreur de discrétisation. Soit

$$v_n = (n-1)\Delta v \quad (5.28)$$

avec $n = 1, \dots, N$. De plus, soit

$$k_u = k_1 + (u-1)\Delta k, \quad (5.29)$$

avec $u = 1, \dots, N$, la discrétisation du domaine du strike k . La constante k_1 est choisie de sorte que la grille soit centrée autour de certaines valeurs de strike. Plus précisément, si nous sommes intéressés par des options avec des strikes autour d'une valeur K , nous fixons $k_1 = \log(K/S_0) - \frac{N}{2}\Delta v$. En substituant (5.28) et (5.29) dans (5.27),

$$\begin{aligned} \int_0^\infty g_{k_u}(v) dv &\approx \Delta v \left[\sum_{n=1}^N e^{-i[(n-1)\Delta v][k_1+(u-1)\Delta k]} \hat{c}_\alpha(v_n) - \frac{1}{2} [g_{k_u}(v_1) + g_{k_u}(v_N)] \right] \\ &= \Delta v \left[\sum_{n=1}^N e^{-i\Delta v \Delta k (n-1)(u-1)} g_{k_1}(v_n) - \frac{1}{2} [g_{k_u}(v_1) + g_{k_u}(v_N)] \right]. \end{aligned}$$

En posant

$$\Delta v \Delta k = \frac{2\pi}{N},$$

nous avons

$$\int_0^\infty g_{k_u}(v) dv \approx \Delta v \left[\sum_{n=1}^N \omega_N^{(n-1)(u-1)} g_{k_1}(v_n) - \frac{1}{2} [g_{k_u}(v_1) + g_{k_u}(v_N)] \right]. \quad (5.30)$$

La somme dans (5.30) prend la forme (5.24) avec $f_n = g_{k_1}(v_n)$. Ainsi, la FFT peut être appliquée pour calculer cette somme. Le résultat final de la méthode d'inversion de Carr-Madan est donc

$$C(S_0 e^{k_u}) \approx \frac{e^{-\alpha k_u}}{\pi} \operatorname{Re} \left\{ \Delta v \left[\sum_{n=1}^N \omega_N^{(n-1)(u-1)} g_{k_1}(v_n) \right] - \frac{1}{2} [g_{k_u}(v_1) + g_{k_u}(v_N)] \right\}. \quad (5.31)$$

La méthode des trapèzes peut être remplacée par la méthode de Simpson pour discrétiser l'intégrale. En appliquant la même démarche, nous montrons facilement que dans ce cas le prix de l'option est calculé selon la formule

$$C(S_0 e^{k_u}) \approx \frac{e^{-\alpha k_u}}{\pi} \operatorname{Re} \left\{ \frac{\Delta v}{3} \left[\sum_{n=1}^N \omega_N^{(n-1)(u-1)} g_{k_1}(v_n) (3 + (-1)^n - \delta_{n-1}) - [g_{k_u}(v_{N-1}) + 4g_{k_u}(v_N)] \right] \right\}, \quad (5.32)$$

avec δ_{j-1} la fonction delta de Kronecker qui vaut 1 lorsque $j = 1$ et 0 sinon.

Pour pouvoir appliquer l'algorithme de FFT, N doit être une puissance de 2. Dans nos applications, nous choisissons $N = 4096$ et $\Delta v = 0.25$. Cela donne $\Delta k = 6.13 \cdot 10^{-3}$. D'autre part, nous sommes intéressés par des valeurs de strike autour de $K = S_0$. Nous fixons alors $k_1 = -\frac{N}{2} \Delta k$.

5.3 Exemples des modèles exp-Lévy

Les modèles financiers avec sauts se regroupent en deux catégories : les *modèles de diffusion-sauts* et les *modèles à activité infinie*. La première catégorie consiste à modéliser l'évolution du prix de l'actif par une diffusion à laquelle nous ajoutons des sauts à des intervalles aléatoires de temps. Ainsi les sauts représentent des événements rares tel que les krachs ou les pertes et gains extrêmes. Une telle évolution peut être décrite en modélisant les log-prix de l'actif par une composante gaussienne non-nulle et une partie sauts, définie par un processus de Poisson composé avec un nombre fini de sauts dans chaque intervalle de temps. Les deux exemples les plus connus sont le modèle de Merton, où la taille des sauts suit une loi gaussienne [120], et le modèle de Kou, où la taille des sauts suit une double exponentielle [97]. Dans ces modèles, la structure dynamique du processus est facile à comprendre et à décrire vu que la loi de la taille des sauts est connue. La deuxième catégorie est celle des modèles avec un nombre infini de sauts dans chaque intervalle de temps. Dans ces modèles, nous n'avons pas besoin d'introduire une composante brownienne car la dynamique de sauts est assez riche pour générer un comportement non-trivial de petites variations [31]. Il était défendu dans [31, 72] que de tels modèles donnent une description plus réelle du processus de prix de l'actif, avec des différentes échelles de temps. De plus, plusieurs modèles de cette classe peuvent être construits via la subordination d'un mouvement brownien, ce qui leur donne une traçabilité analytique

additionnelle en comparaison avec les modèles de diffusion-saut. Deux exemples importants de cette catégorie sont le modèle variance-gamma [32, 111] et le modèle gaussienne inverse [8, 9, 142]. Cependant, puisque le processus du prix réel n'est observé que sur une discrétisation de temps, il est difficile, voire impossible, de décider empiriquement à quelle catégorie le processus de prix appartient. Ainsi, le choix est plus une question de conviction qu'un constat empirique. Dans cette section, nous commençons par rappeler le modèle standard de Black et Scholes. Ensuite, nous appliquons la méthode d'évaluation de prix de l'option par FFT sur un modèle de diffusion-sauts : le modèle de Merton. Nous finissons par le modèle variance gamma : un modèle à activité infinie. Sauf contre-indication, nous utilisons les paramètres résumés dans la table 5.1 pour les applications numériques.

Marché	B&S	Merton	VG	FFT
$S_0 = 100$	$\mu = 0.145$	$\gamma = 0.1$	$\gamma = 0.1$	$N = 4096$
$r = 0.02$	$\sigma = 0.3$	$\sigma = 0.3$	$m = -0.01$	$\Delta v = 0.25$
$T = 0.5$		$\lambda = 1$	$\delta = 1$	$\Delta k = \frac{2\pi}{N\Delta v}$
		$m = -0.1$	$\kappa = 0.2$	$k_1 = -\frac{N}{2}\Delta k$
		$\delta = 0.2$		

TABLE 5.1 – Sommaire de différentes valeurs numériques des paramètres utilisés dans la section 6.3.4.

5.3.1 Modèle de Black et Scholes

Le modèle de Black et Scholes, proposé en [25], est l'un des modèles les plus populaire en finance. Le prix de l'actif $(S_t)_{t \in [0, T]}$ est solution de l'EDS

$$dS_t = \mu S_t dt + \sigma S_t dB_t; \quad S_0 > 0, \quad (5.33)$$

avec $\mu \in \mathbb{R}$, $\sigma > 0$ et $(B_t)_{t \in [0, T]}$ un mouvement brownien standard. En appliquant la formule d'Itô, nous décrivons $(S_t)_{t \in [0, T]}$ par un modèle exp-Lévy (5.11) avec

$$X_t = \left(\mu - \frac{\sigma^2}{2} \right) t + \sigma B_t, \quad (5.34)$$

pour $t \in [0, T]$.

Le processus $(X_t)_{t \in [0, T]}$ est un processus de Lévy de triplet $(\mu - \frac{\sigma^2}{2}, \sigma^2, 0)$. En effet, pour tout $t \in [0, T]$, comme X_t suit une loi gaussienne $\mathcal{N}((\mu - \frac{\sigma^2}{2})t, \sigma^2 t)$, sa fonction caractéristique est donnée par

$$\Phi_t(u) = \mathbb{E}[e^{iuX_t}] = \exp \left(t \left(i \left(\mu - \frac{\sigma^2}{2} \right) u - \frac{\sigma^2}{2} u^2 \right) \right).$$

Dans ce modèle, le processus de prix ne présente pas de sauts. Ainsi, nous sommes dans le cas du marché complet et il existe une unique mesure risque-neutre.

Mesure risque-neutre d'Esscher.

Le moment exponentiel $\mathbb{E}[e^{\theta X_1}]$ est fini pour tout $\theta \in \mathbb{R}$. Pour montrer l'existence de la mesure risque-neutre d'Esscher $\mathbb{P}^*$, nous cherchons un réel $\theta \in \mathbb{R}$ solution de

$$\left(\mu - \frac{\sigma^2}{2}\right) + \sigma^2\theta + \frac{\sigma^2}{2} = r. \quad (5.35)$$

La solution est donnée par $\theta = (r - \mu)/\sigma^2$ et la mesure $\mathbb{P}^*$ construite via la transformée d'Esscher par rapport à $(X_t)_{t \in [0, T]}$ et avec le paramètre θ est l'unique mesure risque-neutre. Le processus $(X_t)_{t \in [0, T]}$ est encore un processus de Lévy sous $\mathbb{P}^*$ de triplet $(r - \frac{\sigma^2}{2}, \sigma, 0)$.

Cela implique que $(X_t)_{t \in [0, T]}$ peut être écrit

$$X_t = \left(r - \frac{\sigma^2}{2}\right) t + \sigma W_t,$$

avec $(W_t)_{t \in [0, T]}$ donné par $W_t = B_t + \frac{\mu - r}{\sigma} t$, un mouvement brownien standard sous $\mathbb{P}^*$. Le prix actualisé

$$\tilde{S}_t = S_0 e^{\sigma W_t - \frac{\sigma^2}{2} t}$$

est solution de l'équation

$$d\tilde{S}_t = \tilde{S}_t dW_t; \quad S_0 > 0.$$

Ainsi, nous retrouvons les mêmes résultats du théorème de Girsanov pour le changement de probabilité.

Evaluation de prix de l'option par FFT.

Pour le modèle de Black et Scholes avec les paramètres de la table 5.1, le paramètre d'Esscher, qui correspond à la mesure risque-neutre d'Esscher, est $\theta = (r - \mu)/\sigma^2 = -1.39$. Le calcul numérique du prix d'une option d'achat est présenté dans la figure 5.2 en fonction du strike K . D'autre part, la comparaison de la valeur calculée numériquement par la méthode de Carr-Madan avec celle calculée par la formule fermée de Black et Scholes nous donne une erreur absolue de la méthode numérique de l'ordre de $6 \cdot 10^{-7}$ indépendamment du strike.

5.3.2 Modèle diffusion-sauts de Merton

Ce modèle a été proposé dans [120]. Le prix est décrit par un processus exp-Lévy (5.11) où

$$X_t = \gamma t + \sigma B_t + \sum_{i=1}^{N_t} Y_i, \quad (5.36)$$

avec $\gamma \in \mathbb{R}$, $(B_t)_{t \in [0, T]}$ un mouvement brownien standard, $(N_t)_{t \in [0, T]}$ un processus de Poisson d'intensité λ et $(Y_i)_{i \geq 1}$ des variables aléatoires i.i.d. qui suivent une loi

gaussienne de moyenne m et de variance δ^2 . Pour tout $0 \leq t \leq T$, la fonction caractéristique de X_t est donnée par

$$\Phi_t(u) = \mathbb{E}[e^{iuX_t}] = e^{t\Psi(u)},$$

et

$$\Psi(u) = i\gamma u - \frac{\sigma^2}{2}u^2 + \lambda \left(\exp \left(imu - \frac{\delta^2}{2}u^2 \right) - 1 \right).$$

Définissons

$$\nu(x) = \lambda \times \frac{1}{\sqrt{2\pi}\delta} \exp \left(-\frac{(x-m)^2}{2\delta^2} \right), \quad x \in \mathbb{R}$$

et

$$b = \gamma + \int_{|x| \leq 1} x \nu(dx).$$

Alors, Ψ s'écrit sous la forme

$$\begin{aligned} \Psi(u) &= i\gamma u - \frac{\sigma^2}{2}u^2 + \int_{\mathbb{R}} (e^{iux} - 1) \nu(dx) \\ &= ibu - \frac{\sigma^2}{2}u^2 + \int_{\mathbb{R}} (e^{iux} - 1 - iux\mathbf{1}_{|x| \leq 1}) \nu(dx). \end{aligned}$$

Nous concluons que $(X_t)_{t \in [0, T]}$ est un processus de Lévy sous $\mathbb{P}$ de triplet (b, σ^2, ν) . Avec ce modèle, le marché est incomplet. Avec ses deux sources d'aléa, l'instant et la taille des sauts, le processus de Poisson composé rend le risque incontrôlable et comme conséquence il existe une infinité de mesures risque-neutres.

Mesure risque-neutre d'Esscher.

Le moment exponentiel $\mathbb{E}[e^{\theta X_1}]$ est fini pour tout $\theta \in \mathbb{R}$. Pour montrer l'existence de la mesure risque-neutre d'Esscher $\mathbb{P}^*$, nous cherchons un réel $\theta \in \mathbb{R}$ solution de

$$\gamma + \sigma^2\theta + \frac{\sigma^2}{2} + \int_{\mathbb{R}} e^{\theta x} (e^x - 1) \nu(x) = r,$$

ou encore

$$\gamma + \sigma^2\theta + \frac{\sigma^2}{2} + \lambda \left(e^{m(\theta+1) + \frac{\delta^2}{2}(\theta+1)^2} - e^{m\theta + \frac{\delta^2}{2}\theta^2} \right) = r.$$

La résolution numérique, avec les paramètres donnés dans la table 5.1, nous donne $\theta \approx -0.352$.

La fonction caractéristique de $(X_t)_{t \in [0, T]}$ sous $\mathbb{P}^*$ est donnée par

$$\Phi_t^*(u) = e^{t\Psi^*(u)},$$

avec

$$\Psi^*(u) = i(\gamma + \sigma^2\theta)u + \frac{\sigma^2}{2}u^2 + \lambda \left(e^{m(\theta+iu) + \frac{\delta^2}{2}(\theta+iu)^2} - e^{m\theta + \frac{\delta^2}{2}\theta^2} \right).$$

Ainsi, sous la mesure risque-neutre d'Esscher, la fonction caractéristique de $(X_t)_{t \in [0, T]}$ admet une forme explicite. Cette dernière sera inversée numériquement par la transformée de Fourier rapide pour donner une évaluation du prix de l'option.

En appliquant la proposition 5.4, $(X_t)_{t \in [0, T]}$ est un processus de Lévy sous $\mathbb{P}^*$ de triplet (b^*, σ^2, ν^*) défini par

$$b^* = \gamma + \sigma^2 \theta + \int_{|x| \leq 1} x \nu^*(dx)$$

et

$$\nu^*(x) = e^{\theta x} \nu(x) = \lambda^* \times \frac{1}{\sqrt{2\pi\delta}} \exp(x - (m + \delta^2 \theta))^2 / (2\delta^2),$$

où $\lambda^* = \lambda \exp(m\theta + \frac{\delta^2 \theta^2}{2})$. Nous constatons donc que $(X_t)_{t \in [0, T]}$ est de nouveau une diffusion-sauts, sous $\mathbb{P}^*$, avec un processus de Poisson composé comme partie sauts :

$$X_t = (\gamma + \sigma^2 \theta)t + \sigma B_t^* + \sum_{i=1}^{N_t^*} Y_i^*,$$

avec $(B_t^*)_{t \in [0, T]}$ un $\mathbb{P}^*$ -mouvement brownien standard, $(N_t^*)_{t \in [0, T]}$ un $\mathbb{P}^*$ -processus de Poisson d'intensité λ^* et $(Y_i^*)_{i \geq 1}$ des variables aléatoires i.i.d. qui suivent une loi gaussienne de moyenne $m + \delta^2 \theta$ et de variance δ^2 .

Mesure risque-neutre d'entropie minimale

Dans le contexte du modèle diffusion-sauts de Merton, l'existence de la mesure risque-neutre d'entropie minimale $\mathbb{P}^\dagger$ est équivalente à l'existence d'un réel β solution de

$$\gamma + \sigma^2 \beta + \frac{\sigma^2}{2} + \int_{\mathbb{R}} ((e^x - 1)e^{\beta(e^x - 1)}) \nu(dx) = r.$$

Avec les mêmes paramètres de la table 5.1, nous obtenons numériquement $\beta \approx -0.365$.

Par le théorème 5.3, $(X_t)_{t \in [0, T]}$ est un processus de Lévy sous $\mathbb{P}^\dagger$ de triplet $(b^\dagger, \sigma^2, \nu^\dagger)$ défini par

$$b^\dagger = \gamma + \sigma^2 \beta + \int_{|x| \leq 1} x \nu^\dagger(dx)$$

et

$$\nu^\dagger(x) = e^{\beta(e^x - 1)} \nu(x) = \lambda^\dagger \rho^\dagger(x)$$

où $\lambda^\dagger = \int \nu^\dagger(dx)$ et $\rho^\dagger(x) = \frac{\nu^\dagger(x)}{\int \nu^\dagger(dx)}$. Nous constatons donc que $(X_t)_{t \in [0, T]}$ est de nouveau une diffusion-sauts, sous $\mathbb{P}^\dagger$, avec un processus de Poisson composé comme partie sauts :

$$X_t = (\gamma + \sigma^2 \beta)t + \sigma B_t^\dagger + \sum_{i=1}^{N_t^\dagger} Y_i^\dagger,$$

avec $(B_t^\dagger)_{t \in [0, T]}$ un $\mathbb{P}^\dagger$ -mouvement brownien standard, $(N_t^\dagger)_{t \in [0, T]}$ un $\mathbb{P}^\dagger$ -processus de Poisson d'intensité $\lambda^\dagger$ et $(Y_i^\dagger)_{i \geq 1}$ des variables aléatoires i.i.d. qui suivent une loi de densité de probabilité $\rho^\dagger$.

Les facteurs d'amortissement $e^{\theta x}$, dans le cas de la mesure risque-neutre d'Esscher $\mathbb{P}^*$, et $e^{\beta(e^x - 1)}$, dans le cas de la mesure risque-neutre d'entropie minimale $\mathbb{P}^\dagger$, permettent d'atténuer la tendance générale du logarithme de prix. Dans un modèle de Merton, la tendance générale est donnée par $\mathbb{E}[X_t] = (\gamma + \lambda m)t$. Une pente positive de l'espérance du log prix va être compensée par des paramètres θ et β négatifs, ce qui permettra de donner plus de poids aux sauts négatifs et moins de poids aux sauts positifs, dans les densités de Lévy ν^* et $\nu^\dagger$. Cela conduit à rééquilibrer le marché dans le modèle risque-neutre et à avoir la propriété de martingale. Inversement, si la pente de l'espérance du log prix est négative, les facteurs d'amortissement vont fortement diminuer la queue gauche de la densité de Lévy sous les mesures risque-neutres.

La densité de Lévy initiale ν , et les deux densités de Lévy risque-neutres, ν^* et $\nu^\dagger$, ainsi que les trois densités de la taille des sauts correspondantes, sont présentées dans la figure 5.1 pour différentes valeurs de l'espérance du log prix. Nous varions la taille moyenne des sauts $m = \{-0.1, -0.3, 0.3\}$ pour avoir différents signes de $\mathbb{E}[X_t]$ et par conséquent différents signes de θ et β . Les autres paramètres du modèle sont pris tels que donnés dans la table 5.1. En fonction du signe du paramètre d'Esscher, nous avons un affaiblissement de la queue gauche ou droite de la densité de Lévy sous la mesure risque-neutre. Cependant, nous constatons dans les figures de droite que les densités de la taille des sauts ne sont pas affectées par ce changement de mesure de probabilité et que l'effet principal est une augmentation ou une diminution des intensités de saut λ^* et $\lambda^\dagger$.

D'autre part, si nous comparons ν^* et $\nu^\dagger$, nous trouvons que les deux fonctions sont presque égales avec des différences qui n'excèdent pas $1.6 \cdot 10^{-2}$. Nous expliquons ce constat par le fait que X et $\hat{X}$ représentent respectivement les rendements logarithmiques et les rendements arithmétiques du prix de l'actif. nous considérons alors des modèles avec des petits sauts x pour lesquels $e^x - 1 \approx x$. Ainsi, la mesure risque-neutre d'Esscher et la mesure risque neutre d'entropie minimale sont presque les mêmes.

Bien que nous avons le triplet de $(X_t)_{t \in [0, T]}$ sous $\mathbb{P}^\dagger$, nous ne disposons toujours pas d'une expression explicite de la fonction caractéristique sous cette mesure. Ainsi, l'application de la FFT pour évaluer le prix de l'option sous la mesure risque-neutre d'entropie minimale n'est pas possible.

Évaluation de prix de l'option sous la mesure risque-neutre d'Esscher.

Dans la figure 5.2, nous comparons le prix de l'option sous un modèle de Black et Scholes et sous un modèle de Merton ayant la même espérance de X_T . Nous choisissons les paramètres tel que $\mu - \frac{\sigma^2}{2} = \gamma + \lambda m$. Le résultat numérique du prix de l'option est présenté dans la figure 5.2 en fonction du strike K . Comme le modèle

de Merton permet des sauts négatifs plus grands et une volatilité plus importante, le prix d'une option d'achat est plus conservatif « atour de la monnaie ».

Afin de discuter l'effet des sauts sur le prix de l'option, nous étudions la sensibilité du prix de l'option pour les différents paramètres du modèle de Merton. Nous traçons dans un premier temps le graphe suivant : en abscisse le strike K et en ordonnée le prix de l'option d'achat $C(K)$, pour différentes valeurs d'intensité de sauts. Rappelons que le modèle de Black et Scholes et le modèle diffusion-sauts de Merton sont deux modèles exponentielle-Lévy de triplets respectivement $(\mu - \frac{\sigma^2}{2}, \sigma^2, 0)$ et $(\gamma + \int_{|x| \leq 1} x\nu(dx), \sigma^2, \nu)$, avec

$$\nu(x) = \lambda \times \frac{1}{\sqrt{2\pi}\delta} \exp\left(-\frac{(x-m)^2}{2\delta^2}\right).$$

Ainsi, le modèle de Black et Scholes n'est qu'un modèle diffusion-sauts de Merton de paramètres particuliers $\gamma = \mu - \frac{\sigma^2}{2}$ et $\lambda = 0$. Ce modèle ne contient pas de sauts vu que l'intensité de sauts λ est nulle. Pour illustrer l'effet de l'intensité de sauts sur le prix de l'option, nous fixons les paramètres μ , σ et γ aux valeurs données par la table 5.1. La condition $\gamma = \mu - \frac{\sigma^2}{2}$ est ainsi respectée par ce choix de paramètres. Nous faisons par la suite varier λ dans $\{2, 4, 6\}$. Les différentes courbes de prix de l'option d'achat sont présentées dans la figure 5.3. Une variation de l'intensité de saut λ modifie le prix de l'option d'achat pour les strikes « around the money ». Lorsque λ augmente, nous avons plus de sauts dans l'intervalle de temps $[0, T]$ et la variance $\text{Var}(X_t) = (\sigma^2 + \lambda(m^2 + \delta^2))t$ augmente. Par conséquent, le prix de l'option d'achat augmente à cause du risque supplémentaire que les sauts ont apporté au vendeur de l'option. La même observation est constatée lorsque la volatilité de la taille des sauts augmente, comme présenté dans la figure 5.4. Le prix de l'option en fonction de la taille moyenne des saut est illustré dans la figure 5.5. Avec une moyenne négative des tailles de saut, nous avons des options « out-the-money » plus intéressantes, tandis que les options « in-the-money » sont moins intéressantes. Inversement, une moyenne positive de la taille des sauts conduit à l'observation opposée.

5.3.3 Modèle variance gamma

Le processus variance gamma a été proposé dans [112] afin de décrire la dynamique du prix de l'actif en remplaçant le mouvement brownien dans le modèle de Black et Scholes. Deux nouveaux paramètres : m et κ ont été introduits pour contrôler l'asymétrie et les queues lourdes constatés dans les distributions empiriques. Un processus de variance gamma est obtenu en évaluant un mouvement brownien avec une dérive à un temps aléatoire donné par un processus gamma.

