

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Faculté des Sciences Exactes
Département de Recherche Opérationnelle



Mémoire de fin de cycle

En vue de l'obtention du diplôme de Master

Spécialité : Mathématiques Financières

Thème

Sur l'Estimation dans les Modèles de Rendements Financiers

Présenté par :

Mlle AZZOUG SONIA

Mme RAHMANI RYMA

Soutenu le 08 juillet 2019 devant le jury composé de :

Présidente	Mme Y. ZIANE	MCB
Promotrice	Mlle K. BOUCHEBBAH	MAA
Examineur	Mr F. MOUHOUS	Doctorant
Examinatrice	Mme L. HARFOUCHE	Docteur

Promotion 2018-2019



Tout d'abord, Nous remercions Dieu le tout puissant de nous avoir accordé la santé, le courage et la volonté pour accomplir ce modeste travail.

Un grand merci pour nos familles, surtout nos parents qui nous ont épaulés, soutenus et suivis, qui nous ont toujours orientés vers le droit chemin.

Nous tenons à remercier Mlle **BOUCHEBBAH** qui nous a fait l'honneur d'être notre promotrice, d'avoir proposé ce travail, son soutien, sa patience et ses conseils qui nous ont beaucoup aidées dans la réalisation et la finalisation de ce travail.

Nos vifs remerciements aux membres de l'unité de recherche Lamos pour leurs aides et leurs disponibilités pendant toute la réalisation de notre travail.

Enfin nous tenons aussi à remercier également tous les membres de jury pour avoir accepté d'évaluer notre travail.

Nous tenons tout simplement à exprimer notre profonde gratitude à tous ceux qui nous ont soutenus de près ou de loin durant tout notre cursus et espérant que ce mémoire servira de guide pour les promotions à venir.

Dédicace

A cœur veillant, rien d'impossible.

A conscience tranquille, tout est accessible.

Quand il y a la soif d'apprendre.

Tout vient à point à qui sait attendre.

Les études sont avant tout notre unique et seul atout.

Souhaitant que le fruit de nos efforts fournis jour et nuit nous mènera vers le bonheur fleuri

Je dédie ce modeste travail :

A celle qui ma donné la vie, le symbole de tendresse, qui s'est sacrifiée pour mon bonheur. et ma réussite, à ma mère.

A mon père, école de mon enfance, qui a été mon ombre durant toutes les années des études, Que dieu les gardes et les protège.

A mes très chers frères .

A mon futur mari

A toute ma famille.

A tous mes amis avec lesquels j'ai partagé mes moments de joie et de bonheur

SONIA

Dédicace

Je dédie ce modeste travail

A mes chers parents, pour tous leurs sacrifices, leur amour, leur tendresse, leur soutien et leurs prières tout au long de mes études.

A ma chère sœur pour son encouragement permanent et son soutien moral

A mes chers frères pour leur appui et leur encouragement.

A mon mari

A toute ma famille pour leur soutien tout au long de mon parcours universitaire.

RYMA

Table des matières

Table des Matières	iv
Table des Figures	v
Liste des Tableaux	vi
Introduction Générale	1
1 Quelques modèles de rendements financiers	4
1.1 Différents types de modèles	4
1.2 Les rendements financiers	5
1.2.1 Définitions et notations	5
1.2.2 Rendements cumulés	6
1.3 Données disponibles	6
1.4 Quelques modèles de rendements	7
1.4.1 Le modèle de Bachelier	7
1.4.2 Proposition de Mandelbrot	9
1.4.3 Le modèle de Press	10
1.4.4 Le modèle de Praetz	12
1.5 Conditions essentielles de Madan et Seneta	13
1.6 Modèle de Laplace asymétrique généralisé	14
1.6.1 Le processus de Laplace	14
1.6.2 Le processus de Laplace est un processus subordonné	15
1.7 Distribution de Laplace asymétrique généralisée	16
1.7.1 Fonction caractéristique	16
1.7.2 Invariance d'échelle	17
1.7.3 Rôle des paramètres	18
1.7.4 Changement d'échelle et de localisation	18

1.8	Cas particuliers	19
2	Rappel sur les approches fréquentistes	22
2.1	La méthode du maximum de vraisemblance	22
2.2	La méthode des moments	23
2.3	La méthode des moments généralisée	24
2.4	Méthode de l'équation d'estimation optimale	25
2.4.1	Équation d'estimation optimale	27
2.4.2	Équation d'estimation optimale modifiée	30
2.5	Test de Shapiro-Wilk	30
2.5.1	Description	31
3	Le paradigme bayésien	32
3.1	Modèle bayésien	32
3.1.1	Modélisation de l'information a priori	33
3.1.2	Calcul de la loi a posteriori	35
3.2	Inférence bayésienne décisionnelle	35
3.2.1	Fonction de perte et risque bayésien	36
3.2.2	Estimateur de Bayes	37
3.3	Les méthodes de Monte Carlo par Chaîne de Markov (MCMC)	37
3.3.1	Algorithme de Metropolis-Hastings	38
3.3.2	Echantillonneur de Gibbs	39
3.3.3	Propriétés :	39
4	Application sur des données réelles	41
4.1	Collecte et faits stylisés des données	41
4.1.1	Description des données réelles	41
4.2	Test d'ajustement de la non normalité des rendements	42
4.3	Estimation des paramètres du modèle de Laplace asymétrique généralisé	44
4.4	Estimation des paramètres du modèle de Laplace symétrique	46
4.4.1	Estimateur du maximum de vraisemblance	46
4.4.2	Approche bayésienne	47
4.5	Résultats obtenus	49
4.6	Discussion et interprétation des résultats :	51
A	Notions fondamentales utilisées	55
A.1	Produit de convolution	55
A.2	Processus de Lévy	55

A.2.1	Définition et propriétés	55
A.2.2	Représentation de Lévy-Khintchine	56
A.3	Processus subordonné [30]	56
A.3.1	Le processus gamma [30]	57
A.3.2	Le processus de Wiener [30]	58
A.4	Lemme d'Itô	59
A.5	Chaîne de Markov à espace d'état discret E	59
A.5.1	Noyau d'une chaîne de Markov	60
B	Données utilisées	62
C	Code source sous R	63

Table des figures

1.1	Modèle de Bachelier : probabilité composée	7
1.2	Distribution des rendements annuels de 40 titres boursiers, de 1890 à 1915, Table XVIII de Mitchell [29].	9
1.3	Densité de probabilité de la distribution de Laplace symétrique.	21
4.1	Histogramme et distribution de l'échantillon R_1	43
4.2	Graphique Quantile-Quantile	44
4.3	Comparaison entre la densité théorique et celle estimée par bayes dans le 1 ^{er} cas.	51

Liste des tableaux

1.1	Respect des conditions émises par Madan et Seneta pour les différents modèles présentés	14
1.2	Domaine et rôle des paramètres de la distribution de Laplace asymétrique généralisée	18
2.1	Méthodes d'estimation représentables par la forme quadratique (2.21)	27
3.1	Exemple de lois a priori conjuguées naturelles pour quelques familles exponentielles usuelles	34
4.1	La distribution des rendements financiers	42
4.2	Statistiques d'ordre de l'échantillon R_1	42
4.3	Valeurs des paramètres estimée	45
4.4	Estimation des paramètres du modèle par l'approche bayésienne	50
B.1	Prix du titre Abbey National (penny sterling) du 31 juillet au 8 octobre 1991 [7][30]	62

Introduction Générale

La théorie de la finance concerne l'étude des prix des produits financiers et plus précisément de leurs évolutions temporelles. Un modèle financier repose le plus souvent sur une représentation des prix d'actifs financiers (ou des niveaux de taux d'intérêt). Une telle représentation peut être recherchée dans une perspective de meilleure compréhension des marchés financiers, ou encore dans une perspective d'élaboration d'outils permettant d'améliorer la gestion financière : gestion des risques, allocation d'actifs, création de nouveaux produits financiers.

Une modélisation financière réalise toujours un compromis entre adéquation aux observations dans les marchés financiers et commodité d'utilisation. Une modélisation qui cherche à reproduire toutes les propriétés statistiques des observations conduit généralement à un modèle compliqué qui se révèle souvent difficile à utiliser ; car les calculs théoriques étant généralement difficile à mener. A l'inverse, une simplification excessive du modèle permet, certes, de mener à bout de nombreux calculs, par exemple, pricing des options et gestion du portefeuille, mais risque de ne pas être cohérent avec la réalité financière. Une caractéristique des prix des produits financiers est que leur évolution ne peut s'expliquer par des modèles déterministes simples. La plupart des modélisations financières reposent sur une représentation probabiliste, ce qui revient à déterminer la loi de probabilité du prix de l'actif, ce dernier étant vu comme une variable aléatoire.

La première modélisation des cours boursiers, proposée en 1900 par Bachelier [3], qui consiste à modéliser les cours financiers par des mouvements browniens. Il faut attendre plus de soixante ans pour qu'un autre modèle voit le jour et s'installe. En 1960, Samuelson [36] propose de retenir cette modélisation pour les rendements, plutôt que pour les cours eux-mêmes. Il est important de souligner que cette modélisation de Samuelson est à la base de la théorie très populaire de Black et Scholes [6]. Il faut noter que les propriétés très spécifiques de la modélisation gaussienne permettent de faciliter beaucoup de calculs théoriques dans la théorie financière. Dans les années 60, Mandelbrot étudie des fluctuations boursières, pour lesquelles il était tout à fait

clair que le modèle gaussien ne convenait pas. En effet, l'analyse des queues de distribution des rendements des actifs financiers, a priori indispensable pour une mesure fine des risques, montre que les valeurs des exposants de queues sont beaucoup trop grandes pour correspondre à des distributions normales. Mandelbrot [25] utilise alors les lois de Pareto pour mettre en évidence un nouveau modèle de variation des prix basé sur les "lois α -stables" dont la variance est infinie. Dans son article, Mandelbrot affirme que son modèle décrit de façon réaliste la variation des prix observées sur certaines bourses des valeurs. Deux ans plus tard, Fama [12] valide ce modèle pour le marché des actions.

De façon générale, il existe beaucoup de modèles différents pour les distributions des rendements. Certains supposent que les actifs financiers ont des rendements assez proches de distributions stables, d'autres, de distributions de Student, qui ne font pas partie de la classe des distributions stables. Il existe également d'autres modèles qui se basent sur un processus de Poisson composé, processus de Lévy, processus de Laplace avec et sans dérive, etc. Ce dernier fait l'objet de la naissance du modèle de Laplace asymétrique généralisée [30] qui s'accorde, comme tous les autres modèles sur le caractère à queue épaisse des distributions des rendements.

Dans ce travail, nous passons en revue, d'abord, la synthèse présentée dans le texte de développer certains outils qui permettent l'estimation des paramètres de la distribution de Laplace asymétrique généralisée [30] et l'approximation de celle-ci. Cependant, la littérature actuelle présente peu d'outils qui facilitent l'utilisation en pratique de cette distribution. Par la suite, la performance accrue des ordinateurs a permis de tenter d'appliquer le paradigme bayésien à l'estimation des modèles développés, donc en supposant des informations a priori sur les paramètres de ceux-ci. Ce qui nous permettra au passage d'appliquer ce paradigme au cas particulier de celui de Laplace asymétrique généralisé, dit modèle de Laplace symétrique.

Après une introduction générale, ce manuscrit est organisé en quatre chapitres.

Dans le premier chapitre, nous effectuons une brève revue des modèles d'évolution des prix des actifs financiers. Pour chaque modèle, nous présentons les faits stylisés¹ et en mettant l'accent sur ses principales faiblesses. Ce bref historique ayant mené à l'utilisation de la distribution de Laplace asymétrique généralisée dans le cadre de la modélisation des rendements financiers.

Dans le deuxième chapitre, nous adressons quelques approches fréquentistes sophistiquées développées dans la littérature afin d'estimer les paramètres des modèles statistiques. L'intérêt

1. Les propriétés statistiques observées dans la plus part des marchés financiers.

s'est focalisé sur l'application de ses approches dans le cas de la distribution de Laplace asymétrique généralisée.

Dans le troisième chapitre, nous donnons attention à l'estimation bayésienne et c'est elle qui fera l'objet de ce mémoire. Cette méthode, qui sera expliquée un peu plus en détail par la suite, consiste à générer un échantillon pour chaque paramètre conditionnellement à l'ensemble des autres paramètres en utilisant la technique de Monte Carlo par Chaînes de Markov (*MCMC*). L'application de ce paradigme dans le cas de la distribution de Laplace symétrique sera discutée dans le chapitre qui suit.

Le quatrième chapitre est une étude empirique d'une série financière considérée comme un échantillon formé de l'ensemble des prix $S_1(t)$ à la fermeture du titre Abbey National entre le 31 juillet et le 8 octobre 1991. Dans ce chapitre, nous présentons, d'abord, les résultats de l'estimation des paramètres de la distribution de Laplace asymétrique généralisée obtenus en utilisant les approches fréquentistes citées dans le deuxième chapitre. Par la suite, nous discutons les résultats de l'estimation bayésienne obtenus après implantation de la méthode en considérant le cas particulier "modèle de Laplace symétrique".

Enfin, nous clôturons ce travail par une conclusion générale. Toute notion fondamentale citée dans ce manuscrit est définie dans la partie Annexe.

Quelques modèles de rendements financiers

Introduction

Le but de ce chapitre est d'effectuer une brève revue des modèles d'évolution des prix des actifs financiers. Pour chaque modèle, nous présentons les faits stylisés qui sont reproduits ainsi ses principales faiblesses. Pour ce faire, nous nous référons à la note «Model Risk» publiée par Derman [10] .

1.1 Différents types de modèles

Toujours selon Derman, un modèle financier peut être classé parmi au moins trois catégories :

1. **Le modèle fondamental** : basé sur un système de données, entre lesquels on peut établir différentes relations. Le modèle de Black-Scholes en est un exemple.
2. **Le modèle phénoménologique** : C'est un modèle moins fondamental, basé sur des liens de cause à effet il présente une description ou une analogie afin d'illustrer quelque chose qui ne peut être directement observé. Un modèle qui chercherait à expliquer l'impact du retrait du porteur de parts majoritaire d'une entreprise sur la valeur des actions de celle-ci serait phénoménologique.
3. **Le modèle statistique** : basé sur une régression ou un réglage optimal entre différents ensembles de données. On cherche à décrire une tendance ou une corrélation. Le modèle d'évaluation des actifs financiers et celui des trois facteurs de Fama et French [13] en sont des exemples.

Un modèle financier est en partie basé sur des variables qui représentent des opinions et des anticipations. Ces variables peuvent être, le rendement et la volatilité future espérés

1.2 Les rendements financiers

Le rendement correspond au retour financier ou retour sur investissement obtenu par rapport à l'investissement précédemment effectué et s'exprime en pourcentage. Il permet d'évaluer la performance d'un investissement.

1.2.1 Définitions et notations

En finance, un titre, ou titre financier est un droit de propriété. Il représente une reconnaissance du fait qu'une personne ou une organisation est propriétaire d'une partie du capital d'une société ou d'une partie de la dette d'une société ou d'un état. On parle aussi de valeur mobilière pour parler d'un titre financier.

On définit le prix $S(t) > 0$ d'un titre financier observé au temps t . Implicitement, le prix considéré est celui à la fermeture. On définit aussi le taux de rendement effectif $R(t)$ sur une période comprise dans l'intervalle de temps $[t - 1, t]$. C'est le taux composé continument, aussi appelé force d'intérêt, qui aurait occasionné les mêmes gains ou pertes sur un montant déposé en banque au cours de la période concernée. Le taux de rendement est la variable d'intérêt dans le contexte de la modélisation financière.

On associe le taux de rendement effectif à la différence entre le logarithme du prix initial et final. Dans la situation où le taux de rendement est déterministe et non aléatoire, on obtient l'équation différentielle suivante :

$$\frac{dS(t)}{dt} = R(t).S(t).$$

On peut interpréter cette équation en affirmant que la variation du prix $dS(t)$ sur un intervalle de temps infiniment petit dt est proportionnelle à la valeur actuelle $S(t)$. Cette équation différentielle a pour solution générale :

$$S(t) = S(0)e^{R(t).t}. \tag{1.1}$$

Afin de définir les propriétés de l'échantillon sélectionné, on pose que le rendement $R(t)$ est constant durant la période définie par l'intervalle de temps $[t - 1, t]$, mais il est différent d'une à l'autre : $R(s) \neq R(t)$ ($s \neq t$). On peut alors représenter le rendement $R(t)$ comme étant le logarithme des prix observés au temps t et $t - 1$:

$$R(t) = \ln \left(\frac{S(t)}{S(t-1)} \right) \tag{1.2}$$

On définit aussi le rendement cumulé $L(t)$. Il correspond à la somme des rendements effectifs observés sur l'intervalle $[0, t]$:

$$\begin{aligned} L(t) &= \sum_{i=1}^t R(i) \\ &= \sum_{i=1}^t [\ln S(i) - \ln(S(i-1))] \\ &= \ln \left(\frac{S(t)}{S(0)} \right) \end{aligned}$$

1.2.2 Rendements cumulés

Si Les rendements $R(i)$, $i \in 1, \dots, t$ sont indépendants, mais pas nécessairement identiquement distribués. Alors on peut obtenir la distribution du rendement cumulé $L(t)$ en utilisant le produit de convolution¹. Considérons $\phi_{R(i)}(\xi)$ la fonction caractéristique d'un rendement et $\phi_{L(t)}(\xi)$ celle du cumulé $L(t)$. Alors la forme de cette dernière sur l'intervalle $[0, t]$ est donnée par :

$$\phi_{L(t)}(\xi) = \prod_{i=1}^t \phi_{R(i)}(\xi) \quad (1.3)$$

On considère la situation où les rendements $R(i)$, $i \in 1, \dots, t$ sont à la fois indépendants et identiquement distribués. Alors, la fonction caractéristique des rendements est égale pour chaque période :

$$\phi_R(\xi) = \phi_{R(1)}(\xi) = \dots = \phi_{R(t)}(\xi) \quad (1.4)$$

On peut donc simplifier l'expression (1.3) pour obtenir la fonction caractéristique :

$$\phi_{L(t)}(\xi) = [\phi_R(\xi)]^t \quad (1.5)$$

Considérer une distribution qui est fermée sous la convolution pour modéliser les rendements sur une période, $R(i)$ peut alors être intéressant. Le rendement cumulé $L(t)$ pourra aussi être modélisé à l'aide de la même distribution. Pour ce faire, on modifie un paramètre d'échelle en fonction de la longueur t de l'intervalle de temps considéré.

1.3 Données disponibles

Les données disponibles auprès des fournisseurs d'informations financières prennent habituellement la forme de séries chronologiques discontinues. Celles-ci incluent les prix à l'ouver-

1. Voir Annexe A.1

ture, le plus bas et le plus élevé au courant de la journée ainsi qu'à la fermeture, pour chaque jour où les marchés financiers sont en activité. Afin de mesurer le rendement quotidien d'un titre, seuls les prix à la fermeture seront considérés.

