

HEC MONTRÉAL  
École affiliée à l'Université de Montréal

Aziz SORÉ

**Réplication et analyse de performance des fonds de couverture**

Mémoire présenté à HEC Montréal en vue de l'obtention du grade de  
Maîtrise ès sciences de la gestion, option Ingénierie Financière

Mars 2007

©Aziz SORÉ, 2007

*Merci à Bruno et Nicolas.*

# Table des matières

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Revue de littérature</b>	<b>3</b>
2.1	Réplication des fonds de couverture . . . . .	3
2.1.1	Les fonds de couverture . . . . .	3
2.1.2	Analyse de la performance des investissements . . . . .	5
2.1.3	Réplication de fonds de couverture . . . . .	7
2.2	Distribution des rendements . . . . .	8
2.2.1	Dépendance et utilisations . . . . .	8
2.2.2	Mélanges de distributions . . . . .	9
2.3	Réplication de contrats conditionnels . . . . .	11
<b>3</b>	<b>Tarification et couverture des contrats</b>	<b>13</b>
3.1	Procédure de réplication . . . . .	13
3.2	Construction des contrats conditionnels . . . . .	16
3.3	Tarification et couverture des contrats . . . . .	18
3.3.1	Hypothèses . . . . .	18
3.3.2	Résultats théoriques . . . . .	21
3.3.3	Réplication dans le cadre markovien . . . . .	23

---

3.4	Réplication par la méthode de Monte Carlo . . . . .	25
<b>4</b>	<b>Validation pratique de la procédure de réplication</b>	<b>27</b>
4.1	Introduction . . . . .	27
4.2	Construction du processus de rendements . . . . .	28
4.2.1	Loi de Laplace . . . . .	28
4.2.2	Loi des rendements journaliers . . . . .	29
4.3	Application . . . . .	29
4.3.1	Le processus de prix et le contrat conditionnel . . . . .	30
4.3.2	Analyse des résultats . . . . .	31
4.4	Conclusion . . . . .	33
<b>5</b>	<b>Application : Approche par les mélanges de lois normales</b>	<b>35</b>
5.1	Introduction . . . . .	35
5.2	Mélanges de lois normales . . . . .	36
5.2.1	Définition . . . . .	36
5.2.2	Estimations des paramètres . . . . .	39
5.2.3	Résultats . . . . .	40
5.3	Construction de fonds synthétiques . . . . .	41
5.3.1	Système de distributions de Johnson . . . . .	41
5.3.2	Construction de la distribution de Johnson SU . . . . .	43
5.4	Sélection de la dépendance entre le fond et le portefeuille . . . . .	44
5.4.1	Copule Gaussienne . . . . .	44
5.4.2	Copule de Clayton . . . . .	45
5.5	Analyse des résultats . . . . .	47

---

5.5.1	Construction de fonds synthétiques . . . . .	47
5.5.2	Sensibilité à la dépendance . . . . .	50
5.6	Conclusion du chapitre . . . . .	52
<b>6</b>	<b>Conclusion générale</b>	<b>54</b>
	<b>Annexes</b>	<b>59</b>
<b>A</b>	<b>Preuve de la proposition 1</b>	<b>60</b>
<b>B</b>	<b>Détermination des paramètres de la distribution de Johnson-SU</b>	<b>62</b>
<b>C</b>	<b>Copules</b>	<b>64</b>
C.1	Introduction aux copules . . . . .	64
C.2	Quelques mesures de dépendance . . . . .	65

# Liste des tableaux

4.1	Statistiques de la réplication : cas de la distribution de Laplace . . . .	33
5.1	Estimation des paramètres du modèle des rendements journaliers . . .	40
5.2	Estimation des paramètres du modèle des rendements mensuels . . .	41
5.3	Statistiques de la réplication : marge Johnson-SU, copule d'indépendance . . . . .	48
5.4	Statistiques de la réplication : marge Johnson-SU, copule gaussienne .	48
5.5	Statistiques de la réplication : marge Johnson-SU, copule de Clayton	49
5.6	Tableau comparatif des performances . . . . .	50
5.7	Performance de la réplication en fonction du paramètre de la copule gaussienne . . . . .	51
5.8	Performance de la réplication en fonction du paramètre de la copule de Clayton . . . . .	52

# Chapitre 1

## Introduction

Les fonds de couverture ont connu un développement important depuis le début des années 1980. Ce sont des produits financiers spéculatifs à rendements élevés. Cependant les gestionnaires de ces fonds exigent d'important frais de gestion. Au début des années 1990, les fonds de couverture ont connu des rendements décroissants au fil des années. Depuis la fin des années 1990, ils sont de plus en plus utilisés dans la diversification de risques de portefeuilles sous gestion car ils présentent quelques fois, une faible corrélation aux marchés. Ils profitent ainsi de l'intérêt croissant pour la gestion du risque.

Les conditions imposées aux investisseurs, en plus des frais de gestion, sont cependant très contraignantes. D'une manière générale, ces fonds imposent des conditions temporelles importantes pour les nouveaux investissements. Certains d'entre eux exigent des frais de compensation en cas de retrait. En somme, les fonds de couverture ont une faible liquidité. Certains auteurs ont donc entrepris d'étudier le comportement des rendements de ces investissements et de les répliquer à partir de titres plus liquides.

Le cadre théorique général de la réplification de la distribution des rendements des investissements a été développé par Dybvig (1988a). Dans une problématique de diversification de risque et d'analyse de performance, Kat and Palaro (2005), en se

basant sur les travaux de Dybvig (1988a) et Dybvig (1988b), a développé une méthode de réplcation des rendements joints du fond de couverture et du portefeuille. En se basant sur les hypothèses de Dybvig, ils montrent que la distribution jointe du portefeuille et du fond de couverture pourrait être obtenue à partir d'une fonction de la distribution jointe du portefeuille et d'un autre actif. Ils déduisent ainsi un contrat contingent qui réplique le comportement du portefeuille et du fond de couverture. Enfin, ils évaluent le contrat et proposent une stratégie de couverture. Cependant, leur technique de réplcation du contrat est inadaptée.

L'objectif du présent mémoire est de répliquer l'article Papageorgiou et al. (2007).

# Chapitre 2

## Revue de littérature

### 2.1 Réplication des fonds de couverture

#### 2.1.1 Les fonds de couverture

Le fond de couverture est un véhicule d'investissement utilisant des stratégies agressives comprenant, entre autre, les ventes à découvert, les leviers financiers, les swaps, les stratégies d'arbitrage et les produits dérivés. Les réglementations de ces fonds sont différentes de celles des fonds classiques. Les gestionnaires de fonds de couverture ne sont par exemple pas tenus de reporter périodiquement la valeur globale de leurs portefeuilles sous gestion. Ils ne sont pas non plus tenus de préciser leurs positions dans les marchés. Cependant, la réglementation américaine limite le nombre d'investisseurs dans un fond de couverture à 99. Les fonds de couvertures exigent, en général, des frais de gestion et des frais incitatifs sur les profits réalisés.

Les fonds de couvertures ont été initiés pour la première fois par Alfred Winslow Jones en 1949. Celui-ci a développé des stratégies neutres au risque du marché ainsi que des stratégies de leviers. Les fonds de couverture ont connu une importante croissance en volume et en capital sous gestion depuis 25 ans. Leur succès était initialement dû à leurs rendements élevés comme le précise Liang (1999). Ce dernier

étudie la performance des fonds de couvertures en utilisant des outils traditionnellement utilisés pour l'analyse des performances des fonds mutuels. Il constate que le risque spécifique de ces fonds est plus faible. Ces fonds performant mieux lorsque les frais incitatifs sont importants. Les ratios de Sharpe sont, en outre, plus élevés pour les fonds de couvertures que pour les mutuels classiques. Cependant, ces mesures présentent des insuffisances, lorsqu'elles sont utilisées dans le cas des fonds de couvertures. Quant à l'explication de ces rendements, plusieurs auteurs ont proposé des solutions d'analyse factorielles, qui sont utilisées quelques fois pour interpréter les performances des fonds. Elton et al. (1995) proposent un modèle rajoutant deux facteurs (le momentum et les obligations) aux facteurs explicités dans Fama and French (1993). Agarwal and Naik (2003) proposent l'utilisation des performances d'options pour expliquer la non linéarité des rendements de certaines stratégies.

Les fonds de couvertures sont, de nos jours, utilisés dans le cadre de la diversification du risque des portefeuilles. En effet, comme le précise Liang (1999), ces fonds ont en général, une corrélation faible au marché. Cependant, ils présentent quelques fois des coefficients d'asymétrie et d'aplatissement qui ne sont pas toujours très adaptés et qui pourrait influencer fortement ceux du portefeuille global. Pour palier à ce problème, certains gestionnaires ont créé des portefeuilles de fonds de couverture appelés, fonds de fonds. Les fonds de fonds sont des portefeuilles de fonds de couverture. Dave et al. (2004), présentent une technique permettant de déterminer la composition du fond de fonds selon les préférences (des quatre premiers moments de la distribution) des investisseurs. Les fonds de fonds présentent deux structures de frais : les frais (fixes et incitatifs) liés à la gestion de chaque fonds de couverture et ceux liés à la gestion des fonds de fonds. Ces deux niveaux présentent de situations quelques fois désavantageuses pour les investisseurs, qui peuvent se retrouver à payer des frais de performance pour certains fonds de couvertures pendant que, globalement, le fond de fonds ne performe pas. Pour remédier à cette situation, Brown et al. (2004) suggèrent une stratégie de restructuration des frais en proposant des

options sur des actifs fortement corrélés au fond de fonds. Cette méthode permet de couvrir les frais des fonds de couverture.

Les fonds de couverture présentent un risque qui n'est pas toujours bien maîtrisé. Selon Lo (2001), l'utilisation d'outils traditionnels dans le cadre de l'analyse moyenne-variance comme le bêta du marché, la Valeur à Risque, ne traduisent que partiellement les risques des fonds de couvertures.

Les études des fonds de couvertures se sont, en général, faites à partir d'outils traditionnellement utilisés pour les fonds mutuels. Pourtant les distributions de ces fonds sont, en général, différentes de celle des fonds mutuels. Il n'est donc pas toujours possible d'effectuer des études dans le cadre de l'analyse moyenne-variance. La section suivante intègre la distribution des rendements dans le cadre de l'analyse de la performance du portefeuille.