Définition 5.5. *Un processus variance gamma (VG) $(Y_t)_{t \in [0, T]}$ de paramètres (m, δ, κ) est défini par*

$$Y_t = m\gamma_t + \delta B_{\gamma_t},$$

avec $(B_t)_{t \in [0, T]}$ un mouvement brownien standard et $(\gamma_t)_{t \in [0, T]}$ un processus gamma de taux de la moyenne (mean rate) égal à 1 et de taux de la variance (variance rate) κ .

Proposition 5.7. *Le processus variance gamma $(Y_t)_{t \in [0, T]}$ est un processus de Lévy process de triplet caractéristique $(\int_{|x| \leq 1} x \nu(dx), 0, \nu)$ avec*

$$\nu(x) = \frac{1}{\kappa|x|} \exp \left(\frac{m}{\delta^2} x - \frac{\sqrt{\frac{m^2}{\delta^2} + \frac{2}{\kappa}}}{\delta} |x| \right)$$

et de fonction caractéristique

$$\Phi_t(u) = \mathbb{E}[e^{iuY_t}] = \left(\frac{1}{1 - im\kappa u + (\delta^2 \kappa / 2) u^2} \right)^{t/\kappa}.$$

À partir de l'expression de la mesure de Lévy ν , nous constatons qu'il existe une infinité de petits sauts. Un tel processus est dit processus à activité infinie. Il n'admet pas une loi de la taille des sauts puisque les sauts arrivent infiniment.

Pour modéliser le prix de l'actif boursier, nous ajoutons une composante dérive au processus VG. Une composante brownienne n'est pas nécessaire puisque la dynamique du processus VG est assez riche et le prix bouge essentiellement par les sauts. Ainsi, le prix de l'actif est décrit par un modèle exp-Lévy (5.11) avec

$$X_t = \gamma t + Y_t.$$

Le processus $(X_t)_{t \in [0, T]}$ est un processus de Lévy de triplet $(\gamma + \int_{|x| \leq 1} x \nu(dx), 0, \nu)$ et de fonction caractéristique donnée par

$$\Phi_t(u) = \mathbb{E}[e^{iuX_t}] = \frac{e^{i\gamma u t}}{(1 - im\kappa u + (\delta^2 \kappa / 2) u^2)^{t/\kappa}}.$$

Mesure risque-neutre d'Esscher.

Dans un premier temps, nous déterminons le domaine de définition de la fonction génératrice des moments de X_t par la résolution de l'inégalité $1 - im\kappa u + (\delta^2 \kappa / 2) u^2 > 0$ en u . Ainsi, le moment exponentiel $\mathbb{E}[\exp(uX_t)]$ existe pour tout

$$u \in \left(-m/\delta^2 - \sqrt{m^2/\delta^2 + 2/\kappa} \delta, -m/\delta^2 + \sqrt{m^2/\delta^2 + 2/\kappa} \delta \right).$$

Pour appliquer le théorème 5.2, nous supposons que $2\sqrt{m^2/\delta^2 + 2/\kappa} \delta > 1$. Sous cette hypothèse, nous cherchons un réel

$$\theta \in \left(-m/\delta^2 - \sqrt{m^2/\delta^2 + 2/\kappa} \delta, -m/\delta^2 + \sqrt{m^2/\delta^2 + 2/\kappa} \delta - 1 \right)$$

solution de

$$\mathbb{E}[e^{(\theta+1)X_t}] = e^{rt} \mathbb{E}[e^{\theta X_t}].$$

Pour les paramètres de la table 5.1, la fonction génératrice des moments de X_t est bien définie sur $(-3.15, 3.17)$. Numériquement, nous obtenons $\theta = -0.57$.

La fonction caractéristique de X_t sous $\mathbb{P}^*$ est donnée par

$$\Phi_t^*(u) = \frac{\mathbb{E}[e^{(\theta+iu)X_t}]}{\mathbb{E}[e^{\theta X_t}]} = \frac{e^{i\gamma ut}}{(1 - im^*\kappa u + (\delta^{*2}\kappa/2)u^2)^{t/\kappa}}$$

avec $m^* = (m + \delta^2\theta)/A$, $\delta^* = \delta/\sqrt{A}$ et $A = 1 - m\kappa\theta - \frac{\delta^2\kappa}{2}\theta^2$. Ainsi, sous $\mathbb{P}^*$, $(X_t)_{t \in [0, T]}$ est de nouveau un processus variance gamma de paramètres (m^*, δ^*, κ) auquel nous ajoutons une dérive γt . Cela implique que $(X_t)_{t \in [0, T]}$ est un processus de Lévy sous $\mathbb{P}^*$ de triplet caractéristique $(\gamma + \int_{|x| \leq 1} x\nu^*(dx), 0, \nu^*)$ avec

$$\nu^*(x) = e^{\theta x} \nu(x) = \frac{1}{\kappa|x|} \exp\left(\frac{m^*}{\delta^{*2}}x - \frac{\sqrt{\frac{m^{*2}}{\delta^{*2}} + \frac{2}{\kappa}}}{\delta^*}|x|\right).$$

La densité initiale de Lévy ν et la densité risque-neutre de Lévy ν^* sont illustrées dans la figure 5.6. Avec notre choix de paramètres, la tendance générale $\mathbb{E}[X_t] = (\gamma + m)t$ du log prix de l'actif est positif, ce qui conduit à une valeur négative de θ . Les grands sauts positifs vont être presque sans effet dans l'évaluation du prix de l'option, tandis que les sauts négatifs y vont largement contribuer.

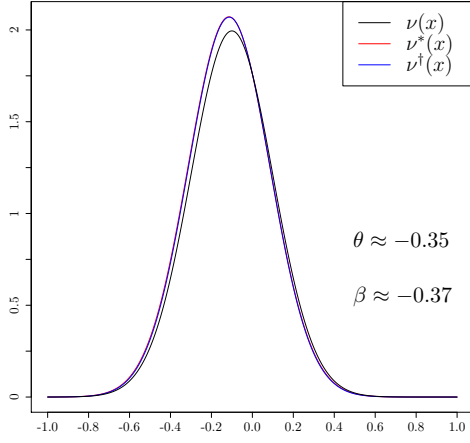
Evaluation de prix de l'option sous la mesure risque-neutre d'Esscher.

L'expression explicite de la fonction caractéristique sous la mesure risque-neutre d'Esscher $\mathbb{P}^*$ est exploitée pour calculer le prix d'une option d'achat avec la méthode d'inversion de Carr et Madan. Les figures 5.7 et 5.8 illustrent la sensibilité du prix de l'option à une variation des paramètres δ et κ respectivement. Nous prenons les paramètres de la table 5.1 à l'exception du δ qui varie dans $\{0.6, 0.8, 1\}$ pour la figure 5.7 et κ qui varie dans $\{0.02, 0.2\}$ pour la figure 5.8. Les paramètres δ et κ permettent respectivement de contrôler la volatilité et le kurtosis. Augmenter δ conduit à une plus grande volatilité, ce qui à son tour augmente le prix de l'option d'achat. D'autre part, lorsque κ augmente, les queues de distribution du log prix deviennent plus épaisses et le prix des options d'achat « out-the-money » augmente, contrairement au prix des options d'achat « in-the-money » qui va décroître.

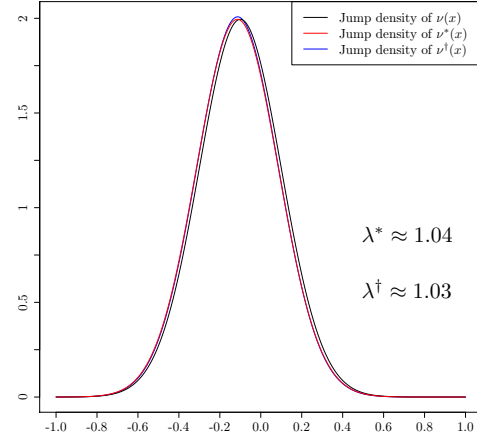
5.4 Conclusion

Ce chapitre a permis d'introduire les différentes notions de l'évaluation des options sous des modèles exponentielle-Lévy. Le fait de considérer des modèles discontinus, où nous permettons au prix de l'actif de sauter, rend la modélisation plus réelle, mais en contrepartie plus complexe, à savoir l'existence d'une infinité de mesures risque-neutres et l'absence d'une expression explicite de la densité du prix de l'actif sous-jacent dans le cas général, et par conséquent l'absence d'une formule fermée du prix de l'option. Cependant, en construisant une mesure risque-neutre et en exploitant

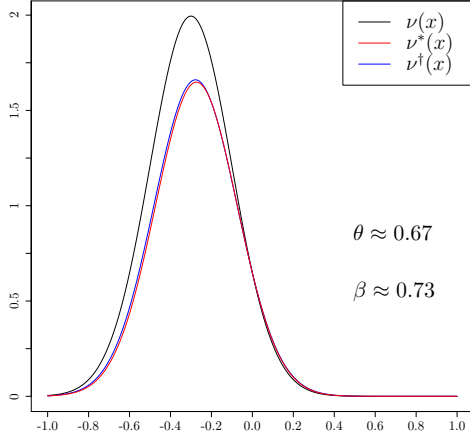
l'expression explicite de la fonction caractéristique de X sous cette mesure, nous arrivons à évaluer numériquement le prix de l'option en utilisant l'algorithme de la transformée de Fourier rapide.



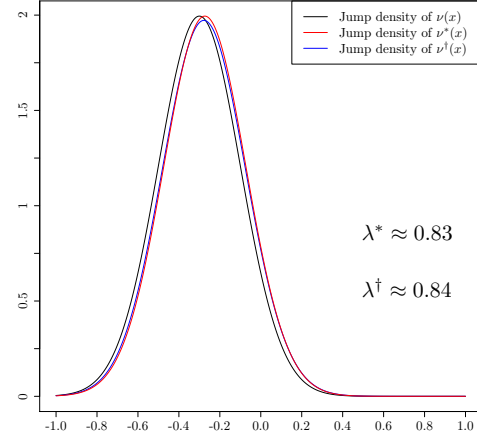
(a) $m = -0.1$



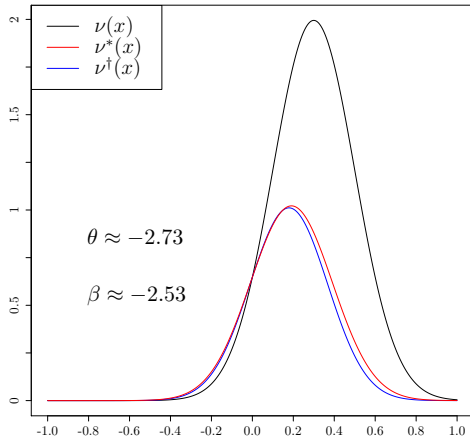
(b) $m = -0.1$



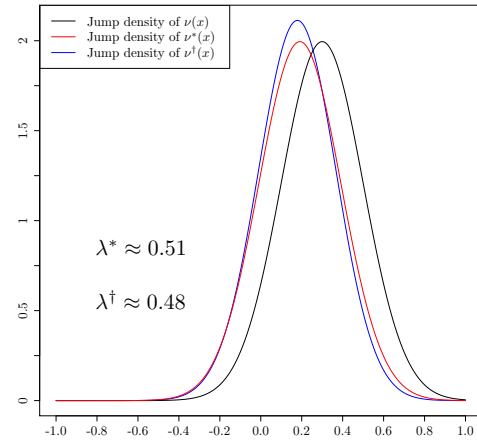
(c) $m = -0.3$



(d) $m = -0.3$



(e) $m = 0.3$



(f) $m = 0.3$

FIGURE 5.1 – Densités de Lévy et densités de la taille des sauts sous la mesure de probabilité historique $\mathbb{P}$, la mesure risque-neutre d'Esscher $\mathbb{P}^*$ et la mesure risque-neutre d'entropie minimale $\mathbb{P}^\dagger$ sous des modèles de Merton.

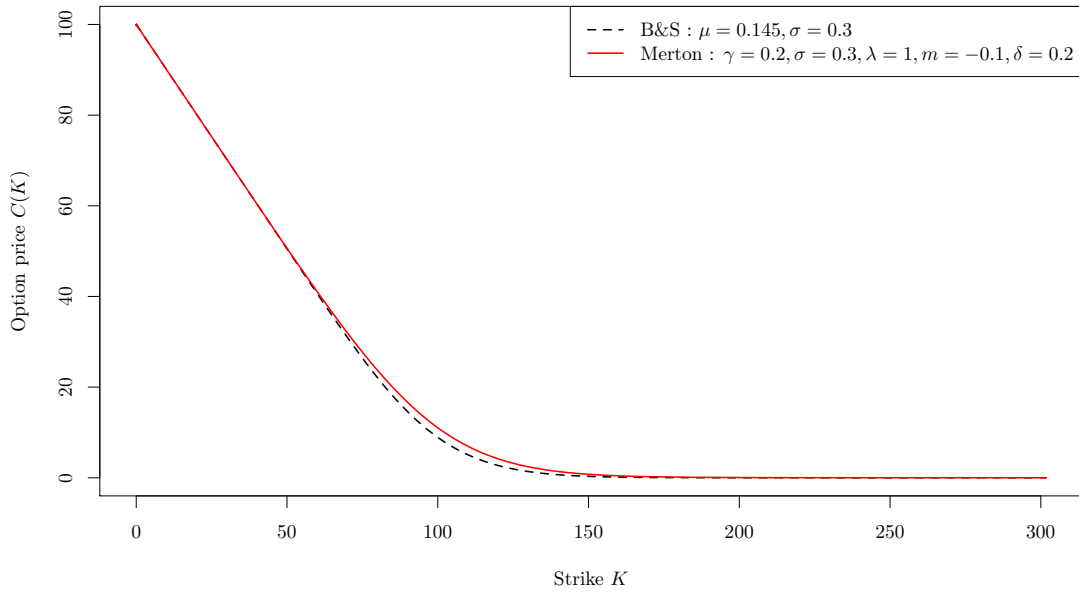


FIGURE 5.2 – Prix d’une option d’achat en fonction du strike K . Une comparaison entre un modèle de B&S et un modèle de Merton ayant la même espérance de X_T : $\mu - \sigma^2/2 = \gamma + \lambda m$.

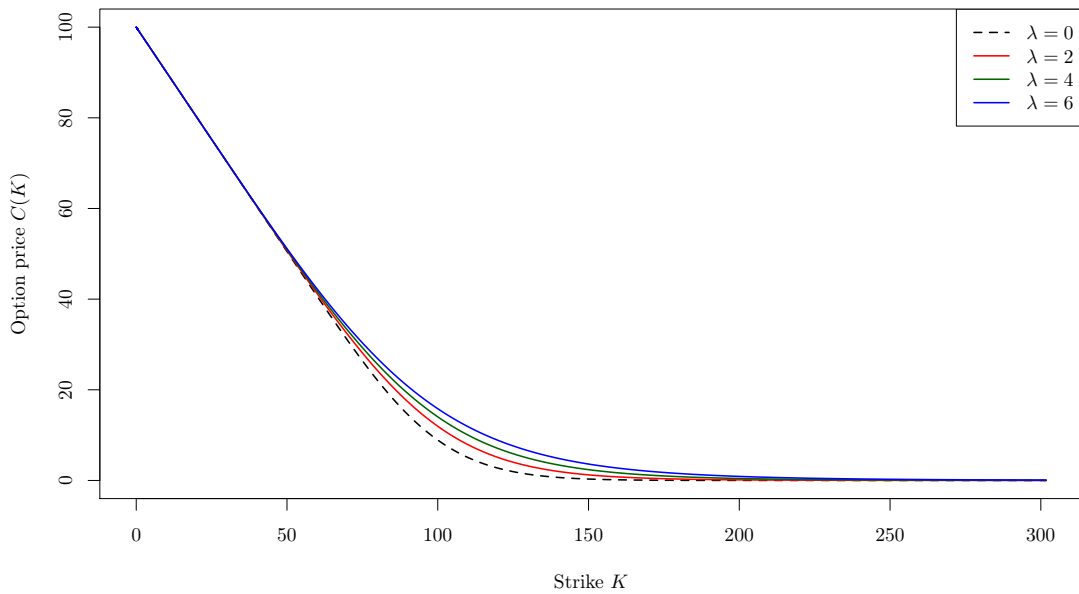


FIGURE 5.3 – Sensibilité du prix de l’option par rapport à l’intensité de saut dans le modèle de Merton.

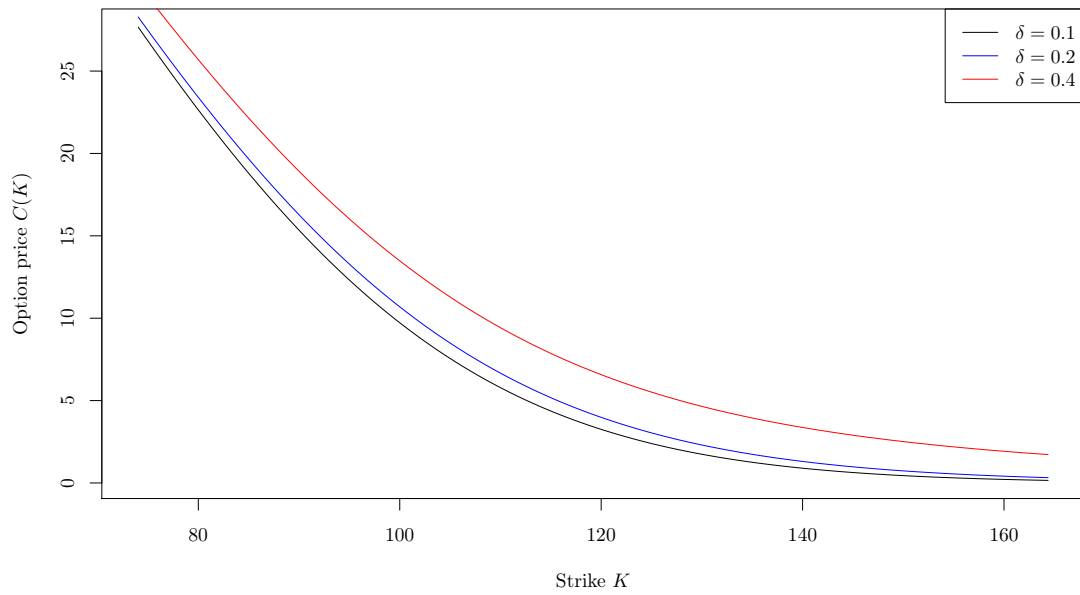


FIGURE 5.4 – Sensibilité du prix de l’option par rapport à la variance de la taille des sauts dans le modèle de Merton.

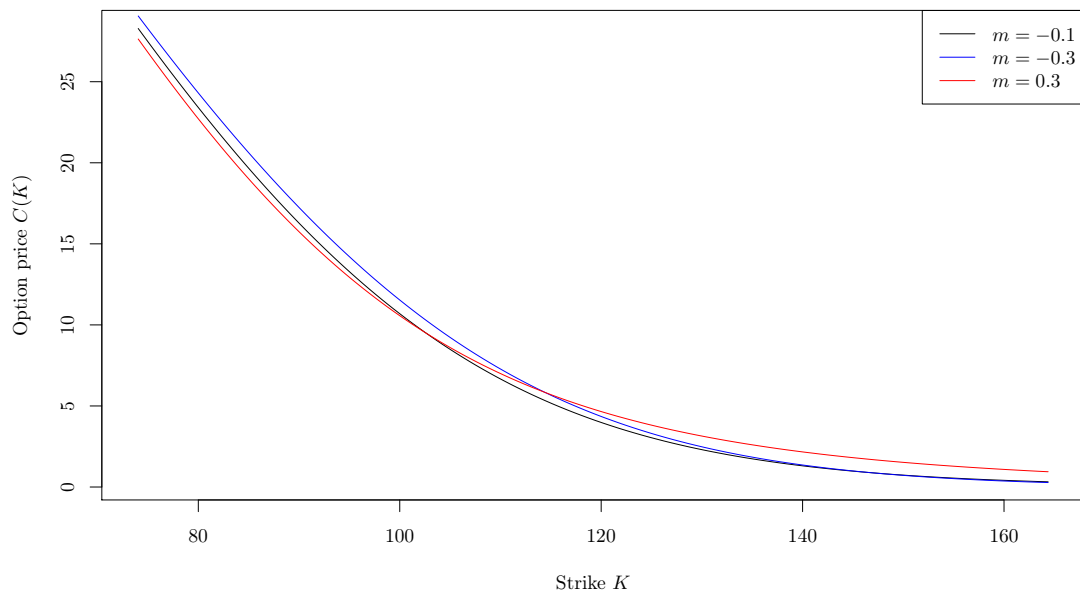


FIGURE 5.5 – Sensibilité du prix de l’option par rapport à la taille moyenne de sauts dans le modèle de Merton.

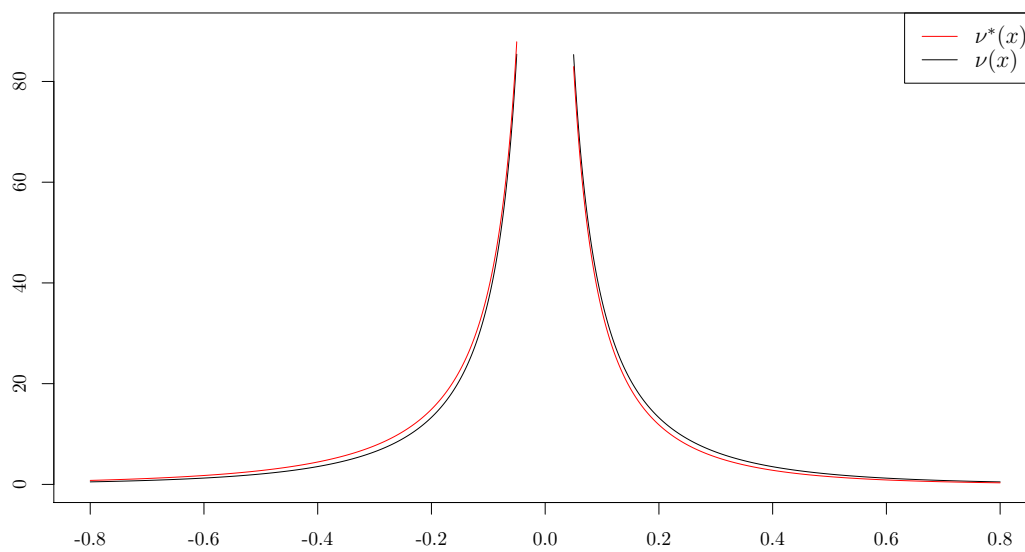


FIGURE 5.6 – La densité initiale de Lévy $\nu(x)$ et la densité risque-neutre de Lévy $\nu^*(x)$ dans le modèle variance gamma.

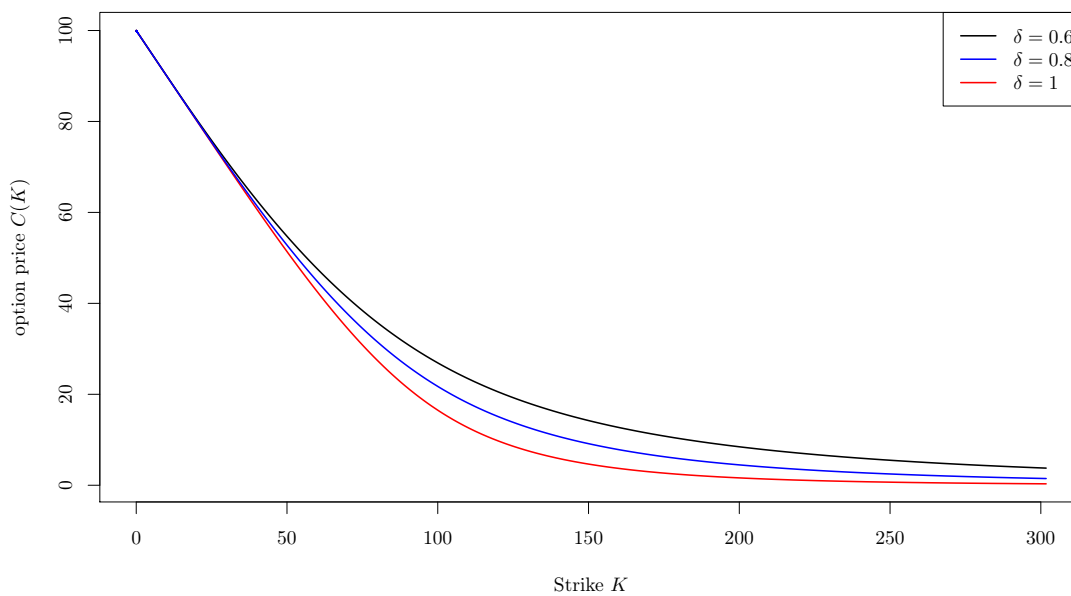


FIGURE 5.7 – Sensibilité du prix de l’option par rapport au paramètre de la volatilité dans un modèle variance gamma.