1.4 Quelques modèles de rendements

1.4.1 Le modèle de Bachelier

Un des premiers modèles proposés afin de représenter les rendements financiers a été celui de Bachelier [3]. Il propose qu'un intervalle $[-S(t), \infty]$ soit remplacé par l'ensemble du domaine réel \mathbb{R} car le prix d'un titre peut varier, durant une période, de n'importe quelle valeur comprise dans cette intervalle. La probabilité que le titre atteigne une valeur nulle ou négative ou que celle-ci double devrait donc être négligeable. Il ajoute aussi que la variation est indépendante du prix actuel du titre $S(t)$ et que la distribution de probabilités de celle-ci est symétrique et centrée en ce point.

Il utilise le principe selon lequel la probabilité que deux événements indépendants consécutifs aient lieu est le produit de celles que chacun d'entre eux se réalise, pour établir la distribution des variations du prix. Par exemple, la variation du prix sur une première période prend la valeur x et celle sur une seconde, $z - x$, comme illustré à la figure 1.1.

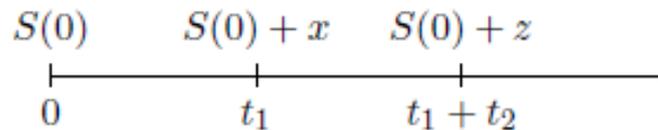


FIGURE 1.1 – Modèle de Bachelier : probabilité composée

On définit $f(x, t)$ la fonction de densité de la variation du prix $S(t)$ par rapport au niveau initial $S(0)$. Alors, selon le principe précédent, on obtient l'expression :

$$f(z, t_1 + t_2) = \int_{-\infty}^{+\infty} f(x, t_1) \cdot f(z - x, t_2) \cdot dx. \quad (1.6)$$

La solution proposée est que la densité de probabilité soit de la forme

$$f(x, t) = A \cdot \exp\{-B^2 x^2\}$$

Afin que la fonction $f(x, t)$ soit une densité de probabilité, la condition suivante doit être respectée :

$$\int_{-\infty}^{+\infty} A \cdot \exp\{-B^2 x^2\} dx = 1 \quad (1.7)$$

En reprenant l'intégrale (1.6), on obtient que la densité de probabilité $f(z, t_1 + t_2)$ soit aussi de la forme (1.7) :

$$f(z, t_1 + t_2) = \frac{f(x, 1)f(z - x, 2)}{\sqrt{f(x, 1)^2 + f(z - x, 2)^2}} \exp \left\{ -\pi \frac{f(x, 1)f(z - x, 2)}{f(x, 1)^2 + f(z - x, 2)^2} z^2 \right\} \quad (1.8)$$

Cette densité est, à un changement de variable près, une loi normale. La démarche suggère qu'il recherchait une distribution qui était fermée sous la convolution, une propriété souhaitable pour un modèle cohérent des rendements financiers.

Ce modèle implique un processus de Wiener-Bachelier selon lequel les incréments, ou les changements de prix, suivent une distribution normale :

$$S(T) - S(t) \rightsquigarrow N(0, \sigma^2(T - t)). \quad (1.9)$$

Ce modèle implique que la variance des fluctuations n'est pas proportionnelle au prix initial. Une première correction sera apportée au modèle afin de considérer le logarithme du prix. Ce changement permettra d'obtenir un modèle où elle est désormais proportionnelle au prix initial. Le processus du prix suivra alors un mouvement brownien géométrique :

$$S(T) - S(t) \rightsquigarrow LN(0, \sigma(T - t)). \quad (1.10)$$

Le logarithme du prix suivra alors un processus de Wiener-Bachelier :

$$\ln(S(T)) - \ln(S(t)) \rightsquigarrow N(0, \sigma(T - t)). \quad (1.11)$$

Un des principaux avantages du processus de Bachelier modifié est que le rendement cumulé $L(t)$ est aussi une variable aléatoire gaussienne. Cette propriété est appelée L-stabilité ou invariance sous l'addition. La distribution gaussienne est la seule ayant cette propriété où le second moment est fini.

Quelques années après sa publication, ce modèle est l'objet de critiques de la part d'économistes et de financiers. En se référant à Mitchell [29], on observe que, sur une base annuelle, les variations négatives par rapport à la moyenne (149) sont plus fréquentes que celles qui sont positives (126), pour un ensemble de 40 titres boursiers, entre 1890 et 1915 (figure 1.2).

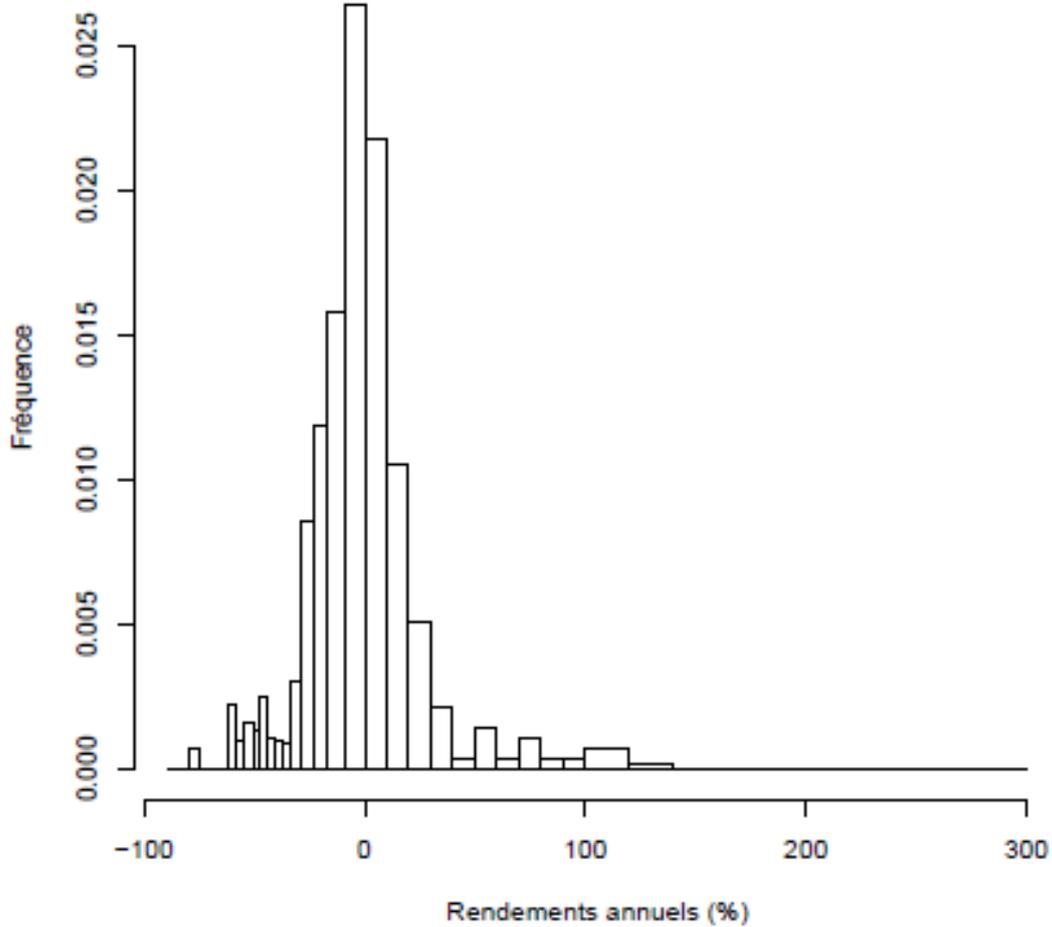


FIGURE 1.2 – Distribution des rendements annuels de 40 titres boursiers, de 1890 à 1915, Table XVIII de Mitchell [29].

Une asymétrie négative des rendements sera alors présente. De plus, les variations extrêmes sont plus fréquentes que ne pourrait le prédire un modèle basé sur un mouvement brownien. La distribution des rendements aurait donc des queues plus épaisses que la normale. Pour cela on doit trouver un modèle qui permet de tenir compte de ces particularités.

1.4.2 Proposition de Mandelbrot

Le modèle proposé par Mandelbrot [25] vise à combler les lacunes du processus brownien géométrique (1.10). Il explique que les distributions empiriques des changements de prix sont habituellement trop pointues pour être considérées comme des échantillons d'une population gaussienne. Il identifie différentes caractéristiques qu'un bon modèle des rendements financiers devrait posséder :

1. Il doit tenir compte de la fréquence des grands changements de prix. Il doit donc être basé sur une distribution leptokurtique, plus pointue au centre que la normale.
2. Il doit permettre des changements instantanés et imprévisibles de toute amplitude.
3. Il doit admettre une probabilité non nulle que plusieurs changements consécutifs semblent corrélés.
4. Il doit admettre un processus de prix non stationnaire, car la variance échantillonnale prend différentes valeurs à travers le temps.

La famille de distributions L stables semble être celle qui répond le mieux à l'ensemble de ces conditions (Walter [42]). L'équation suivante définit la propriété de L -stabilité de la distribution de la variable aléatoire des rendements sur une période R :

$$(a_1 R_1 + b_1) + (a_2 R_2 + b_2) \stackrel{d}{=} aR + b. \quad (1.12)$$

$$\forall a_1, a_2 > 0, \forall b_1, b_2.$$

La solution générale de cette équation a été découverte par Lévy en 1925. Le logarithme de la fonction caractéristique de celle-ci prend la forme suivante :

$$\ln(\phi_R(\xi)) = i\delta\xi - \gamma|\xi|^\alpha \left[1 + \frac{i\beta\xi}{|\xi|} \tan \frac{\alpha\pi}{2} \right]. \quad (1.13)$$

L'approche classique, selon Mandelbrot, pour expliquer les grands changements de prix a été de considérer un mélange de deux distributions normales, dont une pour les fluctuations régulières et une qui a une variance plus importante, pour les discontinuités. Il remarque que pour expliquer adéquatement le comportement des données empiriques, on doit introduire un mélange de plusieurs distributions normales, ce qui rendrait le modèle plus complexe.

1.4.3 Le modèle de Press

Press [32] propose un modèle statistique basé sur un processus de Poisson composé auquel on ajoute un mouvement brownien $B(t)$. Press utilise aussi la transformation logarithmique (1.10) afin que la variation soit proportionnelle au prix. Il remarque aussi que le modèle logarithmique de Bachelier est inadéquat, car il ne tient pas compte des queues de la distribution empirique des rendements qui sont plus épaisses que celles de la normale. Il ajoute que le modèle proposé par Mandelbrot est discutable, car il ne trouve aucune évidence, à partir des données observées, que la distribution de la population aurait une variance infinie.

Le processus de Poisson $\{N(t)\}$ de paramètre λt est un processus de comptage qui détermine les occurrences des sauts $Y_k, k = 1, \dots, N(t)$. Ces sauts surviennent généralement lorsqu'une information importante est rendue publique par rapport à un titre. Ceux-ci sont aussi de distribution normale, mais leur espérance n'est pas nulle et leur variance est différente de celle du processus $B(t)$. Cette composante que l'on ajoute au modèle de Bachelier modifié permet d'expliquer les variations plus importantes et moins fréquentes observées empiriquement.

Le processus du logarithme du prix $\{s(t)\} \equiv \{\ln(S(t))\}$ est donc représenté par l'équation suivante :

$$s(t) = s(0) + \sum_{k=1}^{N(t)} Y_k + B(t) \quad (1.14)$$

On définit les différentes variables aléatoires composant le processus comme suit :

$$\begin{aligned} Y_k &\rightsquigarrow N(\theta, \sigma_2^2) \\ B(t) &\rightsquigarrow N(0, \sigma_1^2 t) \\ N(t) &\rightsquigarrow \text{Poisson}(\lambda t) \end{aligned}$$

Comme pour la plupart des processus de Lévy, on ne peut obtenir une forme explicite pour la fonction de densité, car celle-ci se présente sous la forme d'une série infinie. On représente alors ces processus par leur fonction caractéristique, formée par le produit de celles de leurs différentes composantes.

La distribution du logarithme du prix $s(t)$ est définie par la fonction caractéristique $\phi_{s(t)}(\xi)$, qui est le produit de celle de la constante et celles des processus de Wiener et de Poisson composé :

$$\begin{aligned} \phi_{s(t)}(\xi) &= E \left[e^{i\xi s(t)} \right] \\ &= \exp\{i\xi \cdot s(0)\} \times \exp \left\{ \frac{-t\sigma_1^2 \xi^2}{2} \right\} \times \exp \left\{ \lambda t \left[e^{i\theta\xi - (\sigma_2^2 \xi^2 / 2)} - 1 \right] \right\} \\ &= \exp \left\{ i\xi \cdot s(0) - \frac{t}{2} \sigma_1^2 \xi^2 + \lambda t \left[e^{i\theta\xi - (\sigma_2^2 \xi^2 / 2)} - 1 \right] \right\}. \end{aligned} \quad (1.15)$$

Afin d'estimer le modèle, on s'intéressera plutôt à la distribution d'un incrément $\Delta s(t) = s(t) - s(t-1)$ de ce processus. La fonction caractéristique $\phi_{\Delta s(t)}$ de cette variable aléatoire peut être facilement identifiée à partir de celle du processus (1.15). Essentiellement, on pose $s(0) = 0$ et $t = 1$, pour obtenir :

$$\begin{aligned} \phi_{\Delta s(t)}(\xi) &= E \left[e^{i\xi \Delta s(t)} \right] \\ &= \exp \left\{ -\frac{\sigma_1^2 \xi^2}{2} + \lambda \left[e^{i\theta\xi - (\sigma_2^2 \xi^2 / 2)} - 1 \right] \right\} \end{aligned} \quad (1.16)$$

L'utilisation de ce modèle présente deux désavantages. L'estimation du modèle est difficile lorsque la moyenne s'approche de zéro². De plus, contrairement à d'autres modèles, il est difficile d'identifier le rôle des paramètres par rapport à un moment en particulier (classification de Pearson), contrairement à ce qu'on pourra observer avec la distribution de Laplace asymétrique généralisée.

1.4.4 Le modèle de Praetz

Praetz [31] propose un modèle inspiré par la physique des particules. Il pose comme hypothèse que deux intervalles qui ne se chevauchent pas forment une marche aléatoire, et que les éléments qui composent la séquence des rendements financiers $\{R(t)\}$ sont mutuellement indépendants. Il considère qu'un état stable existe où les rendements suivent une loi normale de paramètres μ et σ^2 .

Cependant, cet état stable n'est jamais réellement atteint, et la fonction de densité empirique généralement observée suppose une distribution symétrique concave, pointue au centre et ayant des queues épaisses. Il fait une analogie entre la température d'un gaz et le niveau d'activité sur les marchés, où la variance du mouvement brownien est proportionnelle à ces deux quantités. Il propose que le paramètre de variance de la normale σ^2 suive une distribution $g(\sigma^2)$ ayant un support positif. La distribution conditionnelle est normale lorsque ce paramètre est connu.

$$h_{R(t)}(r) = \int_0^\infty f_{R(t)}(r | \sigma^2) g(\sigma^2) d\sigma^2 \quad (1.17)$$

$$f_{R(t)}(r | \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(r - \mu)^2}{2\sigma^2}\right\} \quad (1.18)$$

Il propose comme solution acceptable pour la densité $g(\sigma^2)$, la distribution gamma inverse de paramètres m et s^2 :

$$g(\sigma^2) = \frac{s^{2m}(m-1)^m e^{-\frac{(m-1)s^2}{\sigma^2}}}{\sigma^{2(m-1)}\Gamma(m)} \quad (1.19)$$

Cette distribution a pour moyenne s^2 et variance $\frac{s^2}{m-2}$. La distribution non conditionnelle des rendements $h_{R(t)}$ est approximativement une Student avec $2m$ degrés de liberté à un facteur d'échelle de $\left(\frac{m}{m-1}\right)^{1/2}$ près :

2. Voir la référence [32]

$$h_{R(t)}(r) = \frac{\Gamma(m) [2(m-1)\pi]^{1/2} s}{\left[1 + \frac{(y-\mu)^2}{s^2(2m-2)}\right]^{m+1/2}}. \quad (1.20)$$

D'autres distributions pourraient être utilisées au lieu de la gamma inverse. En utilisant la loi gamma, on obtient la distribution de Laplace asymétrique généralisée, qui sera l'objet d'une étude approfondie aux chapitres suivants. Il propose enfin d'utiliser aussi la distribution a priori gamma inverse pour le paramètre μ . Par contre, il remarque qu'il obtient aussi une distribution similaire à celle de Student. Cette généralisation n'est donc pas nécessaire.

1.5 Conditions essentielles de Madan et Seneta

Inspirés par les travaux de Mandelbrot, Press et Praetz, Madan et Seneta [24] présentent un ensemble de conditions considérées essentielles dans l'élaboration d'un modèle de rendements financiers. Ils se baseront sur celles-ci pour proposer le modèle Variance Gamma :

1. La distribution des rendements R doit avoir une queue épaisse. Ainsi, la probabilité que cette variable aléatoire ait une valeur supérieure à $r + t$ avec un t petit, sachant qu'elle est supérieure à r , doit tendre vers 1, ce qui signifie que la fonction de survie converge lorsque cette quantité est grande.

$$\lim_{r \rightarrow \infty} P[R > r + t \mid R > r] = 1 \quad (1.21)$$

$$\bar{F}(r + t) \rightsquigarrow \bar{F}(r), \quad r \rightarrow \infty$$

2. La distribution doit posséder des moments finis pour les n premières puissances des rendements R . Étant donné que l'on cherche à modéliser la queue de la distribution, on fixe $n = 4$.

$$E[R^k] < \infty, \quad k \in \{1, 2, 3, 4\} \quad (1.22)$$

3. a) Le modèle doit proposer un processus de temps continu ayant des accroissements stationnaires et indépendants.
b) Les distributions des accroissements doivent appartenir à la même famille, quelle que soit leur longueur. Cette condition est essentielle afin de permettre l'échantillonnage et l'analyse des séries chronologiques.
4. Le modèle doit permettre une extension multivariée avec une distribution elliptique afin de conserver la validité du modèle d'évaluation des actifs financiers.

Chacun des modèles présentés précédemment respecte la majorité ou toutes ces conditions. Les résultats se retrouvent à la table 1.1.

Modèles	1	2	3	4
Mouvement brownien de Bachelier		*	*	*
Distribution stable symétrique de Mandelbrot	*	*		
Processus de Poisson composé de Press	*	*	*	*
Mélange gaussien/inverse gamma de Praetz	*	*	*	
Modèle Variance Gamma de Madan et Seneta	*	*	*	*

TABLE 1.1 – Respect des conditions émises par Madan et Seneta pour les différents modèles présentés

Ce qu'on peut remarquer d'après la table 1.1 que le modèle de Press vérifie toutes les conditions émises par Madan et Seneta. Cependant, ils remarqueront que ce n'est pas un processus de sauts, car il contient aussi une composante de diffusion³, ce qui va à l'encontre de l'intuition derrière la continuité de la trajectoire du prix. C'est cette dernière observation qui les incitera à proposer le modèle Variance Gamma, qui est un processus de sauts. Ce modèle, aussi étudié sous le nom de distribution de Laplace asymétrique généralisée par Kotz et al [23] qui sera présenté dans la section suivante.