### 2.1.2 Analyse de la performance des investissements

L'analyse de Dybvig (1988a) s'articule essentiellement sur la distribution des portefeuilles et des actifs ainsi que de leurs prix à des dates antérieures. C'est un problème classique d'évaluation de produit dérivés.

Le CAPM (Capital Asset Pricing Model), construit à partir de l'analyse moyenne-variance, repose sur des hypothèses simples et conduit à des conclusions élégantes. Cependant, il n'est pas adapté à toutes les classes d'actifs. De plus, comme l'a montré Dybvig and Ingersoll (1982), l'application du CAPM aux options présente des opportunités d'arbitrage. Le PDPM (Payoff Distribution Pricing Model) est une extension de ce modèle. Les hypothèses de ce modèle sont :

- La préférence de l'agent dépend uniquement de la distribution terminale du bien.

- L'agent préfère la meilleure à la moins bonne distribution terminale du bien.
- La complétude des marchés est sans imperfection (pas de taxes, pas de coûts de transactions, pas d'asymétrie de l'information).

La dernière hypothèse est la plus importante. La complétude est caractérisée par l'équiprobabilité des états dans le cas discret ou la continuité des états. Les états caractérisent les situations économiquement réalisables dans un marché.

À partir de ces hypothèses, il déduit des résultats très importants, qui motivent la procédure de réplication des fonds de couverture. Le premier résultat est la caractérisation de la distribution des gains par rapport à la densité des prix des états. Le second résultat est la conformité de ce résultat avec la maximisation de la fonction d'utilité de von Neumann-Morgenstern.

Le troisième résultat précise le prix le plus abordable d'une distribution terminale de rendement donnée dans un marché. Ce prix est une fonction de la distribution inverse du payoff terminal et de la distribution inverse de la densité des prix.

Le quatrième résultat, qui est l'un des plus importants, permet de mesurer d'une manière assez simple l'efficacité d'un investissement. Ce résultat indique que si le prix de la distribution est supérieur à la valeur initiale de l'investissement, alors l'investissement est optimal dans la cadre de ce marché. Si le coût de l'investissement est supérieur au prix de la distribution, alors il est plus optimal de prendre une position sur un contrat conditionnel répliquant la distribution de l'investissement dans ce marché. Il propose également, dans ce cadre, des outils utilisés pour l'évaluation de la performance de l'investissement.

Il souligne, en conclusion, l'analyse de l'importance des erreurs dans la modélisation, tant de la distribution terminale de l'investissement que de la distribution des états (le marché). Cependant, son analyse n'a concerné que des distributions unidimensionnelles. La section suivante propose une extension de ces analyses à des

distributions multidimensionnelles.

### 2.1.3 Réplication de fonds de couverture

Dans l'analyse unidimensionnelle, Dybvig (1988a) et Dybvig (1988b) ont déterminé le prix d'une distribution terminale d'un investissement dans un marché complet. Kat and Palaro (2005) proposent une extension de cette analyse au cas multidimensionnel. Les hypothèses sont conservées, à la seule modification de la première : les préférences des investisseurs dépendent de la distribution jointe terminale de l'investissement et du portefeuille. À partir de ces nouvelles hypothèses, il est possible de retrouver la valeur de la distribution terminale jointe .

Kat and Palaro (2005) se sont principalement intéressés à la procédure de répliation de la fonction payoff du marché qui est la suivante :

1. Collecte de données du fond de couverture, du portefeuille et de l'actif de réserve.
2. Modélisation de la distribution jointe des rendements du fond et du portefeuille (appelée la distribution désirée).
3. Modélisation de la distribution jointe des rendements du portefeuille et de la réserve.
4. Détermination du contrat conditionnel le moins coûteux.
5. Evaluation ce contrat conditionnel.
6. Recherche de stratégies optimales d'allocation du portefeuille et de l'actif de réserve.

La valeur du contrat conditionnel (qui correspond également à la performance de l'investissement) est appelée "KP-Measure".

La modélisation des portefeuilles, fonds de couverture et réserves se fait en deux étapes : la modélisation des lois marginales et celle de la dépendance. Pour modéliser les lois marginales, Kat and Palaro (2005) utilisent des distributions gaussiennes, de Student et de Johnson. Les dépendances sont modélisées par des copules méta-elliptiques et des copules archimédiennes.

La détermination du contrat conditionnel le moins coûteux s'effectue à l'aide de la dépendance qui existe entre les rendements du portefeuille et le fonds de couverture, celle du portefeuille et de la réserve et enfin les distributions marginales des trois actifs. La méthode générale d'évaluation développée par Harrison and Kreps (1979) est utilisée pour évaluer le contrat contingent. La réplication du contrat est faite dans le cadre du modèle de Black et Scholes. La réplication en temps discret produit inexorablement des erreurs.

Plusieurs critiques ont été faites pour l'approche de Kat and Palaro (2005). La principale porte essentiellement sur la technique de réplication des droits contingents issus de leur modèle. En effet, la taille des erreurs de réplication n'est pas précisée dans leur travail. De plus, leur procédure de réplication pose une hypothèse de normalité des rendements journaliers du portefeuille et de la réserve. Pourtant, ils construisent des contrats conditionnels en ne tenant pas compte de ce que cela impliquera. Ils ont toutefois, admis dans leur article, que cette procédure est imparfaite. L'utilisation de méthodes bien plus adaptées permettrait certainement d'obtenir de meilleurs résultats. Papageorgiou et al. (2007) proposent une méthode plus adaptée pour l'évaluation des contrats.

## 2.2 Distribution des rendements

### 2.2.1 Dépendance et utilisations

La modélisation de la dépendance est très utilisée de nos jours dans les disciplines scientifiques, particulièrement en finance. Les copules furent introduites pour la première fois par Sklar (1959). Cependant, elles restèrent peu utilisées jusqu'en 1975 où on les retrouve dans les travaux de Deheuvels (1979). Mais dans les années 1980, Genest and MacKay (1986) ont été les précurseurs de la modélisation de la dépendance. Dans leur article, ils construisent une nouvelle classe de copules appelée copules archimédiennes. Les copules sont devenues un outil essentiel dans tous les domaines utilisant des modèles probabilistes.

L'utilisation des copules en finance est encore plus récente tant en milieu académique qu'en entreprise. L'apparition de produits financiers de plus en plus sophistiqués, notamment des produits dérivés sur plusieurs sous-jacents, rendit leurs valorisations plus difficiles. Rosenberg (1999) est le premier à introduire les copules en finance. Dans cet article, il utilise la copule de Plackett pour calculer la fonction de densité risque neutre bidimensionnelle. Par la suite, plusieurs méthodes proposées dans Rosenberg (2003) et Cherubini and Luciano (2002) ont été développées à partir de cette idée.

Les copules ont aussi été d'un apport important dans la gestion de risques financiers et dans le développement des modèles de portefeuilles de crédit. La fonction copule a été introduite en gestion de risque par Embrechts et al. (1999). Dans leur article, ils ont explicité l'application du concept de la dépendance dans la valorisation du risque. Les travaux essentiels de Roncalli (2004) ont permis la vulgarisation des modèles de dépendances dans la gestion de risques. Berrada et al. (2005) ont étudié l'évaluation des migrations de crédit et des dérivés de crédit en utilisant des

copules.

L'utilisation des copules dans le domaine de l'investissement alternatif est très récente. Elle est amorcée par Kat and Palaro (2005).

### 2.2.2 Mélanges de distributions

Les distributions standards comme les lois normales ne sont pas toujours adaptées dans la modélisation des phénomènes aléatoires. Les mélanges de distributions constituent des solutions alternatives intéressantes.

Les mélanges de lois sont généralement utilisées dans les techniques de reconnaissance et les méthodes d'apprentissage. Ce sont des lois importantes utilisées de la technique de classification automatique des données. Les mélanges de lois normales sont les plus utilisées. Elles ont des propriétés asymptotiques intéressantes.

Les mélanges de distributions furent introduites pour la première fois par Pearson (1894) pour des mesures effectuées sur des crabes. Il utilise à cette occasion, une mixture unidimensionnelle de deux distributions normales. Cependant, ces types de modèles présentèrent des difficultés dans l'estimation des paramètres. L'estimation des paramètres par la méthode des moments est largement étudiée. Des extensions ont aussi été étudiées pour les distributions bivariées par Charlier and Wicksell (1924). Ce n'est qu'à partir des années 1970, que les études de Tan and Chang (1972) et celles de Fryer and Robertson (1972) sur les estimateurs du maximum de vraisemblance ont permis de prouver la supériorité de ces derniers sur les estimateurs issus de la méthode des moments. Il n'existe pourtant pas de formule explicite donnant ces estimateurs. Dempster et al. (1977) a développé un algorithme récursif convergent vers les estimateurs du maximum de vraisemblance. Ce résultat, très important, a été étendu et amélioré, tant dans la précision que dans la vitesse de convergence.

Les mélanges de lois, généralement utilisées en finance, sont les mélanges de lois normales. Ces modèles sont intervenus en finance à travers des études du comportement de la volatilité, le volume des transactions et du prix. Jones et al. (1994), à travers des études sur les prix quotidiens d'actifs sur le Nasdaq ont conclu que seul le nombre de transactions et non le volume génèrent la volatilité. Plusieurs approches ont alors été proposées pour étendre l'hypothèse de normalité des rendements. Berger and Mandelbrot (1963) ont introduit les classes de processus stables pour modéliser les déviations de rendements du mouvement brownien. Clark (1973) propose, quant à lui, de relier les changements de normalité à l'existence du changement en volume pendant différentes périodes de trading. Il modélise alors les titres par des processus subordonnés. Les processus subordonnés, comme le définit Bochner (1955), sont des processus constitués à partir la composition d'un processus des rendements et d'un processus caractérisant le temps aléatoire appelé subordinator. Le subordinator est un processus continu à droite, et ayant des incréments homogènes et indépendants. En exemple, les processus arrêtés sont des processus subordonnés. Ané and Geman (2000) constatent que les propriétés (indépendance et homogénéité des incréments) sont trop fortes et assez contraignantes dans le contexte financier actuel. Ils proposent une généralisation des subordinateurs, avec des propriétés bien moins restrictives en utilisant, selon les cas, des processus discrets ou continus. Les subordinateurs discrets permettent d'obtenir des processus à sauts. Ils en arrivent à la conclusion suivante : tout processus modélisant le rendement des actifs peut être représenté comme un mouvement brownien à temps aléatoires. Le changement de régime est un élément important dans le flux d'information et l'activité des marchés.