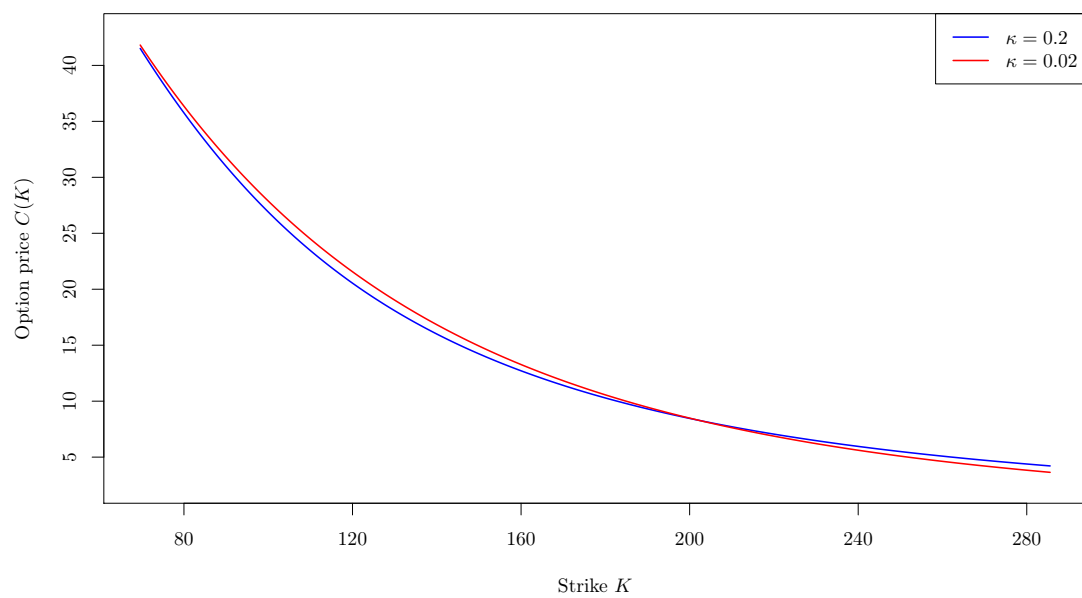


FIGURE 5.8 – Sensibilité du prix de l’option par rapport au paramètre de kurtosis dans un modèle variance gamma.

Chapitre 6

La Value-at-Risk et la Value-at-Risk Conditionnelle sous le modèle exp-Lévy

Dans ce chapitre, nous développons deux méthodes pour le calcul de la VaR et de la CVaR sous les modèles de type exponentielle de Lévy. La première méthode permet d'évaluer la VaR et la CVaR monétaires, tandis que la deuxième s'intéresse à la VaR et la CVaR en termes de rendement. Le principe de ces méthodes est d'exprimer la VaR et la CVaR comme, respectivement l'argument du minimum et le minimum, d'une certaine fonction que nous évaluons numériquement par des techniques d'inversion de Fourier. Le premier avantage de notre méthodologie est qu'elle ne nécessite que la connaissance de la fonction caractéristique du processus de Lévy. Le deuxième est qu'elle évalue simultanément la VaR et la CVaR, contrairement à d'autres méthodes de la littérature [28, 94, 95, 105, 106] qui les calculent séparément en évaluant d'abord la fonction de densité du processus de Lévy par inversion de Fourier.

Nous commençons le chapitre par introduire la Value-at-Risk Conditionnelle (CVaR), une notion complémentaire à la VaR, ainsi que l'algorithme de Rockafellar et Uryasev [136] qui servira à exprimer la VaR et la CVaR comme solutions d'un problème d'optimisation convexe. Nous présentons ensuite, dans la section 6.2, la méthode du calcul de la VaR et de la CVaR monétaires. Sous le modèle de Black et Scholes, les valeurs exactes de la VaR et la CVaR peuvent être obtenues et nous les comparons avec nos résultats numériques. Nous illustrons aussi la VaR et la CVaR sous le modèle de Merton et le modèle variance gamma. Nous finirons par développer la méthode de calcul de la VaR et la CVaR en termes de rendement, dans la section 6.3, ainsi que son application dans le cadre des modèles exponentielle-Lévy mentionnés ci-dessous.

6.1 La CVaR et l'algorithme de Rockafellar et Uryasev

Nous considérons toujours le modèle exp-Lévy sous lequel le prix de l'actif est défini par le processus $S_t = S_0 \exp(X_t)$ avec $(X_t)_{t \in [0, T]}$ un processus de Lévy sur $\mathbb{R}$ de triplet (b, σ^2, ν) . Avant de commencer avec les définitions formelles, il faut noter que le concept de la Value-at-Risk a toujours été une norme en matière de mesure de risque au sein des établissements financiers et dans les accords de Bâle. En effet, elle représente un indicateur simple et global permettant d'allouer un capital disponible et de fixer des limites de négociations. Néanmoins, le fait que la VaR donne un seuil de perte sans fournir la valeur de la perte, rend cet indicateur incomplet pour répondre à certaines questions d'optimisation de portefeuille. Cela a induit l'apparition d'une mesure de risque complémentaire, à savoir la CVaR. Cette dernière commence récemment à intéresser l'industrie bancaire et notamment celle du comité de Bâle.

Définition 6.1. *La Value-at-Risk Conditionnelle (CVaR) d'une fonction perte L , et à un seuil de confiance a , est définie par*

$$\text{CVaR}_a(L) = \frac{1}{1-a} \int_a^1 \text{VaR}_x(L) dx. \quad (6.1)$$

Nous rappelons que la VaR est le quantile de niveau a de la variable aléatoire L .

Une autre façon de définir la CVaR est :

$$\text{CVaR}_a(L) = \frac{1}{1-a} \left(\mathbb{E} \left[L \mathbf{1}_{\{L \geq \text{VaR}_a(L)\}} \right] + \text{VaR}_a(L) (1-a - \mathbb{P}(L \geq \text{VaR}_a(L))) \right).$$

Si L est une variable aléatoire continue, ou plus généralement si elle est continue stochastiquement (la probabilité d'apparition d'une valeur particulière est nulle), comme c'est le cas du processus de Lévy, alors la CVaR coïncide avec l'espérance conditionnelle de la perte dépassant la VaR, appelée souvent la TVaR. Néanmoins, la différence entre la CVaR et la TVaR peut apparaître dès qu'on considère une variable aléatoire discrète pour décrire la loi de L . Généralement, la VaR et d'autres mesures de risque telles que la TVaR ne sont pas continues par rapport au niveau de confiance a , d'où un léger changement du niveau de confiance peut engendrer des grands sauts dans leurs valeurs. Au contraire, la CVaR est continue par rapport à a , indépendamment de la loi de la perte.

Proposition 6.1 ([1]). *Pour toute v.a. réelle L satisfaisant $\mathbb{E}[|L|] < \infty$, la fonction $a \mapsto \text{CVaR}_a(L)$ est continue dans $(0, 1)$.*

Une propriété intéressante de la CVaR est le fait qu'elle est une mesure de risque cohérente, au sens de la définition donnée par [6] (des généralisations de cette définition ont été développées ultérieurement [51, 92]).

Définition 6.2 ([6]). Soit $\mathcal{L}$ l'ensemble des v.a. réelles dans $(\Omega, \mathcal{F}, \mathbb{P})$ tel que $\mathbb{E}[|L|] < \infty$ pour toute $L \in \mathcal{L}$. Une mesure de risque $\rho : \mathcal{L} \rightarrow \mathbb{R}$ est dite cohérente si elle est :

1. monotone : $L_1, L_2 \in \mathcal{L}, L_1 \geq L_2 \Rightarrow \rho(L_1) \geq \rho(L_2)$;
2. invariante par translation : $L \in \mathcal{L}, a \in \mathbb{R} \Rightarrow \rho(L + a) = \rho(L) + a$;
3. homogène positivement : $L \in \mathcal{L}, \lambda > 0 \Rightarrow \rho(\lambda L) = \lambda \rho(L)$;
4. sous additive : $L_1, L_2 \in \mathcal{L} \Rightarrow \rho(L_1 + L_2) \leq \rho(L_1) + \rho(L_2)$.

Proposition 6.2 ([1]). Considérons un niveau de confiance $a \in (0, 1)$ fixé et l'ensemble $\mathcal{L}$ des v.a. réelles dans $(\Omega, \mathcal{F}, \mathbb{P})$ tel que $\mathbb{E}[|L|] < \infty$ pour toute $L \in \mathcal{L}$. Alors, $\rho : \mathcal{L} \rightarrow \mathbb{R}$ définie par $\rho(L) = \text{CVaR}_a(L)$ est une mesure de risque cohérente.

Notons que la VaR n'est pas une mesure de risque cohérente du fait qu'elle n'est pas sous-additive [7]. La violation de la sous-additivité contredit le principe de diversification. D'autre part, nous pouvons montrer que la CVaR est la plus petite mesure de risque majorant la VaR qui est cohérente [1, 102]. Pour cela, dans le chapitre suivant de ce document, qui s'intéresse à la gestion du portefeuille, nous adoptons la CVaR comme mesure de risque.

Rockafellar et Uryasev [136] proposent un algorithme pour calculer simultanément la VaR et la CVaR par la résolution d'un certain problème unidimensionnel d'optimisation convexe, présenté par le théorème ci-dessous. Soit $\mathcal{A}$ l'ensemble des stratégies admissibles et notons $L(x)$ la fonction perte sous la stratégie $x \in \mathcal{A}$.

Théorème 6.1 ([136]). Soit $a \in (0, 1)$. La fonction

$$z \mapsto F_a(x, z) = z + \frac{1}{1-a} \mathbb{E}[(L(x) - z)_+], \quad (6.2)$$

est finie et convexe (donc continue), et

$$\text{CVaR}_a(L(x)) = \min_{z \in \mathbb{R}} F_a(x, z), \quad (6.3)$$

$$\text{VaR}_a(L(x)) = \min\{y : y \in \underset{z \in \mathbb{R}}{\text{argmin}} F_a(x, z)\}. \quad (6.4)$$

En particulier, nous avons toujours

$$\text{VaR}_a(L(x)) \in \underset{z \in \mathbb{R}}{\text{argmin}} F_a(x, z), \quad (6.5)$$

$$\text{CVaR}_a(L(x)) = F_a(x, \text{VaR}_a(L(x))). \quad (6.6)$$

Un corollaire évident, mais important, est que le problème de minimisation de la CVaR sur l'ensemble des stratégies possibles peut être exprimé en fonction de celui de la minimisation de $F_a(x, z)$.

Corollaire 6.1 ([136]). *La minimisation de la $\text{CVaR}_a(x)$ sur l'ensemble des stratégies X est équivalente à la minimisation de $F_a(x, z)$ sur $\mathcal{A} \times \mathbb{R}$:*

$$\min_{x \in \mathcal{A}} \text{CVaR}_a(x) = \min_{(x, z) \in \mathcal{A} \times \mathbb{R}} F_a(x, z). \quad (6.7)$$

En outre,

$$(x^*, z^*) \in \underset{(x, z) \in \mathcal{A} \times \mathbb{R}}{\text{argmin}} F_a(x, z) \Leftrightarrow \begin{cases} x^* \in \underset{x \in \mathcal{A}}{\text{argmin}} \text{CVaR}_a(L(x)), \\ z^* \in \underset{z \in \mathbb{R}}{\text{argmin}} F_a(x^*, z). \end{cases} \quad (6.8)$$

En introduisant la fonction

$$f(z) = z + \frac{1}{1-a} \min_{x \in \mathcal{A}} \mathbb{E}[(L(x) - z)_+], \quad (6.9)$$

le problème d'optimisation de la CVaR s'écrit

$$\min_{x \in \mathcal{A}} \text{CVaR}_a(L(x)) = \min_{z \in \mathbb{R}} f(z).$$

6.2 VaR et CVaR monétaires

Dans cette section, nous nous intéressons à la VaR et la CVaR monétaires. Nous considérons ainsi la perte en termes de monnaie, c'est-à-dire,

$$L = S_0 - S_T.$$

Nous développons une méthode pour calculer simultanément ces deux mesures de risque. Nous nous plaçons ensuite dans le cadre du modèle de Black et Scholes, où des formules fermées, pour la VaR et la CVaR, existent. Nous comparons nos résultats numériques avec les valeurs exactes. Finalement, nous illustrons les résultats de notre méthode dans le cadre de trois modèles exponentielle-Lévy.

6.2.1 Méthode de calcul

En utilisant l'algorithme de Rockafellar et Uryasev, la CVaR monétaire peut être exprimée par

$$\text{CVaR}_a(S_0 - S_T) = \min_{z \in \mathbb{R}} \left\{ z + \frac{1}{1-a} \mathbb{E}[(S_0 - S_T - z)_+] \right\}.$$

L'enjeu est de pouvoir calculer $\mathbb{E}[(S_0 - S_T - z)_+]$, pour tout $z \in \mathbb{R}$. Remarquons que pour $z \geq S_0$, cette espérance est nulle, ainsi

$$\text{CVaR}_a(S_0 - S_T) = \min_{z \in \mathbb{R}} \begin{cases} z + \frac{1}{1-a} \mathbb{E}[(S_0 - S_T - z)_+] & \text{pour } z < S_0, \\ z & \text{pour } z \geq S_0. \end{cases}$$

Or, la fonction $z \mapsto z$ pour $z \geq S_0$ est croissante, donc le minimum global sera atteint dans $] -\infty, S_0]$. Pour $z < S_0$, posons $w = \log \left(1 - \frac{z}{S_0}\right)$ et définissons

$$f_0(w) \equiv \mathbb{E}[(S_0 e^w - S_T)_+] = \mathbb{E}[(S_0 - S_T - z)_+].$$

Notre idée est d'appliquer la méthode de Fourier inverse pour évaluer cette quantité numériquement. Nous rencontrons un problème d'intégrabilité lorsque $w \rightarrow +\infty$ ($z \rightarrow -\infty$). Pour le contourner, nous appliquons la transformée de Fourier à la fonction modifiée

$$f_\alpha(w) = e^{-\alpha w} f_0(w), \quad w \in \mathbb{R},$$

avec $\alpha > 1$. Le choix $\alpha > 1$ permet de répondre à des conditions d'intégrabilité dans le calcul de la transformée de Fourier de f_α , comme montré dans la proposition suivante.

Proposition 6.3. *Soit $\alpha > 1$ tel que $\mathbb{E}[e^{-(\alpha-1)X_T}] < \infty$. La transformée de Fourier de $f_\alpha(w)$ est bien définie. Elle est donnée par*

$$\hat{f}_\alpha(u) = \frac{S_0 \Phi_T(u + i(\alpha - 1))}{\alpha(\alpha - 1) - u^2 - iu(2\alpha - 1)}, \quad (6.10)$$

où Φ_T est la fonction caractéristique de X_T .

Preuve. Supposons, dans un premier temps, que $\hat{f}_\alpha$ est bien définie. Nous avons

$$\begin{aligned} \hat{f}_\alpha(u) &= \int_{\mathbb{R}} e^{i u w} f_\alpha(w) dw \\ &= S_0 \int_{\mathbb{R}} e^{(i u - \alpha) w} \int_{-\infty}^w (e^w - e^x) \rho_T(x) dx dw, \end{aligned}$$

avec ρ_T la fonction de densité de X_T (qui n'est pas forcément explicite). En échangeant les intégrales, nous obtenons

$$\begin{aligned} \hat{f}_\alpha(u) &= S_0 \int_{\mathbb{R}} \rho_T(x) \int_x^{+\infty} e^{(i u - \alpha) w} (e^w - e^x) dw dx \\ &= S_0 \int_{\mathbb{R}} \rho_T(x) \left(\int_x^{+\infty} e^{(i u - \alpha + 1) w} dw - e^x \int_x^{+\infty} e^{(i u - \alpha) w} dw \right) dx. \end{aligned}$$

La condition $\alpha > 1$ assure l'existence de l'intégrale $\int_x^{+\infty} e^{(i u - \alpha + 1) w} dw$ dans la ligne ci-dessous et le calcul des intégrales donne

$$\hat{f}_\alpha(u) = S_0 \int_{\mathbb{R}} \rho_T(x) \frac{e^{(i u - \alpha + 1)x}}{(i u - \alpha + 1)(i u - \alpha)} dx = \frac{S_0 \Phi_T(u + i(\alpha - 1))}{\alpha(\alpha - 1) - u^2 - iu(2\alpha - 1)}.$$

Ainsi, nous obtenons l'égalité (6.10).

Nous prouvons maintenant l'existence de $\hat{f}_\alpha$. Tout d'abord, notons que

$$\hat{f}_\alpha(0) = \frac{S_0 \Phi_T(i(\alpha - 1))}{\alpha(\alpha - 1)} = \frac{S_0 \mathbb{E}[e^{-(\alpha-1)X_T}]}{\alpha(\alpha - 1)}.$$

Ainsi, la condition $\mathbb{E}[e^{-(\alpha-1)X_T}] < \infty$ implique que $\hat{f}_\alpha(0) < \infty$. D'autre part, comme $f_\alpha(0)$ est positive par définition, nous avons

$$|\hat{f}_\alpha(u)| = \left| \int_{-\infty}^{\infty} e^{i u w} f_\alpha(w) dw \right| \leq \int_{-\infty}^{\infty} f_\alpha(w) dw = \hat{f}_\alpha(0).$$

En combinant cela avec le fait que $\hat{f}_\alpha(0) < \infty$, l'existence de $\hat{f}_\alpha$ est prouvée. $\square$

La CVaR monétaire s'obtient par inversion de Fourier de $\hat{f}_\alpha(u)$. Le résultat est donné dans le théorème suivant.

Théorème 6.2. *La CVaR monétaire est donnée par*

$$\text{CVaR}_a(S_0 - S_T) = \max_{w \in \mathbb{R}} \left\{ S_0(1 - e^w) + \frac{e^{\alpha w}}{(1-a)\pi} \text{Re} \left(\int_0^\infty e^{-i u w} \hat{f}_\alpha(u) du \right) \right\}, \quad (6.11)$$

où l'expression de $\hat{f}_\alpha$ est donnée par (6.10). De plus, la VaR monétaire est donnée par

$$\text{VaR}_a(S_0 - S_T) = S_0(1 - e^{w_{\max}}) \quad (6.12)$$

avec

$$w_{\max} = \operatorname{argmax}_{w \in \mathbb{R}} \left\{ S_0(1 - e^w) + \frac{e^{\alpha w}}{(1-a)\pi} \text{Re} \left(\int_0^\infty e^{-i u w} \hat{f}_\alpha(u) du \right) \right\}.$$

La transformée de Fourier inverse $\int_0^\infty e^{-i u w} \hat{f}_\alpha(u) du$ peut être évaluée par l'algorithme de la transformée de Fourier rapide (FFT). L'avantage de cet algorithme est qu'en une seule opération, nous évaluons cet intégrale pour N valeurs de w .

6.2.2 Étude de l'erreur numérique dans le cadre du modèle de Black et Scholes

Le modèle de Black et Scholes est un modèle d'exponentielle de Lévy pour lequel le rendement total $X_T = S_T/S_0$ suit une loi normale de moyenne $(\mu - \frac{\sigma^2}{2})T$ et de variance $\sigma^2 T$. Ainsi, des formules explicites pour la VaR et la CVaR existent et sont données par la proposition suivante.

Proposition 6.4. *Sous le modèle de Black et Scholes, la VaR et la CVaR monétaires sont données par*

$$\text{VaR}_a(S_0 - S_T) = S_0 \left(1 - \exp \left(\left(\mu - \frac{\sigma^2}{2} \right) T + \mathcal{N}^{-1}(1-a) \sigma \sqrt{T} \right) \right)$$

et

$$\text{CVaR}_a(S_0 - S_T) = S_0 \left(1 - \frac{e^{\mu T} \mathcal{N} \left(\mathcal{N}^{-1}(1-a) - \sqrt{\sigma^2 T} \right)}{1-a} \right),$$

où $\mathcal{N}$ est la fonction de répartition d'une loi normale centrée réduite.

Preuve. Soit Z une variable gaussienne centrée réduite. Ainsi $X_T \equiv \log(S_T/S_0)$ est égal en loi à $(\mu - \frac{\sigma^2}{2})T + \sigma\sqrt{T}Z$.

Par définition, la $\text{VaR}_a(S_0 - S_T)$ est solution en x de $\mathbb{P}(S_0 - S_T \leq x) = a$, donc solution en x de

$$\mathbb{P}\left(Z \leq \frac{\log(1 - \frac{x}{S_0}) - (\mu - \frac{\sigma^2}{2})T}{\sigma\sqrt{T}}\right) = 1 - a.$$

D'où, en notant $\mathcal{N}^{-1}(1 - a)$ le quantile d'ordre $1 - a$ de Z , nous obtenons facilement

$$\text{VaR}_a(S_0 - S_T) = S_0 \left(1 - \exp\left(\left(\mu - \frac{\sigma^2}{2}\right)T + \mathcal{N}^{-1}(1 - a)\sigma\sqrt{T}\right)\right).$$

Comme le modèle de Black et Scholes est continu, la $\text{CVaR}_a(S_0 - S_T)$ peut être exprimée par

$$\begin{aligned}\text{CVaR}_a(S_0 - S_T) &= \frac{1}{1 - a} \mathbb{E}[(S_0 - S_T) \mathbf{1}_{\{S_0 - S_T \geq \text{VaR}_a(S_0 - S_T)\}}] \\ &= S_0 - \frac{1}{1 - a} \mathbb{E}[S_T \mathbf{1}_{\{S_0 - S_T \geq \text{VaR}_a(S_0 - S_T)\}}].\end{aligned}$$

D'après l'expression de $\text{VaR}_a(S_0 - S_T)$, les ensembles $\{S_0 - S_T \geq \text{VaR}_a(S_0 - S_T)\}$ et $\{Z \leq \mathcal{N}^{-1}(1 - a)\}$ sont égaux en loi. Ainsi,

$$\begin{aligned}\text{CVaR}_a(S_0 - S_T) &= S_0 - \frac{S_0}{1 - a} \int_{-\infty}^{\mathcal{N}^{-1}(1-a)} e^{(\mu - \frac{\sigma^2}{2})T + \sigma\sqrt{T}x} \cdot \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx \\ &= S_0 - \frac{S_0 e^{\mu T}}{1 - a} \int_{-\infty}^{\mathcal{N}^{-1}(1-a)} \frac{e^{-(x - \sigma\sqrt{T})^2/2}}{\sqrt{2\pi}} dx \\ &= S_0 - \frac{S_0 e^{\mu T}}{1 - a} \int_{-\infty}^{\mathcal{N}^{-1}(1-a) - \sigma\sqrt{T}} \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx \\ &= S_0 \left(1 - \frac{e^{\mu T} \mathcal{N}\left(\mathcal{N}^{-1}(1 - a) - \sqrt{\sigma^2 T}\right)}{1 - a}\right),\end{aligned}$$

avec $\mathcal{N}$ la fonction de répartition d'une loi normale centrée réduite. Nous obtenons ainsi l'expression explicite de la CVaR monétaire sous le modèle de Black et Scholes. $\square$

Nous cherchons maintenant à distinguer entre les erreurs numériques dues à l'inversion de Fourier et celles dues à la minimisation. Pour cela, nous calculons d'une façon explicite la fonction à minimiser dans l'algorithme de Rockafellar et Uryasev. Ainsi, les seules erreurs qui résident sont celles de la minimisation.