1.6 Modèle de Laplace asymétrique généralisé

1.6.1 Le processus de Laplace

Le processus de Laplace est défini comme étant un processus de Wiener évalué à des temps aléatoires déterminés par un processus gamma. Selon Kotz et al. [21], ce dernier est à la distribution de Laplace ce que le mouvement brownien est à la loi normale. Il est aussi un cas particulier des processus de Lévy, et en conserve donc la principale propriété, celle d'être infiniment divisible. Mais, il possède une représentation alternative qui n'implique aucun processus de Wiener. Il peut en fait être représenté comme la différence de deux processus gamma indépendants. On peut le représenter en utilisant la forme générale de processus de Lévy.

3. Voir Annexe A.2.2

1.6.2 Le processus de Laplace est un processus subordonné

On considère un processus gamma⁴ $G(t; \tau, \beta = 1)$ et un processus de Wiener⁵ $W(t; \mu, \sigma^2)$. On se rappelle que la variance d'un incrément (1.13) de ce dernier est proportionnelle à la longueur de l'intervalle de temps. On peut modifier l'échelle de temps du processus de Wiener En utilisant une propriété appelée la subordination, de sorte que la variance soit aléatoire pour tout intervalle. Tout processus de Lévy peut être utilisé comme subordonnant pour définir cette échelle de temps. Si on utilise le processus gamma, on obtiendra le processus de Laplace sans dérive $\{Y(t; \sigma, \mu, \tau)\}$ défini comme suit :

$$\{Y(t; \sigma, \mu, \tau)\} \equiv \{W(G(t; \tau, \beta = 1); \mu, \sigma^2)\}. \quad (1.23)$$

On obtient l'exposant caractéristique $\Psi(\xi, t = 1; \sigma, \mu, \tau)$ d'un incrément $Y(t + 1; \sigma, \mu, \tau) - Y(t; \sigma, \mu, \tau)$ en utilisant la propriété de subordination définie par l'équation (A.5), où $\Xi(\zeta, t = 1; \tau, \beta = 1)$ est l'exposant caractéristique du processus gamma et $\Lambda(\xi, t = 1; \mu, \sigma^2)$ celui du processus de Wiener :

$$\begin{aligned} \Psi(\xi, t = 1; \sigma, \mu, \tau) &= \Xi(i\Lambda(\xi, t = 1; \mu, \sigma^2), t = 1; \tau, \beta = 1) \\ &= \tau \ln(1 - i(i\Lambda(\xi))) \\ &= \tau \ln \left(1 + \left(\frac{\sigma^2 \xi^2}{2} - i\mu\xi \right) \right) \end{aligned} \quad (1.24)$$

Le processus de Laplace sans dérive est donc, par définition, un processus de Lévy et par conséquent infiniment divisible. En utilisant l'exposant caractéristique (1.24) et la définition (A.6), on obtient sa fonction caractéristique :

$$\begin{aligned} \phi_{Y(t; \sigma, \mu, \tau)}(\xi) &= \exp\{-t \cdot \Psi(\xi, t = 1; \sigma, \mu, \tau)\} \\ &= \exp \left\{ -t \cdot \left(\tau \ln \left(1 + \left(\frac{\sigma^2 \xi^2}{2} - i\mu\xi \right) \right) \right) \right\} \\ &= \left(1 + \frac{\sigma^2 \xi^2}{2} - i\mu\xi \right)^{-\tau t}. \end{aligned} \quad (1.25)$$

Un processus stochastique qui représente le comportement du prix d'un titre financier doit inclure une composante de dérive. Celle-ci exprime le rendement moyen réalisé et est indépendante du processus de sauts. Pour cette raison, on ajoute un coefficient de dérive *alpha* au processus de Laplace sans dérive, pour obtenir sa forme générale. Comme ce coefficient est constant, on peut multiplier la fonction caractéristique (1.25) par la transformée de Fourier

4. Voir Annexe A.3.1

5. Voir Annexe A.13

inverse du produit de celui-ci et de la longueur de l'intervalle de temps t , $F^{-1}(\theta.t) = \exp^{i\xi\alpha.t}$, pour obtenir celle du processus de Laplace :

$$\phi_{y(t,\alpha,\sigma,\mu,\tau)}(\xi) = \left(\frac{e^{i\xi\theta}}{\left(1 + \frac{\sigma^2\xi^2}{2} - i\mu\xi\right)^\tau} \right)^t \quad (1.26)$$

Le processus $\{Y(t; \theta, \sigma, \mu, \tau)\}$ définit, dans le contexte financier, l'évolution du logarithme du prix, tel que présenté par Kotz et al. [21]. Pour des fins de simplification, on fixe le prix initial à 1 : $Y(0; \alpha, \sigma, \mu, \tau) = 0$.

La fonction caractéristique (1.26) constituera la principale représentation du processus de Laplace pour la suite de ce texte. La construction du modèle «Variance Gamma» de Madan et Seneta [24] est similaire, à l'exception que la paramétrisation et le processus gamma utilisés sont différents, ce qui rend leur approche moins intuitive, bien que le résultat soit équivalent.

1.7 Distribution de Laplace asymétrique généralisée

La distribution de Laplace asymétrique généralisée a été principalement étudiée par Kozubowski et Podgórski [22]. Cependant, elle a été introduite près d'une décennie auparavant par Madan et Seneta [24], sous le nom de distribution Variance Gamma. La différence entre les approches des deux auteurs est majeure. Madan et Seneta [24] développent un modèle financier à partir du mouvement brownien géométrique, qu'ils généralisent en proposant que la variance suive une distribution gamma. Kozubowski et Podgórski [22] généralisent la distribution de Laplace asymétrique. Leur approche est plus générale, car ils ne cherchent pas à développer un modèle financier, mais une nouvelle classe de distributions utilisable dans divers domaines scientifiques.

La distribution de Laplace asymétrique généralisée caractérise un intervalle du processus de Laplace avec dérive. Aussi appelée distribution de Bessel, elle a été introduite par Karl Pearson en 1929, en lien avec la covariance d'un échantillon tiré d'une population normale à deux variables. C'est aussi une généralisation de la distribution de Laplace asymétrique.

1.7.1 Fonction caractéristique

On définit cette distribution principalement par sa fonction caractéristique. Celle-ci s'obtient facilement à partir de la fonction caractéristique du processus de Laplace avec dérive, en considérant un incrément de longueur $t = 1$. À partir de la définition de ce dernier (1.26), on

déduit qu'elle est en fait un mélange de la loi normale dont le paramètre de variance suit une distribution gamma.

La variable aléatoire Y est définie comme suit : $Y = \alpha + \mu W + \sigma\sqrt{W}Z$ où

α est un paramètre de translation, $Z \rightsquigarrow \mathcal{N}(0, 1)$ et $W \rightsquigarrow \Gamma(\tau, \beta = 1)$

Alors, la variable aléatoire Y , sachant que $W = w$, suit une distribution normale de moyenne $w\mu$ et de variance $w\sigma^2$:

$$(Y|W = w) \rightsquigarrow N(w\mu, w\sigma^2).$$

La fonction caractéristique de la variable aléatoire Y de distribution Laplace asymétrique généralisée est donnée par :

$$\begin{aligned} \phi_Y(t; \alpha, \sigma, \mu, \tau) &= \frac{e^{i\alpha t}}{\left(1 + \frac{1}{2}\sigma^2 t^2 - i\mu t\right)^\tau} \int_0^{+\infty} \frac{\left(1 + \frac{1}{2}\sigma^2 t^2 - i\mu t\right)^\tau}{\Gamma(\tau)} w^{\tau-1} e^{-w\left(1 + \frac{1}{2}\sigma^2 t^2 - i\mu t\right)} dw \\ &= \frac{e^{i\alpha t}}{\left(1 + \frac{1}{2}\sigma^2 t^2 - i\mu t\right)^\tau}. \end{aligned}$$

1.7.2 Invariance d'échelle

Cette propriété permet d'appliquer un changement d'échelle à une variable aléatoire en modifiant un seul paramètre sans que la valeur des autres ne soit affectée. Ici on remplace μ par un paramètre d'invariance d'échelle κ . Le paramètre κ est obtenu à l'aide de la transformation suivante :

$$\kappa = \frac{\sqrt{2\sigma^2 + \mu^2} - \mu}{\sqrt{2}\sigma}.$$

À l'inverse, on peut retrouver le paramètre μ en l'isolant dans l'équation précédente. On obtient donc :

$$\mu = \frac{\sigma}{\sqrt{2}} \left(\frac{1}{\kappa} - \kappa \right)$$

1.7.3 Rôle des paramètres

Paramètre	Domaine	Rôle	Observations
α	R	Localisation	N'influence que la moyenne. Équivaut au mode lorsque $\mu = 0$.
σ	$R^+ \setminus \{0\}$	Échelle	Vrai paramètre d'échelle lorsque κ est utilisé.
μ	R	Asymétrie	Distribution asymétrique à gauche lorsque négatif et à droite lorsque positif. Déplace la moyenne dans la même direction. Corrélation positive avec la variance et le coefficient d'aplatissement.
κ	$R^+ \setminus \{0\}$	Asymétrie	Valeur dans l'intervalle $[0, 1]$ lorsque $\mu < 0$, dans $[1, \infty]$ lorsque $\mu \geq 0$.
τ	$R^+ \setminus \{0\}$	Aplatissement	Négativement corrélé avec le coefficient d'aplatissement. Une petite valeur donne une distribution pointue.

TABLE 1.2 – Domaine et rôle des paramètres de la distribution de Laplace asymétrique généralisée

1.7.4 Changement d'échelle et de localisation

Parfois, on doit modifier un ensemble de données afin d'effectuer des comparaisons ou pour bénéficier des avantages d'une méthode numérique. Les principales transformations utilisées sont :

1. un changement de localisation, où l'on additionne une constante à l'ensemble des données.
2. un changement d'échelle, où l'on multiplie l'ensemble des données par un facteur.
3. une combinaison des deux transformations précédentes.

Soit deux constantes $a, b \neq 0$ et une variable aléatoire X qui suit une distribution Laplace asymétrique généralisée : $X \rightsquigarrow GAL(\alpha, \sigma, \kappa, \tau)$. Alors :

$$Y = aX + b \rightsquigarrow GAL(a\alpha + b, a\sigma, \kappa^{sgn a.1}, \tau). \quad (1.27)$$

On estime les paramètres $\hat{\alpha}, \hat{\sigma}, \hat{\mu}$ et $\hat{\tau}$ sur un ensemble de données centrées et réduites X_t à partir d'un échantillon original Y_t .

\hat{m} et $\hat{s} > 0$ sont respectivement la moyenne et l'écart-type de l'échantillon Y_t . En utilisant

l'équation de transformation (1.27) avec les paramètres estimés précédemment, on peut alors obtenir ceux correspondants pour l'échantillon :

$$Y_t = \hat{s}X_t + \hat{m} \rightsquigarrow GAL(\hat{s}\alpha + \hat{m}, \hat{s}\sigma, \kappa, \tau)$$

Le paramètre κ n'est pas modifié puisque l'écart-type \hat{s} est strictement positif. Cette propriété permettra de pratiquer l'estimation sur des données centrées et réduites, ce qui diminue le risque d'erreurs numériques sans nuire à sa précision, puisqu'aucune information contenue dans l'échantillon n'est perdue.

Fonction de Bessel et densité

La densité de la distribution de Laplace asymétrique généralisée peut être exprimée à l'aide de la fonction de Bessel modifiée du troisième type Abramowitz et Stegun [1] :

$$K_\lambda(u) = \frac{(u/2)^\lambda \Gamma(1/2)}{\Gamma(\lambda + 1/2)} \int_1^\infty e^{-ut} (t^2 - 1)^{\lambda-1/2} dt, \quad \lambda \geq -1/2 \quad (1.28)$$

Le modèle de Laplace Asymétrique généralisé est donné sous la forme suivante (Kotz et al. [23]) :

$$f(x) = \frac{\sqrt{2} e^{\frac{\sqrt{2}}{2\sigma}(1/\kappa - \kappa)(x - \alpha)}}{\sqrt{\pi} \sigma^{\tau+1/2} \Gamma(\tau)} \left(\frac{\sqrt{2}|x - \alpha|}{\kappa + 1/\kappa} \right)^{\tau-1/2} K_{\tau-1/2} \left(\frac{\sqrt{2}}{2\sigma} \left(\frac{1}{\kappa} + \kappa \right) |x - \alpha| \right), \quad (1.29)$$

Avec $\lambda = \tau - \frac{1}{2}$

Cette fonction a été implémentée dans un algorithme de maximum de vraisemblance pour le logiciel *GNU R* par Scott et Dong [40]. Elle reste, cependant, sensible numériquement en plus de demander un temps de calcul important pour de grands échantillons. De plus, comme la fonction de densité n'est pas différentiable, on ne peut pas obtenir les estimateurs du maximum de vraisemblance et comme on ne peut pas appliquer l'approche bayésien ordinaire. C'est une des principales raisons qui ont motivé la recherche d'autres méthodes d'estimation plus efficaces pour cette distribution.

1.8 Cas particuliers

La distribution de Laplace asymétrique généralisée peut être considérée comme une extension de plusieurs cas particuliers :

1. Distribution Exponentielle

pour $\alpha = 0$, $\sigma = 0$, $\mu > 0$ et $\tau = 1$ on obtient la distribution Exponentielle de moyenne $\mu > 0$, sa fonction de densité est donnée par :

$$f(x) = \frac{1}{\mu} \exp^{-x/\mu}, \quad (x > 0)$$

2. Distribution de Gamma

pour $\alpha = 0$, $\sigma = 0$ et $\mu > 0$ on obtient la distribution de gamma de paramètres $\nu = \tau$, $\beta = \mu > 0$, sa fonction de densité est donnée par :

$$\frac{x^{\tau-1} \exp^{-x/\mu}}{\mu^\tau \Gamma(\tau)}, \quad (x > 0)$$

3. Distribution de Laplace Asymétrique

La distribution de Laplace asymétrique a été introduite par Hinkley et Revankar [20] et étudiée en profondeur par Kozubowski et Podgórski [23]. Cette distribution est notée $AL(\alpha, \sigma, \kappa)$.

Pour $\sigma > 0$, $\mu \neq 0$ et $\tau = 1$ on obtient la distribution Distribution de Laplace Asymétrique ($\sigma > 0$), sa fonction de densité est donnée par :

$$f(x; \alpha, \sigma, \kappa) = \frac{\sqrt{2}}{\sigma} \frac{\kappa}{1 + \kappa^2} \begin{cases} \exp\left(\frac{-\sqrt{2}\kappa}{\sigma}(x - \alpha)\right), & x \geq \alpha \\ \exp\left(\frac{\sqrt{2}}{\alpha\kappa}(\alpha - x)\right), & x < \alpha \end{cases}$$

4. Distribution de Laplace symétrique

La distribution de Laplace symétrique à deux paramètres est un cas particulier de la distribution de Laplace asymétrique quand $\sigma > 0$, $\mu = 0$, et $\tau = 1$, sa fonction de densité est donnée par :

$$f(x; \alpha, \sigma) = \frac{1}{\sqrt{2}\sigma} e^{-\sqrt{2}|x-\alpha|/\sigma}, \quad (x \in \mathbb{R}) \quad (1.30)$$

La Figure (1.3) représente les courbes des densités de probabilité de la loi de Laplace symétrique pour différentes valeurs des deux paramètres.

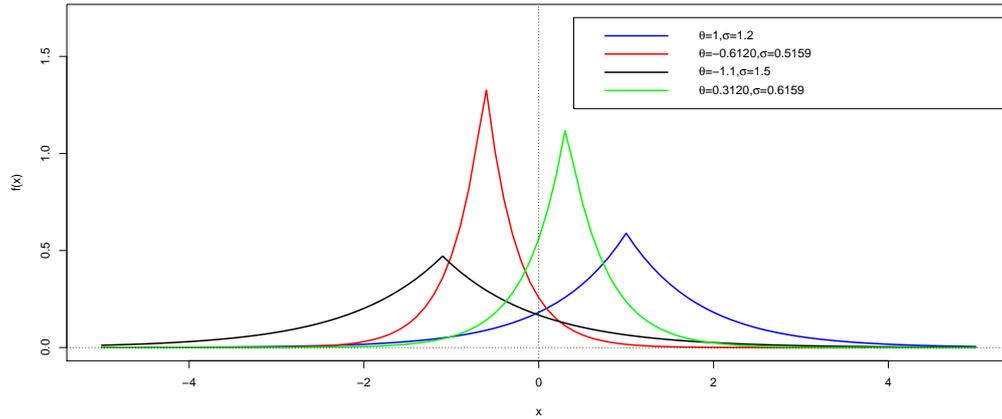


FIGURE 1.3 – Densité de probabilité de la distribution de Laplace symétrique.

5. Distribution Dégénérée

pour $\sigma = 0$, $\mu = 0$, $\tau = 1$ on obtient la distribution dégénérée à (α) , sa fonction de densité est donnée par :

$$f(x) = \mathbf{1}_{\{x=\alpha\}}$$

Conclusion

Dans ce chapitre, un bref historique sur les modèles classiques d'évolutions des rendements des actifs financiers, ayant mené à l'utilisation de la distribution de Laplace Asymétrique Généralisé, montrée comme une alternative intéressante de ceux-ci.

Rappel sur les approches fréquentistes

Introduction

Dans un modèle statistique, nous supposons qu'un échantillon issu d'une population a une certaine distribution de probabilité identifiée par certains paramètres. Ces paramètres peuvent être estimés à l'aide de différentes méthodes statistiques. Dans ce chapitre, nous passons en revue quelque approches fréquentistes d'estimation.

2.1 La méthode du maximum de vraisemblance

Soit X_1, X_2, \dots, X_n un n -échantillon indépendant issu d'une variable aléatoire X dont la loi dépend d'un paramètre inconnu θ .

La fonction du maximum de vraisemblance est définie comme suit :

$$L(x_1, x_2, \dots, x_n, \theta) = L(x, \theta) = \prod_{i=1}^n f(x_i; \theta) \quad (2.1)$$

L'estimateur de maximum de vraisemblance (*EMV*) de θ est la valeur $\hat{\theta}$ de θ qui rend maximales la fonction de vraisemblance $L(x, \theta)$.

Donc $\hat{\theta}_n$ sera, en général, calculé en maximisant la vraisemblance

$$\hat{\theta} = \arg \max_{\theta} L(x, \theta). \quad (2.2)$$

Puisque $L(x, \theta)$ et $\log L(x, \theta)$ atteignent leur maximum en un même point $\hat{\theta}_n$, et pour des raisons de simplification de calculs, $\hat{\theta}_n$ sera déterminé par :

$$\hat{\theta}_n = \arg \max_{\theta} \log L(x, \theta).$$

Quand $\theta = (\theta_1, \dots, \theta_d) \in \mathbb{R}^d$ et que toutes les dérivées partielles ci-dessus existent, $\hat{\theta}$ est la solution du système d'équations appelées équations de vraisemblance :

$$\forall j \in \{1, \dots, d\}, \frac{\partial}{\partial \theta_j} \log L(x; \theta) = 0.$$

Pour assurer l'existence du maximum, on vérifie l'inégalité

$$\frac{\partial^2 \log L(x, \theta)}{\partial \theta^2} < 0.$$

Donc, pour $\theta = (\theta_1, \dots, \theta_d) \in \mathbb{R}^d$, on vérifie

$$\frac{\partial^2 \log L(x, \theta_j)}{\partial \theta_j^2} < 0.$$

2.2 La méthode des moments

La méthode des moments permet d'estimer une espérance mathématique par une moyenne empirique, une variance par une variance empirique, etc.