## 2.3 Réplication de contrats conditionnels

L'évaluation des contrats conditionnels peut être réalisée par plusieurs méthodes : la méthode martingale, la méthode des équations aux dérivées partielles, et la mé-

thode de couverture des contrats (la valeur d'un contrat étant égale à la valeur de la stratégie autofinancée permettant de le répliquer). La méthode martingale est essentiellement construite suite sur le postulat de l'absence d'arbitrage. Cette méthode a été formalisée dans Harrison and Kreps (1979) et Harrison and Pliska (1981).

La méthode de réplication des droits conditionnels est relativement récente. Elle consiste à créer, avec les actifs transigés dans le marché, une stratégie admissible et autofinancée reproduisant exactement le contrat conditionnel. Par exemple, dans le cadre d'un contrat conditionnel de type européen, l'objectif est de déterminer les paramètres de la stratégie c'est-à-dire le prix initial du droit et la composition du portefeuille de couverture évoluant dans le temps, permettant d'obtenir une distribution terminale proche du contrat. Cette méthode permet, non seulement de déterminer le prix du contrat, mais aussi de retrouver une stratégie de couverture.

Schweizer (1995) est l'un des précurseurs de l'approche par la méthode de réplication des droits conditionnels. Dans son article, il a formalisé cette méthode, en la décrivant comme la solution d'un problème d'optimisation stochastique liée à l'erreur quadratique de la couverture.

Cette dernière hypothèse consiste en la minimisation de l'espérance de l'erreur quadratique existant entre le contrat conditionnel et la distribution terminale de la stratégie de réplication. Cette méthode pose plusieurs interrogations. Existe-t-il une solution optimale à ce problème ? Si oui, existe-t-il une technique permettant de déterminer une solution ? Sous certaines conditions sur le processus modélisation des actifs, il existe au moins une solution optimale. La solution optimale se fait, si besoins est, en deux étapes : la détermination de la stratégie de couverture pour n'importe quel prix du contrat initial et ensuite la détermination du prix minimisant la fonction objective. Il propose également une formule explicite, dans ces conditions, d'une stratégie optimale de couverture.

# Chapitre 3

## Tarification et couverture des contrats

Ce chapitre reproduit la méthode de tarification développée par Papageorgiou et al. (2007)

### 3.1 Procédure de réplcation

Kat and Palaro (2005), en se référant aux travaux de Dybvig (1988a) et Dybvig (1988b), ont démontré qu'à partir de deux titres risqués  $S_1$  et  $S_2$ , il toujours possible de reproduire, à la fois la distribution des rendements d'un troisième titre  $S_3$ , mais aussi sa dépendance d'avec le titre  $S_1$ . Il est à noter qu'il ne s'agit pas d'une reproduction exacte de la distribution des rendements du titre  $S_3$  mais d'un titre ayant la même loi de probabilité et la même dépendance d'avec le titre  $S_1$ . Cela se traduit par l'équation suivante :

$$P(S_{1,T} \leq x, \tilde{g}(S_{1,T}, S_{2,T}) \leq y) = P(S_{1,T} \leq x, S_{3,T} \leq y), \forall x, y > 0$$

Après avoir explicité la fonction  $\tilde{g}$ , en utilisant les fonctions de dépendance et de répartition, il faudra donc générer le flux  $\tilde{g}(S_{1,T}, S_{2,T})$ , sans investir dans le titre  $S_3$ .

on utilise donc les titres  $S_0$  (titre non risque),  $S_1$  et  $S_2$  rebalancés périodiquement à  $t = 0, \dots, T - 1$ .

Dans le contexte des fonds de couverture et fonds synthétiques,

- $S_0$  représente l'actif sans risque
- $S_1$  le portefeuille ( $P$ )
- $S_2$  l'actif de réserve, utilise dans la répliation ( $R$ )
- $S_3$  le fond à répliquer  $I$

Les étapes de la répliation sont les suivantes :

1. Collecter les rendements(ou prix) journaliers, historiques, du portefeuille initial de l'investisseur et de l'actif de réserve.
2. Modéliser, à partir de ces données journalières, la loi jointe de la distribution du portefeuille et de la réserve. Cette modélisation peut se faire de deux manières. La première méthode est la modélisation dite directe de la distribution jointe du portefeuille et de la réserve. La seconde méthode est la modélisation des distributions marginales et les dépendances entre les deux actifs. L'avantage de la première méthode est l'obtention directe, à la fois de lois marginales du portefeuille et de la réserve et la dépendance entre les deux actifs. L'inconvénient de la première méthode est la limitation des familles de distributions candidates. La seconde méthode, de ce fait, offre une plus large famille de distributions. Cependant elle conduit à de plus larges erreurs dans l'estimation des paramètres.
3. Si l'on désire répliquer un fond de couverture, il faudra collecter également les rendements historiques de ce fond.
4. Modéliser la loi jointe de la distribution mensuelle du portefeuille et de la réserve. C'est la partie la plus importante de la procédure. En effet, l'une des propriétés fondamentales de la procédure de replication (que Kat and Palaro (2005) n'ont pas intégré dans leur démarche) est que la somme des lois des

rendements journaliers doit être égale à loi des rendements mensuels.

$$\sum_{i=1}^T R_i = R_{0,T}$$

Du fait de cette contrainte fondamentale dans la modélisation, si l'on souhaite modéliser la loi des rendements journaliers et celle des rendements mensuels, il faudrait, en théorie, que cette égalité soit faite en *loi*. Pourtant, dans la pratique, il est difficile d'estimer séparément les paramètres de lois journalières et mensuels tout en respectant cette contrainte, qui de plus, ne peut qu'être qu'une égalité en loi, dans le meilleur des cas. Alors, deux solutions sont à envisager dans ce cas : On déduit analytiquement la distribution mensuelle à partir des distributions journalières. Dans le présent cas, l'erreur intrinsèque du modèle est nulle, par construction. La seconde méthode, approximative, consiste à sélectionner les familles de distributions pouvant avoir des propriétés statistiques proches de celle de la somme des distributions journalières. Cette dernière méthode a l'avantage d'être assez pratique. Cependant, elle induit une erreur presque certaine du modèle qui influencera certainement l'erreur de répliation.

5. On déduit la fonction  $g$  permettant de calculer le contrat conditionnel construit à partir des actifs  $S_1$  et  $S_2$ . Tout en restant dans la logique de Dybvig (1988a), on recherche la fonction de profit la moins chère, qui serait à même de reproduire les propriétés statistiques du fond à réplier.
6. On réplie le contrat conditionnel  $\tilde{g}(S_{1,T}, S_{2,T})$  à partir des actifs  $S_0$ ,  $S_1$  et  $S_2$ , selon le critère suivant : la minimisation du carré de l'erreur existant entre le portefeuille de répliation et le contrat conditionnel. Dans le cadre d'un marché complet, cette erreur est nulle.
7. Analyse de la performance du fond : A partir d'un investissement initial de 100\$ dans le fond, la valeur initiale du portefeuille de répliation permet de

confronter les deux stratégies d'investissement. Si la valeur du portefeuille de réplication est inférieure à 100, elle est plus optimale, en terme d'investissement que le fond. Dans le cas contraire, le fond est optimal.

## 3.2 Construction des contrats conditionnels

Conformément à la procédure de réplication de la distribution jointe du portefeuille de l'investisseur et du fond de couverture, il existe une fonction  $\tilde{g}$ , à deux variables, qui, appliquée au portefeuille de l'investisseur et à la réserve, ayant la même distribution. Cela se traduit par la relation

$$P(S_{1,T} \leq x, \tilde{g}(S_{1,T}, S_{2,T}) \leq y) = P(S_{1,T} \leq x, S_{3,T} \leq y), \quad \forall x, y > 0.$$

Si on dénote par  $S_{1,0}$ ,  $S_{2,0}$  et  $S_{3,0}$ , les valeurs initiales des trois actifs, le rendement de la stratégie la moins chère dans le marché constitué par les titres  $S_1$ ,  $S_2$  et l'actif sans risque  $S_0$  est alors

$$\tilde{g}(x, y) = 100 \exp \{g(\log(x/S_{1,0}), \log(y/S_{2,0}))\},$$

où

$$P(X_{1,T} \leq x, g(X_{1,T}, X_{2,T}) \leq y) = P(X_{1,T} \leq x, X_{3,T} \leq y),$$

avec  $X_{1,T} = \log(\frac{S_{1,T}}{S_{1,0}})$ ,  $X_{2,T} = \log(\frac{S_{2,T}}{S_{2,0}})$ , et  $X_{3,T} = \log(\frac{S_{3,T}}{S_{3,0}})$ .

Alors

$$P(g(X_{1,T}, X_{2,T}) \leq y | X_{1,T} = x) = P(X_{3,T} \leq y | X_{1,T} = x), \quad \forall x, y > 0.$$

Si on dénote par  $F_{2|1}(y|x) = P(X_{2,T} \leq y | X_{1,T} = x)$  et  $F_{3|1}(y|x) = P(X_{3,T} \leq$

$y|X_{1,T} = x$ ), les fonctions de répartition conditionnelles respectives du rendement des titres  $S_2$  et  $S_3$ , sachant celui du titre  $S_1$ , la fonction  $g$  est alors donnée par

$$g(x, y) = F_{3|1}^{-1}(F_{2|1}(y|x)|x)$$

où  $F_{3|1}^{-1}(y|x)$  est la pseudo inverse de  $F_{3|1}(y|x)$ .