Proposition 6.5. *Sous le modèle de Black et Scholes, la CVaR monétaire est donnée par*

$$\begin{aligned} \text{CVaR}(-X_T) = \max_{w \in \mathbb{R}} & \left\{ S_0(1 - e^w) \right. \\ & \left. + \frac{S_0}{1-a} \times \left(e^w \mathcal{N}\left(\frac{w - (\mu - \frac{\sigma^2}{2})T}{\sigma\sqrt{T}}\right) - e^{\mu T} \mathcal{N}\left(\frac{w - (\mu + \frac{\sigma^2}{2})T}{\sigma\sqrt{T}}\right) \right) \right\}, \quad (6.13) \end{aligned}$$

où $\mathcal{N}$ est la fonction de répartition d'une loi normale centrée réduite. La VaR monétaire est l'argument du minimum de cette fonction.

Preuve. La preuve consiste à calculer explicitement l'espérance $\mathbb{E}[(S_0 e^w - S_T)_+]$ dans le cadre du modèle de Black et Scholes. Sous ce modèle, la variable aléatoire X_T est égale en loi à $(\mu - \frac{\sigma^2}{2})T - \sigma\sqrt{T}Z$ où Z est une variable normale centrée réduite. Nous avons donc

$$\begin{aligned} \mathbb{E}[(S_0 e^w - S_T)_+] &= S_0 \int_{\mathbb{R}} (e^w - e^{(\mu - \frac{\sigma^2}{2})T + \sigma\sqrt{T}x})_+ \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx \\ &= S_0 e^w \int_{-\infty}^{\frac{w - (\mu - \frac{\sigma^2}{2})T}{\sigma\sqrt{T}}} \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx - S_0 e^{\mu T} \int_{-\infty}^{\frac{w - (\mu + \frac{\sigma^2}{2})T}{\sigma\sqrt{T}}} \frac{e^{-(x - \sigma\sqrt{T})^2/2}}{\sqrt{2\pi}} dx \\ &= S_0 e^w \mathcal{N}\left(\frac{w - (\mu - \frac{\sigma^2}{2})T}{\sigma\sqrt{T}}\right) - S_0 e^{\mu T} \mathcal{N}\left(\frac{w - (\mu + \frac{\sigma^2}{2})T}{\sigma\sqrt{T}}\right). \end{aligned}$$

Cette dernière ligne est obtenue par changement de variables dans la deuxième intégrale. Ainsi, la preuve est achevée en injectant cette expression dans le problème de minimisation. $\square$

Les résultats des trois méthodes de calcul des VaR/CVaR sont résumés dans la table 6.1 pour différentes périodes de temps et différents niveaux de risque. La méthode 1 est celle qui utilise les formules fermées données par la proposition 6.4. La méthode 2 consiste à appliquer la proposition 6.5, tandis que la troisième méthode est celle proposée par le théorème 6.2 où nous calculons numériquement l'inverse de la transformée de Fourier évoquée par application de l'algorithme de transformée de Fourier rapide, avec une discrétisation de l'intégrale de taille $N = 2^{14}$ et de pas 0.25. Ainsi, le pas de discrétisation en w qui résulte est de $\Delta w = \frac{2\pi}{0.25N} = 1.53 \cdot 10^{-3}$. La valeur du paramètre d'amortissement est $\alpha = 1.75$. Notons que la différence entre les résultats de la première et la deuxième méthode représente l'erreur liée à la discrétisation de la fonction à minimiser en w , et que la différence entre les résultats de la deuxième et la troisième méthode représente l'erreur de l'inversion numérique de la transformée de Fourier qui, elle aussi, est la somme des erreurs induites par la troncature et par la discrétisation de l'intégrale (voir la section 5.2.3 pour plus de détails). Ainsi, la différence entre la troisième et la première méthode donne l'ensemble des erreurs de la méthode de calcul de la VaR et la CVaR par inversion de Fourier. La table 6.1 montre que les erreurs de la méthode que nous

venons de proposer est de l'ordre de 10^{-2} (quelques centimes d'euro si les actions sont cotées en euro) pour la VaR monétaire et inférieures à 10^{-4} pour le calcul de la CVaR monétaire.

		$T = 0.5$		$T = 1$	
		$a = 0.95$	0.99	$a = 0.95$	$a = 0.99$
VaR_a(S₀ - S_T)					
	Meth. 1	25.8385	35.8209	32.5279	45.0037
	Meth. 2	25.8388	35.8207	32.5281	45.0034
	Meth. 3	25.8534	35.8098	32.4768	45.0228
CVaR_a(S₀ - S_T)					
	Meth. 1	31.9261	40.1466	40.1257	50.1131
	Meth. 2	31.9261	40.1466	40.1257	50.1131
	Meth. 3	31.9261	40.1466	40.1258	50.1131

TABLE 6.1 – Comparaison des différentes méthodes de calcul de la VaR et de la CVaR monétaires sous le modèle de Black et Scholes.

6.2.3 Application aux modèles exponentielle-Lévy

Nous calculons la VaR et la CVaR monétaires par la méthode d'inversion de Fourier présentée dans le théorème 6.2. Nous appliquons la transformée de Fourier rapide pour calculer l'intégrale avec $N = 4096$ et un pas de discrétisation de l'intégrale $\Delta u = 0.25$. Nous obtenons alors une évaluation de la fonction-objectif (fonction à minimiser) sur l'intervalle $[w_1, w_1 + (N - 1)\Delta w] = [-12.55, 11.55]$ avec un pas de discrétisation $\Delta w = 6.13 \cdot 10^{-3}$ (voir section 5.2.3 et équation (5.28) pour plus de détails). Nous illustrons les résultats pour les trois modèles exponentielle-Lévy suivants : le modèle de Black et Scholes, le modèle diffusion-sauts de Merton et le modèle de variance gamma. Les différents paramètres des modèles, ainsi que ceux de la transformée de Fourier rapide, sont résumés dans la table 6.2.

Marché	B&S	Merton	VG	FFT
$S_0 = 100$	$\mu = 0.145$	$\gamma = 0.1$	$\gamma = 0.1$	$N = 4096$
	$\sigma = 0.3$	$\sigma = 0.3$	$m = -0.01$	$\Delta u = 0.25$
		$\lambda = 1$	$\delta = 1$	
		$m = -0.1$	$\kappa = 0.2$	$\Delta w = \frac{2\pi}{N\Delta u}$
		$\delta = 0.2$		$w_1 = -\frac{N}{2}\Delta w$
				$\Delta z = \frac{2\pi}{N\Delta u}$
				$z_1 = -\frac{N}{2}\Delta z$

TABLE 6.2 – Sommaire de différentes valeurs numériques des paramètres utilisés dans les sections 6.2.3 et 6.3.3.

La figure 6.1a illustre la VaR et la CVaR monétaires sous le modèle de Black et Scholes, calculées par la méthode d'inversion de Fourier. Les résultats sous le modèle diffusion-sauts de Merton sont donnés par la figure 6.1b. Finalement, la figure 6.1c montre les courbes de la VaR et la CVaR en fonction du temps sous le modèle variance gamma.

6.3 VaR et CVaR en termes de rendement

Dans cette section, nous considérons la perte en termes de rendement, c'est-à-dire, l'inverse du rendement logarithmique

$$L = -\log(S_T/S_0) = -X_T.$$

Nous développons une méthode pour calculer simultanément la VaR et la CVaR en termes de rendement. Nous nous plaçons ensuite dans le cadre du modèle de Black et Scholes, où des formules fermées, pour ces mesures de risque existent. Nous comparons nos résultats numériques avec les valeurs exactes. Finalement, nous illustrons les résultats de notre méthode dans le cadre de trois modèles exponentielle-Lévy.

6.3.1 Méthode de calcul

En appliquant l'algorithme de Rockafellar et Uryasev, la CVaR en termes de rendement peut être exprimée par

$$\text{CVaR}_a(-X_T) = \min_{z \in \mathbb{R}} \left\{ z + \frac{1}{1-a} \mathbb{E}[(-X_T - z)_+] \right\}. \quad (6.14)$$

Définissons la fonction

$$g_0(z) \equiv \mathbb{E}[(-X_T - z)_+].$$

Pour $z \rightarrow -\infty$, $g_0(z)$ diverge vers $+\infty$. Ainsi, g_0 n'est pas intégrable et nous ne pouvons pas appliquer directement la transformée de Fourier. Pour cela, nous introduisons la fonction modifiée

$$g_\alpha(z) = e^{\alpha z} g_0(z),$$

où $\alpha > 0$.

Proposition 6.6. *Soit $\alpha > 0$ tel que $\mathbb{E}[e^{-\alpha X_T}] < \infty$. La transformée de Fourier de $g_\alpha(z)$ est bien définie et elle est donnée par*

$$\hat{g}_\alpha(u) = \frac{\Phi_T(-u + i\alpha)}{\alpha^2 - u^2 + 2i\alpha u}, \quad (6.15)$$

avec Φ_T est la fonction caractéristique de X_T .

Preuve. Supposons pour l'instant que $\hat{g}_\alpha$ est bien définie. Nous avons

$$\begin{aligned}\hat{g}_\alpha(u) &= \int_{\mathbb{R}} e^{iuz} g_\alpha(z) dz \\ &= - \int_{\mathbb{R}} e^{(iu+\alpha)z} \int_{-\infty}^{-z} (x+z) \rho_T(x) dx dz\end{aligned}$$

avec ρ_T est la fonction de densité de X_T ,

$$\begin{aligned}&= - \int_{\mathbb{R}} \rho_T(x) \int_{-\infty}^{-x} (x+z) e^{(iu+\alpha)z} dz dx \\ &= - \int_{\mathbb{R}} \rho_T(x) \left(x \int_{-\infty}^{-x} e^{(iu+\alpha)z} dz + \int_{-\infty}^{-x} z e^{(iu+\alpha)z} dz \right) dx.\end{aligned}$$

Par intégration par parties, nous obtenons

$$\begin{aligned}&= - \int_{\mathbb{R}} \rho_T(x) \left(\frac{x e^{-(iu+\alpha)x}}{iu+\alpha} - \frac{x e^{-(iu+\alpha)x}}{iu+\alpha} - \int_{-\infty}^{-x} \frac{e^{(iu+\alpha)z}}{iu+\alpha} dz \right) dx \\ &= \int_{\mathbb{R}} \rho_T(x) \frac{e^{-(iu+\alpha)x}}{(iu+\alpha)^2} dx = \frac{\Phi_T(-u+i\alpha)}{\alpha^2 - u^2 + 2i\alpha u}.\end{aligned}$$

Ainsi, nous prouvons l'égalité (6.15).

Nous prouvons maintenant l'existence de $\hat{g}_\alpha$. Tout d'abord, notons que

$$\hat{g}_\alpha(0) = \frac{\Phi_T(i\alpha)}{\alpha^2} = \frac{\mathbb{E}[e^{-\alpha X_T}]}{\alpha^2}.$$

Ainsi, la condition $\mathbb{E}[e^{-\alpha X_T}] < \infty$ implique que $\hat{g}_\alpha(0) < \infty$. D'autre part, comme $g_\alpha(0)$ est positive par définition,

$$|\hat{g}_\alpha(u)| = \left| \int_{-\infty}^{\infty} e^{iuz} g_\alpha(z) dz \right| \leq \int_{-\infty}^{\infty} g_\alpha(z) dz = \hat{g}_\alpha(0).$$

En combinant cela avec le fait que $\hat{g}_\alpha(0) < \infty$, l'existence de $\hat{g}_\alpha$ est prouvée. □

La CVaR en termes de rendement s'obtient par inversion de Fourier de $\hat{g}_\alpha(u)$. Le résultat est donné dans le théorème suivant.

Théorème 6.3. *La CVaR en termes de rendement est donnée par*

$$\text{CVaR}_a(-X_T) = \min_{z \in \mathbb{R}} \left\{ z + \frac{e^{-\alpha z}}{(1-a)\pi} \text{Re} \left(\int_0^\infty e^{iuz} \hat{g}_\alpha(u) du \right) \right\}, \quad (6.16)$$

où $\hat{g}_\alpha$ est donnée par (6.15). De plus, la VaR en termes de rendement est donnée par

$$\text{VaR}_a(-X_T) = \underset{z \in \mathbb{R}}{\text{argmin}} \left\{ z + \frac{e^{-\alpha z}}{(1-a)\pi} \text{Re} \left(\int_0^\infty e^{iuz} \hat{g}_\alpha(u) du \right) \right\}. \quad (6.17)$$

L'intégrale dans (6.16) peut être calculée en appliquant la transformée de Fourier rapide tel que montré dans le chapitre précédent.

6.3.2 Étude de l'erreur numérique dans le cadre du modèle de Black et Scholes

Dans ce modèle le rendement total $X_T = \log(S_T/S_0)$ suit une loi normale de moyenne $(\mu - \frac{\sigma^2}{2})T$ et de variance $\sigma^2 T$. Ainsi, des formules explicites pour la VaR et la CVaR existent et sont données par la proposition suivante.

Proposition 6.7. *Sous le modèle de Black et Scholes, la VaR et la CVaR en termes de rendement total $X_T = (\mu - \frac{\sigma^2}{2})T + \sigma B_T$ sont données par*

$$\text{VaR}_a(-X_T) = -(\mu - \frac{\sigma^2}{2})T - \mathcal{N}^{-1}(1-a)\sqrt{\sigma^2 T}$$

et

$$\text{CVaR}_a(-X_T) = \frac{1}{1-a} \sqrt{\frac{\sigma^2 T}{2\pi}} \exp\left(-\frac{(\mathcal{N}^{-1}(1-a))^2}{2}\right) - (\mu - \frac{\sigma^2}{2})T,$$

où $\mathcal{N}^{-1}(1-a)$ est le quantile d'ordre $1-a$ d'une loi normale centrée réduite.

Preuve. Soit Z une variable gaussienne centrée réduite. Ainsi, $X_T = \log(S_T/S_0)$ est égal en loi à $(\mu - \frac{\sigma^2}{2})T + \sigma\sqrt{T}Z$.

Par définition, la $\text{VaR}_a(-X_T)$ est solution en x de $\mathbb{P}(-X_T \leq x) = a$, donc solution en x de

$$\mathbb{P}\left(Z \leq -\frac{x + (\mu - \frac{\sigma^2}{2})T}{\sigma\sqrt{T}}\right) = 1-a.$$

D'où, en notant $\mathcal{N}^{-1}(1-a)$ le quantile d'ordre $1-a$ de Z , nous obtenons

$$\text{VaR}_a(-X_T) = -(\mu - \frac{\sigma^2}{2})T - \mathcal{N}^{-1}(1-a)\sqrt{\sigma^2 T}.$$

Comme le modèle de Black et Scholes est continu, la $\text{CVaR}_a(-X_T)$ est exprimée par

$$\begin{aligned} \text{CVaR}_a(-X_T) &= \frac{1}{1-a} \mathbb{E}[-X_T \mathbf{1}_{\{-X_T \geq \text{VaR}_a(-X_T)\}}] \\ &= \frac{1}{1-a} \mathbb{E}\left[-\left((\mu - \frac{\sigma^2}{2})T + \sigma\sqrt{T}Z\right) \mathbf{1}_{\{Z \leq \mathcal{N}^{-1}(1-a)\}}\right] \\ &= -(\mu - \frac{\sigma^2}{2})T - \frac{1}{1-a} \sqrt{\frac{\sigma^2 T}{2\pi}} \int_{-\infty}^{\mathcal{N}^{-1}(1-a)} x e^{-\frac{x^2}{2}} dx \\ &= -(\mu - \frac{\sigma^2}{2})T + \frac{1}{1-a} \sqrt{\frac{\sigma^2 T}{2\pi}} \exp\left(-\frac{(\mathcal{N}^{-1}(1-a))^2}{2}\right). \end{aligned}$$

Ainsi, nous finissons la preuve. □

D'autre part, dans un tel modèle, nous pouvons expliciter la fonction à minimiser de l'algorithme de Rockafellar et Uryasev et calculer ainsi la CVaR par minimisation sans avoir recours à l'inversion de Fourier.

Proposition 6.8. *Sous le modèle de Black et Scholes, la CVaR en termes de rendement total $X_T = (\mu - \frac{\sigma^2}{2})T + \sigma B_T$ est donnée par*

$$\text{CVaR}(-X_T) = \min_{z \in \mathbb{R}} \left\{ z + \frac{1}{1-a} \times \left(\sqrt{\frac{\sigma^2 T}{2\pi}} \exp\left(-\frac{((\mu - \frac{\sigma^2}{2})T + z)^2}{2\sigma^2 T}\right) - \left((\mu - \frac{\sigma^2}{2})T + z\right) \mathcal{N}\left(-\frac{(\mu - \frac{\sigma^2}{2})T + z}{\sigma\sqrt{T}}\right) \right) \right\}, \quad (6.18)$$

où $\mathcal{N}$ est la fonction de répartition d'une loi normale centrée réduite.

Preuve. L'idée est de calculer directement la quantité $\mathbb{E}[(-X_T - z)_+]$. Soit Z une variable gaussienne centrée réduite. Ainsi, $X_T = \log(S_T/S_0)$ est égal en loi à $(\mu - \frac{\sigma^2}{2})T + \sigma\sqrt{T}Z$.

$$\begin{aligned} \mathbb{E}[(-X_T - z)_+] &= \mathbb{E}\left[\left(-(\mu - \frac{\sigma^2}{2})T - \sigma\sqrt{T}Z - z\right) \mathbf{1}_{\left\{Z \leq -\frac{(\mu - \frac{\sigma^2}{2})T + z}{\sigma\sqrt{T}}\right\}}\right] \\ &= -\left((\mu - \frac{\sigma^2}{2})T + z\right) \times \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-\frac{(\mu - \frac{\sigma^2}{2})T + z}{\sigma\sqrt{T}}} e^{-\frac{x^2}{2}} dx - \sqrt{\frac{\sigma^2 T}{2\pi}} \int_{-\infty}^{-\frac{(\mu - \frac{\sigma^2}{2})T + z}{\sigma\sqrt{T}}} x e^{-\frac{x^2}{2}} dx \\ &= -\left((\mu - \frac{\sigma^2}{2})T + z\right) \mathcal{N}\left(-\frac{(\mu - \frac{\sigma^2}{2})T + z}{\sigma\sqrt{T}}\right) + \sqrt{\frac{\sigma^2 T}{2\pi}} \exp\left(-\frac{((\mu - \frac{\sigma^2}{2})T + z)^2}{2\sigma^2 T}\right). \end{aligned}$$

En injectant cette expression dans (6.14), nous finissons la preuve. $\square$

Les résultats de la CVaR et VaR obtenus par les formules explicites (Méthode 1, proposition 6.7), par la méthode de minimisation directe (Méthode 2, proposition 6.8) et par la méthode d'inversion (Méthode 3, théorème 6.3) de Fourier sont résumés dans la table 6.3. La comparaison avec les résultats des formules explicites obtenues sous le modèle de Black et Scholes nous donne des erreurs numériques pour notre méthode d'inversion de Fourier de l'ordre de 10^{-3} pour la VaR et inférieures à 10^{-4} pour la CVaR.

6.3.3 Application aux modèles exponentielle-Lévy

Nous calculons la VaR et la CVaR en termes de rendement par la méthode d'inversion de Fourier présentée dans le théorème 6.3. Nous appliquons la transformée de Fourier rapide pour calculer l'intégrale avec $N = 4096$ et un pas de discrétisation de l'intégrale $\Delta u = 0.25$. Nous obtenons alors une évaluation de la fonction-objectif (fonction à minimiser) sur l'intervalle $[z_1, z_1 + (N-1)\Delta z] = [-12.55, 11.55]$ avec un pas de discrétisation $\Delta z = 6.13 \cdot 10^{-3}$ (équation (5.28)). Nous illustrons les résultats pour les trois modèles exponentielle-Lévy suivants : le modèle de Black et Scholes, le modèle diffusion-sauts de Merton et le modèle de variance gamma. Les différents

		$T = 0.5$		$T = 1$	
		$a = 0.95$	0.99	$a = 0.95$	$a = 0.99$
VaR_a(-X_T)					
	Meth. 1	0.29892	0.44349	0.39345	0.59790
	Meth. 2	0.29893	0.44349	0.39346	0.59790
	Meth. 3	0.29912	0.44332	0.39269	0.59825
CVaR_a(-X_T)					
	Meth. 1	0.38756	0.51537	0.51881	0.69956
	Meth. 2	0.38756	0.51537	0.51881	0.69956
	Meth. 3	0.38756	0.51537	0.51881	0.69956

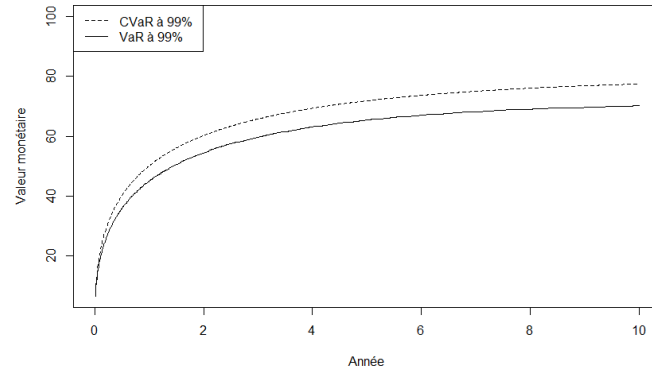
TABLE 6.3 – Comparaison des différentes méthodes de calcul de la VaR et de la CVaR en termes de rendement sous le modèle de Black et Scholes.

paramètres des modèles, ainsi que ceux de la transformée de Fourier rapide sont résumés dans la table 6.2.

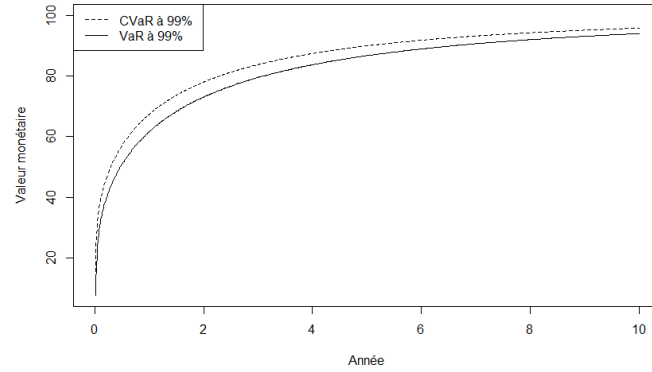
La figure 6.2a présente la VaR et la CVaR en termes de rendement sous le modèle de Black et Scholes, calculées par la méthode d'inversion de Fourier du théorème 6.3. Les résultats sous le modèle diffusion-sauts sont illustrés dans la figure 6.2b. Finalement, la figure 6.2c donne l'évolution de la VaR et la CVaR au cours du temps, sous le modèle variance gamma.

6.4 Conclusion

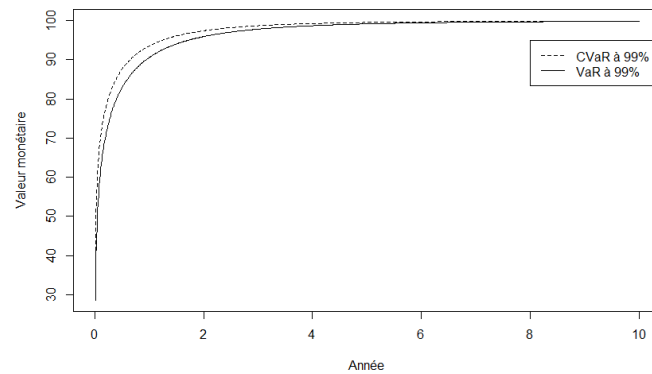
Ce chapitre a permis d'introduire la notion de la Value-at-Risk conditionnelle (CVaR) qui fournit, avec la VaR, une meilleure compréhension du risque extrême. Nous avons aussi introduit l'algorithme de Rockafellar et Uryasev, un algorithme particulièrement efficace dans le calcul de la VaR et la CVaR et qui a ouvert la porte pour l'optimisation de ces indicateurs de risque, comme nous allons le voir dans le chapitre suivant. En exploitant cet algorithme et des techniques d'inversion de Fourier, nous avons développé des méthodes numériques pour calculer simultanément la VaR et la CVaR dans le cadre des modèles exp-Lévy. Ces méthodes nécessitent la seule connaissance de la forme explicite de la fonction caractéristique du processus de Lévy. Finalement, les résultats numériques ont été illustrés pour différents modèles d'exponentielle de Lévy.