Si le paramètre à estimer est l'espérance de la loi de X_i , alors on peut l'estimer par la moyenne empirique de l'échantillon. Autrement dit, si $\theta = E(X)$, alors l'estimateur de θ par la méthode des moments (*EMM*) est

$$\hat{\theta} = \frac{1}{n} \sum_{i=1}^n X_i \quad (2.3)$$

Plus généralement, pour $\theta \in \mathbb{R}$, si $E(\theta) = \varphi(\theta)$, ou φ est une fonction inversible, alors l'estimateur de θ par la méthode des moments est

$$\hat{\theta} = \varphi^{-1}(\bar{X}). \quad (2.4)$$

De la même manière, on estime la variance de la loi des X_i par la variance empirique de l'échantillon

$$S^2 = \frac{1}{n} \sum_{i=1}^n X_i^2 - \bar{X}^2. \quad (2.5)$$

Plus généralement, si la loi de deux paramètres θ_1 et θ_2 tels que $(E(X), V(X)) = \varphi(\theta_1, \theta_2)$

où φ est une fonction inversible alors les estimateurs de θ_1 et θ_2 par la méthode des moments sont

$$(\hat{\theta}_1, \hat{\theta}_2) = \varphi^{-1}(\bar{X}, S^2)$$

.

Ce principe peut naturellement se généraliser aux moments de tous ordres, centrés ou non centrés : $E((X - E(X))^k)$ et $E(X^k)$, $k \geq 1$.

2.3 La méthode des moments généralisée

En statistique, la méthode des moments généralisée (en anglais *generalized method of moments* ou *GMM*) est une méthode générique pour estimer les paramètres d'un modèle statistique qui s'appuie sur un certain nombre de conditions sur les moments d'un modèle.

Elle a été initialement proposé par Hansen dans [18], mais l'idée de base remonte à Sargan [37].

Nous considérons le processus $\{Y_{\Theta}[\kappa]\}_{\kappa}$ caractérisé par le vecteur de paramètres Θ de dimension p . Étant donné une observation $\{Y_{\Theta_0}[\kappa]\}_{\kappa}$, notons $f(Y_{\Theta_0}[\kappa], \Theta)$ la fonction de dimension r , supérieure à p , qui vérifie la condition sur les moments :

$$\mathbb{E}[f(Y_{\Theta_0}[\kappa], \Theta)] = 0, \quad \text{si et seulement si } \Theta = \Theta_0 \quad (2.6)$$

La condition sur les moments peut être approximée en utilisant la moyenne empirique

$$g_N(\Theta) = \frac{1}{N} \sum_{\kappa=1}^N f(Y_{\Theta_0}[\kappa], \Theta) \quad (2.7)$$

Définition 2.3.1. *L'estimateur GMM est défini par :*

$$\hat{\Theta} = \arg \min_{\Theta} (g_N^T W_N g_N), \quad (2.8)$$

où W_N est une suite de matrices de pondération qui converge vers une matrice W_{∞} symétrique définie positive lorsque N tend vers $+\infty$.

Théorème 2.3.1. *Si les propriétés suivantes sont vérifiées*

- le processus $\{Y[\kappa]\}_{\kappa}$ est ergodique.
- la série $\{f(Y_{\Theta_0}[\kappa], \Theta)\}_{\kappa}$ vérifie le théorème central limite, c'est-à-dire

$$\frac{1}{\sqrt{N}} \sum_{\kappa=1}^N f(Y_{\Theta_0}[\kappa], \Theta) \longrightarrow \mathcal{N}(0, V_{\Theta}) \quad (2.9)$$

où la matrice V_{Θ} est définie par :

$$V_{\Theta} = \lim_{K \rightarrow +\infty} \sum_{\kappa=-K}^K \mathbb{E} [f(Y_{\Theta_0}[\kappa], \Theta) f(Y_{\Theta_0}[\kappa], \Theta)^T] \quad (2.10)$$

- la matrice $(r \times p)$ -dimensionnelle $D_{g_N} = \frac{\partial g_N}{\partial \Theta}$ a un rang p et elle converge vers

$$Df = \mathbb{E} \left[\frac{\partial f(Y_{\Theta_0}[\kappa], \Theta)}{\partial \Theta} \right], \quad (2.11)$$

alors l'estimateur GMM $\hat{\Theta}$ est consistant et

$$\sqrt{N}(\hat{\Theta} - \Theta) \longrightarrow \mathcal{N}(0, \Sigma), \quad (2.12)$$

où

$$\Sigma = (Df^T W_\infty Df)^{-1} Df^T W_\infty V_{\Theta_0} W_\infty Df (Df^T W_\infty Df)^{-1} \quad (2.13)$$

Preuve 2.3.1. La démonstration est détaillée dans [18].

proposition 2.3.1. L'estimateur GMM est optimal si la limite W_∞ des matrices de pondération W_N (2.7) est égale à $V_{\Theta_0}^{-1}$, définie par l'équation (2.9). Dans ce cas la matrice de variance-covariance de l'estimateur GMM est donnée par :

$$\Sigma_{opt} = (Df^T V_{\Theta_0}^{-1} Df)^{-1} \quad (2.14)$$

Preuve 2.3.2. La démonstration est détaillée dans [18].

En pratique, il est bien évidemment impossible d'utiliser la matrice de pondération optimale $V_{\Theta_0}^{-1}$, puisque le vecteur Θ_0 est inconnu. On suit donc généralement la procédure récursive suivante :

1. On choisit la matrice de pondération W_N quelconque, on peut prendre par exemple la matrice \mathbb{I}_{d_N} ou encore la matrice $V_{\hat{\Theta}}^{-1}$ où $\hat{\Theta}$ est un estimateur a priori de Θ_0 .
2. On calcule (2.13) l'estimateur GMM $\hat{\Theta}$ en utilisant la matrice de pondération W_N .
3. On fixe la matrice de pondération $W_N = V_{\hat{\Theta}}^{-1}$ ou $\hat{\Theta}$ est l'estimateur GMM obtenu à l'étape 2.
4. On répète les étapes 2 et 3 jusqu'à ce que les valeurs successives de $\hat{\Theta}$ soient suffisamment proches l'une de l'autre.

Les intervalles de confiance de l'estimation $\hat{\Theta}$ peuvent être déterminés à l'aide de l'équation (2.11). Sous condition que le régime asymptotique est atteint. Si ce n'est pas le cas, seule une méthode de Monte Carlo permet d'obtenir ces intervalles de confiance.

2.4 Méthode de l'équation d'estimation optimale

Crowder [8] définit l'équation d'estimation sous une forme générale :

$$g(\theta; Y) = S^{-1}(\theta; Y) \sum_{t=1}^T u(\theta; Y) W(\theta; Y) \quad (2.15)$$

$S(\theta; Y)$: une matrice d'échelle.

$u(\theta; Y)$: un vecteur de fonctions aléatoires.

$W(\theta; Y)$: un vecteur de pondération.

En faisant abstraction du vecteur de pondération $W(\theta; Y)$, on retrouve une forme qui rappelle la méthode des moments généralisée. La fonction d'estimation $u(\theta; Y)$ définit, dans ce cas, des conditions de moment. Toutefois, en faisant abstraction de la matrice $S(\theta; Y)$, on obtient une classe d'équations d'estimation $g(\theta; Y)$ qui généralise plusieurs méthodes connues, comme développé par Crowder [9] :

$$g(\theta; Y) = \sum_{t=1}^T u(\theta; Y)W(\theta; Y). \quad (2.16)$$

L'équation d'estimation (2.15) , et donc (2.16) , c'est un produit des estimateurs sans biais, alors :

$$\mathbb{E}[g(\theta; Y)] = 0 \quad (2.17)$$

Par analogie avec la méthode du maximum de vraisemblance, on pourra alors la considérer comme une équation de quasi-score. On peut en effet représenter l'équation précédente par celle du score (2.18) donnée par l'espérance de la dérivée de la fonction de log-vraisemblance par rapport au vecteur de paramètres θ :

$$\mathbb{E}[g(\theta; Y)^{EMV}] = \mathbb{E} \left[\frac{\partial \ln L(\theta; Y)}{\partial \theta} \right] = 0 \quad (2.18)$$

Les dérivées premières de la moyenne $\mu(\theta)$ et de l'écart-type $\sigma^2(\theta)$ par rapport au vecteur de paramètres sont données par :

$$\mu'(\theta) = \frac{\partial \mu(\theta)}{\partial \theta} \quad (2.19)$$

$$\sigma'(\theta) = \frac{\sqrt{\partial \sigma^2(\theta)}}{\partial \theta} \quad (2.20)$$

Dans cette section, on considère une équation d'estimation quadratique de la forme suivante :

$$g(\theta; Y) = \sum_{i=1}^n [a(\theta; y_t)(y_t - \mu(\theta)) + b(\theta; y_t)(y_t - \mu(\theta))^2 - \sigma^2(\theta)] \quad (2.21)$$

Le vecteur de pondération $W(\theta; Y)$ de l'équation (2.16) est composé de deux fonctions déterministes : $a(\theta; Y)$ et $b(\theta; Y)$:

$$W(\theta; Y) = \begin{bmatrix} a(\theta; Y) \\ b(\theta; Y) \end{bmatrix} \quad (2.22)$$

La forme quadratique regroupe certaines méthodes d'estimation bien établies, recensées à la table 2.1.

La méthode d'estimation gaussienne de Whittle présente un avantage par rapport aux modèles linéaires conventionnels (moindres carrés pondérés et maximum de quasi-vraisemblance) du fait qu'elle prend en considération la variance comme critère d'optimisation et non seulement comme pondération.

Méthode	$a(\theta; Y_t)$	$b(\theta; Y_t)$
Moindres carrés non pondérés	$\mu'(\theta)$	0
Maximum de quasi-vraisemblance	$\frac{\mu'(\theta)}{\sigma^2(\theta)}$	0
Estimation gaussienne de Whittle (Fox et Taqqu, [15])	$\frac{\mu'(\theta)}{\sigma^2(\theta)}$	$\sigma^2(\theta)^{3/2}$

TABLE 2.1 – Méthodes d'estimation représentables par la forme quadratique (2.21)

Crowder [9] développe, à partir de cette dernière, une équation d'estimation optimale qui vise à remplacer la méthode de quasi-vraisemblance basée sur la moyenne et la variance jusqu'alors utilisée.

2.4.1 Équation d'estimation optimale

Soit une équation d'estimation de la forme générale (2.16) . On a, par définition :

$$\mathbb{E}[u(\theta_0; Y_t)] = 0 \quad (2.23)$$

$$\mathbb{E}[u(\theta_0; Y_s)u(\theta_0; Y_t)] = \delta_{s,t} Var[u(\theta_0; Y_t)]. \quad (2.24)$$

Les matrices non singulières $M(\theta_0)$ et $V(\theta_0)$ représentent respectivement, pour l'équation d'estimation, l'espérance de sa dérivée par rapport à θ et l'inverse de sa variance-covariance, évaluée à la vraie valeur des paramètres θ_0 :

$$\begin{aligned}
M(\theta_0; Y) &= \sum_{t=1}^T w(\theta_0) \mathbb{E}[u'(\theta_0; Y_t)] \\
V(\theta_0; Y) &= \sum_{t=1}^T w(\theta_0) \text{Var}[u(\theta_0; Y_t)] w^T(\theta_0)
\end{aligned} \tag{2.25}$$

La variance asymptotique de l'estimateur $\hat{\theta}$ prend alors la forme suivante, en utilisant la méthode delta multivariée :

$$\text{Var}[\hat{\theta}] = M^{-1}(\theta_0) V(\theta_0) [M^{-1}(\theta_0)]^T \tag{2.26}$$

On désire obtenir des équations d'estimation optimales de telle sorte que la variance soit minimale.

Selon Crowder [8], lorsqu'on utilise le vecteur de pondération optimal (2.27), on obtient que la dérivée de la fonction d'estimation $M^*(\theta)$ soit égale à l'inverse de la matrice de variance covariance $V^*(\theta)$. Le vecteur de pondération utilisé est :

$$w^*(\theta) = \{\mathbb{E}[u'(\theta; Y_t)]\}^T \{\text{Var}[u(\theta; Y_t)]\}^{-1}. \tag{2.27}$$

On considère la fonction $u(\theta; Y_t)$ de l'équation quadratique (2.21) , dont on cherche le vecteur de pondération optimal correspondant selon la proposition (2.26) :

$$u(\theta; Y_t) = \begin{bmatrix} Y_t - \mu(\theta) \\ (Y_t - \mu(\theta))^2 - \sigma^2(\theta) \end{bmatrix} \tag{2.28}$$

On évalue l'espérance, sous le vecteur de vrais paramètres θ_0 , de la dérivée par rapport à θ de la fonction $u(\theta; Y)$:

$$\mathbb{E}[u'(\theta; Y_t)] = \begin{bmatrix} -u'(\theta) \\ -2\sqrt{\sigma^2(\theta)}\sigma'(\theta) \end{bmatrix} \tag{2.29}$$

Puis, on évalue la variance-covariance du vecteur $u(\theta)$. Pour ce faire, on devra évaluer séparément chaque élément composant cette matrice :

$$\text{Var}[u(\theta; Y_t)] = \begin{bmatrix} \text{Var}[Y_t - \mu(\theta)] & \text{Cov}[Y - \mu(\theta), (Y - \mu(\theta))^2 - \sigma^2(\theta)] \\ \text{Cov}[Y - \mu(\theta), (Y - \mu(\theta))^2 - \sigma^2(\theta)] & \text{Var}[(Y_t - \mu(\theta))^2 - \sigma^2(\theta)] \end{bmatrix} \tag{2.30}$$

On définit les coefficients d'asymétrie et d'aplatissement normalisés, ainsi qu'une constante $\gamma_3(\theta)$ qui en découle, afin de simplifier les expressions qui seront obtenues conséquemment :

$$\begin{aligned}\gamma_1(\theta) &= \frac{m_3(\theta)}{m_2(\theta)^{3/2}} \\ &= \frac{E[(Y - \mu(\theta))^3]}{\sigma^2(\theta)^{3/2}}\end{aligned}\quad (2.31)$$

$$\begin{aligned}\gamma_2(\theta) &= \frac{m_4(\theta)}{m_2(\theta)^2} - 3 \\ &= \frac{E[(Y - \mu(\theta))^4]}{\sigma^2(\theta)^2} - 3\end{aligned}\quad (2.32)$$

$$\gamma_3(\theta) = \gamma_2(\theta) + 2 - \gamma_1(\theta)^2 \quad (2.33)$$

On obtient alors la matrice de variance-covariance de la fonction $u(\theta)$, dont on évalue par la suite l'inverse :

$$Var[u(\theta); Y_t] = \begin{bmatrix} \sigma^2(\theta) & \sigma^2(\theta)^{3/2}\gamma_1(\theta) \\ \sigma^2(\theta)^{3/2}\gamma_1(\theta) & \sigma^2(\theta)^2(\sigma_2 + 2) \end{bmatrix} \quad (2.34)$$

On peut enfin évaluer l'expression (2.27) à l'aide des résultats précédents :

$$w^*(\theta) = \left[-\mu'(\theta) - 2\sqrt{\sigma^2(\theta)}\sigma'(\theta) \right] \times \frac{1}{\sigma^2(\theta)^3\gamma_3(\theta)} \begin{bmatrix} \sigma^2(\theta)^2(\gamma_2(\theta) + 2) & -\sigma^2(\theta)^{3/2}\gamma_1(\theta) \\ -\sigma^2(\theta)^{3/2}\gamma_1(\theta) & \sigma^2(\theta) \end{bmatrix} \quad (2.35)$$

On obtient alors le vecteur de pondération optimal (2.22) formé des fonctions $a^*(\theta)$ et $b^*(\theta)$:

$$a^*(\theta) = \frac{-(\gamma_2(\theta) + 2)\mu'(\theta) + 2\gamma_1(\theta)\sigma'(\theta)}{\sigma^2(\theta)\gamma_3(\theta)} \quad (2.36)$$

$$b^*(\theta) = \frac{\gamma_1(\theta)\mu'(\theta) - 2\sigma'(\theta)}{\sigma^2(\theta)^{3/2}\gamma_3(\theta)}. \quad (2.37)$$

On obtient aisément les matrices $M^*(\theta)$ et $V^*(\theta)$ à partir de leur définition (2.22) :

$$\begin{aligned}M^*(\theta) &= V^*(\theta) \\ &= T \cdot \sigma(\theta)^{-2} \left\{ (u'(\theta))(u'(\theta))^T + \gamma_3(\theta)^{-1}(\gamma_1(\theta)\mu'(\theta) - 2\sigma'(\theta))(\gamma_1(\theta)\mu'(\theta) - 2\sigma'(\theta))^T \right\}\end{aligned}\quad (2.38)$$

Afin d'obtenir la solution optimale θ^* , on définit la fonction objectif $\Lambda(\theta; y_T)$:

$$\Lambda(\theta; y_T) = g(\theta; y_T)W(\theta; y_T)g(\theta; y_T)^T \quad (2.39)$$

$W(\theta; y_T)$ est une matrice définie positive. En premier lieu, on suggère d'utiliser la matrice identité. Puis, on peut raffiner l'estimation en lui substituant la matrice $V^*(\hat{\theta})$ de manière itérative, de la même manière qu'avec la méthode des moments généralisée itérative .

2.4.2 Équation d'estimation optimale modifiée

Il est possible que, pour certaines distributions, les expressions formant le vecteur de pondération optimal $w^*(\theta)$ soient particulièrement complexes. On pourra toujours les utiliser, cependant, il peut être intéressant de développer une approximation de celles-ci. Pour ce faire, on substitue les valeurs théoriques des coefficients d'asymétrie et d'aplatissement par une valeur estimée à partir de l'échantillon :

$$a_{mod}^*(\theta) = \frac{\{-(\hat{\gamma}_2(y_T) + 2)\mu'(\theta) + 2\hat{\gamma}_1(y_T)\sigma'(\theta)\}}{\sigma^2(\theta)\hat{\gamma}_3(y_T)} \quad (2.40)$$

$$b_{mod}^*(\theta) = \frac{\hat{\gamma}_1(y_T)\mu'(\theta) - 2\sigma'(\theta)}{\sigma^{3/2}(\theta)\hat{\gamma}_3(y_T)} \quad (2.41)$$

où

$$\hat{\gamma}_1(y_T) = \frac{\sum_{t=1}^T (y_t - \bar{y})^3}{\left(\sum_{t=1}^T (y_t - \bar{y})^2\right)^{3/2}}$$

$$\hat{\gamma}_2(y_T) = \frac{\sum_{t=1}^T (y_t - \bar{y})^4}{\left(\sum_{t=1}^T (y_t - \bar{y})^2\right)^2}$$

$$\hat{\gamma}_3(y_T) = \hat{\gamma}_2(y_T) + 2 - \hat{\gamma}_1(y_T)^2$$

On pourra ainsi réduire considérablement la taille des expressions à évaluer tout en conservant les propriétés asymptotiques des estimateurs, puisque l'espérance de chacune des statistiques utilisées est égale à la valeur que l'on a remplacée.