En effet,

$$\begin{aligned} P(g(X_{1,T}, X_{2,T}) \leq y | X_{1,T} = x) &= P(g(x, X_{2,T}) \leq y | X_{1,T} = x) \\ &= P(F_{3|1}^{-1}(F_{2|1}(X_{2,T}|x)|x) \leq y | X_{1,T} = x) \\ &= P(F_{3|1}^{-1}(U|x) \leq y) = F_{3|1}(y|x), \end{aligned}$$

avec  $U$  suivant une distribution uniforme sur  $[0, 1]$ , d'après Rosenblatt (1952).

Afin de calculer la fonction  $g$ , une petite réflexion s'impose. Si les fonctions de répartition conditionnelles  $F_{2|1}$  et  $F_{3|1}$  sont des distributions connues, alors dans ce cas, la fonction  $\tilde{g}$  se déduit aisément. Un développement supplémentaire n'est pas nécessaire.

Dans le cas où la fonction de répartition jointe du portefeuille de l'investisseur et de la réserve est construite à partir d'une fonction copule et des distributions marginales du portefeuille et de la réserve, la procédure est différente. Il faudrait établir une relation entre  $F_{2|1}$ , la copule caractérisant la dépendance entre les deux actifs notée  $C_{1,2}$ , et les distributions marginales  $F_1$  et  $F_2$ .

En utilisant les propriétés des copules, e.g., Nelsen (2005), on a

$$F_{2|1}(y|x) = \frac{\partial}{\partial u} C_{1,2}(u, v) \Big|_{u=F_1(x), v=F_2(y)}.$$

De même,

$$F_{3|1}(y|x) = \frac{\partial}{\partial u} C_{1,3}(u, v)|_{u=F_1(x), v=F_3(y)}.$$

On détermine ainsi le pseudo-inverse  $Q$  de la loi conditionnelle de  $X_3$  sachant  $X_1$ , est donné par

$$Q(x, \alpha) = F_3^{-1}[\mathcal{Q}(F_1(x), \alpha)],$$

où  $\mathcal{Q}(F_1(x), \alpha)$  est le pseudo-inverse de  $\frac{\partial}{\partial u} C_{1,3}(u, v)$ .

Ainsi, pour un investissement initial de 100\$,

$$\tilde{g}(x, y) = 100e^{(g(\log(\frac{x}{s_{1,0}}), \log(\frac{y}{s_{2,0}})))}.$$

Par conséquent,  $(S_{1,T}, \tilde{g}(S_{1,T}, S_{2,T}))$  a la même distribution de probabilité que  $(S_{1,T}, S_{3,T})$ . On réplique donc, non seulement la distribution du fond, mais aussi sa dépendance avec le portefeuille. Ce dernier aspect met en concurrence la stratégie de réplication et les fonds de couvertures en ce qui concerne l'argument de la diversification du portefeuille.

### 3.3 Tarification et couverture des contrats

Nous arrivons au point central du présent mémoire : la définition d'une stratégie optimale de couverture des contrats contingents en temps discret. Nous introduirons les hypothèses pour lesquelles sera établie la procédure de couverture de contrats conditionnels, ensuite les principaux résultats seront exposés. Enfin, dans le cadre markovien, un algorithme permettant de décrire la stratégie de couverture sera développé. Tous ces résultats proviennent de Papageorgiou et al. (2007).

### 3.3.1 Hypothèses

Considérons deux actifs dont les prix sont *actualisés* et notés  $S_{1,t}$  et  $S_{2,t}$ ,  $t = 0, \dots, n$ .  $S_{0,t}$  représente l'actif sans risque. La filtration  $\mathcal{F} = (\mathcal{F}_k)_{k \in 0, \dots, n}$  représente l'information disponible dans le temps.

On dit que  $(S_{1,t}, S_{2,t})$  est un processus adapté à la filtration  $\mathcal{F}$  si

$$(S_{1,t}, S_{2,t}) \in \mathcal{F}_t, \quad \forall t \in 0, \dots, n.$$

Soit  $\phi$  un processus adapté à la filtration  $\mathcal{F}$ . On dit que  $\phi$  est une stratégie prévisible par rapport à la filtration  $\mathcal{F}$  lorsque

$$\phi_0 \in \mathcal{F}_0 \text{ et } \phi_t \in \mathcal{F}_{t-1}, \quad \forall t \in 1, \dots, n.$$

L'actif sans risque est supposé prévisible. On suppose que les actifs  $S_{1,t}$  et  $S_{2,t}$  sont actualisés et on conserve la même notation.

La stratégie  $\phi$  est dite autofinancée si

$$\phi_t^\top S_t - \phi_{t-1}^\top S_{t-1} = \phi_t^\top (S_t - S_{t-1}).$$

Alors la valeur de la stratégie au temps  $n$  est

$$V_n = V_0 + \sum_{t=1}^n \phi_t^\top (S_t - S_{t-1}),$$

où  $V_0$  représente la valeur initiale de l'investissement.

Lorsque le marché est complet, pour tout contrat conditionnel,  $C$ , il existe une stratégie autofinancée  $\phi$  telle que

$$C = V_0 + \sum_{t=1}^n \phi_t^\top (S_t - S_{t-1}) \text{ presque sûrement.}$$

Dans le cas des marchés incomplets, cette égalité n'est pas vérifiée. Dans ce cas, on définit la variable aléatoire  $G$  par

$$G = C - V_0 - \sum_{t=1}^n \phi_t^\top (S_t - S_{t-1}).$$

Dans le contexte de la réplication des fonds de couverture,  $G$  représente l'erreur due à la réplication du contrat conditionnel. Ainsi, plusieurs stratégies autofinancées reproduisent le contrat conditionnel.

Nous recherchons dans cette étude, la ou les stratégies optimales (si elles existent) qui minimisent la moyenne du carré de l'erreur  $G$ . En d'autres termes, il s'agit de rechercher, parmi les stratégies autofinancées non arbitrées, celles dont la valeur de  $E(G^2) = E(G^2 | \mathcal{F}_0)$  est minimale.

Dans la recherche de la stratégie optimale, une hypothèse s'impose :

La matrice  $\Sigma_t$ , définie par

$$\Sigma_t = E[(S_t - E(S_t | \mathcal{F}_{t-1}))(S_t - E(S_t | \mathcal{F}_{t-1}))^\top], \quad \forall t \in 1, \dots, n, \quad (3.1)$$

doit être inversible.

Comment interpréter cette hypothèse en ingénierie financière ?

Tout d'abord, il faut noter que  $\Sigma_t$  représente la matrice de variance-covariance de l'erreur de prédiction des prix du marché au temps  $t - 1$ .

Ensuite, on remarque que si  $\Sigma_t$  n'est pas inversible, cela signifie qu'il existe une

stratégie prévisible  $\phi$  telle que

$$\phi_t^\top S_t = \phi_t^\top E(S_t | \mathcal{F}_{t-1}),$$

ce qui montrerait que  $\phi_t^\top S_t$  est prévisible. Dans ce cas, cela prouverait qu'il n'y a qu'un seul véritable actif, ce qui est contraire à la réalité.

### 3.3.2 Résultats théoriques

On considère le modèle de marché défini dans la section précédente. Alors on a :

**Théorème 1** *On suppose que  $\Sigma_t$ , définie précédemment est inversible pour tout  $t = 1, \dots, T$ .*

*Alors  $E[G^2(v_0, \phi)]$  est minimisée en définissant récursivement la stratégie  $\phi$  de la manière suivante :*

$$\phi_t = (\Sigma_t)^{-1} E[\{S_t - E(S_t | \mathcal{F}_{t-1})\} C_t | \mathcal{F}_{t-1}], \quad t = T, \dots, 1,$$

où  $C_T, \dots, C_0$  sont définies récursivement par :

$$C_T = C$$

et

$$C_{t-1} = E(C_t | \mathcal{F}_{t-1}) - \phi_t^\top E(S_t - S_{t-1} | \mathcal{F}_{t-1}),$$

pour  $t = T, \dots, 1$ .

De plus, la valeur  $v_0$  du portefeuille de réplication est  $C_0$  et

$$E[G^2] = \sum_{t=1}^T E(G_t^2),$$

où  $G_t = \phi_t^\top S_t - E(S_t | \mathcal{F}_{t-1}) - C_t - E(C_t | \mathcal{F}_{t-1})$ ,  $1 \leq t \leq T$ .

*Du fait du théorème précédent, et comme  $v_0 = C_0$ , on peut interpréter  $C_t$  comme étant la valeur de l'investissement au temps  $t$  permettant de répliquer le contrat conditionnel  $C$  au temps  $T$ . Dans le contexte d'évaluation d'actifs contingents, cette valeur représente un prix du contrat conditionnel au temps  $t$ .*

Quelle pourrait être le lien entre le prix donné par le théorème précédent et celui donné par la mesure martingale? La réponse est donnée par le corollaire suivant :

**Corollaire 1** *Pour tout  $t = 1, \dots, T$  on pose*

$$U_t = 1 - \Delta_t^\top (\Sigma_t)^{-1} E(S_t - S_{t-1} | \mathcal{F}_{t-1}),$$

*où le vecteur  $\Delta$  est défini par*

$$\Delta_t = E[S_t - E(S_t | \mathcal{F}_{t-1})], \quad \forall t \in 1, \dots, n. \quad (3.2)$$

*Ensuite, on définit le processus  $M$  par  $M_0 = 1$  et  $M_t = U_t M_{t-1}$ ,  $1 \leq k \leq n$ . Alors  $M$  est une  $\mathcal{F}$ -martingale (non nécessairement positive) et*

$$C_{t-1} = E(C_t U_t | \mathcal{F}_{t-1}).$$

*En particulier,  $CM$  est une  $\mathcal{F}$ -martingale et  $C_0 = E(C_T M_T | \mathcal{F}_0)$  ainsi,*

$$E(S_t M_t | \mathcal{F}_{t-1}) = S_{t-1}$$

*et donc  $SM$  est une  $\mathcal{F}$ -martingale.*

Les preuves de ces résultats sont faites dans Papageorgiou et al. (2007).