(a) La VaR et la CVaR monétaires sous un modèle de Black et Scholes de paramètres $\mu = 0.145$ et $\sigma = 0.3$.

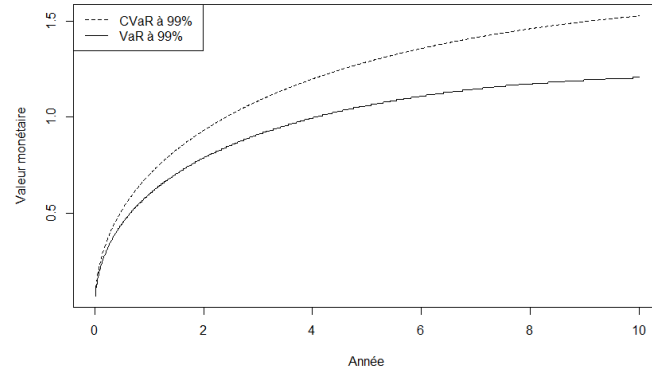


(b) La VaR et la CVaR monétaires sous un modèle de Merton de paramètres $\gamma = 0.1$, $\sigma = 0.3$, $\lambda = 1$, $m = -0.1$ et $\delta = 0.2$.

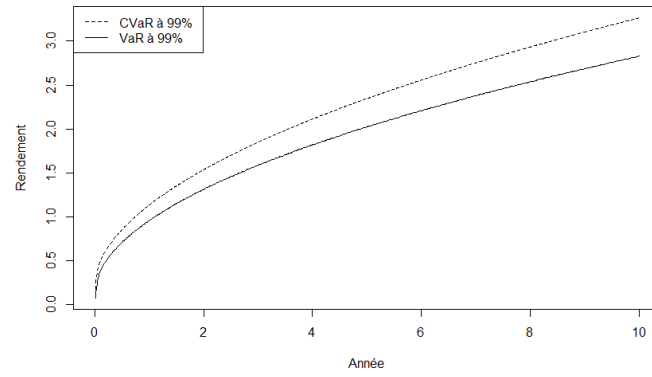


(c) La VaR et la CVaR monétaires sous un modèle variance gamma de paramètres $\gamma = 0.1$, $m = -0.01$, $\delta = 1$ et $\kappa = 0.2$.

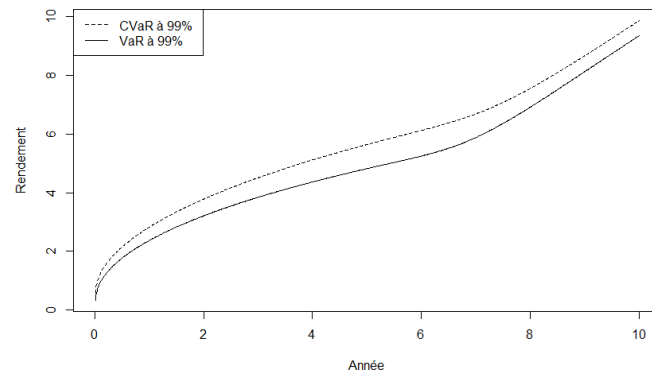
FIGURE 6.1 – Calcul de la VaR et la CVaR monétaires, à un niveau de confiance 99%, sous des modèles exponentielle-Lévy.



(a) La VaR et la CVaR en termes de rendement sous un modèle de Black et Scholes de paramètres $\mu = 0.145$ et $\sigma = 0.3$.



(b) La VaR et la CVaR en termes de rendement sous un modèle de Merton de paramètres $\gamma = 0.1$, $\sigma = 0.3$, $\lambda = 1$, $m = -0.1$ et $\delta = 0.2$.



(c) La VaR et la CVaR en termes de rendement sous un modèle variance gamma de paramètres $\gamma = 0.1$, $m = -0.01$, $\delta = 1$ et $\kappa = 0.2$.

FIGURE 6.2 – Calcul de la VaR et la CVaR en termes de rendement, à un niveau de confiance 99%, sous des modèles exponentielle-Lévy.

Chapitre 7

Minimisation de la CVaR pour le risque de couverture des options dans le cadre des modèles exp-Lévy

Dans ce chapitre, nous nous intéressons à la couverture des options sous une contrainte budgétaire. Nous considérons la CVaR pour mesurer le risque de la couverture et nous cherchons la CVaR minimale sur l'ensemble des stratégies autofinancées. Pour cela, nous proposons une approximation de la CVaR minimale fondée sur la relaxation de la contrainte du problème dual. Cette approximation est évaluée numériquement par des techniques d'inversion de Fourier. Nous commençons le chapitre par un aperçu des différentes approches de couverture des options dans le cadre des marchés incomplets. Ensuite, nous présentons, dans la section 7.2, la formulation de notre problème d'optimisation. Nous déterminons son problème dual dans la section 7.3 et nous proposons ensuite une approximation de ce dernier dans la section 7.4. La section 7.5 traite le cas d'une option d'achat européenne, où nous développons les calculs et nous illustrons les résultats pour des exemples des modèles exponentielle-Lévy.

7.1 Couverture des options dans les marchés incomplets

Un problème de base en mathématiques financières est de construire des stratégies optimales pour couvrir le risque de l'option par des opérations d'achat et de vente de l'actif sous-jacent. Cette question est bien comprise dans le cadre des marchés complets et complètement résolue par Harrison et Kreps [87]. Dans ce cas, il suffit d'acheter le portefeuille de réplication afin d'éliminer totalement le risque. Cette approche s'applique dans les cas du modèle standard de Black et Scholes ou de celui de Cox-Ross-Rubinstein, mais elle ne s'applique pas pour des modèles plus complexes, tel que les modèles à volatilité stochastique [137] et les modèles à sauts,

et qui donnent lieu à des marchés incomplets.

Sous les modèles à sauts, les problèmes deviennent plus difficiles à traiter. Lorsque la taille des sauts est constante, comme pour le processus de Poisson, le marché peut être complété par l'introduction d'autres actifs pour que les titres risqués couvrent le marché. Mais une fois que les prix des actifs ont des sauts avec des tailles imprévisibles, le marché sera toujours incomplet. Ceci se produit dès que le prix peut avoir plus d'une seule taille des sauts. L'aléa n'est plus contrôlable et les marchés ne peuvent être complétés par aucun nombre de titres négociés.

Dans les marchés incomplets, sous l'hypothèse d'absence d'opportunité d'arbitrage, il existe une infinité de mesures risque-neutres. Ceci induit un intervalle des prix risque-neutres. La réplication du produit dérivé n'existe pas dans ce contexte. Une des approches consiste alors à étudier l'intervalle des prix risque-neutres. Ceci est le concept de la sur-couverture [59, 64, 93, 98, 155]. L'idée est d'utiliser le plus petit montant pour construire un portefeuille autofinancé dont la valeur finale est plus grande que le payoff du produit dérivé dans tous les scénarios possibles. Par cette stratégie, l'opérateur a suffisamment de fonds pour couvrir toutes ses obligations futures. L'inconvénient est que la stratégie est trop coûteuse pour être d'un intérêt pratique [81]. Depuis, d'autres approches se sont développées. Elles consistent à introduire des critères subjectifs qui conduisent à des stratégies de couverture plus raisonnables.

Une première direction consiste à supprimer la contrainte que les stratégies doivent être autofinancées, tout en imposant que la valeur terminale du portefeuille doit être égale au payoff du produit dérivé. Comme les stratégies ne sont plus auto-financées, nous avons un processus de coût. La couverture par minimisation locale du risque vise à minimiser localement ce coût de manière séquentielle. Ce problème a été étudié par [69] dans le cas où le processus de prix actualisé de l'actif sous-jacent est une martingale sous $\mathbb{P}$, et puis étendu au cas des semimartingales par [67]. Il s'est avéré que les stratégies de minimisation locale du risque sont *autofinancées en moyenne*. Ce qui signifie que le processus de coût est une martingale. De là, il est possible d'obtenir des conditions nécessaires et suffisantes pour l'existence et l'unicité de ces stratégies. Ces conditions concernent essentiellement une décomposition du payoff actualisé $\tilde{H}$ en composantes orthogonales. Ceci est quelque part similaire à la décomposition de Kunita-Watanabe [100].

Alternativement, d'autres études suggèrent de garder la condition que les stratégies soient auto-financées, et essayer de minimiser, dans un certain sens, le risque terminal de couverture

$$L(V_0, \xi) = \tilde{H} - V_0 - \int_0^T \xi_t d\tilde{S}_t,$$

où $\tilde{H}$ est le payoff actualisé de l'option, $(\tilde{S}_t)_{0 \leq t \leq T}$ est le prix actualisé de l'actif sous-jacent et (V_0, ξ) est le capital initial et le processus de stratégie de couverture. Une des méthodes les plus étudiées est la couverture moyenne-variance. Elle suggère de réduire l'espérance $\mathbb{E}[L^2]$ au minimum [80, 89, 104, 159].

Malgré le développement de la couverture moyenne-variance et l'existence de so-

lutions explicites pour des modèles complexes [89], cette méthode présente des inconvénients évidents. Entre autres, la couverture moyenne-variance ne distingue pas entre les pertes et les profits et pénalise également les deux. Ainsi, d'autres critères qui se concentrent uniquement sur la perte L_+ ont été développés. L'optimisation de ces critères sur toutes les stratégies admissibles (V_0, ξ) conduit naturellement à la stratégie de sur-couverture sous laquelle L est négatif pour tous les scénarios possibles. Ainsi, ces critères ont été étudiés dans la littérature sous une contrainte budgétaire en fixant un montant maximal V_0^m pour le capital initial. Ce montant maximal est généralement beaucoup plus petit que le coût de la couverture parfaite dans le cadre des marchés complets et le coût de la sur-couverture dans le cas des marchés incomplets. Le premier critère proposé, dans ce contexte de « couverture sous une contrainte budgétaire », est la couverture par maximisation de la probabilité de succès [48, 65] qui garantit que, pour une contrainte budgétaire donnée, la probabilité d'avoir une couverture réussie (c'est-à-dire $\mathbb{P}(L \leq 0)$) est maximale. Föllmer et Leurkert résolvent ce problème dans le marché complet quand V_0^m est inférieur au prix risque-neutre [65]. Ils appliquent le lemme de Neyman-Pearson [166] pour résoudre le problème dual et donner une solution semi-explicite. Leur article analyse également le problème dans le contexte difficile des marchés incomplets.

D'autres variantes de stratégies optimales de couverture, sous une contrainte budgétaire, sont proposées dans la littérature. Par exemple, l'approche de la couverture par maximisation de la probabilité de succès a été généralisée dans [66] au problème de minimisation de la perte moyenne $\mathbb{E}[L_+]$ et plus généralement, $\mathbb{E}[\ell(L_+)]$ pour une fonction de perte $\ell(\cdot)$. Nakado [125] considère le critère de minimisation de certaines mesures de risque cohérentes, qui peuvent essentiellement être représentées comme la perte moyenne sous une mesure de probabilité particulière. Ces résultats ont été étendus pour certaines mesures de risque convexes [139]. Plus récemment, Smirnov et Melnikov [119] adoptent, comme critère, la minimisation de la CVaR dans un marché complet. Ils proposent une solution semi-explicite au problème de couverture d'une option d'achat européenne sous une contrainte budgétaire, en appliquant l'approche de Föllmer et Leurkert [66] et l'algorithme de Rockafellar et Uryasev [135]. En minimisant la VaR du risque terminal de la couverture, l'article [40] montre analytiquement que, sous ce critère de la VaR, la stratégie de couverture optimale d'une option d'achat européenne est soit une stratégie d'appel knock-out ou d'une stratégie de bull spread, en fonction des classes admissibles de stratégies de couverture. Ensuite, les résultats de [40] ont été étendus au contexte de la CVaR dans [41]. D'autres résultats connexes, sur la couverture sous une contrainte budgétaire, peuvent être trouvés dans [46, 124, 151] et leurs références. En outre, la couverture par maximisation de la probabilité de succès a été appliquée avec succès à de nombreux contrats financiers et d'assurance ; voir, par exemple, [96, 118, 150, 163].

Bien que la principale motivation derrière le développement de la couverture sous une contrainte budgétaire soit le coût élevé des stratégies de sur-couverture dans les marchés incomplets, la majorité des études, traitant ce sujet, se consacrent au cas du marché complet. Pour les études qui traitent le cas des marchés incomplets, nous citons [138–140] qui donnent la structure de la solution optimale sous différentes mesures de risque. Dans ce chapitre, nous nous intéressons à la minimisation de la

CVaR sous des modèles d'exponentielle de Lévy. Par ce travail, nous développons un calcul numérique complet et donnons une approximation de la CVaR minimale dans le cadre d'un marché incomplet.

7.2 Présentation du problème

Soit $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0, T]}, \mathbb{P})$ un espace de probabilité filtré qui satisfait les conditions usuelles. Nous considérons toujours un modèle exp-Lévy pour décrire l'évolution du prix $(S_t)_{t \in [0, T]}$ d'un actif financier. Ainsi, pour tout $t \in [0, T]$, $S_t = S_0 e^{X_t}$ avec $S_0 > 0$ et $(X_t)_{t \in [0, T]}$ est un processus de Lévy de triplet caractéristique (b, σ^2, ν) . Le prix actualisé est donné par $\tilde{S}_t = e^{-rt} S_t$ où r est le taux d'intérêt sans risque.

Soit $\mathcal{P}$ l'ensemble de toutes les mesures de probabilité risque-neutre. L'hypothèse de d'absence d'opportunité d'arbitrage peut être exprimée par l'existence d'au moins une mesure risque-neutre [52]. Cela est vérifié dès que $\sigma \neq 0$ ou $\nu \neq 0$. Sous les modèles exp-Lévy nous sommes principalement dans le cas du marché incomplet c'est-à-dire avec une infinité de mesures de probabilité risque-neutre. Tout au long de ce chapitre, nous supposons que nous sommes dans un tel cas. De plus, nous utilisons $(\mathcal{F}_t)_{t \in [0, T]}$ comme la filtration par rapport à laquelle $(S_t)_{t \in [0, T]}$ est adapté. D'autre part, les équations et les inégalités entre les variables aléatoires doivent être toujours comprises $\mathbb{P}$ presque sûrement.

Définition 7.1. Une stratégie autofinancée (V_0, ξ) est définie par un capital initial $V_0 \geq 0$ et par un processus prévisible ξ tel que le processus de la valeur actualisée du portefeuille

$$\tilde{V}_t = e^{-rt} V_t = V_0 + \int_0^t \xi_s d\tilde{S}_s, \quad \forall t \in [0, T] \quad (7.1)$$

est bien défini.

Définition 7.2. Une stratégie autofinancée (V_0, ξ) est dite admissible si le processus de valeur du portefeuille correspondant satisfait

$$V_t \geq 0, \quad \forall t \in [0, T]. \quad (7.2)$$

Nous noterons $\mathcal{A}$ l'ensemble de toutes les stratégies autofinancées admissibles.

Considérons un produit dérivé dont le payoff à l'échéance T est donné par une variable aléatoire non-négative et $\mathcal{F}_T$ -mesurable $H \in L^1$. Notons son payoff actualisé par $\tilde{H} = e^{-rT} H$. Nous supposons que

$$U_0 = \sup_{\mathbb{P}^* \in \mathcal{P}} \mathbb{E}^*[\tilde{H}] < \infty, \quad (7.3)$$

où $\mathbb{E}^*$ représente l'espérance sous la probabilité risque-neutre $\mathbb{P}^*$. La valeur U_0 représente le plus petit montant V_0 tel qu'il existe une stratégie admissible (V_0, ξ) de processus de valeur de portefeuille qui satisfait $V_T \geq H$. La *sur-couverture*

consiste à fournir une décomposition optionnelle [64, 98] de la surmartingale $U_t = \operatorname{esssup}_{\mathbb{P}^* \in \mathcal{P}} \mathbb{E}^*[\tilde{H} | \mathcal{F}_t]$ en

$$U_t = U_0 + \int_0^t \xi_s d\tilde{S}_s - C_t, \quad 0 \leq t \leq T,$$

où $(C_t)_{t \in [0, T]}$ est un processus croissant adapté avec $C_0 = 0$. La stratégie (U_0, ξ) est celle de la sur-couverture qui assure que $V_T \geq H$.

Néanmoins, la sur-couverture des produits dérivés est souvent très coûteuse dans le marché incomplet [81] et elle peut même atteindre le prix de l'actif sous-jacent dans certains cas, tel que le modèle de Merton [19, 58]. Ainsi, nous cherchons la meilleure couverture qu'un investisseur peut réaliser avec un montant initial V_0 , inférieur à U_0 , en fixant une contrainte budgétaire V_0^m , c'est-à-dire

$$V_0 \leq V_0^m < U_0.$$

Dans ce chapitre, nous nous intéressons à la minimisation de la CVaR comme critère d'optimalité de la couverture, dite partielle. Nous choisissons la CVaR car elle est une mesure cohérente du risque et elle est devenue très populaire parmi les institutions financières et les compagnies d'assurance.

Définition 7.3. Soient $\tilde{H}$ le payoff actualisé et $\tilde{V}_T$ la valeur actualisée du portefeuille de couverture à la maturité. Le coût actualisé de la stratégie de couverture est défini par la variable aléatoire $L(V_0, \xi) = \tilde{H} - \tilde{V}_T$. Si $L \geq 0$, nous parlons d'une perte et nous parlons d'un gain dans le cas contraire.

En notant $\mathcal{A}_{V_0^m}$ l'ensemble de toutes les stratégies autofinancées admissibles qui satisfait la contrainte budgétaire :

$$\mathcal{A}_{V_0^m} = \{(V_0, \xi) \mid (V_0, \xi) \in \mathcal{A}, V_0 \leq V_0^m\}, \quad (7.4)$$

notre problème d'optimisation s'écrit

$$\min_{(V_0, \xi) \in \mathcal{A}_{V_0^m}} \operatorname{CVaR}_a(L(V_0, \xi)), \quad (7.5)$$

où la fonction perte $L(V_0, \xi) = \tilde{H} - \tilde{V}_T$ correspond au coût actualisé de la couverture.

Nous appliquons par la suite le corollaire 6.1 du chapitre précédent. En introduisant la fonction

$$f(z) = z + \frac{1}{1-a} \min_{(V_0, \xi) \in \mathcal{A}_{V_0^m}} \mathbb{E}[(L(V_0, \xi) - z)_+], \quad (7.6)$$

le problème d'optimisation (7.5) se réécrit sous la forme

$$\min_{(V_0, \xi) \in \mathcal{A}_{V_0^m}} \operatorname{CVaR}_a(L(V_0, \xi)) = \min_{z \in \mathbb{R}} f(z).$$

La résolution du problème se fait en deux étapes :

1. Expliciter $f(z)$: Pour tout $z \in \mathbb{R}$, calculer $c(z)$ en trouvant $(V_0^*(z), \xi^*(z))$, tel que

$$(V_0^*(z), \xi^*(z)) \in \underset{(V_0, \xi) \in \mathcal{A}_{V_0^m}}{\operatorname{argmin}} \mathbb{E}[(L(V_0, \xi) - z)_+]. \quad (7.7)$$

2. Minimiser la CVaR : Nous résolvons le problème d'optimisation uni-dimensionnel

$$z^* \in \underset{z \in \mathbb{R}}{\operatorname{argmin}} f(z). \quad (7.8)$$

Pour conclure, nous observons que la stratégie $(V_0^*(z^*), \xi^*(z^*))$ est optimale au sens où

$$\min_{(V_0, \xi) \in \mathcal{A}_{V_0^m}} \operatorname{CVaR}_a(L(V_0, \xi)) = \operatorname{CVaR}_a(L(V_0^*(z^*), \xi^*(z^*))).$$

7.3 Problème dual

Dans cette section, nous nous concentrons sur la minimisation de $\mathbb{E}[(L(V_0, \xi) - z)_+]$ dans $\mathcal{A}_{V_0^m}$ afin d'avoir une évaluation numérique de la fonction $c(z)$. Pour traiter la contrainte $V_0 \leq V_0^m$, nous écrivons le problème (7.7) sous la forme d'un problème de test d'hypothèses qui peut être résolu en utilisant l'approche de Neyman-Pearson [166]. Nous introduisons $\mathcal{R}$ la classe des tests aléatoires définie par

$$\mathcal{R} = \{\varphi : \Omega \rightarrow [0, 1] \mid \varphi \text{ est } \mathcal{F}_T\text{-mesurable}\}.$$

L'application de la proposition 3.1 de [66] au payoff actualisé $(\tilde{H} - z)_+$ nous donne le résultat suivant.

Proposition 7.1. *Pour tout $z \in \mathbb{R}$, il existe une solution $\varphi^*(z) \in \mathcal{R}$ au problème d'optimisation*

$$\begin{aligned} \varphi^*(z) \in \underset{\varphi(z) \in \mathcal{R}}{\operatorname{argmin}} \mathbb{E}[(1 - \varphi(z))(\tilde{H} - z)_+] \\ \text{s. c. } \sup_{\mathbb{P}^* \in \mathcal{P}} \mathbb{E}^*[\varphi(z)(\tilde{H} - z)_+] \leq V_0^m. \end{aligned} \quad (7.9)$$

Théorème 7.1. *Pour tout $z \in \mathbb{R}$, nous avons l'égalité suivante*

$$\begin{aligned} \min_{(V_0, \xi) \in \mathcal{A}_{V_0^m}} \mathbb{E}[(L(V_0, \xi) - z)_+] &= \min_{\varphi(z) \in \mathcal{R}} \mathbb{E}[(1 - \varphi(z))(\tilde{H} - z)_+] \\ \text{s. c. } \sup_{\mathbb{P}^* \in \mathcal{P}} \mathbb{E}^*[\varphi(z)(\tilde{H} - z)_+] &\leq V_0^m. \end{aligned} \quad (7.10)$$

De plus, les arguments des minimums $(V_0^*(z), \xi^*(z))$ et $\varphi^*(z)$ sont reliés de la façon suivante

— Étant donné la stratégie optimale $(V_0^*(z), \xi^*(z))$, alors son ratio de succès, défini par

$$\varphi(V_0^*(z), \xi^*(z)) = \mathbf{1}_{\{\tilde{V}_T^*(z) \geq (\tilde{H} - z)_+\}} + \frac{\tilde{V}_T^*(z)}{(\tilde{H} - z)_+} \mathbf{1}_{\{\tilde{V}_T^*(z) < (\tilde{H} - z)_+\}} \quad (7.11)$$

est égal à $\varphi^*(z)$.

- Inversement, étant donné $\varphi^*(z)$, la décomposition optionnelle de $\text{esssup}_{\mathbb{P}^* \in \mathcal{P}} \mathbb{E}^*[\varphi^*(z)(\tilde{H} - z)_+ | \mathcal{F}_t]$ donne la stratégie optimale $(V_0^*(z), \xi^*(z))$.

Idée de Preuve. Comme H et V_T sont positives, nous avons

$$(L(V_0, \xi) - z)_+ = (\tilde{H} - \tilde{V}_T - z)_+ = ((\tilde{H} - z)_+ - \tilde{V}_T)_+.$$

Ainsi, la minimisation de $\mathbb{E}[(L(V_0, \xi) - z)_+]$ peut être traitée comme un problème de minimisation de la perte moyenne pour un produit dérivé de payoff actualisé $(\tilde{H} - z)_+$ qui dépend d'un paramètre réel z .

Ce dernier est résolu par application de l'approche de Föllmer et Leukert [66] à un produit dérivé de payoff actualisé $(\tilde{H} - z)_+$. $\square$

Ainsi, la fonction $f(z)$, définie par (7.6), se réécrit sous la forme

$$f(z) = z + \frac{1}{1 - \alpha} \mathbb{E}[(1 - \varphi^*(z))(\tilde{H} - z)_+], \quad (7.12)$$

où $\varphi^*(z)$ est solution de (7.9).