2.5 Test de Shapiro-Wilk

Nous utilisons un test de normalité de Shapiro-Wilk [41] afin de vérifier si l'hypothèse de normalité d'une distribution empirique peut être rejetée.

2.5.1 Description

Très populaire, le test de Shapiro-Wilk [41] est basé sur la statistique W . En comparaison des autres tests, il est particulièrement puissant pour les petits effectifs ($n \leq 50$). La statistique du test s'écrit :

$$W = \frac{\left[\sum_{i=1}^{\lfloor \frac{n}{2} \rfloor} a_i (x_{(n-i+1)} - x_{(i)}) \right]^2}{\sum_i (x_i - \bar{x})^2} \quad (2.42)$$

où

- $x_{(i)}$ correspond à la série des données triées.
- $\lfloor \frac{n}{2} \rfloor$ est la partie entière du rapport $\frac{n}{2}$.
- a_i sont des constantes générées à partir de la moyenne et de la matrice de variance co-variance des quantiles d'un échantillon de taille n suivant la loi normale.

La statistique W peut donc être interprétée comme le coefficient de détermination (le carré du coefficient de corrélation) entre la série des quantiles générées à partir de la loi normale et les quantiles empiriques obtenues à partir des données. Plus W est élevé, plus la compatibilité avec la loi normale est crédible. La région critique, rejet de la normalité, s'écrit :

$$R.C. : W < W_{crit}$$

Les valeurs seuils W_{crit} pour différents risques α et effectifs n sont lues dans la table de Shapiro-Wilk [41].

Conclusion

Afin de faciliter l'utilisation en pratique de certaines distributions de probabilité, en faisant recours à apporter des modifications de certains outils classiques qui permettent l'estimation des paramètres de celle-ci.

Ce chapitre a eu objectif d'illustrer des démarches de la prise en comptes des propriétés statistiques¹ afin d'y apporter des modifications aux approches classiques d'estimation.

1. faits stylisés

Le paradigme bayésien

Introduction

la démarche statistique classique repose essentiellement sur un principe de vraisemblance qui consiste à considérer que ce qui a été observé rend compte de manière exhaustive du phénomène . L'analyse bayésienne des problèmes statistiques propose d'introduire dans la démarche d'inférence, l'information dont dispose a priori le praticien. Dans le cadre de la statistique paramétrique, ceux-ci ce traduira par le choix d'une loi sur le paramètre d'intérêt.

Ce chapitre présente brièvement le développement historique de l'approche bayésienne ainsi le principe de la méthode *MCMC*.

3.1 Modèle bayésien

Le but de l'analyse statistique est de faire de l'inférence sur θ (utiliser les observations pour améliorer l'information sur le paramètre) afin de pouvoir prendre une décision. Si toute l'information à propos de θ , est celle qui est apportée par les observations, on peut appliquer la méthode du maximum de vraisemblance, et on interprète la loi des observations (vraisemblance) comme la loi conditionnelle des observations $x = (x_1, x_2, \dots, x_n)$ sachant θ . On note sa densité par $L(x|\theta)$.

L'approche bayésienne consiste à utiliser cette information (vraisemblance) pour l'estimation du paramètre inconnu θ . On considère le cas où on a des connaissances a priori sur le paramètre θ inconnu.

Définition 3.1.1. [34]

Un modèle statistique bayésien est la double donnée d'un modèle paramétrique $f_\theta(x), \theta \in \Theta$

et une loi de probabilité, de densité, π , dite aussi loi a priori qui est la loi marginale de la variable aléatoire θ .

3.1.1 Modélisation de l'information a priori

L'une des caractéristiques fondamentales de l'analyse bayésienne est l'utilisation des densités de probabilité décrivant l'état de connaissance ou d'ignorance concernant les paramètres avant la prise en compte des observations. Le choix d'une telle densité $\pi(\theta)$ qui est dite a priori est le problème le plus délicat et le plus critique de l'analyse Bayésienne. En effet, il est très rare que l'information a priori disponible soit suffisamment précise pour pouvoir déterminer exactement cette densité. Il faut donc faire des approximations dont le choix aura des répercussions sur les inférences statistiques bayésiennes.[46]

Le choix des lois a priori est une étape fondamentale en statistique bayésienne et constitue une différence notable avec la statistique fréquentiste. Les différents choix possibles peuvent être motivés par différents points de vue :

- Choix basé sur des expériences du passé ou sur une intuition du statisticien.
- Choix basé sur la faisabilité des calculs.
- Choix basé sur la volonté de n'apporter aucune information nouvelle pouvant biaiser l'estimation.

On introduira deux types de lois a priori qui sont les plus courants : les densités a priori informatives et les densités a priori non informatives .

• Distribution a priori informative (conjuguées naturelles)

La modélisation a priori informative est le point le plus délicat de l'analyse bayésienne. Il existe plusieurs procédés pour obtenir des lois informatives. On donne la description de l'un des plus intéressants des procédés qui est celui des familles naturelles conjugues.

Définition 3.1.2. *Une famille F de distributions sur Θ est dite conjuguee pour la loi $f(x|\theta)$ si pour tout $\pi \in F$, la distribution a posteriori $\pi(\theta|x)$ appartient également à F .*

L'avantage des familles conjugues est avant tout de simplifier les calculs. Avant le développement des outils de calcul numérique, ces familles étaient pratiquement les seules qui permettaient de faire aboutir des calculs. Un autre intérêt est que la mise à jour la loi se fait à travers les paramètres de la loi et donc l'interprétation est souvent bien plus facile.

$f(x \theta)$	$\pi(\theta)$	$\pi(\theta x)$
$N(\theta, \sigma^2)$	$N(\mu, \tau^2)$	$N(\varrho(\mu\sigma^2 + \tau^2x), \varrho\sigma^2\tau^2)$ $\varrho^{-1} = \sigma^2 + \tau^2$
$N(\mu, 1/\theta)$	$G(\alpha, \beta)$	$G(\alpha + 1/2, \beta + (\mu - x)^2/2)$
$\rho(\theta)$	$G(\alpha, \beta)$	$G(\alpha + x, \beta + 1)$
$G(\nu, \theta)$	$G(\alpha, \beta)$	$G(\alpha + \nu, \beta + x)$
$B(n, \theta)$	$B(\alpha, \beta)$	$B(\alpha + x, \beta + n - x)$
$Bneg(m, \theta)$	$B(\alpha, \beta)$	$B(\alpha + m, \beta + x)$

TABLE 3.1 – Exemple de lois a priori conjuguées naturelles pour quelques familles exponentielles usuelles

• Distribution a priori non informative

Dans le cas où on dispose que de peu d'informations sur θ , on peut choisir des loi a priori dites peu ou non informatives. On souhaite que l'a priori intervienne de façon minimale dans la loi a posteriori ; c'est à dire que les données parlent d'elles mêmes.

a) A priori de Laplace

Historiquement, Laplace¹ fut le premier à utiliser des techniques non informatives. Il munit les paramètres d'une loi a priori qui prend en compte son ignorance en donnant la même vraisemblance à chaque valeur du paramètre, donc en utilisant une loi uniforme. Ainsi, la densité a priori d'un paramètre θ est défini par :

$$\pi(\theta) = k, \quad \text{où } k \text{ est une constante.}$$

b) Mesures a priori de Jeffreys [11]

Le mode de spécification d'une loi a priori non informative connue sous le nom de la mesure a priori de Jeffreys consiste à assigner à un modèle d'échantillonnage caractérisé par sa vraisemblance $L(x|\theta)$, la mesure a priori de densité :

$$\pi_j(\theta) \propto [\det I(\theta)]^{\frac{1}{2}}$$

où θ est un vecteur de paramètres inconnus et $I(\theta)$ est la matrice de l'information de Fischer définie par :

$$I(\theta) = -E \left[\frac{\partial^2}{\partial \theta_i \partial \theta_j} \ln L(x|\theta) \right].$$

1. Nous devons envisager l'état présent de l'Univers comme un effet de l'état antérieur et comme la cause de l'état suivant, 1795.

c) A priori impropre

La loi a priori peut être impropre c'est à dire :

$$\int_{\Theta} \pi(\theta) d\theta = \infty.$$

Ce choix de type de loi n'a donc plus d'intérêt que calculatoire et s'interprète difficilement. La construction de lois non informatives peut conduire à des lois a priori de ce type.

3.1.2 Calcul de la loi a posteriori

Soit $X = (X_1, \dots, X_n)$ un n -échantillon de variables aléatoires indépendantes et identiquement distribuées (*iid*) de densité $f(x|\theta)$. On suppose que θ est une valeur d'une variable aléatoire θ de densité $\pi(\theta)$ et soit la fonction de vraisemblance des observations.

$$L(x|\theta) = \prod_{i=1}^n f(x_i|\theta).$$

La distribution a posteriori de θ notée $\pi(\theta|x)$ représente la distribution conditionnelle de θ sachant les observations $x = (x_1, \dots, x_n)$.

$$\pi(\theta|x) = \frac{\prod_{i=1}^n f(x_i|\theta)\pi(\theta)}{\int_{\Theta} \prod_{i=1}^n f(x_i|\theta)\pi(\theta) d\theta}$$

Ou encore, on peut écrire :

$$\pi(\theta|x) \propto f(x_i|\theta)\pi(\theta),$$

où \propto signifie "proportionnelle à".

La distribution a posteriori $\pi(\theta|x)$ donne alors l'information dont on dispose sur θ après observation. Elle représente un compromis entre l'information a priori (donnée par π) et l'information tirée de l'échantillon X (donnée par $L(x|\theta)$).

3.2 Inférence bayésienne décisionnelle

L'ensemble des décisions possibles D est appelé espace des décisions et la plupart des exemples théoriques se concentrent sur le cas $D = \Theta$ qui représente le cadre d'estimation standard.

Une inférence bayésienne est fondée sur la détermination rigoureuse de trois facteurs :

- (a) La loi des observations $L(x|\theta)$

(b) La distribution a priori des paramètres $\pi(\theta)$.

(c) Le coût associé aux décisions $l(\theta, d)$.

On note que les critiques fréquentes de l'approche bayésienne se font souvent sur le point b , alors que, conceptuellement, les points a et c se trouvent sur le même rang et que le point 3, qui est un point commun entre l'approche classique et l'approche bayésienne, est encore plus subtil. Le choix d'une fonction de coût dépend de l'application concernée. Lorsque le contexte ne conduit pas à la détermination de la fonction d'utilité qui permet d'en déduire la fonction de coût, on peut avoir recours à des fonctions de coûts classiques, qui sont à la fois simples et bien étudiées.

3.2.1 Fonction de perte et risque bayésien

• Fonction de perte

C'est la fonction notée l de $\Theta \times D$ dans $[0, \infty]$

$$\begin{aligned} l : \Theta \times D &\longrightarrow \mathbb{R}^+ \\ (\theta, d) &\longrightarrow l(\theta, d) \end{aligned}$$

$l(\theta, d)$ évalue le coût d'une décision d lorsque le paramètre vaut θ .

• Coût quadratique

Parmi les fonctions de perte les plus utilisées on peut noter le coût quadratique

$$l(\theta, d) = (\theta - d)^2$$

qui pénalise trop fortement les grandes erreurs, du fait de sa convexité stricte.

• Fonction de risque

La fonction de risque est l'espérance mathématique de fonction de perte composée avec une fonction de décision [34].

$$R(\theta, d) = \mathbb{E}_\theta[l(\theta, d(x))]$$

• Risque bayésien

Définition 3.2.1. Soit $T = T(x_1, \dots, x_n)$ un estimateur de θ et $l(t, \theta)$, la fonction de perte associée. On appelle risque bayésien moyen noté $R_T(\theta)$, la quantité :

$$R_T(\theta) = \mathbb{E}_\theta[l(\theta, t)] = \int_{\Theta} l(\theta, t) L(x|\theta) dx$$

3.2.2 Estimateur de Bayes

Définition 3.2.2. La règle de Bayes associée à la loi a priori π est donnée par la solution du système :

$$\min_d \mathbb{E}^\pi [l(\theta, d)|X] \quad (3.1)$$

On appelle estimateur de Bayes toute fonction δ^π qui associe à x une solution de (3.1). L'estimateur de bayes est donc défini par :

$$\hat{\theta} = \delta^\pi = \mathbb{E}[\theta|X]$$

Définition 3.2.3. [11] **Admissibilité** Un estimateur δ est inadmissible s'il existe un estimateur δ' tel que $R(\theta, \delta') \leq R(\theta, \delta)$ pour tout $\theta \in \Theta$ et $R(\theta_0, \delta') < R(\theta_0, \delta)$ pour au moins un $\theta \in \Theta$. Dans le cas contraire, δ est admissible .

• Propriétés de l'estimateur de Bayes [11]

- L'estimateur de Bayes est admissible.
- L'estimateur de Bayes est biaisé.
- L'estimateur de Bayes converge en probabilité pour la taille de l'échantillon $n \rightarrow \infty$.

3.3 Les méthodes de Monte Carlo par Chaîne de Markov (MCMC)

La méthode *MCMC* est publiée en 1953 par Metropolis et ses coauteurs [27], elle est étendue en 1970 par Hastings [19]. En 1984, les frères Geman proposent l'échantillonneur de Gibbs pour la restauration bayésienne d'images. Cet échantillonneur est développé par Gelfand et Smith [16]. C'est au début des années 90, après le développement et la démocratisation de l'outil informatique, que ces outils rencontrent un important succès [44] .

En statistique bayésienne, θ est considéré comme un vecteur aléatoire de densité $\pi(\theta)$ (loi a priori sur θ). L'estimation bayésienne est basée sur le calcul de la loi a posteriori de θ dont la densité est notée $\pi(\theta|y)$. Les estimateurs bayésiens sont de la forme $\mathbb{E}_\pi(h(\theta))$ où \mathbb{E}_π est l'opérateur des espérances mathématiques pour la loi $\pi(\theta | y)$.

$$I_h = \mathbb{E}_\pi(h(\theta)) = \int h(\theta)\pi(\theta|y)d\theta$$

Souvent, le calcul explicite de la densité $\pi(\theta|y)$ n'est pas envisageable ; ainsi de nombreuses méthodes d'approximation ont été proposées et la plus utilisée est celle de *MCMC* dont le

principe est de construire une chaîne de Markov ergodique de réalisations $\theta^{(1)}, \dots, \theta^{(t)}, \dots$, qui a pour distribution stationnaire $\pi(\theta|y)$, ainsi l'estimateur qui sera obtenu est \hat{I}_h :

$$\hat{I}_h = \frac{1}{T} \sum_{t=1}^T h(\theta^{(t)}).$$

La convergence de \hat{I}_h vers I_h est assurée par l'ergodicité de la chaîne de Markov. Parmi les méthodes *MCMC*, l'algorithme de Hastings-Metropolis [19] et l'algorithme de Gibbs [17] sont les plus utilisés et ont donné lieu à de nombreux algorithmes dérivés.

3.3.1 Algorithme de Metropolis-Hastings

Le principe est de construire une chaîne de Markov ergodique² de loi stationnaire $\pi(\theta|y)$ partant de θ^0 , on génère une chaîne θ^t à partir d'un noyau de transition de loi stationnaire $\pi(\theta|y)$, qui garantit de plus la convergence en loi vers $\pi(\theta|y)$. Pour T "assez grand" on peut considérer θ^T comme distribué suivant $\pi(\theta|y)$ et obtenir ainsi un échantillon θ^T qui est effectivement distribué suivant $\pi(\theta|y)$.

Algorithme 1 [45]

L'algorithme de Metropolis-Hastings consiste à simuler un échantillon selon une distribution objectif $\pi(\theta|y)$, à partir d'une distribution conditionnelle $q(y|\theta)$. Pour la mise en œuvre de l'algorithme, on doit pouvoir simuler facilement à partir de la distribution $q(\cdot|\theta)$, et sa densité $q(y|\theta)$ doit être disponible analytiquement (au moins à une constante multiplicative près) ou à défaut, le rapport $\pi(\theta|y)/q(y|\theta)$ doit être fini.

- **Étape 1** : Pour $t = 0$, on initialise $\theta^{(0)}$ le premier élément de la chaîne ;
- **Étape 2** : Pour $t \in \{1, \dots, T\}$;
 - On génère $u \rightsquigarrow U_{[0;1]}$;
 - y_t selon $q(\cdot | \theta^{t-1})$;
 - On calcul $\rho(\theta^t, y_t) = \text{Min} \left\{ 1, \frac{\pi(y_t) q(\theta^{(t)}|y_t)}{\pi(\theta^{(t)}) q(y_t|\theta^{(t)})} \right\}$
 - Puis, on prend

2. voir Annexe A.5

$$\theta^{t+1} = \begin{cases} y_t, & \text{avec la probabilité } \rho(\theta^t, y_t) & \text{si } u < \rho(\theta^t, y_t) \\ \theta^t, & \text{avec la probabilité } 1-\rho(\theta^t, y_t) & \text{sinon.} \end{cases}$$

– **Étape 3** : $t = t + 1$ et retourner à l'étape 2 ;

– **Étape 4** : finalement calculer l'estimateur de Bayes $\hat{I}_h = \frac{1}{T - T_0} \sum_{t=1}^T h(\theta^{(t)})$.

3.3.2 Echantillonneur de Gibbs

Plus généralement, si l'on peut écrire $\theta_1, \dots, \theta_p$ une partition du vecteur des paramètres θ pour le modèle étudié, et si l'on peut spécifier complètement les lois de comportement a posteriori conditionnelles alors on peut facilement simuler, pas par pas, des réalisations conditionnelles de $\theta_1, \dots, \theta_p$. En itérant le procédé un grand nombre de fois, la chaîne de Markov produite par ces simulations répétées a pour distribution stationnaire $\pi(\theta|y)$.

Algorithme 2 [4]

– **Étape 1** : Initialiser $\theta^{(0)} = (\theta_1^{(0)}, \theta_2^{(0)}, \dots, \theta_p^{(0)})$;

– **Étape 2** : A l'itération i : simuler

$$\theta_1^{(i+1)} \rightsquigarrow \pi_1(\theta_1 | \theta_2^{(i)}, \dots, \theta_p^{(i)}) ;$$

$$\theta_2^{(i+1)} \rightsquigarrow \pi_2(\theta_2 | \theta_1^{(i+1)}, \dots, \theta_p^{(i)}) ;$$

$$\theta_p^{(i+1)} \rightsquigarrow \pi_p(\theta_p | \theta_1^{(i+1)}, \dots, \theta_{p-1}^{(i+1)}) ;$$

– **Étape 3** : $i \rightarrow i + 1$ et retourner à l'étape 2.

Les densités conditionnelles π_i sont appelées conditionnelles complètes.

3.3.3 Propriétés :

- Taux d'acceptation égal à 1. Toutes les valeurs simulées sont acceptées.
- Nécessite de connaître les lois conditionnelles de π , d'où une connaissance préalable de certaines propriétés probabilistes ou analytiques de π .