**Remarque 1** *Dans le contexte de l'évaluation d'actifs contingents, le corollaire précédant pourrait s'interpréter de la manière suivante : contrairement à la technique habituelle utilisée dans l'évaluation des actifs contingents, la valeur optimale de la stratégie de couverture n'est pas nécessairement donnée par l'espérance sous une mesure martingale équivalente. L'absence d'opportunité d'arbitrage assure l'existence d'une telle mesure.*

**Remarque 2** *Dans le cas d'un marché complet, la mesure martingale équivalente étant unique, le risque associé la stratégie optimale d'exercice est nulle, presque sûrement. Ainsi, la martingale  $M_t$  est positive par cette méthode et l'on a une représentation explicite de la densité de la mesure martingale  $Q$  sachant  $P$ .*

### 3.3.3 Réplication dans le cadre markovien

Dans le cadre markovien, la méthode de réplication est pratique et se décrit aisément. Soit  $S$  un processus de prix adapté à la filtration  $\mathcal{F}$ . On dit que  $S$  est un processus markovien si la loi de  $S_t$  étant donné  $\mathcal{F}_{t-1}$  ne dépend que de  $S_{t-1}$ , i.e., qu'elle est de la forme  $\nu_t(S_{t-1}, dx)$ . Toute l'information passée, accumulée jusqu'au temps  $t - 1$  est contenue dans l'information disponible uniquement au temps  $t - 1$ .

D'autre part, si on considère un contrat conditionnel dépendant uniquement du prix terminal du processus, c'est-à-dire  $C = f_T(S_T)$ , alors d'après le théorème 1, pour tout  $1 \leq t \leq T$ , il existe deux fonctions  $C_{t-1}$  et  $\phi_t$  dépendant uniquement de  $S_{t-1}$ , représentant respectivement, la valeur et la composition du portefeuille de réplication au temps  $t - 1$  du contrat dans le cadre de la stratégie optimale de couverture définie précédemment.

Ainsi, en posant

$$\begin{aligned}
 L_{1t}(s) &= E(S_t | S_{t-1} = s) \\
 &= \int x \nu_t(s, dx), \\
 L_{2t}(s) &= E(S_t S_t^\top | S_{t-1} = s) \\
 &= \int x x^\top \nu_t(s, dx), \\
 A_t(s) &= L_{2t}(s) - L_{1t}(s) L_{1t}(s)^\top,
 \end{aligned}$$

on a,

$$\begin{aligned}
 \phi_t(s) &= A_t^{-1}(s) E(\{S_t - L_{1t}(s) L_{1t}(s)^\top\} f_t(S_t) | S_{t-1} = s) \\
 &= A_t^{-1}(s) \int (x - L_{1t}(s)) f_t(x) \nu_t(s, dx).
 \end{aligned}$$

En posant de plus,

$$U_t(s, x) = 1 - (L_{1t}(s) - s)^\top A_t^{-1}(s) (x - L_{1t}(s)),$$

on obtient

$$\begin{aligned}
 f_{t-1}(s) &= E(U_t(s, S_t) f_t(S_t) | S_{t-1} = s) \\
 &= \int U_t(s, x) f_t(x) \nu_t(s, dx).
 \end{aligned}$$

Il est à noter que

$$E[S_t | \mathcal{F}_{t-1}] = L_{1t}(S_{t-1}) \quad \text{et} \quad \Sigma_t = A_t(S_{t-1}).$$

De plus, dans plusieurs modèles,  $S_t = \omega_t(S_{t-1}, \xi_t)$  où  $\xi_t$  est indépendant de  $\mathcal{F}_{t-1}$ . Lorsque la loi de  $\xi_t$  est de support infini, il est possible d'approximer  $\phi_t$  et  $f_t$ .

### 3.4 Réplication par la méthode de Monte Carlo

Dans le cadre markovien, il est possible d'utiliser la méthodologie développée par Papageorgiou et al. (2007). Soit  $S$  un processus markovien de prix tel que  $S_t = \omega_t(S_{t-1}, \xi_t)$ ,  $\xi_t \sim \mu_t$ , avec  $\xi_t$  indépendant de  $\mathcal{F}_{t-1}$ ,  $t = 1, \dots, T$ , pour une certaine loi de probabilité  $\mu_t$ . Autrement dit, on suppose ici que  $\nu_t(s, dx) = \omega_t(s, x)\mu_t(dx)$ .

On appelle partition  $\mathcal{P}$  d'un ensemble convexe et compact  $K$ , tout ensemble fini  $\{S_1, \dots, S_m\}$  de simplexes disjoints, d'intérieurs non vides tels que  $K = \bigcup_{j=1}^m S_j$ . L'ensemble des vertices de la partition  $\mathcal{P}$  est noté  $\mathcal{V}(\mathcal{P})$ .  $\mathcal{V}(\mathcal{P})$  représente l'enveloppe convexe de  $K$ .

L'algorithme est construit à partir de simulations de Monte Carlo et d'approximations des ensembles convexes  $K_0, \dots, K_{T-1}$ , déterminés par les partitions  $\mathcal{P}_0, \dots, \mathcal{P}_{T-1}$ .

Le principe de cet algorithme est le suivant : étant donné,  $\tilde{f}_t$ , une approximation de  $f_t$ , on calcule  $\hat{L}_{1t}, \hat{L}_{2t}, \hat{A}_t, \hat{\Delta}_t, \hat{U}_t$  et  $\hat{f}_{t-1}$  pour tout point  $x \in \mathcal{V}(\mathcal{P}_{t-1})$ , en utilisant des simulations de Monte Carlo, afin de remplacer l'intégration par rapport à  $\mu_t$  par une moyenne, et ensuite en effectuant une interpolation linéaire sur tout l'ensemble  $K_{t-1}$ .

1. Poser  $\tilde{f}_T = f_T$ ;
2. Pour tout  $t = T, \dots, 1$  générer  $\xi_{1,t}, \dots, \xi_{N_t,t}$  selon  $\mu_t$  ;
  - Pour tout  $s \in \mathcal{V}(\mathcal{P}_{t-1})$ , calculer :

$$\begin{aligned}
\hat{L}_{1t}(s) &= \frac{1}{N_t} \sum_{i=1}^{N_t} \omega_t(s, \xi_{i,t}), \\
\hat{L}_{2t}(s) &= \frac{1}{N_t} \sum_{i=1}^{N_t} \omega_t(s, \xi_{i,t}) \omega_t(s, \xi_{i,t})^\top, \\
\hat{A}_t(s) &= \hat{L}_{2t}(s) - \hat{L}_{1t}(s) \hat{L}_{1t}(s)^\top, \\
\phi_t(s) &= \hat{A}_t(s)^{-1} \frac{1}{N_t} \sum_{i=1}^{N_t} \{\omega_t(s, \xi_{i,t}) - \hat{L}_{1t}(s)\} \hat{f}_t\{\omega_t(s, \xi_{i,t})\}, \\
\hat{U}_t(s, x) &= 1 - \{\hat{L}_{1t}(s) - s\}^\top \hat{A}_t(s)^{-1} \{x - \hat{L}_{1t}(s)\}, \\
f_{t-1}(s) &= \frac{1}{N_t} \sum_{i=1}^{N_t} \hat{U}_t\{s, \omega_t(s, \xi_{i,t})\} \tilde{f}_t\{\omega_t(s, \xi_{i,t})\}.
\end{aligned}$$

- Interpoler linéairement  $\hat{f}_{t-1}$  sur  $K_{t-1}$  et retourner à l'étape 2 jusqu'à ce que  $t = 1$ .

# Chapitre 4

## Validation pratique de la procédure de répliation

### 4.1 Introduction

Ce chapitre est essentiellement consacré à la validation pratique de la procédure de répliation. On construit artificiellement un processus de rendements. Les processus de rendements journaliers sont construits de tel sorte que les propriétés les reliant aux rendements mensuels soient théoriquement valides. Ainsi, à partir de rendements mensuels, on construit des lois de rendements journaliers, indépendantes et identiquement distribuées de tel sorte que la loi de leur somme ait la même loi que celle des rendements mensuels. La procédure de construction du modèle est donc inversée. L'objectif est double. Il s'agit, tout d'abord, de tester la validité de la procédure en évaluant un contrat contingent dont la valeur théorique est connue. Ensuite, on évaluera la performance de la répliation elle-même.

## 4.2 Construction du processus de rendements

La construction du processus de prix se fait de la manière suivante. À partir d'une loi de rendements mensuels, on détermine des distributions de lois journalières indépendantes et identiquement distribuées de tel sorte que leur somme soit égale en loi à la distribution mensuelle.

### 4.2.1 Loi de Laplace

La loi de Laplace ou double exponentielle est une loi continue à un paramètre, construite à partir de deux lois exponentielles. La différence entre deux variables aléatoires exponentielles indépendantes et identiquement distribuées est une loi de Laplace. La densité cette loi est donnée par :

$$f(x; \sigma) = \frac{1}{2\sigma} \exp\left(-\frac{|x|}{\sigma}\right).$$

La fonction de répartition de la loi est donnée par

$$F(x; \sigma) = \begin{cases} \frac{1}{2}e^{x/\sigma} & \text{si } x < 0; \\ 1 - \frac{1}{2}e^{-x/\sigma} & \text{si } x \geq 0. \end{cases}$$

La fonction de répartition inverse est, quant à elle,

$$F^{-1}(x; \sigma) = \begin{cases} -\sigma(\log(2 - 2x)) & \text{si } x \geq 0.5; \\ \sigma(\log(2x)) & \text{si } x < 0.5. \end{cases}$$

D'autre part, la moyenne et la variance de cette distribution sont respectivement 0 et  $2\sigma^2$ . Du plus, on remarque que la distribution est symétrique par rapport à

zéro, le coefficient d'asymétrie est nul, et le coefficient d'aplatissement est 6. Cette distribution servira à modéliser les rendements mensuels.

### 4.2.2 Loi des rendements journaliers

La loi des rendements journaliers est construite à partir de la loi des rendements mensuels. On a, de ce fait, la proposition suivante :

**Proposition 1** *Soit  $m$ , le nombre de périodes pour la composition. On considère  $m$  variables aléatoires indépendantes et identiquement distribuées telles que la loi de la somme des périodes suit une distribution de Laplace. Alors la distribution de chaque période est donnée par*

$$R_i \sim \sqrt{2}\sigma Z\sqrt{X},$$

avec  $Z \sim N(0, 1)$  et  $X$  une loi  $\Gamma(\frac{1}{m}, 1)$ ,  $Z$  et  $X$  étant des variables indépendantes. Les lois des rendements journaliers s'interprètent donc comme des mouvements browniens évalués en un temps aléatoire égal à  $2\sigma^2 X$ .