Sous certaines hypothèses, des généralisations du lemme de Neyman-Pearson ont été développées pour résoudre des problèmes de la forme (7.9). Nous citons par exemple [47, 138–141, 148]. Ces études montrent que la solution n'est pas unique et trouvent uniquement la structure de la solution sans pouvoir aller plus loin. Sous l'hypothèse que l'ensemble des densités $Z_{\mathbb{P}^*} = \{\frac{d\mathbb{P}^*}{d\mathbb{P}} : \mathbb{P}^* \in \mathcal{P}\}$ est compact, l'application des résultats de [141] fournit la structure du test aléatoire optimal $\varphi^*(z)$ du problème (7.9)

$$\varphi^*(z) = \begin{cases} 1 & \text{si } (\tilde{H} - z)_+ \frac{d\tilde{\mathbb{Q}}}{d\mathbb{P}} > (\tilde{H} - z)_+ \int_{\mathcal{P}} \frac{d\mathbb{P}^*}{d\mathbb{P}} d\tilde{\lambda}(\mathbb{P}^*) \\ 0 & \text{si } (\tilde{H} - z)_+ \frac{d\tilde{\mathbb{Q}}}{d\mathbb{P}} < (\tilde{H} - z)_+ \int_{\mathcal{P}} \frac{d\mathbb{P}^*}{d\mathbb{P}} d\tilde{\lambda}(\mathbb{P}^*) \end{cases} \quad (7.13)$$

et

$$\mathbb{E}^*[\varphi^*(z)(\tilde{H} - z)_+] = V_0^m, \quad \tilde{\lambda} - p.s., \quad (7.14)$$

où $(\tilde{Q}, \tilde{\lambda})$ est solution d'un certain problème dual. La construction des solutions numériques n'est pas possible à partir de ce résultat. De plus, [154] montre que l'ensemble $Z_{\mathbb{P}^*}$ n'est jamais compact, sauf si $\mathcal{P}$ est un singleton, ce qui correspond au marché complet. Nous prouvons que cette hypothèse de compacité n'est pas vérifiée par les modèles exp-Lévy et nous proposons dans la section 7.3 de relaxer la contrainte dans (7.9) afin d'avoir une approximation semi-explicite de la solution.

Proposition 7.2. Soit $(X_t)_{t \in [0, T]}$ un processus de Lévy sous $\mathbb{P}$ de triplet caractéristique (b, σ^2, ν) . L'ensemble des densités $Z_{\mathbb{P}^*} = \{\frac{d\mathbb{P}^*}{d\mathbb{P}} : \mathbb{P}^* \in \mathcal{P}\}$ est donné par

$$Z_{\mathbb{P}^*} = \left\{ e^{U_T^{(\theta, \phi)}} : \theta \in \mathbb{R} \text{ et } \phi : \mathbb{R} \mapsto \mathbb{R} \text{ tel que } \int_{\mathbb{R}} (e^{\phi(x)/2} - 1)^2 \nu(dx) < \infty \text{ et } \right. \\ \left. b + \sigma^2 \theta + \frac{\sigma^2}{2} + \int_{\mathbb{R}} (e^{\phi(x)}(e^x - 1) - x \mathbf{1}_{|x| \leq 1}) \nu(dx) = r \right\}$$

où $\left(U_t^{(\theta, \phi)}\right)_{t \in [0, T]}$ est un processus de Lévy sous $\mathbb{P}$ de triplet caractéristique

$$\left(-\frac{\sigma^2 \theta^2}{2} - \int_{\mathbb{R}} (e^x - 1 - x \mathbf{1}_{|x| \leq 1}) \nu \circ \phi^{-1}(dx), \sigma^2 \theta^2, \nu \circ \phi^{-1}\right).$$

Preuve. D'après la proposition 5.3 du chapitre 5 sur le changement de probabilité, nous pouvons décrire l'ensemble des mesures équivalentes à $\mathbb{P}$ par

$$\left\{ \mathbb{P}_{(\theta, \phi)}^* : \theta \in \mathbb{R} \text{ et } \phi : \mathbb{R} \mapsto \mathbb{R} \text{ tel que } \int_{\mathbb{R}} (e^{\phi(x)/2} - 1)^2 \nu(dx) < \infty \right\},$$

avec $\frac{d\mathbb{P}_{(\theta, \phi)}^*}{d\mathbb{P}} = e^{U_T^{(\theta, \phi)}}$ et $(U_t^{(\theta, \phi)})_{t \in [0, T]}$ un processus de Lévy sous $\mathbb{P}$ de triplet caractéristique

$$\left(-\frac{\sigma^2 \theta^2}{2} - \int_{\mathbb{R}} (e^x - 1 - x \mathbf{1}_{|x| \leq 1}) \nu \circ \phi^{-1}(dx), \sigma^2 \theta^2, \nu \circ \phi^{-1}\right).$$

Sous $\mathbb{P}_{(\theta, \phi)}^*$, $(X_t)_{t \in [0, T]}$ est un processus de Lévy de triplet caractéristique $(b_{(\theta, \phi)}^*, \sigma^2, \nu_{(\theta, \phi)}^*)$ tel que

$$\nu_{(\theta, \phi)}^*(dx) = e^{\phi(x)} \nu(dx)$$

et

$$b_{(\theta, \phi)}^* = b + \sigma^2 \theta + \int_{|x| \leq 1} x \nu_{(\theta, \phi)}^*(dx) - \int_{|x| \leq 1} x \nu(dx).$$

Or, d'après (5.4) du corollaire 5.1, $(\tilde{S}_t)_{t \in [0, T]}$ est une martingale sous $\mathbb{P}^*$ si et seulement si

$$b_{(\theta, \phi)}^* + \frac{1}{2} \sigma^2 + \int_{\mathbb{R}} (e^x - 1 - x \mathbf{1}_{|x| \leq 1}) \nu_{(\theta, \phi)}^*(dx) = r,$$

donc

$$b + \sigma^2 \theta + \frac{\sigma^2}{2} + \int_{\mathbb{R}} (e^{\phi(x)}(e^x - 1) - x \mathbf{1}_{|x| \leq 1}) \nu(dx) = r.$$

Nous concluons que

$$Z_{\mathbb{P}^*} = \left\{ e^{U_T^{(\theta, \phi)}} : \theta \in \mathbb{R} \text{ et } \phi : \mathbb{R} \mapsto \mathbb{R} \text{ tel que } \int_{\mathbb{R}} (e^{\phi(x)/2} - 1)^2 \nu(dx) < \infty \text{ et } b + \sigma^2 \theta + \frac{\sigma^2}{2} + \int_{\mathbb{R}} (e^{\phi(x)}(e^x - 1) - x \mathbf{1}_{|x| \leq 1}) \nu(dx) = r \right\}.$$

Ce qui finit la preuve. □

Nous voyons bien d'après cette proposition que si $\nu \neq 0$ alors l'ensemble $Z_{\mathbb{P}^*}$ est infini puisque nous avons une infinité de choix pour la fonction ϕ . Donc il ne peut être compact que dans le cas $\nu = 0$. Inversement, si $\nu = 0$ alors l'ensemble $Z_{\mathbb{P}^*}$ se réduit à un singleton, il est donc compact. Cela est résumé dans le corollaire suivant.

Corollaire 7.1. Soit $(X_t)_{t \in [0, T]}$ un processus de Lévy sous $\mathbb{P}$ de triplet caractéristique (b, σ^2, ν) . L'ensemble des densités $Z_{\mathbb{P}^*} = \left\{ \frac{d\mathbb{P}^*}{d\mathbb{P}} : \mathbb{P}^* \in \mathcal{P} \right\}$ est un compact si et seulement si $\nu = 0$. Dans ce cas, il se réduit à un singleton

$$Z_{\mathbb{P}^*} = \left\{ \exp \left(-\frac{\sigma^2 \theta^2}{2} T + \sigma \theta B_T \right) \right\},$$

avec $\theta = (r - b - \frac{\sigma^2}{2})/\sigma^2$ et $(B_t)_{t \in [0, T]}$ est un mouvement brownien standard.

Ainsi, sous les modèles exponentielle-Lévy, l'ensemble des densités $Z_{\mathbb{P}^*}$ n'est compact que dans le cas des diffusions, à l'instar du modèle de Black et Scholes. Ces modèles correspondent au marché complet et à l'unicité de la mesure risque-neutre et peuvent être résolus par le simple lemme de Neyman-Pearson. Ainsi, les généralisations proposées de ce lemme, qui traitent le cas où il existe une infinité de mesures équivalentes, ne sont pas applicables dans le cadre de la finance où l'hypothèse de compacité requise n'est pas vérifiée dans les marchés incomplets.

7.4 Approximation du problème dual

Dans cette section, nous proposons de relaxer la contrainte $\sup_{\mathbb{P}^* \in \mathcal{P}} \mathbb{E}^*[\varphi(z)(\tilde{H} - z)_+] \leq V_0^m$ dans (7.9). Au lieu de considérer le supremum sur toutes les mesures risque-neutre, nous fixons une mesure risque-neutre $\mathbb{P}^*$ et nous optimisons sous la contrainte $\mathbb{E}^*[\varphi(z)(\tilde{H} - z)_+] \leq V_0^m$ où $\mathbb{E}^*$ est l'espérance sous la mesure choisie $\mathbb{P}^*$. Nous optons pour la mesure risque-neutre d'Esscher pour la possibilité de calcul explicite qu'elle fournit. Nous rappelons que cette mesure est donnée par

$$\frac{d\mathbb{P}^*}{d\mathbb{P}} = \frac{e^{\theta X_T}}{\mathbb{E}[e^{\theta X_T}]},$$

où $\theta \in \mathbb{R}$ est solution de

$$b + \sigma^2 \theta + \frac{\sigma^2}{2} + \int_{\mathbb{R}} (e^{\theta x}(e^x - 1) - x \mathbf{1}_{|x| \leq 1}) \nu(dx) = r.$$

En termes de la proposition 7.2, la mesure risque-neutre d'Esscher correspond au choix $\phi(x) = \theta x$. Une approximation du problème (7.9) est donnée par

$$\begin{aligned} \hat{\varphi}(z) = \underset{\varphi(z) \in \mathcal{R}}{\operatorname{argmin}} \quad & \mathbb{E}[(1 - \varphi(z))(\tilde{H} - z)_+] \\ \text{s. c.} \quad & \mathbb{E}^*[\varphi(z)(\tilde{H} - z)_+] \leq V_0^m, \end{aligned} \tag{7.15}$$

où $\mathbb{E}^*$ est l'espérance sous la mesure risque-neutre d'Esscher $\mathbb{P}^*$.

Hypothèse 7.1. Nous supposons que $V_0^m < \hat{U}_0 = \mathbb{E}^*[\tilde{H}]$, pour ne pas considérer le cas trivial où $\hat{\varphi}(z) = 1$ pour tout $z \leq 0$ et $\widehat{\text{CVaR}}_a$ est égale à zéro.

La solution de ce problème peut être dérivée du lemme de Neyman-Pearson.

Théorème 7.2. *La solution du problème (7.15) est donnée par*

$$\hat{\varphi}(z) = \mathbf{1}_{\{e^{-\theta X_T} > c(z)\}}$$

où

$$c(z) = \inf \left\{ u \geq 0 : \mathbb{E}^* \left[(\tilde{H} - z)_+ \cdot \mathbf{1}_{\{e^{-\theta X_T} > u\}} \right] \leq V_0^m \right\}.$$

Preuve. En définissant les mesures $\mathbb{Q}$ et $\mathbb{Q}^*$ par

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \frac{(\tilde{H} - z)_+}{\mathbb{E}[(\tilde{H} - z)_+]} \quad \text{et} \quad \frac{d\mathbb{Q}^*}{d\mathbb{P}^*} = \frac{(\tilde{H} - z)_+}{\mathbb{E}^*[(\tilde{H} - z)_+]},$$

le problème d'optimisation

$$\begin{aligned} \operatorname{argmin}_{\varphi(z) \in \mathcal{R}} \mathbb{E}[(1 - \varphi(z))(\tilde{H} - z)_+] &= \operatorname{argmax}_{\varphi(z) \in \mathcal{R}} \mathbb{E}[\varphi(z)(\tilde{H} - z)_+] \\ \text{s. c. } \mathbb{E}^*[\varphi(z)(\tilde{H} - z)_+] &\leq V_0^m, \end{aligned}$$

prend la forme

$$\begin{aligned} \operatorname{argmax}_{\varphi(z) \in \mathcal{R}} \int \varphi(z) d\mathbb{Q} \\ \text{s. c. } \int \varphi(z) d\mathbb{Q}^* \leq \alpha(z) &:= \frac{V_0^m}{\mathbb{E}^*[(\tilde{H} - z)_+]}. \end{aligned}$$

Cette dernière forme est celle d'un problème d'optimisation permettant de choisir le meilleur test d'hypothèses entre deux mesures $\mathbb{Q}$ et $\mathbb{Q}^*$. Ce dernier est résolu par le lemme de Neyman-Pearson [47, 166] et la solution est donnée par

$$\hat{\varphi}(z) = \mathbf{1}_{\{\frac{d\mathbb{Q}}{d\mathbb{Q}^*} > c''(z)\}} + \gamma(z) \cdot \mathbf{1}_{\{\frac{d\mathbb{Q}}{d\mathbb{Q}^*} = c''(z)\}}, \quad (7.16)$$

où

$$c''(z) = \inf \left\{ u \geq 0 : \mathbb{Q}^* \left(\frac{d\mathbb{Q}}{d\mathbb{Q}^*} > u \right) \leq \alpha(z) \right\}$$

et

$$\gamma(z) = \frac{\alpha(z) - \mathbb{Q}^* \left(\frac{d\mathbb{Q}}{d\mathbb{Q}^*} > c''(z) \right)}{\mathbb{Q}^* \left(\frac{d\mathbb{Q}}{d\mathbb{Q}^*} = c''(z) \right)},$$

avec la convention $0/0 = 0$.

En posant $c'(z) = c''(z) \frac{\mathbb{E}[(\tilde{H} - z)_+]}{\mathbb{E}^*[(\tilde{H} - z)_+]}$ et $c(z) = \frac{c'(z)}{\mathbb{E}[e^{\theta X_T}]}$, nous avons les égalités suivantes

$$\left\{ \frac{d\mathbb{Q}}{d\mathbb{Q}^*} > c''(z) \right\} = \left\{ \frac{d\mathbb{P}}{d\mathbb{P}^*} > c'(z) \right\} = \{e^{-\theta X_T} > c(z)\}$$

et

$$\begin{aligned}
\mathbb{Q}^* \left(\frac{d\mathbb{Q}}{d\mathbb{Q}^*} > c''(z) \right) &= \int \mathbf{1}_{\left\{ \frac{d\mathbb{Q}}{d\mathbb{Q}^*} > c''(z) \right\}} d\mathbb{Q}^* \\
&= \int \mathbf{1}_{\{e^{-\theta X_T} > c(z)\}} \frac{(\tilde{H} - z)_+}{\mathbb{E}^*[(\tilde{H} - z)_+]} d\mathbb{P}^* \\
&= \frac{\mathbb{E}^* \left[(\tilde{H} - z)_+ \cdot \mathbf{1}_{\{e^{-\theta X_T} > c(z)\}} \right]}{\mathbb{E}^*[(\tilde{H} - z)_+]}.
\end{aligned}$$

On en déduit facilement que

$$c(z) = \inf \left\{ u \geq 0 : \mathbb{E}^* \left[(\tilde{H} - z)_+ \cdot \mathbf{1}_{\{e^{-\theta X_T} > u\}} \right] \leq V_0^m \right\}.$$

D'autre part, la continuité stochastique du processus de Lévy $(X_t)_{t \in [0, T]}$ nous assure que la probabilité d'apparition d'une valeur fixe est nulle. Ainsi,

$$\mathbb{Q}^* \left(\frac{d\mathbb{Q}}{d\mathbb{Q}^*} = c''(z) \right) = \frac{\mathbb{E}^* \left[(\tilde{H} - z)_+ \cdot \mathbf{1}_{\{e^{-\theta X_T} = c(z)\}} \right]}{\mathbb{E}^*[(\tilde{H} - z)_+]} = 0.$$

D'où, $\mathbb{Q}^* \left(\frac{d\mathbb{Q}}{d\mathbb{Q}^*} > c''(z) \right) = \alpha(z)$ et par conséquent $\gamma(z) = 0$.

Nous montrons ainsi que $\hat{\varphi}(z) = \mathbf{1}_{\{e^{-\theta X_T} > c(z)\}}$ et nous concluons la preuve. $\square$

Nous introduisons l'approximation $\hat{f}(z)$ de $f(z)$ par

$$\hat{f}(z) = z + \frac{1}{1-a} \mathbb{E} \left[(1 - \hat{\varphi}(z)) (\tilde{H} - z)_+ \right], \quad (7.17)$$

où $\hat{\varphi}(z)$ est donné par le théorème 7.2, et nous nous intéressons au problème approché de la minimisation de la CVaR

$$\min_{(V_0, \xi) \in \mathcal{A}_{V_0^m}} \widehat{\text{CVaR}}_a(L(V_0, \xi)) = \min_{z \in \mathbb{R}} \hat{f}(z). \quad (7.18)$$

7.5 Cas d'une option d'achat européenne

Dans cette section, nous étudions le cas d'une option d'achat européenne de prix d'exercice K et de date de maturité T . Le payoff de cette option est donnée par $H = (S_T - K)_+$.

7.5.1 Résolution du problème approché

La résolution du problème approché de la minimisation de la CVaR (7.18) est donnée par le théorème suivant, qui ouvre la porte à l'évaluation numérique par inversion de Fourier.

Théorème 7.3. Soit $(X_t)_{t \in [0, T]}$ un processus de Lévy sous $\mathbb{P}$ de triplet caractéristique (b, σ^2, ν) et soit $\mathbb{P}^*$ la mesure risque-neutre d'Esscher de paramètre θ . La solution du problème approximatif de la minimisation de la CVaR (7.18) est donnée, en fonction du signe de θ , par

- Si $\theta < 0$,

$$\min_{(V_0, \xi) \in \mathcal{A}_{V_0^m}} \widehat{\text{CVaR}}_a(\tilde{H} - \tilde{V}_T) = \min_{z \in (0, z_c)} \left\{ z + \frac{1}{1-a} \left(\mathbb{E}[(\tilde{S}_T - K(z))_+] - \mathbb{E}[(\tilde{S}_T - c_1(z))_+] - (c_1(z) - K(z))\mathbb{P}(\tilde{S}_T > c_1(z)) \right) \right\}, \quad (7.19)$$

où $K(z) = Ke^{-rT} + z$, $z_c > 0$ tel que $\mathbb{E}^*[(\tilde{S}_T - K(z_c))_+] = V_0^m$ et

$$c_1(z) = \inf \left\{ u \geq K(z), \mathbb{E}^*[(\tilde{S}_T - u)_+] + (u - K(z))\mathbb{P}^*(\tilde{S}_T > u) \leq V_0^m \right\}. \quad (7.20)$$

- Si $\theta > 0$,

$$\min_{(V_0, \xi) \in \mathcal{A}_{V_0^m}} \widehat{\text{CVaR}}_a(\tilde{H} - \tilde{V}_T) = \min_{z \in (0, z_c)} \left\{ z + \frac{1}{1-a} \left(\mathbb{E}[(\tilde{S}_T - c_1(z))_+] + (c_1(z) - K(z))\mathbb{P}(\tilde{S}_T > c_1(z)) \right) \right\} \quad (7.21)$$

où $K(z) = Ke^{-rT} + z$, $z_c > 0$ tel que $\mathbb{E}^*[(\tilde{S}_T - K(z_c))_+] = V_0^m$ et

$$c_1(z) = \sup \left\{ u \geq K(z), \mathbb{E}^*[(\tilde{S}_T - K(z))_+] - \mathbb{E}^*[(\tilde{S}_T - u)_+] - (u - K(z))\mathbb{P}^*(\tilde{S}_T > u) \leq V_0^m \right\}. \quad (7.22)$$

Preuve. Préliminaires. La fonction $\mathbb{E}^*[(\tilde{H} - z)_+]$ est continue et décroissante en z , égale à $\mathbb{E}^*[\tilde{H}]$ qui est supérieure strictement à V_0^m , pour $z = 0$ et diverge vers $-\infty$ lorsque $z \rightarrow +\infty$. Donc, il existe $z_c > 0$ tel que $\mathbb{E}^*[(\tilde{H} - z_c)_+] = V_0^m$.

Pour $z \geq z_c$, $\mathbb{E}^*[(\tilde{H} - z)_+] \leq V_0^m$, donc $\hat{\varphi}(z) = 1$ et $\hat{f}(z) = z$. Ainsi, $\min_{z \in \mathbb{R}} \hat{f}(z)$ ne peut pas être atteint en $(z_c, +\infty)$ puisque la fonction $\hat{f}$ est croissante sur cet intervalle. De plus, nous avons vu que la VaR approchée est le plus petit $\arg\min_{z \in \mathbb{R}} \hat{f}(z)$. Or la VaR est toujours positive, donc l'ensemble $\arg\min_{z \in \mathbb{R}} \hat{f}(z)$ est inclus dans $(0, z_c)$.

Soit $z \in (0, z_c)$, alors

- En posant $K(z) = Ke^{-rT} + z$, nous avons

$$(\tilde{H} - z)_+ = ((\tilde{S}_T - Ke^{-rT})_+ - z)_+ = (\tilde{S}_T - (Ke^{-rT} + z))_+ = (\tilde{S}_T - K(z))_+.$$

- Pour tout $u \geq 0$, nous avons

$$\mathbb{E} \left[(\tilde{H} - z)_+ \mathbf{1}_{\{\tilde{S}_T > u\}} \right] = \begin{cases} \mathbb{E}[(\tilde{S}_T - K(z))_+] & \text{si } u < K(z), \\ \mathbb{E}[(\tilde{S}_T - u)_+] + (u - K(z))\mathbb{P}(\tilde{S}_T > u) & \text{si } u \geq K(z). \end{cases} \quad (7.23)$$

- Pour tout $u \geq 0$, nous avons

$$\begin{aligned} & \mathbb{E} \left[(\tilde{H} - z)_+ \mathbf{1}_{\{\tilde{S}_T < u\}} \right] \\ &= \begin{cases} 0 & \text{si } u < K(z), \\ \mathbb{E}[(\tilde{S}_T - K(z))_+] - \mathbb{E}[(\tilde{S}_T - u)_+] - (u - K(z))\mathbb{P}(\tilde{S}_T > u) & \text{si } u \geq K(z). \end{cases} \end{aligned} \quad (7.24)$$

Notons que cela ne dépend pas de la mesure $\mathbb{P}$ et nous aurons les mêmes expressions sous $\mathbb{P}^*$.

Cœur de la preuve. Soit $z \in (0, z_c)$. La solution $\hat{\varphi}(z) = \mathbf{1}_{\{e^{-\theta X_T} > c(z)\}}$ du problème approché (7.9), où $c(z) = \inf \left\{ u \geq 0 : \mathbb{E}^* \left[(\tilde{H} - z)_+ \cdot \mathbf{1}_{\{e^{-\theta X_T} > u\}} \right] \leq V_0^m \right\}$, dépend du signe du paramètre d'Esscher θ . Nous devons alors traiter séparément les deux cas $\theta > 0$ et $\theta < 0$.

1^{er} cas : $\theta < 0$.