Conclusion

Dans ce chapitre, nous avons passé en revue le principe d'application de l'approche bayésienne pour l'estimation des paramètres d'une distribution de probabilité. La complexité de la loi a posteriori, dans l'approche bayésienne, nécessite l'utilisation de la méthode *MCMC*.

Application sur des données réelles

Introduction

L'utilisation de la distribution de Laplace asymétrique généralisée $GAL(\alpha, \sigma, \kappa, \tau)$ dans le cadre de la modélisation de rendements financiers est de plus en plus courante dans la pratique. Nous abordons dans ce chapitre une étude empirique d'une série financière considérée comme un échantillon formé de l'ensemble des prix $S_1(t)$ à la fermeture du titre Abbey National entre le 31 juillet et le 8 octobre 1991. Cet échantillon a notamment été utilisé précédemment par Buckle [7], ensuite, par Pelletier [30] qui a considéré l'échantillon des rendements logarithmiques obtenu à l'aide de l'équation 1.2 à partir de ladite série. Cet auteur a modélisé cet échantillon par la distribution de Laplace asymétrique généralisée. D'abord, nous présentons, dans ce chapitre, les résultats de l'estimation obtenu dans le travail de Pelletier. Ensuite, nous nous intéressons au cas particulier Laplace symétrique (par rapport au cas généralisé ($\kappa = 0, \tau = 1$). En effet, une telle loi $SL(\alpha, \sigma)$ dépend seulement de deux paramètres ; un paramètre $\alpha \in \mathbb{R}$ lié à la localisation et un paramètre $\sigma \in \mathbb{R}_+$ lié à l'échelle. Pour estimer ces deux paramètres nous supposons des lois a priori pour chaque paramètre.

4.1 Collecte et faits stylisés des données

4.1.1 Description des données réelles

Nous considérons un échantillon S_1 formé de l'ensemble des prix $S_1(t)$ à la fermeture du titre Abbey National entre le 31 juillet et le 8 octobre 1991. La table B.1 présente l'ensemble des 50 observations.

Nous évaluons tout d'abord les rendements quotidiens R_1 à l'aide de l'équation (1.2). On

obtient alors 49 observations du processus des rendements $R(t)$ qui seront présentés dans la table 4.1.

La table 4.2 présente quelques statistiques d'ordre de l'échantillon R_1 .

R_t	Valeur	R_t	Valeur	R_t	Valeur	R_t	Valeur
R_1	0	R_{14}	0.003478	R_{27}	-0.003263	R_{40}	0.043400
R_2	0.013423	R_{15}	0.030772	R_{28}	-0.006557	R_{41}	-0.009852
R_3	0.006645	R_{16}	0.026580	R_{29}	-0.013245	R_{42}	-0.006623
R_4	-0.006645	R_{17}	0.006536	R_{30}	-0.013423	R_{43}	0.006620
R_5	0.013245	R_{18}	0	R_{31}	0.016751	R_{44}	0.016367
R_6	-0.003295	R_{19}	-0.009820	R_{32}	-0.010017	R_{45}	-0.009790
R_7	-0.013289	R_{20}	-0.003260	R_{33}	-0.010118	R_{46}	-0.009885
R_8	-0.020271	R_{21}	0.003295	R_{34}	0	R_{47}	-0.003317
R_9	0.003407	R_{22}	0	R_{35}	-0.006803	R_{48}	-0.013378
R_{10}	0	R_{23}	0.016314	R_{36}	-0.003419	R_{49}	0.006711
R_{11}	-0.003407	R_{24}	0	R_{37}	0.016978		
R_{12}	0.006803	R_{25}	0	R_{38}	-0.010152		
R_{13}	-0.027500	R_{26}	-0.006494	R_{39}	-0.003407		

TABLE 4.1 – La distribution des rendements financiers

Statistique d'ordre	Valeur
Minimum	-0.027500
1er quartile	-0.009790
Médiane	-0.003260
3e quartile	0.006620
Maximum	0.043400

TABLE 4.2 – Statistiques d'ordre de l'échantillon R_1

4.2 Test d'ajustement de la non normalité des rendements

Nous présentons la distribution et l'histogramme des rendements financiers à la figure 4.1. Pour assurer la stabilité de la routine de la série des rendements journaliers, ceux-ci seront multipliés par le facteur 100.

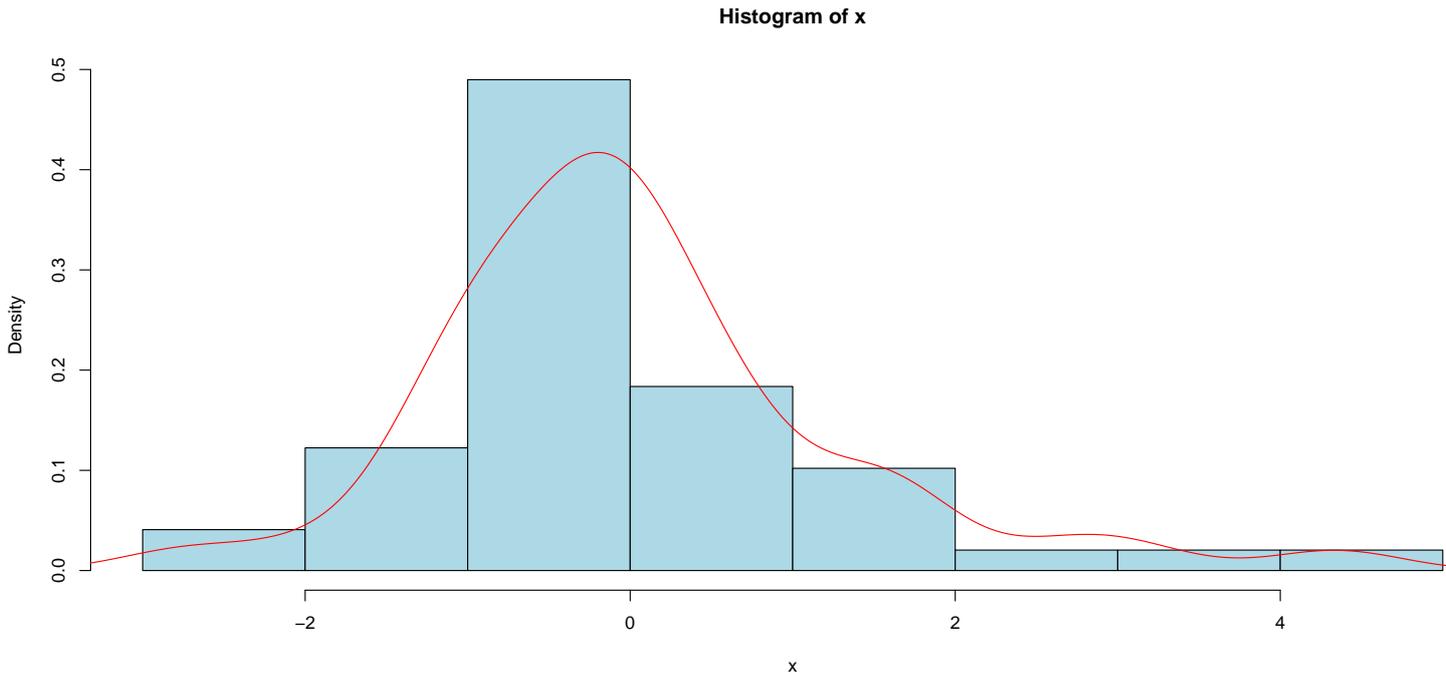


FIGURE 4.1 – Histogramme et distribution de l'échantillon R_1 .

La représentation graphique de l'histogramme et la distribution de l'échantillon (leptokurtique) nous induit à proposer la loi de Laplace Symétrique et pour confirmer que nous pouvons pas ajuster la distribution de cet échantillon R_1 par la loi normal, Nous utiliserons le test de Shapiro-Wilk. La commande utilisée pour tester la normalité sous R est : `shapiro.test`¹. L'ajustement des rendements financiers par la loi normale donne $W = 0.9335$, $p - value = 0.008324$.

Pour un risque = 0.05, le seuil critique lu dans la table pour $n = 49$ est $W_{crit} = 0.947$. Puisque $W < W_{crit}$, alors l'hypothèse de normalité est incompatible avec nos données, c'est à dire l'hypothèse de normalité de l'échantillon R_1 est rejetée. Nous pouvons également vérifier cette affirmation à l'aide du graphique de comparaison des quantiles empiriques avec ceux de la loi normale présenté à la figure (4.2).

1. Voir plus en Annexe C

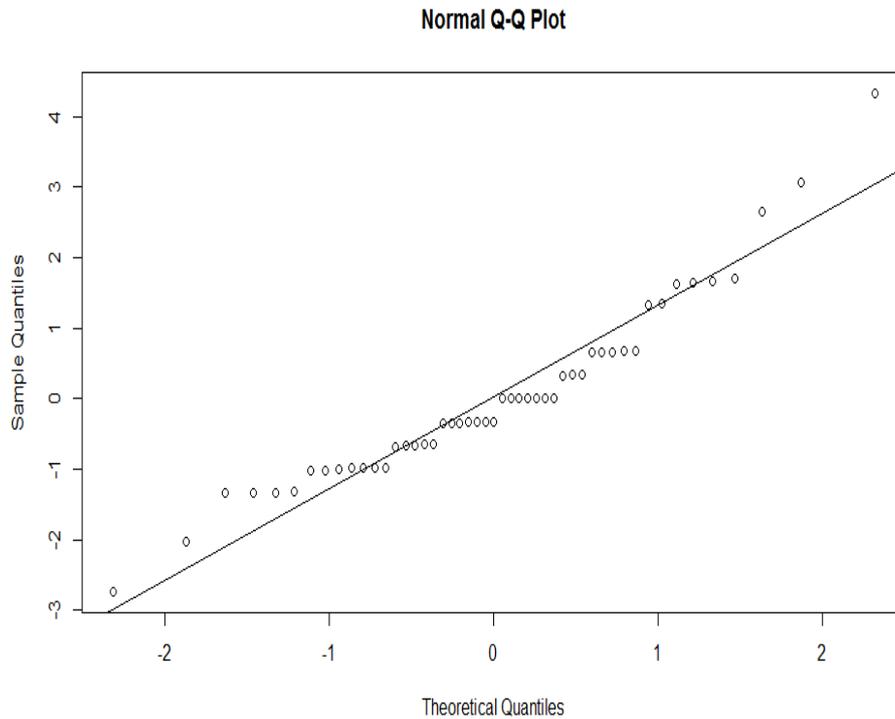


FIGURE 4.2 – Graphique Quantile-Quantile

4.3 Estimation des paramètres du modèle de Laplace asymétrique généralisé

Dans cette section, nous exposons les résultats obtenus dans le travail de Pelletier [30], pour l'échantillon de données centrées et réduites $R_1^* = \frac{R(t) - E[R_1]}{\sqrt{\text{var}[R_1]}}$, à l'aide des méthodes :

- des moments généralisée itérative .
- d'estimation gaussienne de Whittle [15] .
- de l'équation d'estimation optimale de Crowder .
- de l'équation d'estimation optimale modifiée .

Méthode	α	σ	μ	τ
Paramètres initiaux	-0.612081	0.515932	0.325854	1.878388
Moments généralisée	-0.640067	0.625431	0.324311	1.973623
Estimation gaussienne	-0.775071	0.581413	0.384342	2.016619
Équation d'estimation optimale	-0.658697	0.646516	0.376251	1.750685
Équation d'estimation optimale modifiée	-0.712450	0.606193	0.363196	1.961614

TABLE 4.3 – Valeurs des paramètres estimée

Nous exposons aussi les intervalles de confiance pour les paramètres estimés pour chaque méthode :

– Pour la méthode des moments généralisée :

	Borne inférieure	Valeur estimée	Borne supérieure
α	-0.715736	-0.640067	-0.564397
σ	0.441236	0.625431	0.809627
μ	0.245602	0.324311	0.403020
τ	1.913780	1.973623	2.033466

– Pour la méthode d'estimation gaussienne :

	Borne inférieure	Valeur estimée	Borne supérieure
α	-0.8315	-0.7751	-0.7186
σ	0.4152	0.5814	0.7476
μ	0.2866	0.3843	0.4820
τ	1.9896	2.0166	2.0437

– Pour la méthode de l'équation d'estimation optimale :

	Borne inférieure	Valeur estimée	Borne supérieure
α	-0.7230	-0.6587	-0.5943
σ	0.4693	0.6465	0.8237
μ	0.2707	0.3763	0.4818
τ	1.7134	1.7507	1.7880

– Pour la méthode de l'équation d'estimation optimale modifiée :

	Borne inférieure	Valeur estimée	Borne supérieure
α	-0.7663	-0.7124	-0.6586
σ	0.4551	0.6062	0.7573
μ	0.2650	0.3632	0.4614
τ	1.9341	1.9616	1.9892

4.4 Estimation des paramètres du modèle de Laplace symétrique

Dans cette section, nous tentons d'appliquer deux approches d'estimation des paramètres du modèle de Laplace symétrique.

4.4.1 Estimateur du maximum de vraisemblance

Dans cette sous-section, nous montrons la faiblesse de l'algorithme du maximum de vraisemblance .

La fonction de vraisemblance notée par $L(x, \theta)$ est donnée par :

$$L(x, \theta) = \left(\frac{1}{\sqrt{2}\sigma} \right)^n \exp \left(\frac{-\sqrt{2}}{\sigma} \sum_{i=1}^n |x_i - \alpha| \right) \quad (4.1)$$

d'où la log-vraisemblance :

$$\log L(x, \theta) = -n \times \log(\sqrt{2}\sigma) - \frac{\sqrt{2}}{\sigma} \sum_{i=1}^n |x_i - \alpha| \quad (4.2)$$

On calcule les dérivés par rapport à σ et α :

$$\frac{\partial \log L(x, \theta)}{\partial \sigma} = \frac{-n\sigma}{\sigma^2} + \sqrt{2} \sum_{i=1}^n |x_i - \theta| \times \frac{1}{\sigma^2} = 0 \Rightarrow -n\sigma + \sqrt{2} \sum_{i=1}^n |x_i - \theta| = 0 \quad (4.3)$$

$$\frac{\partial \log L(x, \theta)}{\partial \alpha} = \left| -n \frac{\sqrt{2}}{\sigma} \right| = 0 \quad (4.4)$$

Alors on aura les estimateurs suivants :

$$\hat{\sigma} = \frac{\sqrt{2}}{n} \sum_{i=1}^n |x_i - \alpha| \quad (4.5)$$

$$\hat{\alpha} \in \mathbb{R} \quad (4.6)$$

De (4.6) nous pouvons déduire que nous ne pouvons pas obtenir des estimateurs du maximum de la vraisemblance pour raison de la divergence de cette méthode.

Cette principale raison nous motive à appliquer une autre méthode d'estimation. Nous choisissons l'approche bayésienne.

4.4.2 Approche bayésienne

4.4.2.1 Choix des lois a priori

Les lois a priori sur les deux paramètres inconnus α et σ sont données par :

$$\alpha \rightsquigarrow \mathcal{N}(\mu_0, \sigma_0^2) \text{ et } \sigma \rightsquigarrow \Gamma(c, d)$$

avec $\mathcal{N}(\mu_0, \sigma_0^2)$ est la distribution normale de moyenne $\mu_0 \in \mathbb{R}$ et de variance $\sigma_0^2 \in \mathbb{R}_+$ et de fonction de densité donnée par :

$$f(x, \mu_0, \sigma_0) = \frac{1}{\sqrt{2\pi\sigma_0}} e^{-\frac{(x - \mu_0)^2}{2\sigma_0^2}}$$

et $\Gamma(c, d)$ est la distribution gamma de moyenne $\frac{c}{d}$, de variance $\frac{c}{d^2}$ et de fonction de densité donnée par :

$$f(x, c, d) = \frac{d^c x^{c-1} \exp(-xd)}{\Gamma(c)}$$

tel que $x > 0$, $c > 0$, $d > 0$.

4.4.2.2 Calcul des lois a posteriori

En utilisant le théorème de Bayes, la distribution a posteriori pour les paramètres α et σ est donnée par :

$$\pi(\alpha, \sigma/x) \propto \left(\frac{1}{\sqrt{2\sigma}}\right)^n \exp\left(\frac{-\sqrt{2}}{\sigma} \sum_{i=1}^n |x_i - \alpha|\right) \times \pi(\alpha, \sigma) \quad (4.7)$$

Vu la complexité de la loi a posteriori (4.7), il est impossible d'obtenir directement les estimateurs bayésiens. Pour cette raison, la stratégie de calcul consiste d'abord à dériver les lois a posteriori conditionnelles des deux paramètres et d'utiliser les méthodes de Monté Carlo par Chaîne de Markov (*MCMC*) pour estimer les paramètres de Laplace Symétrique. A partir de (4.7), nous pouvons obtenir les lois a posteriori conditionnelles des deux paramètres comme suit :

$$\pi(\alpha/x, \sigma) \propto \exp\left(\frac{-\sqrt{2}}{\sigma} \sum_{i=1}^n |x_i - \alpha|\right) \times \pi(\alpha) \quad (4.8)$$

$$\pi(\sigma/x, \alpha) \propto \left(\frac{1}{\sqrt{2\sigma}}\right)^n \exp\left(\frac{-\sqrt{2}}{\sigma} \sum_{i=1}^n |x_i - \alpha|\right) \times \pi(\sigma) \quad (4.9)$$

4.4.2.3 Application de l'algorithme de Metropolis-Hastings au modèle

Pour simuler suivant les lois a posteriori, nous utilisons l'algorithme de Metropolis-Hasting (M-H). Les étapes de l'algorithme qui permettent de simuler suivant les lois a posteriori sont résumées comme suit :

(a) Initialiser le vecteur des paramètres $\theta^{(1)} = (\alpha^1, \sigma^1)^T$.

(b) Pour $m = \{2, \dots, N\}$.

– Génération des lois candidates

• Générer $\alpha^c \rightsquigarrow \mathcal{N}(\alpha^m, \sigma^2 = 0,002)$; $\alpha^c = |\alpha^c|$.

• Générer $\sigma^c \rightsquigarrow \mathcal{N}(\sigma^m, \sigma^2 = 0,002)$; $\sigma^c = |\sigma^c|$.

– Calculer les probabilités d'acceptation suivantes :

$$\rho_1 = \min \left\{ 1; \frac{\pi(\alpha^c | \sigma^m, t)}{\pi(\alpha^m | \sigma^m, t)} \right\}, \alpha^m = \begin{cases} \alpha^c & \text{si } u < \rho_1, \quad u \rightsquigarrow U_{[0,1]} \\ \alpha^{m-1} & \text{sinon} \end{cases}$$

$$\rho_2 = \min \left\{ 1; \frac{\pi(\sigma^c | \alpha^{m+1}, t)}{\pi(\sigma^m | \alpha^{m+1}, t)} \right\}, \sigma^m = \begin{cases} \sigma^c & \text{si } u < \rho_2, \quad u \rightsquigarrow U_{[0,1]} \\ \sigma^{m-1} & \text{sinon} \end{cases}$$

(c) $m = m + 1$ et aller à 2 .