La preuve de ce résultat est faite dans l'annexe A.

À partir de cette propriété, il est maintenant possible de construire une procédure de réplication dont l'erreur est entièrement dépendante des facteurs de la réplication et de la nature du contrat.

## 4.3 Application

Le but du chapitre étant de valider la procédure de réplication et de mesurer la précision des calculs, nous ne considérons qu'une dynamique de marché relativement simple.

### 4.3.1 Le processus de prix et le contrat conditionnel

On suppose que les lois des rendements journaliers du portefeuille et de la réserve sont indépendantes et identiquement distribuées. Elles suivent la loi définie par la proposition précédente. De ce fait, les lois des rendements mensuels du portefeuille et de la réserve sont indépendantes et identiquement distribuées. Elles sont distribuées selon la loi de Laplace. Toutes ces distributions sont centrées. Ainsi, le seul paramètre à définir dans le modèle est  $\sigma$ , paramètre de la distribution de Laplace. On choisit  $\sigma = 0.3/\sqrt{12}$ .

Pour le choix de la distribution marginale du fond à répliquer, nous considérons également une distribution de Laplace de même paramètre. Quant à la dépendance existant entre le fond et le portefeuille, on choisit une copule d'indépendance. D'autre part, le notionnel des trois actifs est de 100.

Grâce à la procédure de calcul du contrat contingent exposée dans le chapitre précédent, on déduit alors le droit conditionnel suivant :

$$\tilde{g}(S_{1,T}, S_{2,T}) = 100e^{F_{3,T}^{-1}[F_{2,T}(\log(\frac{S_{2,T}}{100}))]} = S_{2,T},$$

car  $F_{3,T} = F_{2,T}$ .

On remarque donc que le contrat est indépendant du portefeuille  $S_1$ . Cet état est tout à fait cohérent avec la logique de construction. En effet, la loi conditionnelle de  $S_2$  étant donné,  $S_1$  est  $S_2$  elle-même car ces deux variables aléatoires sont indépendantes. Le problème pourrait donc être réduit à une dimension. En outre, les rendements du contrat conditionnel suivent, par construction, une loi de Laplace centrée et de paramètre  $\sigma = .30/\sqrt{12}$ . Le fond et la réserve ayant la même distribution, on peut, en utilisant l'analyse de Dybvig prévoir que la valeur du contrat soit de 100 : c'est la réplique d'un fond ayant la même distribution que l'actif utilisé pour la répliquer.

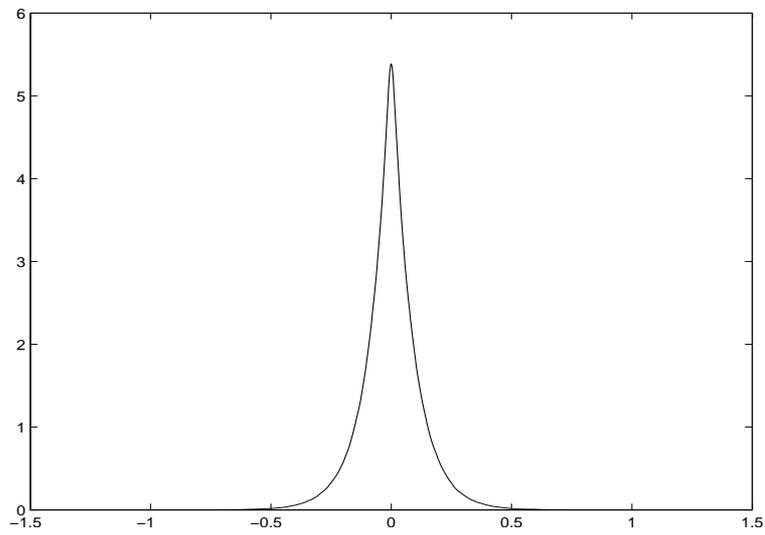
### 4.3.2 Analyse des résultats

Avant de procéder à la réplication en elle-même, il serait intéressant de déterminer une taille des simulations nécessaire à la représentativité de la distribution des rendements mensuels. Une des méthodes simples est de comparer les premiers moments obtenus analytiquement aux estimés de valeurs simulées, ces estimés étant obtenus par la méthode des moments. On pourrait également représenter les densités réelle et estimée de la distribution.

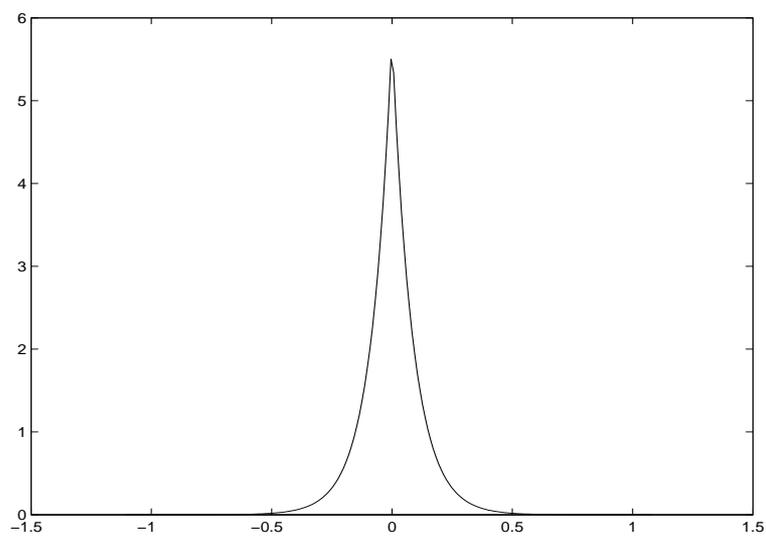
Si on considère une taille de 1000000 simulations, on obtient les estimés suivants : La moyenne empirique de l'échantillon considéré est 0.000179 pendant que la moyenne analytique est nulle. La variance empirique est de 0.014345 tandis que la variance réelle est de 0.0150. La précision des deux moments est alors de l'ordre de  $10^{-4}$ . Les représentations graphiques des densités réelles et estimée illustrent également cette représentativité.

Cette analyse nous permet de déterminer la dimension du calcul nécessaire (taille de la discrétisation des prix et nombre de simulations) pour réduire considérablement l'erreur liée à la représentativité du calcul. Ainsi, l'idéal serait donc d'avoir, à chaque période, 1000000 scénarios de prix pour la période future. Nous choisissons donc d'avoir une discrétisation uniforme et fixe à chaque période de 100 prix. Le nombre de simulations par point de la grille sera donc de 10000. Les bornes supérieures et inférieures des discrétisations à chaque périodes sont fixes et sont respectivement de  $100.0e^{15.0\sigma}$  et  $100.0e^{-15.0*\sigma}$ . Ces bornes sont construites en tenant compte du comportement de la distribution mensuelle des rendements. En effet, la probabilité d'avoir un rendement supérieur à  $10\sigma$  ou inférieur à  $-10\sigma$  étant très faible ( $\ll 1\%$ ). Le nombre de périodes de réplication sur une période d'un mois est supposé fixe et est égal a 22.

Tout d'abord, on remarque que le "KP-Measure", c'est-à-dire la valeur du contrat



(a) Densité estimée à partir d'un échantillon de 1000000 de données



(b) Densité réelle

conditionnel est très proche de 100. Conformément à la remarque faite au paragraphe précédente, le "KP-Measure" est bien celui à quoi nous nous attendions. La sensibilité du portefeuille de réplication par rapport à la réserve est très proche de 1, ce qui confirme également le raisonnement analytique.

Passons à l'analyse de l'efficacité de la procédure de réplication. Pour cela, on considère 1000000 de scénarios de comportement du portefeuille du début à la fin du mois. Ce nombre de trajectoires possibles du portefeuille est en accord avec la taille des simulations utilisées dans la construction de la procédure de réplication. Ensuite, on compare les comportements des portefeuilles de réplication à la date terminale et la valeur du contrat contingent pour chacune des trajectoires. La moyenne des erreurs de réplication est presque nulle et le carré des erreurs, c'est-à-dire la fonction objectif de l'optimisation, est  $2 \times 10^{-10}$ . La faiblesse de ces erreurs traduit l'efficacité de la procédure de réplication. L'efficacité de la réplication est perceptible sur les densités estimées des rendements du portefeuille de réplication et du contrat contingent. Les statistiques de comparaison du portefeuille de réplication et du contrat contingent sont présentées dans le tableau ci-dessous.

	Contrat contingent	Portefeuille de réplication
Moyenne	0.000179	0.000179
Variance	0.014345	0.014345
Coefficient d'asymétrie	0.011142	0.011142
Coefficient d'aplatissement	3.114004	3.114004

TAB. 4.1 – Statistiques de la réplication : cas de la distribution de Laplace

## 4.4 Conclusion

Dans tout procédure numérique d'évaluation, l'enjeu le plus important est le contrôle de l'importance des facteurs utilisés dans la procédure. Ainsi, dans le cas

---

la réplique des fonds de couverture et de la construction des fonds synthétiques, les analyses faites dans ce chapitre, ont permis d'une part de valider d'une manière pratique la procédure de réplique en l'appliquant à un modèle dont les propriétés des distributions sont vérifiées, et d'autre part de vérifier l'importance de la discrétisation et de la taille des simulations nécessaires à son utilisation. Il est alors proposé une procédure relativement simple destinée à l'utilisation efficace de la méthode de réplique, laquelle sera appliquée dans un modèle réel de marché dans le chapitre suivant.

# Chapitre 5

## Application : Approche par les mélanges de lois normales

### 5.1 Introduction

La méthode d'analyse des investissements présentée dans le chapitre 3 permet à la fois d'analyser les performances des investissements existants, mais également de construire des stratégies d'investissement. D'autre part, le chapitre précédant a permis de valider la procédure de tarification et de couverture utilisée dans ce mémoire. Il serait donc intéressant de l'appliquer dans un marché plus réaliste.