Nous montrons facilement que $\hat{\varphi}(z)$ s'écrit en fonction de $\tilde{S}_T$ comme suit

$$\hat{\varphi}(z) = \mathbf{1}_{\{\tilde{S}_T > c_1(z)\}},$$

où

$$c_1(z) = S_0 e^{-rT} (c(z))^{-1/\theta} = \inf \left\{ u \geq 0 : \mathbb{E}^* \left[(\tilde{H} - z)_+ \cdot \mathbf{1}_{\{\tilde{S}_T > u\}} \right] \leq V_0^m \right\}.$$

L'espérance $\mathbb{E}^* \left[(\tilde{H} - z)_+ \cdot \mathbf{1}_{\{\tilde{S}_T > u\}} \right]$ prend la forme de (7.23). Or $\mathbb{E}^*[(\tilde{S}_T - K(z))_+] = \mathbb{E}^*[(\tilde{H} - z)_+] > V_0^m$ pour $z \in (0, z_c)$ vu que cette espérance est décroissante en z et égale à V_0^m en z_c . Ainsi,

$$c_1(z) = \inf \left\{ u \geq K(z) : \mathbb{E}^*[(\tilde{S}_T - u)_+] + (u - K(z))\mathbb{P}^*(\tilde{S}_T > u) \leq V_0^m \right\}.$$

La fonction $\hat{f}(z)$ est égale ainsi à

$$\hat{f}(z) = z + \frac{1}{1-a} \mathbb{E} \left[\mathbf{1}_{\{\tilde{S}_T < c_1(z)\}} \cdot (\tilde{H} - z)_+ \right].$$

Comme $c_1(z) \geq K(z)$, alors d'après (7.24), nous avons

$$\begin{aligned} \hat{f}(z) = z + \frac{1}{1-a} & \left(\mathbb{E}[(\tilde{S}_T - K(z))_+] - \mathbb{E}[(\tilde{S}_T - c_1(z))_+] \right. \\ & \left. - (c_1(z) - K(z))\mathbb{P}(\tilde{S}_T > c_1(z)) \right). \end{aligned}$$

Comme la minimisation $\min_{(V_0, \xi) \in \mathcal{A}_{V_0^m}} \widehat{\text{CVaR}}_a(\tilde{H} - \tilde{V}_T) = \min_{z \in (0, z_c)} \hat{f}(z)$, nous obtenons le résultat (7.19) pour $\theta < 0$.

2^{ème} cas : $\theta > 0$.

Nous montrons facilement que $\hat{\varphi}(z)$ s'écrit en fonction de $\tilde{S}_T$ comme suit

$$\hat{\varphi}(z) = \mathbf{1}_{\{\tilde{S}_T < c_1(z)\}},$$

où

$$c_1(z) = S_0 e^{-rT} (c(z))^{-1/\theta} = \sup \left\{ u \geq 0 : \mathbb{E}^* \left[(\tilde{H} - z)_+ \cdot \mathbf{1}_{\{\tilde{S}_T < u\}} \right] \leq V_0^m \right\}.$$

L'espérance $\mathbb{E}^* \left[(\tilde{H} - z)_+ \cdot \mathbf{1}_{\{\tilde{S}_T < u\}} \right]$ prend la forme de (7.24). Ainsi,

$$c_1(z) = \sup \left\{ u \geq K(z) : \mathbb{E}^*[(\tilde{S}_T - K(z))_+] - \mathbb{E}^*[(\tilde{S}_T - u)_+] - (u - K(z))\mathbb{P}^*(\tilde{S}_T > u) \leq V_0^m \right\}.$$

La fonction $\hat{f}(z)$ est égale ainsi à

$$\hat{f}(z) = z + \frac{1}{1-a} \mathbb{E} \left[\mathbf{1}_{\{\tilde{S}_T > c_1(z)\}} \cdot (\tilde{H} - z)_+ \right].$$

Comme $c_1(z) \geq K(z)$, alors d'après (7.23), nous avons

$$\hat{f}(z) = z + \frac{1}{1-a} \left(\mathbb{E}[(\tilde{S}_T - c_1(z))_+] + (c_1(z) - K(z))\mathbb{P}(\tilde{S}_T > c_1(z)) \right).$$

Comme $\min_{(V_0, \xi) \in \mathcal{A}_{V_0^m}} \widehat{\text{CVaR}}_a(\tilde{H} - \tilde{V}_T) = \min_{z \in (0, z_c)} \hat{f}(z)$, nous obtenons le résultat (7.21) pour $\theta > 0$. □

L'atout du théorème 7.3 est que, si on a des expressions explicites de Φ_T et Φ_T^* , les fonctions caractéristiques de X_T sous $\mathbb{P}$ et sous $\mathbb{P}^*$, les différentes quantités de type $\mathbb{E}[(\tilde{S}_T - x)_+]$, $\mathbb{P}(\tilde{S}_T > x)$, $\mathbb{E}^*[(\tilde{S}_T - x)_+]$ et $\mathbb{P}^*(\tilde{S}_T > x)$ peuvent être calculées par la méthode d'inversion de Fourier, proposée dans le chapitre 5. Ainsi, nous pouvons procéder à une minimisation numérique du problème. Dans la section suivante, nous illustrons les résultats de ce problème dans le cadre du modèle diffusion-sauts de Merton.

7.5.2 Illustration des résultats

Dans cette section nous illustrons les résultats de la minimisation de la CVaR sous le modèle de Black et Scholes et sous le modèle diffusion-sauts de Merton. Les paramètres des différents modèles, ainsi que de la FFT, sont présentés dans la table 7.1.

Pour chaque modèle, la CVaR minimale approchée est illustrée en fonction de la proportion disponible du capital initial U_0 requise pour effectuer une sur-couverture. La figure 7.1 présente le résultat sous le modèle de Black et Scholes. Pour ce modèle,

Marché	B&S	Merton	FFT
$S_0 = 100$	$\mu = 0.145$	$\gamma = 0.1$	$N = 4096$
$r = 0.02$	$\sigma = 0.3$	$\sigma = 0.3$	$\Delta v = 0.25$
$T = 0.5$		$\lambda = 1$	
$K = 110$		$m = -0.1$	$\Delta k = \frac{2\pi}{N\Delta v}$
		$\delta = 0.2$	$k_1 = -\frac{N}{2}\Delta k$

TABLE 7.1 – Sommaire de différentes valeurs numériques des paramètres utilisés dans la section 7.5.2.

le marché est complet et nous obtenons la solution exacte. La figure 7.2 donne la CVaR minimale approchée en fonction de la proportion du U_0 et en fonction du niveau de confiance. Nous remarquons que la CVaR approchée sous le modèle de Merton a la même allure de graphe que la CVaR exacte sous le modèle de Black et Scholes. D'autre part, nous constatons que si le capital initial permis V_0^m est supérieur à 20% de U_0 alors la CVaR minimale approchée est la même quel que soit le niveau de confiance. Cela s'explique par le fait que pour des grandes valeurs de V_0^m , le minimum de $\hat{f}(z)$ est atteint en z_c , et il ne dépend pas de a .

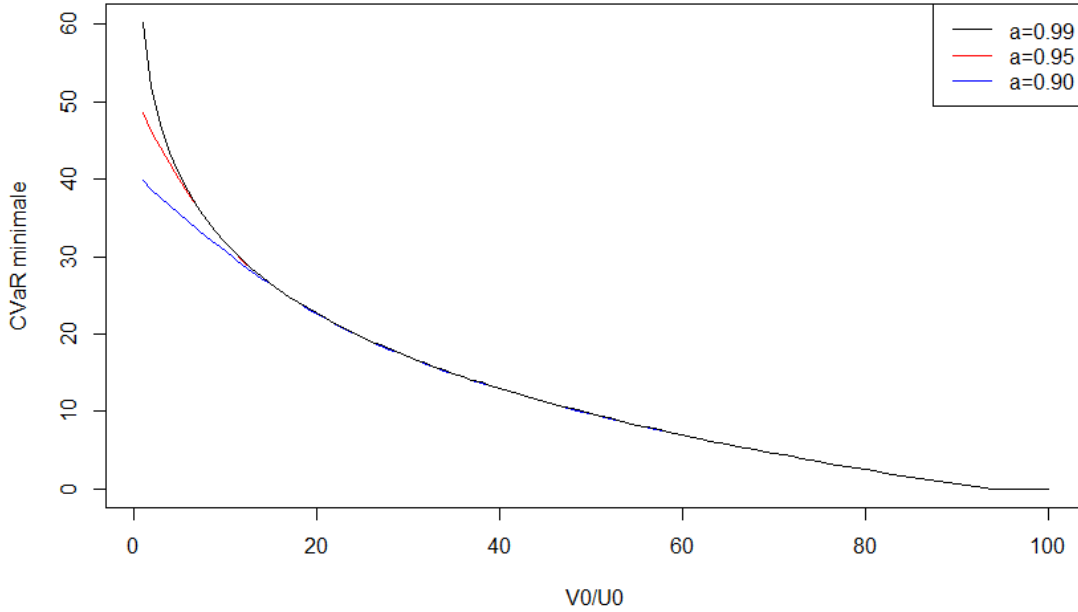


FIGURE 7.1 – CVaR minimale en fonction de la proportion du capital initial permis $V_0^m/\hat{U}_0$, sous le modèle de Black et Scholes.

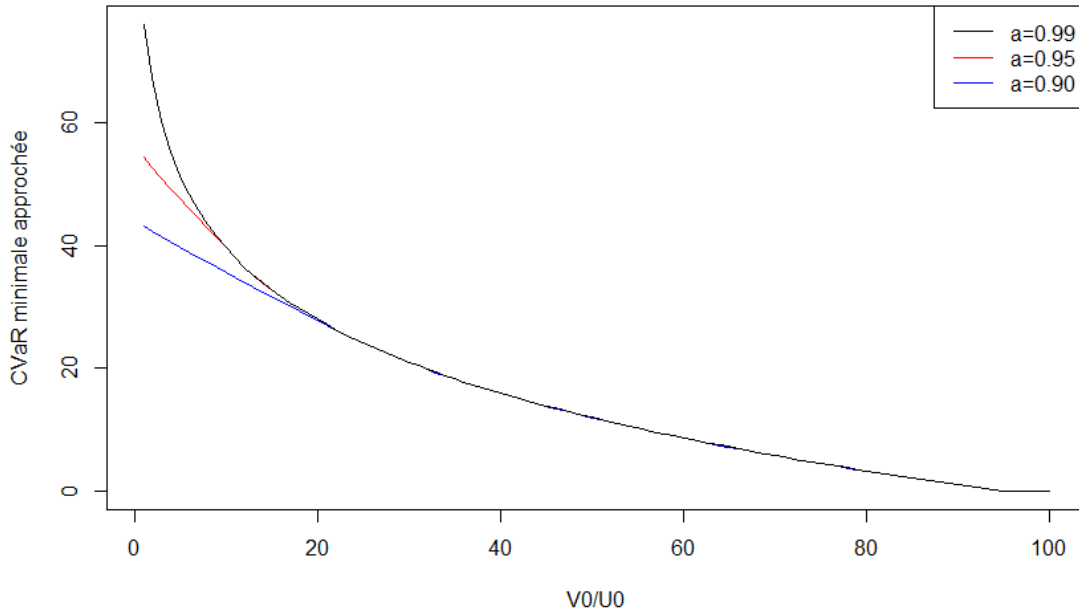


FIGURE 7.2 – CVaR minimale approchée en fonction de la proportion du capital initial permis $V_0^m/\hat{U}_0$, sous le modèle diffusion-sauts de Merton.

7.6 Conclusion

Dans ce chapitre, nous avons étudié la minimisation de la CVaR du risque de couverture des options européennes dans le cadre des modèles de types exp-Lévy. Ces modèles correspondent généralement à des marchés incomplets où la résolution de ce problème de minimisation n'est pas possible avec les généralisations du lemme Neyman-Pearson. Nous avons montré que les hypothèses de ces généralisations ne sont pas vérifiées sous les modèles exp-Lévy et nous avons traité un problème approché en relaxant la contrainte du problème dual et en ne considérant que la mesure risque-neutre d'Esscher. Ce problème approché est résolu et l'évaluation numérique est faite par la méthode d'inversion de Fourier. Les résultats sont illustrés sous le modèle de Black et Scholes et le modèle diffusion-sauts de Merton pour plusieurs niveaux de confiance de l'estimateur de la CVaR.

Conclusion et perspectives

Dans la première partie de la thèse, nous avons construit un modèle pour la queue de distribution des pertes d'un actif boursier, que nous confrontons à des données réelles issues de la bourse de Paris (Euronext Paris). Notre modèle prend en compte les probabilités d'occurrence des pertes extrêmes et les changements de régimes. Dans un premier temps, nous avons montré empiriquement que l'estimation de la Value-at-Risk sous le modèle des lois puissances présente un meilleur résultat que les lois normales et les lois stables. Cela confirme les observations déjà faites par [20, 57, 90]. Cependant, nous avons constaté des erreurs d'estimation à chaque passage d'une période de grandes fluctuations (crise) à une période de petites fluctuations (hors-crise). Pour corriger cela, notre approche a été de détecter les différentes périodes de chaque régime, par la construction d'une chaîne de Markov cachée, et d'estimer la queue de distribution de chaque régime par des lois puissances. La chaîne d'états a été construite avec un apprentissage non-supervisé sans aucun a priori sur la structure des probabilités de transition. Notre modèle a permis de supprimer les erreurs d'estimation liées aux différents changements de régimes et de fournir une meilleure performance de prédiction de la Value-at-Risk comparée à celles des modèles des lois stables, des lois puissances et GARCH.

Comme perspectives de cette partie, nous poursuivrons l'étude de la répartition des périodes de crises. Le but est de pouvoir regrouper les actifs boursiers par bloc de comportements semblables et de proposer une seule chaîne d'états pour chaque groupe d'actifs. Cela apportera un gain énorme en termes de temps de calcul et rendra notre modèle plus adapté à l'utilisation en pratique. Une deuxième direction de recherche est la mesure de risque d'un portefeuille, représenté comme une somme pondérée de plusieurs actifs. Quelques études [126, 127, 168] se sont intéressées à cette question en étudiant la distribution de la somme des variables aléatoires i.i.d. suivant une loi de Pareto. Nous nous proposons d'étudier cette question, dans le but de fournir une estimation de la Value-at-Risk d'un portefeuille sous le modèle des lois puissances à changement de régimes, et par la suite la composition de portefeuille optimal en termes de Value-at-Risk.

Dans la deuxième partie, nous nous sommes placés dans le cadre des modèles exp-Lévy, sous lesquels les prix boursiers sont modélisés par des exponentielles de processus de Lévy. Un premier résultat de cette partie est le développement de deux méthodes numériques de calcul des Value-at-Risk (VaR) et des Value-at-Risk Conditionnelles (CVaR), monétaires et en termes de rendements. Que ce soit en termes

de monnaie ou de rendement, le principe de la méthode est d'exprimer la VaR et la CVaR comme respectivement l'argument du minimum et le minimum d'une certaine fonction que l'on évalue numériquement par des techniques d'inversion de Fourier. Le premier avantage de notre méthodologie est qu'elle nécessite la seule connaissance de la fonction caractéristique du processus de Lévy. Le deuxième est qu'elle évalue simultanément la VaR et la CVaR. Ceci représente un atout par rapport à d'autres méthodes qui les calculent séparément en évaluant d'abord la fonction de densité par inversion de Fourier [28, 94, 95]. Dans un deuxième temps, nous nous sommes intéressés à la minimisation du risque de couverture des options européennes, sous une contrainte budgétaire sur le capital initial. En mesurant ce risque par la CVaR, qui est plus adaptée pour la gestion de portefeuille, nous avons établi une équivalence entre ce problème et des problèmes de type Neyman-Pearson. Comme les marchés sont incomplets sous les modèles exp-Lévy, une solution explicite n'était pas possible. Nous avons proposé ainsi une approximation numérique s'appuyant sur la relaxation de la contrainte du problème d'optimisation dual.

La suite directe des travaux de cette partie est l'analyse approfondie des erreurs d'approximation. Pour les méthodes de calcul de la VaR et la CVaR, nous comptons comparer leurs résultats à ceux d'un calcul par Monte Carlo et d'autres méthodes de calcul mentionnées ci-dessus. Nous envisageons aussi de contrôler les erreurs de la minimisation et celles de l'inversion de Fourier en proposant des inégalités de majoration. Pour le problème de minimisation de la CVaR, nous travaillons sur la quantification des erreurs d'approximation liées au choix de la mesure risque-neutre d'Esscher. Nous réfléchissons aussi aux possibilités de choisir d'autres mesures risque-neutres, voire un ensemble de mesures risque-neutres paramétrées, pour relaxer la contrainte du problème dual.

À plus long terme, nous comptons mettre en pratique ces méthodes. En calibrant des modèles exp-Lévy sur des données réelles, nous pouvons valider empiriquement l'adéquation de ces modèles à sauts pour la mesure de risque par le biais des backtests sur la VaR et la CVaR.

Bibliographie

- [1] C. ACERBI et D. TASCHE : On the coherence of expected shortfall. *Journal of Banking & Finance*, 26(7):1487–1503, 2002.
- [2] A. AGRESTI et B. A. COULL : Approximate is better than “exact” for interval estimation of binomial proportions. *The American Statistician*, 52(2):119–126, 1998.
- [3] S. ALFARANO et T. LUX : Extreme value theory as a theoretical background for power law behavior. Rap. tech., Kiel working paper, 2010.
- [4] A. ANG et G. BEKAERT : Regime switches in interest rates. *Journal of Business & Economic Statistics*, 20(2):pp. 163–182, 2002.
- [5] D. APPLEBAUM : *Lévy processes and stochastic calculus*. Cambridge University Press, 2009.
- [6] P. ARTZNER, F. DELBAEN, J.-M. EBER et D. HEATH : Coherent measures of risk. *Mathematical finance*, 9(3):203–228, 1999.
- [7] P. ARTZNER, F. DELBEAN, J.-M. EBER et D. HEATH : Thinking coherently. *RISK*, 10:68–71, November 1997.
- [8] O. E. BARNDORFF-NIELSEN : Normal inverse Gaussian distributions and stochastic volatility modelling. *Scandinavian Journal of Statistics*, 24(1):1–13, 1997.
- [9] O. E. BARNDORFF-NIELSEN : Processes of normal inverse Gaussian type. *Finance and Stochastics*, 2(1):41–68, 1997.
- [10] J. BARUNIK et L. VACHA : Monte Carlo-based tail exponent estimator. *Physica A : Statistical Mechanics and its Applications*, 389(21):4863 – 4874, 2010.
- [11] L. E. BAUM : An inequality and associated maximization technique in statistical estimation for probabilistic functions of Markov processes. *In Inequalities, III (Proc. Third Sympos., Univ. California, Los Angeles, Calif., 1969; dedicated to the memory of Theodore S. Motzkin)*, p. 1–8. Academic Press, New York, 1972.
- [12] L. E. BAUM et J. A. EAGON : An inequality with applications to statistical estimation for probabilistic functions of Markov processes and to a model for ecology. *Bulletin of the American Mathematical Society*, 73:360–363, 1967.
- [13] L. E. BAUM et T. PETRIE : Statistical inference for probabilistic functions of finite state Markov chains. *Annals of Mathematical Statistics*, 37:1554–1563, 1966.

- [14] L. E. BAUM, T. PETRIE, G. SOULES et N. WEISS : A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. *Annals of Mathematical Statistics*, 41:164–171, 1970.
- [15] L. E. BAUM et G. R. SELL : Growth transformations for functions on manifolds. *Pacific Journal of Mathematics*, 27:211–227, 1968.
- [16] R. A. BECKER, J. M. CHAMBERS et A. R. WILKS : *The New S Language : A Programming Environment for Data Analysis and Graphics*. Wadsworth and Brooks/Cole Advanced Books & Software, Monterey, CA, USA, 1988.
- [17] J. BEIRLANT, Y. GOEGEBEUR, J. TEUGELS et J. SEGERS : *Statistics of Extremes : Theory and Applications*. Wiley Series in Probability and Statistics. John Wiley & Sons Ltd., 2004.
- [18] J. BEIRLANT, P. VYNCKIER et J. L. TEUGELS : Tail index estimation, Pareto quantile plots, and regression diagnostics. *Journal of the American Statistical Association*, 91(436):1659–1667, 1996.
- [19] N. BELLAMY et M. JEANBLANC : Incompleteness of markets driven by a mixed diffusion. *Finance and Stochastics*, 4(2):209–222, 2000.
- [20] J. BERKOWITZ et J. O'BRIEN : How Accurate Are Value-at-Risk Models at Commercial Banks? *The journal of finance*, 57(3):1093–1111, 2002.
- [21] J. BERTOIN : *Lévy processes*, vol. 121. Cambridge University Press, 1998.
- [22] A. BEZAT et A. NIKEGHBALI : La théorie des extrêmes et la gestion des risques de marché. Rapport de groupe de travail, ENSAE, 2000.
- [23] P. BILLINGSLEY : *Probability and measure*. Probability and mathematical statistics. Wiley, 1979.
- [24] M. BILLIO et L. PELIZZON : Value-at-risk : a multivariate switching regime approach. *Journal of Empirical Finance*, 7(5):531 – 554, 2000.
- [25] F. BLACK et M. SCHOLES : The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–654, 1973.
- [26] C. R. BLYTH et H. A. STILL : Binomial confidence intervals. *Journal of the American Statistical Association*, 78(381):108–116, 1983.
- [27] T. BOLLERSLEV : Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, 31(3):307–327, 1986.
- [28] G. BORMETTI, V. CAZZOLA, G. LIVAN, G. MONTAGNA et O. NICROSINI : A generalized Fourier transform approach to risk measures. *Journal of Statistical Mechanics : Theory and Experiment*, 2010(01):P01005, 2010.
- [29] H. BÜHLMANN, F. DELBAEN, P. EMBRECHTS et A. N. SHIRYAEV : No-arbitrage, change of measure and conditional Esscher transforms. *CWI quarterly*, 9(4):291–317, 1996.
- [30] S. D. CAMPBELL : A review of backtesting and backtesting procedures. *Journal of Risk*, 9:1–17, 2007.
- [31] P. CARR, H. GEMAN, D. MADAN et M. YOR : The fine structure of asset returns : An empirical investigation. *The Journal of Business*, 75(2):305–333, 2002.

- [32] P. CARR et D. MADAN : Option valuation using the fast Fourier transform. *Journal of Computational Finance*, 2(4):61–73, 1999.
- [33] N. CHAMPAGNAT, M. DEACONU, A. LEJAY et K. SALHI : Mesure de risque : detection du regime de crise et calcul de la Value-at-Risk. Rapport de collaboration Alphability & Tosca, Inria, 2013.
- [34] T. CHAN : Pricing contingent claims on stocks driven by Lévy processes. *Annals of Applied Probability*, 9(2):504–528, 1999.
- [35] C. W. CHEONG : Optimal choice of sample fraction in univariate financial tail index estimation. *Journal of Applied Statistics*, 37(12):2043–2056, 2010.
- [36] P. F. CHRISTOFFERSEN : Evaluating interval forecasts. *International Economic Review*, 39(4):841–862, 1998. Symposium on Forecasting and Empirical Methods in Macroeconomics and Finance.
- [37] A. CIFTER : Value-at-risk estimation with wavelet-based extreme value theory : Evidence from emerging markets. *Physica A : Statistical Mechanics and its Applications*, 390(12):2356 – 2367, 2011.
- [38] P. K. CLARK : A subordinated stochastic process model with finite variance for speculative prices. *Econometrica : Journal of the Econometric Society*, p. 135–155, 1973.
- [39] D. COCULESCU, M. JEANBLANC et A. NIKEGHBALI : Default times, no-arbitrage conditions and changes of probability measures. *Finance and Stochastics*, 16(3):513–535, 2012.
- [40] J. CONG, K. S. TAN et C. WENG : VaR-based optimal partial hedging. *Astin Bulletin*, 43(03):271–299, 2013.
- [41] J. CONG, K. S. TAN et C. WENG : CVaR-based optimal partial hedging. *The Journal of Risk*, 16(3):49–83, 2014.
- [42] R. CONT : Empirical properties of asset returns : stylized facts and statistical issues. *Quantitative Finance*, 1:223–236, 2001.
- [43] R. CONT et P. TANKOV : *Financial modelling with jump processes*. Chapman & Hall/CRC Financial Mathematics Series. Chapman & Hall/CRC, Boca Raton (Fla.), London, New York, 2004.
- [44] J. W. COOLEY et J. W. TUKEY : An algorithm for the machine calculation of complex Fourier series. *Mathematics of Computation*, 19(90):297–301, 1965.
- [45] J. D. CURTO, J. C. PINTO et G. N. TAVARES : Modeling stock markets’ volatility using GARCH models with Normal, Student’s t and stable Paretian distributions. *Statistical Papers*, 50(2):311–321, juil. 2007.
- [46] J. CVITANIC : Minimizing expected loss of hedging in incomplete and constrained markets. *SIAM Journal on Control and Optimization*, 38(4):1050–1066, 2000.
- [47] J. CVITANIC et I. KARATZAS : Generalized Neyman-Pearson lemma via convex duality. *Bernoulli*, 7(1):79–97, 2001.
- [48] J. CVITANIC et G. SPIVAK : Maximizing the probability of a perfect hedge. *The Annals of Applied Probability*, 9(4):1303–1328, 11 1999.