(d) Calculer le vecteur d'estimation de Bayes $\hat{\theta}_{Bayes} = (\hat{\alpha}, \hat{\sigma})^T$

$$\hat{\alpha}_{Bayes} = \frac{1}{N - N_0} \sum_{m=N_0+1}^N \alpha^m$$

$$\hat{\sigma}_{Bayes} = \frac{1}{N - N_0} \sum_{m=N_0+1}^N \sigma^m.$$

En utilisant nos programmes implémentés sous logiciel R^2 , nous avons estimés les paramètres du modèle en utilisant l'échantillon des rendements R_1 .

Nous initialisons les paramètres de ce modèle et nous étudions trois cas distincts. Ces

2. Voir Annexe C

paramètres ont été choisis comme suit :

- Premier cas : $\alpha = -0.6120$, $\sigma = 0.5159$.
- Deuxième cas : $\alpha = -0.1753$, $\sigma = 0.9159$.
- troisième cas : $\alpha = -1.5120$, $\sigma = 0.4159$.

Nous calculons les estimateurs de Bayes (moyenne a posteriori) en utilisant le programme *MCMC* avec un nombre d'itérations égale : 1000, 10000, 100000, et une période de chauffage égale : 500, 5000, 50000 et les hyper-paramètres de lois a priori, $\mu_0 = -0.4$, $\sigma_0^2 = 0.01$ pour le paramètre α et $c = d = 0.02$ pour le paramètre σ . Nous avons utilisés $n = 40$ réplifications.

4.5 Résultats obtenus

Les résultats sont résumés dans le tableau (4.3) en présentant le nombre d'itérations, la moyenne pour chaque paramètre et les erreurs quadratiques moyennes (*EQM*). La moyenne est donnée comme suit :

$$\bar{\hat{\alpha}} = \frac{1}{n} \sum_{i=1}^n \hat{\alpha}_i, \quad \bar{\hat{\sigma}} = \frac{1}{n} \sum_{i=1}^n \hat{\sigma}_i$$

Les erreurs quadratiques moyennes sont données par :

$$EQM(\alpha) = \frac{1}{n} \sum_{i=1}^n (\hat{\alpha}_i - \bar{\hat{\alpha}})^2$$

$$EQM(\sigma) = \frac{1}{n} \sum_{i=1}^n (\hat{\sigma}_i - \bar{\hat{\sigma}})^2$$

Nombre d'itérations		1 ^{er} cas	2 ^{eme} cas	3 ^{eme} cas
$N = 1000$	$\hat{\alpha}$	-0.639985	-0.207639	-1.535924
	EQM (α)	0.008745	0.006765	0.006023
	$\hat{\sigma}$	0.592756	0.886332	0.50753
	EQM (σ)	0.0073965	0.00079901	0.008396
$N = 10000$	$\hat{\alpha}$	-0.630638	-0.191532	-1.529975
	EQM (α)	0.0030254	0.001068	0.002617
	$\hat{\sigma}$	0.637866	0.825893	0.877540
	EQM (σ)	0.0098043	0.008101	0.051973
$N = 100000$	$\hat{\alpha}$	-0.620086	-0.219056	-1.541966
	EQM (α)	0.000887	0.009946	0.008033
	$\hat{\sigma}$	0.551155	0.772444	0.989765
	EQM (σ)	0.0002461381	0.0026302	0.015341

TABLE 4.4 – Estimation des paramètres du modèle par l'approche bayésienne

Nous pouvons ainsi construire un intervalle de confiance pour les paramètres estimés dans le premier cas après 100000 itérations. Nous utilisons un seuil de tolérance $\alpha = 5\%$:

	Borne inférieure	Valeur estimée	Borne supérieure
α	-0.695755	-0.620086	-0.544417
σ	0.384755	0.551155	0.717555

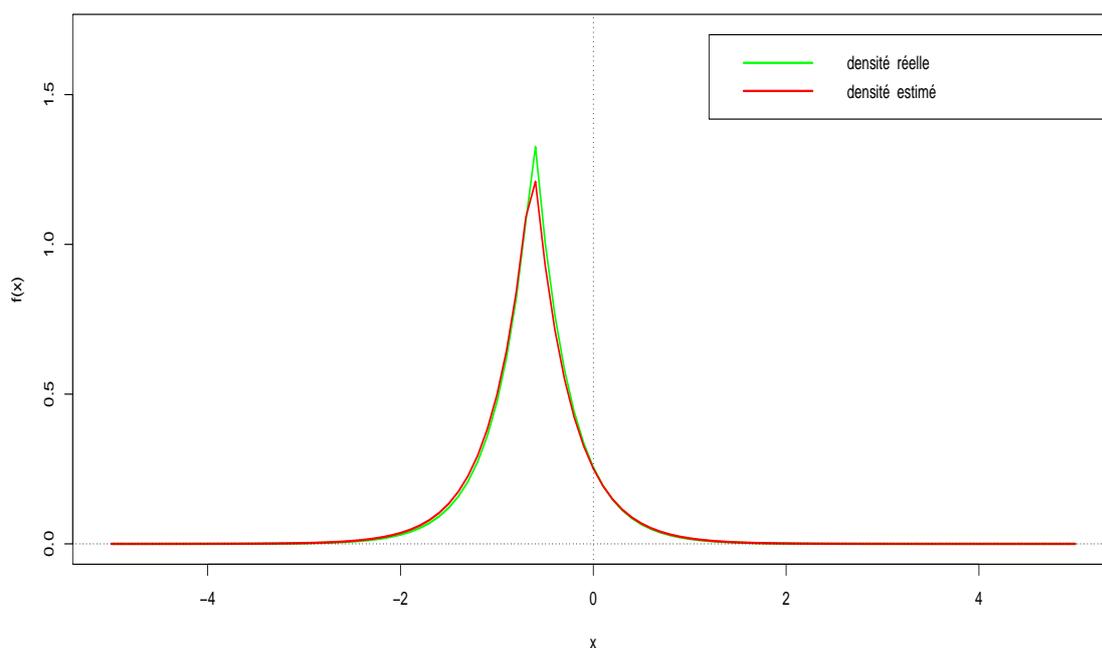


FIGURE 4.3 – Comparaison entre la densité théorique et celle estimée par bayes dans le 1^{er} cas.

4.6 Discussion et interprétation des résultats :

Il existe plusieurs critères pour comparer des estimateurs entre eux, le plus utilisé est la comparaison à partir de leurs variances, cependant le calcul des variances théoriques est souvent difficile, si non impossible. Donc, la comparaison se fait par rapport à l'erreur quadratique moyenne (*EQM*), commise lorsque nous remplaçons la valeur du paramètre par son estimateur. La table (4.4) présente les résultats de la comparaison entre les valeurs des deux estimateurs α , σ obtenus dans les trois cas proposés. Pour chaque estimateur nous lui associons une *EQM*. Nous remarquons que, dans le premier cas ($\alpha = -0.6120$, $\sigma = 0.5159$) plus nous augmentons le nombre d'itérations N , l'*EQM* tend vers zéro.

Par contre, cette propriété ne se présente pas dans le deuxième et le troisième cas. De plus, nous remarquons aussi, pour les mêmes valeurs initiales choisit par Pelletier dans ses estimations, les résultats se rapproche avec une légère différence d'ordre négligeable. Donc, il sera très intéressant d'adapter l'approche bayésienne sur la distribution de Laplace asymétrique généralisée. Sans doute, les résultats seront meilleurs que ceux obtenus par les approches fréquentistes .

Conclusion

Le choix naïf des lois a priori nous a conduit à obtenir des estimateurs de la distribution de Laplace symétrique presque identique que ceux de la distribution de Laplace asymétrique généralisée.

Conclusion Générale

La première modélisation des cours boursiers par un mouvement brownien repose sur l'hypothèse que les rendements (logarithmiques) suivent une loi normale. Cette modélisation reste encore aujourd'hui au cœur de la théorie financière. En pratique, on observe des distributions de rendements logarithmiques ayant des queues plus épaisses que des distributions normales. Ce fait a été remarqué pour la première fois par le mathématicien Mandelbrot dans [25] et a été depuis lors observé pour différents marchés.

Dans ce travail, nous avons présenté quelques modèles classiques de la finance. Ces derniers ont, dès le début, été remis en cause par Mandelbrot qui démontre que la loi normale ne correspond pas à la réalité des marchés parce qu'elle sous estime les risques extrêmes.

L'intérêt de notre travail est double. Le premier objectif était de présenter l'utilisation de la distribution de Laplace asymétrique généralisée comme alternative intéressante des modèles classiques en finance. Secundo, nous avons tenté de rapprocher celle-ci de la réalité en considérant des informations a priori sur les paramètres de la distribution. Ce qui nous a conduit à appliquer le paradigme bayésien sur un cas particulier via l'expression sophistiquée de Laplace asymétrique généralisée. Cette approche demande une grande puissance de calcul : elle n'a donc été rendue possible que grâce au développement relativement récent du calcul numérique (méthodes *MCMC* : Monte Carlo par chaînes de Markov) et de l'outil informatique performant.

Les estimateurs de Bayes sont obtenus en utilisant les méthodes *MCMC* avec des lois a priori conjuguées naturelles. Le choix naïf de ces lois a priori nous a conduit à obtenir des estimateurs de la distribution de Laplace symétrique presque identique que ceux de la distribution de Laplace asymétrique généralisée. Cependant, comme tout modèle paramétrique, il reste dépendant du nombre et de la qualité des données disponibles.

Il sera intéressant d'adapter l'approche bayésienne sur la distribution de Laplace asymétrique généralisée. Sans doute, les résultats seront meilleurs que ceux obtenus par les approches fréquentistes.

Notions fondamentales utilisées

A.1 Produit de convolution

La fonction caractéristique permet d'identifier la distribution d'une somme de variables aléatoires indépendantes $Z = X_1 + \dots + X_n$, appelée produit de convolution et noté :

$$\phi_Z(z) = \phi_{X_1 + \dots + X_n}(z) = \prod_{i=1}^n \phi_{X_i}(z). \quad (\text{A.1})$$

Lorsque les variables aléatoires sommées sont aussi identiquement distribuées, la fonction caractéristique de Z est la $n^{\text{ième}}$ puissance de celle de X :

$$\phi_Z(z) = \phi_{X_1 + \dots + X_n}(z) = [\phi_X(z)]^n. \quad (\text{A.2})$$

Cette fonction est donc une solution de rechange intéressante à utiliser lorsque aucune forme analytique pour la densité ou la fonction de répartition pour une distribution donnée n'existe pas.

A.2 Processus de Lévy

A.2.1 Définition et propriétés

En théorie des probabilités, un processus de Lévy, nommé d'après le mathématicien français Paul Lévy, est un processus stochastique à temps continu, continu à droite limité à gauche, partant de 0, dont les accroissements sont stationnaires et indépendants (cette notion est expliquée ci-dessous). Les exemples les plus connus sont le processus de Wiener et le processus de Poisson.

A.2.2 Représentation de Lévy-Khintchine

Toute variable aléatoire peut être caractérisée par sa fonction caractéristique. Dans le cas d'un processus de Lévy X_t , cette caractérisation pour tout temps t donne la représentation de Lévy-Khintchine (du nom du mathématicien russe Alexandre Khintchine) :

$$\mathbb{E}\left[e^{i\theta X_t}\right] = \exp\left(a i t \theta - \frac{1}{2} \sigma^2 t \theta^2 + t \int_{\mathbb{R} \setminus \{0\}} \left(e^{i\theta x} - 1 - i\theta x 1_{\{|x| < 1\}}\right) W(dx)\right) \quad (\text{A.3})$$

où $a \in \mathbb{R}$, $\sigma \geq 0$ et \mathbf{I} est la fonction indicatrice. La mesure de Lévy W doit vérifier

$$\int_{\mathbb{R} \setminus \{0\}} \min\{x^2, 1\} W(dx) < \infty.$$

Un processus de Lévy est donc caractérisé par trois composantes : une dérive (un drift), un coefficient de diffusion, et une composante de saut. Ces trois composantes, et donc la représentation de Lévy-Khintchine du processus, sont entièrement déterminées par le triplet (a, σ^2, W) . En particulier, un processus de Lévy continu est un mouvement brownien avec dérive.

A.3 Processus subordonné [30]

On considère les processus de Lévy $\{X(t)\}$ et $\{Z(t)\}$. Celui qui suit est défini comme étant un processus subordonné et aussi un processus de Lévy, comme le démontrent Sato [38] et Schoutens [39] :

$$\{Y(t)\} = \{X(Z(t))\} \quad (\text{A.4})$$

Si $\Lambda(\theta)$ est l'exposant caractéristique du processus $\{X(t)\}$ et $\Xi(\theta)$, celui du processus subordonné $\{Z(t)\}$, alors celui du processus $\{Y(t)\}$ prend la forme suivante :

$$\begin{aligned} \Psi(\theta) &= \Xi(\theta) \circ i\Lambda(\theta) \\ &= \Xi(i\Lambda(\theta)). \end{aligned} \quad (\text{A.5})$$

La fonction caractéristique du processus $\{Y(t)\}$ est alors :

$$\phi_{Y(t)}(\xi) = e^{-t\Xi(i\Lambda(\xi))} \quad (\text{A.6})$$

La densité de la variable aléatoire $Y(t)$ s'obtient à l'aide de la formule de l'espérance conditionnelle :

$$\begin{aligned} f_{Y_t}(y) &= E\left[f_{X(t)}(y|Z(t) = z)\right] \\ &= \int_{-\infty}^{+\infty} f_{X(t)}(y|Z(t) = z) \cdot f_{z(t)}(z) dz \end{aligned} \quad (\text{A.7})$$

A.3.1 Le processus gamma [30]

Le processus gamma, noté $\{G(t; \tau, \beta)\}$, est un processus de sauts purs (donc aucune composante de dérive ni de diffusion) dont les incréments $G(t+1; \tau, \beta) - G(t; \tau, \beta)$ suivent une distribution gamma de paramètres de forme τ et d'échelle β , définie par les fonctions de densité $f_{\tau, \beta}(x)$ et caractéristique $\phi_{\tau, \beta}(\xi)$:

$$\begin{aligned} f_{\tau, \beta}(x) &= \frac{\beta^\tau}{\Gamma(\tau)} x^{\tau-1} e^{-\beta x} \mathbf{1}_{x \geq 0} & (A.8) \\ \phi_{\tau, \beta}(\xi) &= E[e^{i\xi X}] \\ &= \int_0^\infty e^{i\xi x} f_{\tau, \beta}(x) dx \\ &= \frac{1}{\left(1 - \frac{i\xi}{\beta}\right)^\tau} \end{aligned}$$

On s'intéresse à la situation où le paramètre d'échelle est de valeur unitaire ($\beta = 1$). Le processus gamma agit alors à titre de compteur et sa valeur $G(t; \tau, \beta = 1)$ au temps t correspondra au nombre de sauts depuis $t = 0$.

La fonction de densité $G(t+1; \tau; \beta = 1) - G(t; \tau, \beta = 1)$ sera alors :

$$f_{\tau, \beta=1}(x) = \frac{1}{\Gamma(\tau)} x^{\tau-1} \exp^{-x} \mathbf{1}_{\{x \geq 0\}} \quad (A.9)$$

Le paramètre τ , qui définit la forme de la distribution, déterminera la fréquence moyenne sauts du processus gamma $\Gamma(t; \tau, \beta = 1)$, étant donné l'espérance $E[G(t)] = \tau.t$

La fonction caractéristique de ce processus sera donc $\phi(\xi, t; \tau, \beta = 1)$ en utilisant la propriété de convolution (A.1) (même si le temps t n'est pas entier, car la distribution est infiniment divisible) :

$$\begin{aligned} \phi(\xi, t; \tau, \beta = 1) &= \left[\frac{1}{\left(1 - \frac{i\xi}{\tau}\right)} \right]^t \\ &= \frac{1}{(1 - i\xi)^{\tau.t}} \end{aligned} \quad (A.10)$$

On peut réécrire la fonction caractéristique d'un incrément de ce processus $\phi(\xi, t; \tau, \beta = 1)$ sous la représentation de Lévy-Khintchine (A.2.2), avec l'exposant caractéristique Ξ :

$$\begin{aligned} \Xi(\zeta; t = 1; \tau, \beta = 1) &= \tau \ln(1 - i\zeta) & (A.11) \\ &= \tau(e^0 - e^{-\infty}) \ln(1 - i\zeta) \\ &= \tau \int_0^{+\infty} \frac{e^{-x} - e^{-(1-i\zeta)x}}{x} dx \end{aligned}$$

(intégrale de Frullani (Spiegel et Liu, [42]))

$$= \tau \int_0^{+\infty} (1 - i\zeta x)^{\frac{1-x}{x}} dx. \quad (\text{A.12})$$

On a donc, par cette représentation, la démonstration que le processus gamma est un processus de sauts purs. Il pourra donc être utilisé comme subordonnant dans la construction d'un processus subordonné (A.3).

A.3.2 Le processus de Wiener [30]

Le processus de Wiener $\{B(t; \mu, \sigma^2)\}$ est un processus de diffusion avec dérive. Il n'a donc pas de composante de saut. Ses incréments suivent une distribution normale :

$$B(t+1; \mu, \sigma^2) - W(t; \mu, \sigma^2) \rightsquigarrow N(\mu, \sigma^2). \quad (\text{A.13})$$

Cette distribution est définie par la fonction de densité $f_{\mu, \sigma}(x)$ et la fonction caractéristique $\phi_{\mu, \sigma}(\xi)$:

$$f_{\mu, \sigma}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp - \left\{ \frac{1}{2} \left(\frac{x - \mu}{\sigma} \right)^2 \right\}$$

$$\phi_{\mu, \sigma}(\xi) = \exp \left\{ i\mu\xi - \frac{\sigma^2\xi^2}{2} \right\}.$$

Notons que la variance d'un incrément est proportionnelle à la longueur de celui-ci. Soit deux incréments indépendants d'un même processus :

$$I_1 = W(t+q; \mu, \sigma^2) - W(t; \mu, \sigma^2) \sim N(q\mu, q\sigma^2)$$

$$I_2 = W(t+q+s; \mu, \sigma^2) - W(t+q; \mu, \sigma^2) \sim N(s\mu, s\sigma^2).$$

La somme de ces incréments suit une distribution normale dont la moyenne et la variance sont respectivement la somme de celles des deux incréments :

$$I_1 + I_2 \rightsquigarrow N((q+s)\mu, (q+s)\sigma^2).$$

Comme la distribution normale est aussi infiniment divisible, on peut obtenir la fonction caractéristique du processus $\phi(\xi; t; \mu, \sigma^2)$ en utilisant la propriété de convolution (A.1) :

$$\phi(\xi; t; \mu, \sigma^2) = \exp \left\{ i\mu t \xi - \frac{\sigma^2 t \xi^2}{2} \right\}$$

On déduit donc facilement l'exposant caractéristique $\Lambda(\xi; t=1; \mu, \sigma^2)$ d'un incrément de ce processus, sous la représentation de Lévy-Khintchine :

$$\Lambda(\xi; t=1; \mu, \sigma^2) = -\left(i\mu\xi - \frac{\sigma^2\xi^2}{2} \right)$$

Ceci démontre que le processus de Wiener est un processus avec dérive et diffusion, mais sans composante de saut. Il pourra donc être utilisé pour construire un processus subordonné (A.3).