Dans ce chapitre, à partir de rendements de deux actifs (Portefeuille et Réserve), nous allons construire un modèle de mélanges de lois normales. Ensuite, nous construirons des fonds synthétiques à partir de l'estimation des quatre premiers moments d'un échantillon de données en utilisant la famille de distribution de Johnson-SU. Enfin, nous analyserons l'impact de la dépendance entre le fond et le portefeuille en utilisant deux familles de dépendance pour la modéliser : la famille des copules normales et celle des copules de Clayton.

## 5.2 Mélanges de lois normales

### 5.2.1 Définition

Une mixture de lois normales est définie de la manière suivante : considérons  $m$  distributions indépendantes de lois normales ayant au moins un des paramètres différents (matrice de variance/covariance et/ou moyennes). Une mixture de loi normale est un choix aléatoire indépendant de ces lois.

La densité d'une mixture de  $m$  lois normales est alors donnée par

$$f_z(x) = \sum_{i=1}^m p_i \phi_d(x; \mu_i, \Sigma_i),$$

où  $\forall i = 1, \dots, m, p_i > 0, \sum_{k=1}^m p_k = 1$ , et  $\phi_d$  est la densité d'une loi normale, c'est-à-dire

$$\phi_d(x; \mu_i, \Sigma_i) = \frac{\exp^{-\frac{1}{2}(x-\mu_i)^\top \Sigma_i^{-1}(x-\mu_i)}}{(2\pi)^{\frac{n}{2}} |\Sigma_i|^{-\frac{1}{2}}}.$$

En dimension 1, le nombre de paramètres d'une mixture de  $m$  lois normales est de  $3m - 1$ , lorsque tous les paramètres des composantes du mélange sont différents. En dimension 2, ce nombre est de  $6m - 1$ .

Les mélanges de lois normales permettent de caractériser les changements aléatoires de régimes de distributions de lois normales. Dans cette optique, elles pourraient caractériser les changements non connus des investisseurs, de stratégie dans la gestion de leurs portefeuilles. Elles permettent alors une certaine flexibilité dans l'analyse des modèles de données.

Dans le cas univarié, la fonction de répartition du mélange de  $m$  variables aléatoire

normales de paramètres  $\mu_k, \sigma_k^2, k = 1, \dots, m$ , est donnée par

$$F(x) = \sum_{k=1}^m \pi_k \Phi\left(\frac{x - \mu_k}{\sigma_k}\right),$$

où  $\pi_1, \dots, \pi_m$  sont les proportions du mélange, et  $\Phi$  est la fonction de répartition de la normale standard.

L'espérance du mélange loi est

$$\tilde{\mu} = \sum_{k=1}^m \pi_k \mu_k$$

et sa variance est

$$\tilde{\sigma}^2 = \sum_{k=1}^m \pi_k \sigma_k^2 + \sum_{k=1}^m \pi_k \mu_k^2 - \tilde{\mu}^2.$$

En outre, il est à noter que lorsque le nombre de composantes du mélange de lois normales est supérieure à 1, il est impossible d'obtenir une approximation analytique de l'inverse de la fonction de répartition de la loi normale. La méthode de Newton est alors utilisée pour déterminer le quantile. Cependant, l'efficacité de la méthode reside dans le choix des points de départ de la méthode. Une propriété importante des mélanges, permet de déterminer des points de départ. En effet, pour tout  $\alpha \in (0, 1)$ , on a

$$F^{-1}(\alpha) \geq \min_{k=1, \dots, m} \{\mu_k + \sigma_k \Phi^{-1}(\alpha)\}$$

et

$$F^{-1}(\alpha) \leq \max_{k=1, \dots, m} \{\mu_k + \sigma_k \Phi^{-1}(\alpha)\}.$$

Une autre propriété intéressante et essentielle dans notre étude est la loi de la somme. En effet la somme de mélanges de lois normales indépendantes et identiquement distribuées est une mixture de lois normales. Cependant, lorsque l'on fait une somme relativement important des mélanges de lois normales, le nombre de com-

posantes de la somme augmente vertigineusement. Cette somme est difficilement manipulable dans la pratique. Une solution alternative consisterait à retrouver par estimation (nous y reviendrons) les paramètres d'une mixture de lois normales ayant un nombre plus faible de composantes.

Il est à noter également que la convergence de la somme des mélanges de lois normales, par le théorème central limite est relativement rapide. En finance, cette propriété correspond au comportement des actifs. La distribution journalière des actifs est difficilement gaussienne. Cependant, leurs distributions mensuelles sont plus proches de la loi normale.

Intéressons-nous maintenant aux mélanges de lois normales en dimension 2.

Soit  $X = (X^{(1)}, X^{(2)})$ , une mixture de lois normales bivariées. On pose  $\beta_k = \rho_k \frac{\sigma_{k2}}{\sigma_{k1}}$  et  $\alpha_k = \mu_{k2} - \beta_k \mu_{k1}$ ,  $k = 1, \dots, m$ .

La distribution conditionnelle de  $X^{(2)}$  sachant  $X^{(1)} = x_1$  est également une mixture de normale à  $m$  régimes de paramètres  $\{\tilde{\pi}_k\}_{k=1}^m$ ,  $\{\tilde{\mu}_k\}_{k=1}^m$ ,  $\{\tilde{\sigma}_k^2\}_{k=1}^m$  tels que

$$\tilde{\pi}_k(x_1) = \frac{\pi_k \phi(x_1; \mu_{k1}, \sigma_{k2}^2)}{\sum_{j=1}^m \pi_j \phi(x_1; \mu_{j1}, \sigma_{j1}^2)},$$

$$\tilde{\mu}_k(x_1) = \alpha_k + \beta_k x_1,$$

et

$$\tilde{\sigma}_k^2 = \sigma_k^2 (1 - \rho_k^2).$$

Cette propriété est essentielle dans l'évaluation du contrat conditionnel.

### 5.2.2 Estimations des paramètres

L'estimation des paramètres du modèle par des mélanges de lois normales utilise une procédure développée pour la première fois par Dempster et al. (1977). Il s'agit d'une méthode récursive permettant de déterminer les estimateurs du maximum de vraisemblance des paramètres du mélange de loi normale. Elle est décrite, en dimension 2, dans ce qui suit.

On considère une série de rendements journaliers et mensuels de deux actifs, celui du portefeuille et de la réserve. Les données en dimension 2 sont notées  $x_1, \dots, x_n$ .

On commence par un estimateur initial  $\theta^{(0)}$ . Pour un estimateur  $\theta^{(l)} = (\pi^{(l)}, \mu^{(l)}, \Sigma^{(l)})$  du paramètre  $\theta = (\pi, \mu, \Sigma)$ , on pose

$$\pi_k(x_i; \theta^{(l)}) = \frac{\pi_k^{(l)} \phi_2(x_i; \mu_k^{(l)}, \Sigma_k^{(l)})}{\sum_{j=1}^m \pi_j \phi_2(x_i; \mu_j^{(l)}, \Sigma_j^{(l)})}, \quad i = 1, \dots, n.$$

On définit alors un nouvel estimateur  $\theta^{(l+1)} = (\pi^{(l+1)}, \mu^{(l+1)}, \Sigma^{(l+1)})$ , de  $\theta$  de la manière suivante :

$$\pi_k^{(l+1)} = \frac{1}{n} \sum_{i=1}^n \pi_k(x_i, \theta^{(l)}),$$

$$\mu_k^{(l+1)} = \frac{1}{n} \sum_{i=1}^n \frac{x_i \pi_k(x_i, \theta^{(l)})}{\pi_k^{(l+1)}},$$

et

$$\Sigma_k^{(l+1)} = \frac{1}{n} \sum_{i=1}^n \frac{(x_i - \mu_k^{(l+1)})(x_i - \mu_k^{(l+1)})^\top \pi_k(x_i, \theta^{(l)})}{\pi_k^{(l+1)}}.$$

Lorsque  $l$  devient de plus en plus grand,  $\{\pi_k(x_i; \theta^{(l)}); k = 1, \dots, m, i = 1, \dots, n\}$  se stabilise. Ainsi, les estimateurs convergent vers une valeur.

Le test d'hypothèse validant la famille de distribution est adopté de Genest and Rémillard (2008). Ce test permet non seulement d'accepter ou de refuser l'hypothèse

nulle pour les mélanges, mais également de déterminer le nombre de ses composantes.

### 5.2.3 Résultats

On considère des séries de rendements d'un portefeuille et d'une réserve. Le test d'hypothèse a permis de conclure qu'il y a 4 composantes dans la mixture, pour les distributions journalières et 1 composante pour les rendements mensuels. Deux remarques peuvent être faites concernant les estimés de la loi mensuelle. La première est la convergence rapide vers la loi normale de la somme des mélanges de lois journalières (le théorème central limite). Cela pourrait s'expliquer par la nature même des distributions journalières : elles sont construites à partir de la loi normale. La seconde remarque, purement financière est basée sur loi des observations empiriques. En effet, il est généralement difficile de valider l'hypothèse de normalité pour les rendements des actifs traditionnels, transigés quotidiennement. En revanche, sa validation se fait pour plusieurs d'entre eux pour les rendements mensuels.

Les estimations des paramètres des rendements journaliers et mensuels sont présentés dans le tableau ci-dessous.

	p	$\mu_1$	$\mu_2$	$\sigma_1$	$\sigma_2$	$\rho$
Composante 1	0.03	0.000222	0.000222	0.000157	0.000157	0.9
Composante 2	0.42	0.000363	0.000351	0.014108	0.002929	0.85
Composante 3	0.10	-0.000689	-0.000573	0.010941	0.005072	0.71
Composante 4	0.45	0.0006110	0.000491	0.004167	0.0028044	0.51

TAB. 5.1 – Estimation des paramètres du modèle des rendements journaliers

Après avoir construit le modèle de marché, nous allons créer des fonds synthétiques dans la section suivante.

	p	$\mu_1$	$\mu_2$	$\sigma_1$	$\sigma_2$	$\rho$
Composante 1	1	0.007819	0.0068086	0.024531	0.014108	0.70

TAB. 5.2 – Estimation des paramètres du modèle des rendements mensuels

## 5.3 Construction de fonds synthétiques

La construction de fonds synthétiques est l'une des innovations importantes qu'a introduit les méthodes complexes de réplication des fonds de couverture. Cette logique permet de construire un modèle d'investissement ou de couverture à partir d'actifs classiques. On élabore ainsi des stratégies à moindre coût, permettant de produire une distribution de rendements particulière.