- [49] L. DE HAAN et L. PENG : Comparison of tail index estimators. *Statistica Neerlandica*, 52(1):60–70, 1998.
- [50] B. de SOUSA et G. MICHAILIDIS : A diagnostic plot for estimating the tail index of a distribution. *Journal of Computational and Graphical Statistics*, 13(4):974–995, 2004.
- [51] F. DELBAEN : Coherent risk measures on general probability spaces. *In Advances in finance and stochastics*, p. 1–37. Springer, 2002.
- [52] F. DELBAEN et W. SCHACHERMAYER : A general version of the fundamental theorem of asset pricing. *Mathematische annalen*, 300(1):463–520, 1994.
- [53] A. P. DEMPSTER, N. M. LAIRD et D. B. RUBIN : Maximum likelihood from incomplete data via the EM algorithm. *Journal of the royal statistical society. Series B (methodological)*, p. 1–38, 1977.
- [54] D. A. DICKEY et W. A. FULLER : Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74(366, part 1):427–431, 1979.
- [55] T. DOGANOGLU, C. HARTZ et S. MITTNIK : Portfolio optimization when risk factors are conditionally varying and heavy tailed. *Computational Economics*, 29(3-4):333–354, jan. 2007.
- [56] H. DREES, L. de HAAN et S. RESNICK : How to make a Hill plot. *The Annals of Statistics*, 28(1):254–274, 2000.
- [57] W. H. DUMOUCHEL : Estimating the stable index α in order to measure tail thickness : A critique. *the Annals of Statistics*, 11(4):1019–1031, 1983.
- [58] E. EBERLEIN et J. JACOD : On the range of options prices. *Finance and Stochastics*, 1(2):131–140, 1997.
- [59] N. EL KAROUI et M.-C. QUENEZ : Dynamic programming and pricing of contingent claims in an incomplete market. *SIAM Journal on Control and Optimization*, 33(1):29–66, 1995.
- [60] R. J. ELLIOTT, L. CHAN et T. K. SIU : Option pricing and Esscher transform under regime switching. *Annals of Finance*, 1(4):423–432, 2005.
- [61] E. F. FAMA : Portfolio analysis in a stable Paretian market. *Management Science*, 11(3):404–419, 1965.
- [62] E. F. FAMA : Risk, return, and equilibrium. *The Journal of Political Economy*, 79(1):30–55, 1971.
- [63] H. FOFACK et J. P. NOLAN : Tail behavior, modes and other characteristics of stable distributions. *Extremes*, 2(1):39–58, 1999.
- [64] H. FÖLLMER et Y. M. KABANOV : Optional decomposition and Lagrange multipliers. *Finance and Stochastics*, 2(1):69–81, 1998.
- [65] H. FÖLLMER et P. LEUKERT : Quantile hedging. *Finance and Stochastics*, 3(3):251–273, 1999.
- [66] H. FÖLLMER et P. LEUKERT : Efficient hedging : cost versus shortfall risk. *Finance and Stochastics*, 4(2):117–146, 2000.

- [67] H. FÖLLMER et M. SCHWEIZER : Hedging by sequential regression : An introduction to the mathematics of option trading. *Astin Bulletin*, 19(S1):29–42, 1989.
- [68] H. FÖLLMER et M. SCHWEIZER : Hedging of contingent claims. *Applied Stochastic Analysis*, 5:389, 1991.
- [69] H. FÖLLMER et D. SONDERMANN : Hedging of non-redundant contingent claims. *Contributions to Mathematical Economics*, 1986.
- [70] M. FRITTELLI : The minimal entropy martingale measure and the valuation problem in incomplete markets. *Mathematical Finance*, 10(1):39–52, 2000.
- [71] T. FUJIWARA et Y. MIYAHARA : The minimal entropy martingale measures for geometric Lévy processes. *Finance and Stochastics*, 7(4):509–531, 2003.
- [72] H. GEMAN : Pure jump Lévy processes for asset price modelling. *Journal of Banking & Finance*, 26(7):1297–1316, 2002.
- [73] R. GENÇAY et F. SELÇUK : Extreme value theory and Value-at-Risk : Relative performance in emerging markets. *International Journal of Forecasting*, 20(2): 287–303, avr. 2004.
- [74] R. GENÇAY, F. SELÇUK et A. ULUGÜLYAGCI : High volatility, thick tails and extreme value theory in value-at-risk estimation. *Insurance : Mathematics and Economics*, 33(2):337 – 356, 2003. Papers presented at the 6th IME Conference, Lisbon, 15-17 July 2002.
- [75] H.-U. GERBER et E. S. W. SHIU : Option pricing by Esscher transforms. *Transactions of the Society of Actuaries*, XLVI:99–191, 1994.
- [76] M. GILLI et E. KELLEZI : An Application of Extreme Value Theory for Measuring Financial Risk. *Computational Economics*, 27(2-3):207–228, 2006.
- [77] Y. GOEGBEUR, J. BEIRLANT et T. DE WET : Linking Pareto-tail kernel goodness-of-fit statistics with tail index estimation at optimal threshold and second order estimation. *Revstat Statistical Journal*, 6:51–69, 2008.
- [78] T. GOLL et J. KALLSEN : Optimal portfolios for logarithmic utility. *Stochastic Processes and their applications*, 89(1):31–48, 2000.
- [79] T. GOLL et L. RÜSCHENDORF : Minimax and minimal distance martingale measures and their relationship to portfolio optimization. *Finance and Stochastics*, 5(4):557–581, 2001.
- [80] C. GOURIEROUX, J. P. LAURENT et H. PHAM : Mean-variance hedging and numéraire. *Mathematical Finance*, 8(3):179–200, 1998.
- [81] A. A. GUSHCHIN et E. MORDECKI : Bounds on option prices for semimartingale market models. *In Proceedings of the Steklov Institute of Mathematics*, p. 73–113, 2001.
- [82] E. HAEUSLER et J. L. TEUGELS : On asymptotic normality of Hill’s estimator for the exponent of regular variation. *The Annals of Statistics*, 13(2):743–756, 1985.
- [83] P. HALL : On some simple estimates of an exponent of regular variation. *Journal of the Royal Statistical Society. Series B (Methodological)*, 44(1):37–42, 1982.

- [84] P. HALL et A. H. WELSH : Adaptive estimates of parameters of regular variation. *The Annals of Statistics*, 13(1):331–341, 03 1985.
- [85] J. D. HAMILTON : Autoregressive conditional heteroskedasticity and changes in regime. *Journal of Econometrics*, 64(1):307–333, 1994.
- [86] J. D. HAMILTON : Specification testing in Markov-switching time-series models. *Journal of Econometrics*, 70(1):127 – 157, 1996.
- [87] J. M. HARRISON et D. M. KREPS : Martingales and arbitrage in multiperiod securities markets. *Journal of Economic Theory*, 20(3):381–408, 1979.
- [88] B. M. HILL : A simple general approach to inference about the tail of a distribution. *The Annals of Statistics*, 3(5):1163–1174, 1975.
- [89] F. HUBALEK, J. KALLSEN et L. KRAWCZYK : Variance-optimal hedging for processes with stationary independent increments. *Annals of Applied Probability*, p. 853–885, 2006.
- [90] D. W. JANSEN et C. G. de VRIES : On the frequency of large stock returns : Putting booms and busts into perspective. *The review of economics and statistics*, p. 18–24, 1991.
- [91] P. JORION : *Value at Risk, 3rd Ed. : The New Benchmark for Managing Financial Risk*. McGraw-Hill Education, 2006.
- [92] E. JOUINI, M. MEDDEB et N. TOUZI : Vector-valued coherent risk measures. *Finance and Stochastics*, 8(4):531–552, 2004.
- [93] I. KARATZAS : *Lectures on the Mathematics of Finance*. CRM monograph series. American Mathematical Society, 1997.
- [94] A. KELANI et F. QUITTARD-PINON : A general approach to compute standard risk measures. In *International Conference of the French Finance Association (AFFI)*, 2011.
- [95] Y. S. KIM, S. RACHEV, M. L. BIANCHI et F. J. FABOZZI : Computing VaR and AVaR in infinitely divisible distributions. *Probability and Mathematical Statistics*, 30(2):223–245, 2010.
- [96] P. KLUSIK et Z. PALMOWSKI : Quantile hedging for equity-linked contracts. *Insurance : Mathematics and Economics*, 48(2):280–286, 2011.
- [97] S. G. KOU : A jump-diffusion model for option pricing. *Management Science*, 48(8):1086–1101, 2002.
- [98] D. O. KRAMKOV : Optional decomposition of supermartingales and hedging contingent claims in incomplete security markets. *Probability Theory and Related Fields*, 105(4):459–479, 1996.
- [99] K. KUESTER, S. MITTNIK et M. S. PAOLELLA : Value-at-Risk prediction : A comparison of alternative strategies. *Journal of Financial Econometrics*, 4(1):53–89, 2006.
- [100] H. KUNITA et S. WATANABE : On square integrable martingales. *Nagoya Mathematical Journal*, 30:209–245, 1967.
- [101] P. H. KUPIEC : Techniques for verifying the accuracy of risk measurement models. *Journal of Derivatives*, 3:73–84, 1995.

- [102] S. KUSUOKA : On law invariant coherent risk measures. *In Advances in mathematical economics*, p. 83–95. Springer, 2001.
- [103] D. KWIATKOWSKI, P. PHILLIPS et P. SCHMIDT : Testing the null hypothesis of stationarity against the alternative of a unit root : How sure are we that economic time series have a unit root ? *Journal of Econometrics*, 54:159–178, 1992.
- [104] J. P. LAURENT et H. PHAM : Dynamic programming and mean-variance hedging. *Finance and Stochastics*, 3(1):83–110, 1999.
- [105] O. LE COURTOIS et C. WALTER : The computation of risk budgets under the Lévy process assumption. *Finance*, 35(2):87–108, 2014.
- [106] O. A. LE COURTOIS et C. P. WALTER : A study on Value-at-Risk and Lévy processes. *SSRN Working Paper Series*, 2010.
- [107] R. W. LEE *et al.* : Option pricing by transform methods : extensions, unification and error control. *Journal of Computational Finance*, 7(3):51–86, 2004.
- [108] V. I. LEVENSHTAIN : Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics. Doklady*, 10:707–710, 1965.
- [109] F. LONGIN : The Asymptotic Distribution of Extreme Stock Market Returns. *The Journal of Business*, 69(3):383–408, 1996.
- [110] M. LORETAN et P. PHILIPS : Testing the covariance stationarity of heavy-tailed time series : An overview of the theory with applications to several financial datasets. *Journal of Empirical Finance*, 1(2):211–248, 1994.
- [111] D. MADAN, P. CARR et E. CHANG : The variance gamma process and option pricing. *European Finance Review*, 2(1):79–105, 1998.
- [112] D. MADAN et E. SENETA : The variance gamma (V.G.) model for share market returns. *The Journal of Business*, 63(4):511–524, 1990.
- [113] Y. MALEVERGNE, V. PISARENKO et D. SORNETTE : On the power of generalized extreme value (GEV) and generalized Pareto distribution (GPD) estimators for empirical distributions of stock returns. *Applied Financial Economics*, 16(3):271–289, 2006.
- [114] B. MANDELBROT : The variation of certain speculative prices. *Journal of business*, XXXVI:392–417, 1963.
- [115] N. MAYO : *Les modèles à variables cachées et leurs applications en finance : risque systématique, détection d'arbitrage et prévision des volumes*. Thèse de doctorat, Université Paris I Panthéon-Sorbonne, 2010.
- [116] J. H. MCCULLOCH : Measuring tail thickness to estimate the stable index α : a critique. *Journal of Business & Economic Statistics*, 15(1):74–81, 1997.
- [117] G. J. MCLACHLAN et T. KRISHNAN : *The EM algorithm and extensions*. Wiley Series in Probability and Statistics. Wiley-Interscience [John Wiley & Sons], Hoboken, NJ, second éd., 2008.
- [118] A. MELNIKOV et V. SKORNYAKOVA : Quantile hedging and its application to life insurance. *Statistics & Decisions*, 23(4/2005):301–316, 2005.
- [119] A. MELNIKOV et I. SMIRNOV : Dynamic hedging of conditional value-at-risk. *Insurance : Mathematics & Economics*, 51(1):182–190, 2012.

- [120] R. C. MERTON : Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, 3(1-2):125–144, 1976.
- [121] J. MINA et J. Y. XIAO : Return to RiskMetrics : The Evolution of a Standard. Rap. tech., RiskMetrics, 2001.
- [122] Y. MIYAHARA : Canonical martingale measures of incomplete assets markets. *Probability Theory and Mathematical Statistics (Tokyo, 1995)*, p. 343–352, 1996.
- [123] Y. MIYAHARA : Geometric Lévy Process & MEMM : pricing model and related estimation problems. *Asia-Pacific Financial Markets*, 8(1):45–60, 2001.
- [124] Y. NAKANO : Minimizing coherent risk measures of shortfall in discrete-time models with cone constraints. *Applied Mathematical Finance*, 10(2):163–181, 2003.
- [125] Y. NAKANO : Efficient hedging with coherent risk measure. *Journal of Mathematical Analysis and Applications*, 293(1):345–354, 2004.
- [126] Q. H. NGUYEN et C. ROBERT : Series expansions for sums of independent Pareto random variables. Rap. tech., Working Paper SAF 2012.16, 2013. <http://docs.isfa.fr/lab0/2013.7.pdf>, 2013.
- [127] Q. H. NGUYEN et C. Y. ROBERT : New efficient estimators in rare event simulation with heavy tails. *Journal of Computational and Applied Mathematics*, 261:39–47, 2014.
- [128] J. NOLAN : *Stable Distributions : Models for Heavy Tailed Data*. Birkhauser, Boston, 2013.
- [129] E. PARZEN : On estimation of a probability density function and mode. *Annals of Mathematical Statistics*, 33:1065–1076, 1962.
- [130] S.-L. PENG, T.-C. SHENG et Y.-W. TSAY : Volatility on stock exchanges and exchange rates based on pairwised alignemnt. *In Proceedings of the 29th Workshop on Combinatorial Mathematics and Computation Theory*, p. 235–238, 2012.
- [131] S. T. RACHEV : *Handbook of Heavy Tailed Distributions in Finance*. Handbooks in Finance. Elsevier/North-Holland, 2003.
- [132] S. RAIBLE : *Lévy processes in finance : Theory, numerics, and empirical facts*. Thèse de doctorat, PhD thesis, Universität Freiburg i. Br, 2000.
- [133] S. RESNICK et C. STARICA : Smoothing the Hill estimator. *Advances in Applied Probability*, 29(1):271–293, 1997.
- [134] S. I. RESNICK : *Extreme values, regular variation, and point processes*, vol. 4 de *Applied Probability. A Series of the Applied Probability Trust*. Springer-Verlag, New York, 1987.
- [135] R. T. ROCKAFELLAR et S. URYASEV : Optimization of Conditional Value-at-Risk. *Journal of Risk*, 2:21–41, 2000.
- [136] R. T. ROCKAFELLAR et S. URYASEV : Conditional Value-at-Risk for general loss distributions. *Journal of Banking and Finance*, 26(7):1443–1471, 2002.
- [137] M. ROMANO et N. TOUZI : Contingent claims and market completeness in a stochastic volatility model. *Mathematical Finance*, 7(4):399–412, 1997.

- [138] B. RUDLOFF : *Hedging in incomplete markets and testing compound hypotheses via convex duality*. Thèse de doctorat, Ph. D. Thesis, 2006.
- [139] B. RUDLOFF : Convex hedging in incomplete markets. *Applied Mathematical Finance*, 14(5):437–452, 2007.
- [140] B. RUDLOFF : Coherent hedging in incomplete markets. *Quantitative Finance*, 9(2):197–206, 2009.
- [141] B. RUDLOFF et I. KARATZAS : Testing composite hypotheses via convex duality. *Bernoulli*, 16(4):1224–1239, 2010.
- [142] T. H. RYDBERG : The normal inverse Gaussian Lévy process : simulation and approximation. *Communications in Statistics. Stochastic Models*, 13(4):887–910, 1997.
- [143] K. SALHI : European options pricing and risk measurement under Exponential Lévy models - a practical guide. prépublication, 2016.
- [144] K. SALHI, M. DEACONU, A. LEJAY, N. CHAMPAGNAT et N. NAVET : Regime switching model for financial data : Empirical risk analysis. *Physica A : Statistical Mechanics and its Applications*, 461:148–157, 2016.
- [145] K.-I. SATO : *Lévy processes and infinitely divisible distributions*. Cambridge University Press, 1999.
- [146] C. SCARROTT et A. MACDONALD : A review of extreme value threshold estimation and uncertainty quantification. *REVSTAT-Statistical Journal*, 10(1):33–60, 2012.
- [147] H. SCHALLER et S. V. NORDEN : Regime switching in stock market returns. *Applied Financial Economics*, 7(2):177–191, 1997.
- [148] A. SCHIED : On the Neyman-Pearson problem for law-invariant risk measures and robust utility functionals. *The Annals of Applied Probability*, 14(3):1398–1423, 2004.
- [149] M. SCHWEIZER : On the minimal martingale measure and the Föllmer-Schweizer decomposition. *Stochastic Analysis and Applications*, 13(5):573–599, 1995.
- [150] J. SEKINE : Quantile hedging for defaultable securities in an incomplete market. *Mathematical Economics (Japanese)*, 1165:215–231, 2000.
- [151] J. SEKINE : Dynamic minimization of worst conditional expectation of shortfall. *Mathematical Finance*, 14(4):605–618, 2004.
- [152] A. SINGH, D. ALLEN et P. ROBERT : Extreme Market Risk and Extreme Value Theory. *Mathematics and computers in simulation*, 94:310–328, 2013.
- [153] R. C. SINGLETON : Mixed radix fast Fourier transform. Rap. tech., Stanford Research Inst., Menlo Park, Calif., 1972.
- [154] I. SMIRNOV : *Dynamic Hedging : CVaR Minimization and Path-wise Comparison*. Thèse de doctorat, University of Alberta, 2013.
- [155] H. M. SONER et N. TOUZI : The problem of super-replication under constraints. In *Paris-Princeton Lectures on Mathematical Finance 2002*, p. 133–172. Springer, 2003.

- [156] SUPERVISION, BASEL COMMITTEE ON BANKING, éd. *International Convergence of Capital Measurement and Capital Standards – A Revised Framework*. Bank of International Settlements, juin 2004.
- [157] P. TANKOV : *Processus de Lévy en finance : problèmes inverses et modélisation de dépendance*. Thèse de doctorat, Centre de Mathématiques Appliquées, Ecole Polytechnique, 2004.
- [158] P. TANKOV : Pricing and hedging in exponential Lévy models : review of recent results. In *Paris-Princeton Lectures on Mathematical Finance 2010*, p. 319–359. Springer, 2011.
- [159] M. S. THORSTEN RHEINLANDER : On L^2 -projections on a space of stochastic integrals. *The Annals of Probability*, 25(4):1810–1831, 1997.
- [160] A. VITERBI : Error bounds for convolutional codes and an asymptotically optimum decoding algorithm. *Information Theory, IEEE Transactions on*, 13(2):260–269, April 1967.
- [161] J. S. WALKER : *Fast Fourier transforms*, vol. 24. CRC press, 1996.
- [162] G.-J. WANG, C. XIE, F. HAN et B. SUN : Similarity measure and topology evolution of foreign exchange markets using dynamic time warping method : Evidence from minimal spanning tree. *Physica A : Statistical Mechanics and its Applications*, 391(16):4136–4146, 2012.
- [163] Y. WANG : Quantile hedging for guaranteed minimum death benefits. *Insurance : Mathematics and Economics*, 45(3):449–458, 2009.
- [164] I. WEISSMAN : Estimation of parameters and large quantiles based on the k largest observations. *Journal of the American Statistical Association*, 73(364): 812–815, 1978.
- [165] R. WERON : Lévy-stable distributions revisited : tail index > 2 does not exclude the Lévy-stable regime. *International Journal of Modern Physics C.*, 12(2):209–223, 2001.
- [166] H. WITTING : Mathematische Statistik I. Stuttgart : BG Teubner. *Mathematical Reviews*, 581, 1985.
- [167] L. YAO, G. YANG et X. YANG : A note on the mean correcting martingale measure for geometric Lévy processes. *Applied Mathematics Letters*, 24(5): 593–597, 2011.
- [168] I. ZALIAPIN, Y. Y. KAGAN et F. P. SCHOENBERG : Approximating the distribution of Pareto sums. *Pure and Applied geophysics*, 162(6-7):1187–1228, 2005.
- [169] Y. ZHANG : *Prediction of financial time series with hidden Markov models*. Thèse de doctorat, Shandong University, China, 2001.

Résumé

Cette thèse étudie la gestion et la couverture du risque en s'appuyant sur la Value-at-Risk (VaR) et la Value-at-Risk Conditionnelle (CVaR), comme mesures de risque. La première partie propose un modèle d'évolution de prix que nous confrontons à des données réelles issues de la bourse de Paris (Euronext PARIS). Notre modèle prend en compte les probabilités d'occurrence des pertes extrêmes et les changements de régimes observés sur les données. Notre approche consiste à détecter les différentes périodes de chaque régime par la construction d'une chaîne de Markov cachée et à estimer la queue de distribution de chaque régime par des lois puissances. Nous montrons empiriquement que ces dernières sont plus adaptées que les lois normales et les lois stables. L'estimation de la VaR est validée par plusieurs backtests et comparée aux résultats d'autres modèles classiques sur une base de 56 actifs boursiers. Dans la deuxième partie, nous supposons que les prix boursiers sont modélisés par des exponentielles de processus de Lévy. Dans un premier temps, nous développons une méthode numérique pour le calcul de la VaR et la CVaR cumulatives. Ce problème est résolu en utilisant la formalisation de Rockafellar et Uryasev, que nous évaluons numériquement par inversion de Fourier. Dans un deuxième temps, nous nous intéressons à la minimisation du risque de couverture des options européennes, sous une contrainte budgétaire sur le capital initial. En mesurant ce risque par la CVaR, nous établissons une équivalence entre ce problème et un problème de type Neyman-Pearson, pour lequel nous proposons une approximation numérique s'appuyant sur la relaxation de la contrainte.

Mots clefs : Value-at-Risk, Value-at-Risk Conditionnelle, lois puissances, modèles de Markov cachés, processus de Lévy, transformée de Fourier rapide, Lemme de Neyman-Pearson.

Abstract

This thesis studies the risk management and hedging, based on the Value-at-Risk (VaR) and the Conditional Value-at-Risk (CVaR) as risk measures. The first part offers a stocks return model that we test in real data from NSYE Euronext. Our model takes into account the probability of occurrence of extreme losses and the regime switching observed in the data. Our approach is to detect the different periods of each regime by constructing a hidden Markov chain and estimate the tail of each regime distribution by power laws. We empirically show that powers laws are more suitable than Gaussian law and stable laws. The estimated VaR is validated by several backtests and compared to other conventional models results on a basis of 56 stock market assets. In the second part, we assume that stock prices are modeled by exponentials of a Lévy process. First, we develop a numerical method to compute the cumulative VaR and CVaR. This problem is solved by using the formalization of Rockafellar and Uryasev, which we numerically evaluate by Fourier inversion techniques. Secondly, we are interested in minimizing the hedging risk of European options under a budget constraint on the initial capital. By measuring this risk by CVaR, we establish an equivalence between this problem and a problem of Neyman-Pearson type, for which we propose a numerical approximation based on the constraint relaxation.

Key words : Value-at-Risk, Conditional Value-at-Risk, power laws, hidden Markov models, Lévy processes, fast Fourier transforms, Neyman-Pearson Lemma.