A.4 Lemme d'Itô

Supposons que la valeur d'une variable X suive un processus d'Itô de paramètres $\mu(X, t)$ et $\sigma(X, t)$

$$dx = \mu(X, t)dt + \sigma(X, t)dz$$

Le lemme d'Itô montre qu'une fonction G de X et t est caractérisée par le processus suivant

$$dG = \left(\frac{\partial G}{\partial X} \mu(X, t) + \frac{\partial G}{\partial t} + \frac{1}{2} \frac{\partial^2 G}{\partial X^2} \sigma^2(X, t) \right) dt + \frac{\partial G}{\partial X} \sigma(X, t) dz$$

ou z est le même processus de Wiener standard que dans dx .

Ainsi, G suit également un processus d'Itô de paramètres

$$\frac{\partial G}{\partial X} \mu(X, t) + \frac{\partial G}{\partial t} + \frac{1}{2} \frac{\partial^2 G}{\partial X^2} \sigma^2(X, t) \text{ et } \frac{\partial G}{\partial X} \sigma(X, t)$$

A.5 Chaîne de Markov à espace d'état discret E

Une chaîne de Markov est une suite X_n de variables aléatoires, telles que la loi conditionnelle de X_n sachant X_{n-1}, X_{n-2}, \dots est la même que celle de X_n sachant X_{n-1} , c'est à dire :

$$Pr[X_n = j \mid X_{n-1} = i_{n-1}, X_{n-2} = i_{n-2}, \dots, X_0 = i_0] = Pr[X_n = j \mid X_{n-1} = i_{n-1}]$$

La chaîne de Markov est homogène si de plus on a :

$$P[X_n = j \mid X_{n-1} = i] = P[X_1 = j \mid X_0 = i]$$

Matrice de transition [35]

Soit :

$$P = \begin{pmatrix} p_{00} & p_{01} & p_{02} & \cdots \\ p_{10} & p_{11} & p_{12} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

P est appelée matrice de passage (ou de transition) de la chaîne et les p_{ij} sont les probabilités de transitions.

Toute matrice de transition $P = (p_{ij})_{i,j \in E}$ vérifie les propriétés suivantes :

- pour tout couple (i, j) , on a $p_{ij} \geq 0$;
- pour tout $i \in E$, on a $\sum_{j \in E} p_{ij} = 1$.

La matrice P est appelée matrice stochastique.

A.5.1 Noyau d'une chaîne de Markov

La première notion utile est celle de noyau car une chaîne de Markov homogène est entièrement définie par :

- Son noyau ;
- La loi initiale de X_0 (ou sa valeur initiale).

Le noyau est une fonction à deux variables qui caractérise les probabilités de transition de X_n à X_{n+1} .

La forme dite intégrale du noyau est :

$$K_n(X, A) = [X_{n+1} \in A \mid X_n = X]$$

A étant un sous-ensemble mesurable du domaine X de X_{n+1} . Pour une chaîne homogène, on notera simplement $K(X, A)$ au lieu de $K_n(X, A)$.

Ainsi, pour tout couple (X, A) , $K(X, \cdot)$ est une mesure de probabilité sur l'espace X indexée par la valeur X et $K(\cdot, A)$ est une fonction mesurable comprise entre 0 et 1, indexée par les sous-ensembles mesurables de X . Dans la plupart des cas, la mesure de probabilité $K(X, \cdot)$ aura une densité $k(X, X')$ par rapport à une mesure support adéquate (indépendante de n pour une chaîne homogène) et telle que :

$$k(X, X') = [X_{n+1} = X' \mid X_n = X]$$

Probabilités invariantes

Une probabilité π sur E est appelée probabilité invariante, ou probabilité stationnaire [1], d'une chaîne de Markov de matrice de transition P si $\pi = P\pi$.

Réversibilité [14]

On dit qu'une chaîne de Markov de matrice de transition P , est réversible par rapport à la probabilité π si on a pour tous $i, j \in E$,

$$\pi(i)P(i, j) = \pi(j)P(j, i)$$

Si une chaîne de Markov est réversible par rapport à la probabilité π , alors π est une probabilité invariante.

Irréductibilité [14]

On dit qu'une chaîne de Markov, ou sa matrice de transition, est irréductible si pour tous $i, j \in E$, la probabilité partant de i d'atteindre j est strictement positive, autrement dit : si pour tous $i, j \in E$, il existe $n = n_{i,j} \geq 1$ telle que $P^n(i; j) > 0$.

Réurrence

Soit $(X_n)_{n \geq 0}$ une chaîne de Markov homogène de matrice de transition P . On définit le temps de retour en i par [16] :

$$T_i = \{\inf n \geq 1; X_n = i\}.$$

- On dit que i est récurrent si $Pr_i(T_i < \infty) = 1$.
- Un état i récurrent est dit récurrent positif si de plus $E_i[T_i] < \infty$
- Un état i récurrent est dit récurrent nul si $E_i[T_i] = \infty$.
- Un état qui n'est pas récurrent est dit transitoire.

Périodicité [14]

Soit X une chaîne de Markov de matrice de transition P . La période d'un état $i \in E$ est le PGCD de $n \in N^*$; $P^n(i, i) > 0$. Un état est dit apériodique si sa période est 1, sinon il est dit périodique. Une chaîne de Markov est dite apériodique si tous ses états sont apériodiques.

Ergodicité [5]

Une chaîne de Markov homogène, irréductible, récurrente positive et apériodique est dite ergodique.

Données utilisées

Ce tableau présente l'échantillon de données utilisées au chapitre 4.

Date	Prix								
31/07	296	14/08	294	28/08	303	11/09	296	25/09	306
01/08	296	15/08	293	29/08	304	12/09	301	26/09	303
02/08	300	16/08	295	30/08	304	13/09	298	27/09	301
05/08	302	19/08	287	02/09	309	16/09	295	30/09	303
06/08	300	20/08	288	03/09	309	17/09	295	01/10	308
07/08	304	21/08	297	04/09	309	18/09	293	02/10	305
08/08	303	22/08	305	05/09	307	19/09	292	03/10	302
09/08	299	23/08	307	06/09	306	20/09	297	04/10	301
12/08	293	26/08	307	09/09	304	23/09	294	07/10	297
13/08	294	27/08	304	10/09	300	24/09	293	08/10	299

TABLE B.1 – Prix du titre Abbey National (penny sterling) du 31 juillet au 8 octobre 1991
[7][30]

Annexe **C**

Code source sous R

Les programmes informatiques ont été implémentés sous R. Dans un premier temps, nous présentons les programmes mis en place pour le calcul des fonctions de répartition et de densité. Dans un second temps, nous donnons les programmes R qui retournent les estimateurs de Bayes pour les paramètres α et σ .

Fonction de densité de Laplace symétrique

```
f = function(x,alpha,sigma)
{
den=(1/(sqrt(2)*sigma))*exp((-sqrt(2))/sigma*abs(x-alpha))
}
```

=====

Graphique Quantile Quantile

```
x=c(0,1.3423,0.6645,-0.6645,1.3245,-0.3295,-1.3289,-2.0271,0.3407,0,-0.3407,0.6803,-
2.7449,0.3478,3.0772,2.6580,0.6536,0,-0.9820,-0.3295,0.3295,0,1.6314,0,0,-0.6494,-0.3263,-
0.6557,-1.3245,-1.3423,1.6751,-1.0017,-1.0118,0,-0.6803,-0.3419,1.6978,-1.0152,-0.3407,4.3412,-
0.9852,-0.6623,0.6623,1.6367,-0.9788,-0.9885,-0.3317,-1.3378,0.6711)

qqnorm(x)

abline(mean(x),sd(x))
```

=====

Test de shapiro-Wilk

```
x=c(0,1.3423,0.6645,-0.6645,1.3245,-0.3295,-1.3289,-2.0271,0.3407,0,-0.3407,0.6803,-
2.7449,0.3478,3.0772,2.6580,0.6536,0,-0.9820,-0.3295,0.3295,0,1.6314,0,0,-0.6494,-0.3263,-
0.6557,-1.3245,-1.3423,1.6751,-1.0017,-1.0118,0,-0.6803,-0.3419,1.6978,-1.0152,-0.3407,4.3412,-
0.9852,-0.6623,0.6623,1.6367,-0.9788,-0.9885,-0.3317,-1.3378,0.6711)

shapiro.test(x)
```

#####Estimation des paramètres de modèle par la méthode MCMC##### »

```
logposterior <- function(x,alpha,sigma){
n=length(x)
postalpha =(-sqrt(2)/sigma)*sum(abs(x-alpha))+dnorm(alpha,-0.4,0.01,log=TRUE)
postalpha}
logposteriorsigma <- function(x,alpha,sigma){
n=length(x)
postsigma = -n*log( sqrt(2)*sigma)- (sqrt(2)/sigma)*sum(abs(x-alpha))
+dgamma(sigma,0.02,0.02,log=TRUE)
postsigma }
MCMC_Gibbs <- function(N,alpha0,sigma0,x){
alpha=sigma=0
z=g=0
r1=r2=0
m=1
alpha[1]<-alpha0   ### Initialisation des paramètres
sigma[1]<-sigma0
while (m<=N){
alphac=rnorm(1,alpha[m],0.001)   ### loi de proposition (candidate)
sigmac=rnorm(1,sigma[m],0.001)
alphac=abs(alphac)
sigmac=abs(sigmac)
```

```

r1[m]=min(0,logposterior(x,alphac,sigma[m])-
logposterior(x,alpha[m],sigma[m]))
u=log(runif(1))
if (u<r1[m]){
alpha[m+1]=alphac }
else{
alpha[m+1]=alpha[m] }
r2[m]=min(0,logposterior(x,alpha[m+1],sigmac)-
logposterior(x,alpha[m+1],sigma[m]))
if (u<r2[m]){
sigma[m+1]=sigmac }
else{
sigma[m+1]=sigma[m]}
m=m+1 print(sigma)
for(m in N0+1:N){
g=g+sigma[m]
z=z+alpha[m]
}
sigma_Bayes=(1/(N-N0)) * g
alpha_bayes= (1/(N-N0)) * z
    print('sigma-chapeau-Bayes est:')
print(sigma_Bayes)
    print('alpha-chapeau-Bayes est:')
print(alpha_bayes)
}
}

```

Bibliographie

- [1] M. Abramowitz et I.A. Stegun, Handbook of Mathematical Functions : with Formulas, Graphs, and Mathematical Tables, volume 55. Dover Publications, 1965.
- [2] D. Applebaum, Lévy Processes : From Probability to Finance and Quantum Groups. Notices of the American Mathematical Society, 51(11) :1336–1347, 2004.
- [3] L. Bachelier, Théorie de la Spéculation. Gauthier-Villars, 1900.
- [4] J. Bégin, Analyse MCMC de Certains Modèles de Diffusion avec Application au Marché Européen du Carbone. Technical report, Université de Montréal, Montréal, Canada, Août 2010.
- [5] M. Benaïm et N. El Karoui, Promenade Aléatoire, Chaînes de Markov et Simulation ; Martingales et Stratégies. 91128 Palaiseau Cedex, École polytechnique edition, Novembre 2007.
- [6] F. Black et M.Scholes : The Pricing of Options and Corporate Liabilities, Journal of Political Economy 3 (1973), pp.637-654.
- [7] D.J. Buckle, Bayesian Inference for Stable Distributions. Journal of the American Statistical Association, 90(430) :605–613, 1995.
- [8] M. Crowder, On Consistency and Inconsistency of Estimating Equations. Econometric Theory, pages 305–330, 1986.
- [9] M. Crowder, On Linear and Quadratic Estimating Functions. Biometrika, 74(3) :591–597, 1987.
- [10] E. Derman, Model Risk. Rapport technique, Goldman Sachs, 1996.
- [11] J.J. Dreesbeke, J.Fine et G.Saporta, Méthodes Bayésiennes en Statistique, Editions TECHNIP, France, 2002.
- [12] E.F. Fama, The Behavior of Stock Market Price, Journal of Business 38, pp.34-195, 1965.

- [13] E.F. Fama et Kenneth R. French, Common Risk Factors in the Returns on Stocks and Bonds. *Journal of Financial Economics*, 33(1) :3–56, 1993.
- [14] D. Flipo, Chaînes de Markov. Université des Sciences et Technologies de Lille, U. F. R. de Mathématiques Pures et Appliquées.
- [15] R. Fox et M.S. Taqqu, Large-sample Properties of Parameter Estimates for Strongly Dependent Stationary Gaussian Time Series. *The Annals of Statistics*, 14(2) :517–532, 1986.
- [16] A. Gelfand et A. Smith, Sampling Based Approaches to Calculating Marginal Densities, *J. American Statist. Assoc*, 85, 398409, 1990.
- [17] S. Geman et D. Geman, Stochastic Relaxation, Gibbs Distributions, and The Bayesian Restoration of images. 6 :721–741, 1984.
- [18] L.P. Hansen, Sample Properties of Generalized Method of Moments Estimators, *Econometrica* 50 (1982), pp. 1029-1054.
- [19] W.K. Hastings, Monte Carlo Sampling Methods using Markov Chains and their Applications, volume 57. *Biometrika*, April 1970. pages 97-109.
- [20] D.V. Hinkley et N.S. Revankar, Estimation of the Pareto Law from Underreported Data, A Further Analysis. *Journal of Econometrics*, 5(1) :1–11, 1977.
- [21] S. Kotz, T.J. Kozubowski et K. Podgórski, The Laplace Distribution and Generalizations, A Revisit with Applications to Communications, Economics, Engineering, and Finance. *Progress in Mathematics Series*. Birkhäuser, 2001.
- [22] T.J. Kozubowski et K. Podgórski, A Class of Asymmetric Distributions. *Actuarial Research Clearing House*, 1 :113–134, 1999.
- [23] T.J. Kozubowski et K. Podgórski, Asymmetric Laplace Laws and Modeling Financial Data. *Mathematical and Computer Modeling*, 34(9) :1003–1021, 2001.
- [24] D.B. Madan et E. Seneta, The Variance Gamma Model for Share Market Returns. *Journal of Business*, pages 511–524, 1990.
- [25] B. Mandelbrot, The Variation of Certain Speculative Prices. *Journal of Business*, pages 394–419, 1963.
- [26] R.C. Merton, Option Pricing when Underlying Stock Returns Are Discontinuous. *Journal of Financial Economics*, 3(1) :125–144, 1976.
- [27] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller, Equation of State Calculations by Fast Computings machines. *Chemical Physics*, 21(6) :1087–1092, June 1953.

- [28] T. Mikosch, Sidney I. Resnick et O.E. Barndorff-Nielsen, *Lévy Processes : Theory and Applications*. Birkhäuser, 2001.
- [29] W.C. Mitchell, A Critique of Index Numbers of the Prices of Stocks. *The Journal of Political Economy*, pages 625–693, 1916.
- [30] F. Pelletier, *Modélisation des rendements financiers à l’aide de la distribution de Laplace asymétrique généralisée*, Québec, Canada ,2014.
- [31] P.D. Praetz, The Distribution of Share Price Changes. *Journal of Business*, pages 49–55, 1972.
- [32] S.j. Press, A Compound Events Model for Security Prices. *Journal of Business*, pages 317–335, 1967.
- [33] H. Raiffa et R. Schlaifer, *Applied Statistical Decision Theory*. Boston, Massachusetts :Harvard Business School, 1961.
- [34] C.P. Robert, *Le Choix Bayésien. Principes et pratique*, INSEE, Paris, 2006.
- [35] A. Ruegg, *Processus Stochastique*. Presses polytechniques romandes edition.
- [36] P. Samuelson, Rational theory of warrant pricing, *Industrial Management Review* 6 (1965), pp.13-32
- [37] J. D. Sargan, The Estimation of Economic Relation Ship Using Instrumental variables, *Econometrica* 26 (1958), pp. 393-415.
- [38] I. Sato, *Lévy Processes and Infinitely Divisible Distributions*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 1999.
- [39] W. Schoutens ? *Lévy Processes in Finance*. Wiley, 2003.
- [40] D. Scott et C. Yang Dong, variance Gamma : The Variance Gamma Distribution, 2012. URL [http ://CRAN.R-project.org/package=VarianceGamma](http://CRAN.R-project.org/package=VarianceGamma). Consulté le 24 janvier 2014.
- [41] S.S. Shapiro et M.B. Wilk, An Analysis of Variance Test for Normality (Complete Samples). *Biometrika*, 52(3/4) :591–611, 1965.
- [42] M.R. Spiegel et J. Liu : *Schaum’s Mathematical Handbook of Formulas and Tables*, volume 1000. McGraw-Hill, 1999.
- [43] I. Tabet Derraz, *Système de Rémunération et Management de la Performance des Entreprises*, université Abou bekr belkaid- Tlemcen, 2013/ 2014, ,p,164.
- [44] L. Tierney, Markov Chains for Exploring Posterior Distributions (with discussion). *Annals of Statistics*, 1994.
- [45] J.Y. Tourneret, *Méthodes de Monte-Carlo par Chaînes de Markov*. Institut de recherche en informatique de Toulouse (IRIT), Novembre 2007.

- [46] M. Vieira, Estimation Bayésienne par des méthodes MCMC, Application à la Surveillance des Moteurs Asynchrones. Thèse de doctorat en science, Université de Nice-Sophia-Antipolis, 41 Bd Napoléon III 06041 Nice Cedex, Janvier 1999.
- [47] C. Walter, Lévy-Stability Under Addition and Fractal Structure of Markets : Implications for the Actuaries and Emphasized Examination of MATIF National Contract. In Proceedings of the 5th AFIR colloquium. International Actuarial Association, 1995.

Résumé

L'intérêt de ce travail est double. Le premier objectif était de présenter l'utilisation de la distribution de Laplace asymétrique généralisée comme alternative intéressante des modèles classiques en finance. Secundo, celle-ci est rapprochée de la réalité en considérant des informations a priori sur les paramètres de la distribution. Ce qui a conduit à appliquer le paradigme bayésien sur un cas particulier vue l'expression sophistiquée de Laplace asymétrique généralisée. Cette approche demande une grande puissance de calcul : elle n'a donc été rendue possible que grâce au développement relativement récent du calcul numérique (méthodes *MCMC* : Monte Carlo par chaînes de Markov) et de l'outil informatique performant.

Mots-clés : modèles de rendements financiers, *GAL*, *SL*, estimation bayésienne, *MCMC*.

Abstract

The objective of this work is twofold. The first one was to present the use of the generalized asymmetric Laplace distribution (*GAL*) as an interesting alternative to the financial classic models. Secundo, this one is approached to the real situation considering a prior informations about the distribution's parameters. It led us to apply the bayesian paradigm considering a particular distribution (*SL*). This approach requires great computing power : it was therefore only possible thanks to the relatively recent development of numerical computing (*MCMC* methods : Monte Carlo via Markov Chains) and the high performance computing tool.

Keywords : financial performance models, *GAL*, *SL*, Bayesian estimation, *MCMC*.