En théorie, toutes les familles de distributions peuvent être reproduites par la méthode présente dans ce mémoire. Cependant, dans une logique de l'investissement, seulement quelques unes sont utilisées pour modéliser les rendements des fonds. En ce qui concerne les investissements alternatifs, les analyses empiriques ont montré que l'enjeu ne se situe pas seulement aux deux premiers moments (moyenne et variance). Les moments d'ordres supérieurs, coefficients d'asymétrie et d'aplatissement, jouent un rôle important dans la décision d'investissement. Dans cette optique, la sélection d'une famille de distributions présentant des flexibilités au niveau de l'asymétrie et l'aplatissement pourrait être intéressante à étudier. Le système de distributions de Johnson offre cette possibilité.

### 5.3.1 Système de distributions de Johnson

Le système de Elderton and Johnson (1969) est une famille de distributions dont la transformation par une fonction monotone produit une distribution normale.

En effet, soit  $Z$ , une variable aléatoire normale standard. Une variable aléatoire  $X$  est une distribution de Johnson s'il existe une fonction monotone  $g$  telle que

$$Z = \gamma + \delta g\left(\frac{X - \xi}{\lambda}\right)$$

Les paramètres  $\gamma$  et  $\delta$  sont les paramètres de tendance. Le paramètre  $\lambda$  est un paramètre d'échelle, et  $\xi$  est un paramètre de position.

Il existe trois types de transformations possibles dans la famille de distributions de Johnson :

- la distribution lognormale :  $g(y) = \log(y)$
- la distribution de Johnson SU :  $g(y) = \log(y + \sqrt{y^2 + 1}) = \sinh^{-1}(y)$
- la distribution de Johnson SB :  $g(y) = \log\left(\frac{y}{1-y}\right)$

La distribution de Johnson-SU est la distribution la plus intéressante car son support est l'ensemble des nombres réels. De plus, elle offre, de par le nombre de paramètres la caractérisant, une large classe d'asymétries et d'aplatissements. Cette distribution pourrait donc modéliser un nombre important de produits financiers.

La densité  $p_X$  de la distribution de Johnson SU se déduit aisément de celle d'une loi normale. En effet, on a

$$p_X(x) = \frac{\delta}{\lambda} f'(\gamma + \delta f(x)) p_Z\left(\frac{x - \xi}{\lambda}\right),$$

où  $p_Z(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}z^2\right)$ .

Les quatre premiers moments de la variable aléatoire centrée déduite de la distributions de Johnson SU  $Y = \frac{X - \xi}{\lambda}$  sont :

$$\mu_1(Y) = \omega^{\frac{1}{2}} \sinh\theta,$$

$$\mu_2(Y) = \frac{1}{2}(\omega - 1)(\omega \cosh(2\theta) + 1),$$

$$\mu_3(Y) = -\frac{1}{4}\omega^{\frac{1}{2}}(\omega - 1)^2(\omega(\omega + 2)\sinh(3\theta) + 3\sinh\theta),$$

$$\mu_4(Y) = -\frac{1}{8}(\omega - 1)^2(\omega^2(\omega^4 + 2\omega^3 + 3\omega^2 - 3))\cosh(4\theta) + 4\omega^2(\omega + 1)\cosh(2\theta) + 3(2\omega + 1),$$

où  $\omega = \exp(\delta^{-2})$  et  $\theta = \gamma/\delta$ .

On remarque donc que le troisième et le quatrième moment dépendent des paramètres  $\gamma$  et  $\delta$ . Dans le contexte financier, cette flexibilité peut être intéressante car cette distribution pourrait être utilisée pour modéliser les rendements asymétriques ou des queues de distribution épaisses.

La fonction de répartition et le quantile se déduisent facilement de ceux de la loi normale. Cette famille de distribution sera celle utilisée pour la création de fond synthétiques.

### 5.3.2 Construction de la distribution de Johnson SU

Le principal handicap de la famille de distribution de Johnson SU est l'interprétation financière que l'on pourrait faire de ses paramètres. En effet, ces paramètres n'ont pas de relations simples avec les statistiques empiriques des données. Cependant, comme l'ont montré Elderton and Johnson (1969), il existe une relation bijective entre les quatre premiers moments de la distribution et ces paramètres. De plus, ils montrent que seuls les paramètres  $\gamma$  et  $\delta$  sont liés à l'asymétrie et à l'aplatissement de la distribution. Une procédure récursive, développée dans l'annexe B, permet ainsi de déterminer les paramètres à partir des quatre premiers moments de la distribution. Cette méthode est très intéressante car elle permet de construire des distributions à partir des quatre premiers moments. Elle permet ainsi de mesurer l'impact de l'asymétrie et de l'aplatissement sur une décision d'investissement.

## 5.4 Sélection de la dépendance entre le fond et le portefeuille

Les fonds de couverture sont, en général, faiblement corrélés au marché. La diversification est due au comportement de la dépendance existant entre le portefeuille et le fond. Il serait donc intéressant d'analyser cette dépendance. Dans cette partie, nous analyserons la dépendance traditionnellement utilisée dans le marché c'est-à-dire la dépendance gaussienne (quelque fois de manière inappropriée). Nous étudierons également une dépendance, plus particulière, mais constatée sur le marché des actions, qui relie les scénarios catastrophes des actifs : la dépendance de Clayton.

### 5.4.1 Copule Gaussienne

La corrélation entre les actifs est un outil utilisé pour caractériser la relation existant entre ceux-ci. Cependant cette mesure n'a de validité que lorsque les actifs sont normaux, ce qui n'est pas toujours le cas. La copule normale est la famille de dépendance validant la corrélation dans le contexte de la normalité des actifs.

La copule normale est la famille de copule extraite de loi normales bivariées. Elle est caractérisée par un paramètre  $\rho$  et se définit de la manière suivante :

$$C_\rho(u_1, u_2) = \Phi_2[\Phi^{-1}(u_1), \Phi^{-1}(u_2); \rho],$$

où  $(u_1, u_2) \in [0, 1]^2$ ,  $\Phi_2$  et  $\Phi_1$  étant respectivement les fonctions de répartition bidimensionnelle et unidimensionnelle de la loi normale standard.

La distribution conditionnelle  $C_{2|1}$  se déduit par la formule suivante :

$$C_{2|1} = \Phi \left( \frac{\Phi^{-1}(u_2) - \rho\Phi^{-1}(u_1)}{\sqrt{1 - \rho^2}} \right).$$

Enfin, le quantile d'ordre  $\alpha$  de  $C_{2|1}$  est

$$\mathcal{Q}(u_1, \alpha) = \Phi[\sqrt{1 - \rho^2}\Phi^{-1}(\alpha) + \rho\Phi^{-1}(u_1)].$$

Quant à l'estimation du paramètre  $\rho$  de la copule, en utilisant la méthode semi-paramétrique développée par Genest et al. (1995), le pseudo-maximum de vraisemblance permet de montrer que le coefficient de van der Waerden (voir annexe C) est un estimateur du paramètre de la copule normale.

### 5.4.2 Copule de Clayton

La principale difficulté de la copule normale est son incapacité à caractériser une dépendance des valeurs extrêmes. La dépendance de queue de distribution est d'une grande importance car elle permet de caractériser l'impact que peut avoir la chute, où la montée vertigineuse d'un actif sur un autre. Deux mesures,  $\lambda_L$  et  $\lambda_U$  permettent de caractériser ces liens :

$$\lambda_L = \lim_{u \rightarrow 0^+} P[U_2 < u | U_1 < u]$$

et

$$\lambda_U = \lim_{u \rightarrow 1^-} P[U_2 > u | U_1 > u].$$

Dans le cas de la copule normale, ces deux valeurs sont nulles. C'est de là qu'apparaît l'intérêt de l'utilisation de la copule de Clayton.

La copule de Clayton en dimension 2 est définie de la manière suivante : pour tout  $\theta \geq -1$

$$C_\theta(u_1, u_2) = \{(u_1^{-\theta} + u_2^{-\theta} - 1)^+\}^{-1/\theta}.$$

Lorsque  $\theta > 0$ , on a

$$C_\theta(u_1, u_2) = (u_1^{-\theta} + u_2^{-\theta} - 1)^{-1/\theta}.$$

Nous restreignons l'étude à ce cas seulement.

La distribution conditionnelle de la copule est

$$C_{2|1}(u_1, u_2) = u_1^{-(\theta+1)} [u_1^{-\theta} + u_2^{-\theta} - 1]^{-\frac{\theta+1}{\theta}}$$

et le quantile conditionnel à la première variable est

$$\mathcal{Q}(u_1, \alpha) = [\alpha^{-\frac{\theta}{\theta+1}} u_1^\theta - u_1^{-\theta} + 1]^{-\frac{1}{\theta}}.$$

Lorsque  $\theta$  tend vers 0, la copule de Clayton tend vers la copule d'indépendance.

Contrairement à la copule normale, la copule de Clayton présente une dépendance non nulle pour les scénarios catastrophes ( $\lambda_L = 2^{-\frac{1}{\theta}}$ ). La dépendance extrême supérieure est quant à elle nulle. Ce comportement est présent dans le marché est actions.

Finalement, l'estimation du paramètre de la copule de Clayton est déduite de l'estimation du  $\tau$  (tau) de Kendall (voir annexe C). Cela est dû au fait que le  $\tau$  de Kendall d'une copule de Clayton de paramètre  $\theta$  est  $\tau = \frac{\theta}{\theta+2}$ . Ainsi, à partir d'un estimateur de  $\tau$ , on peut déduire un estimateur du paramètre  $\theta$ .

Il existe d'autres familles de copules caractérisant une dépendance de valeurs extrêmes.

## 5.5 Analyse des résultats

Cette section est essentiellement destinée à la construction de fonds synthétiques dont la distribution marginale suit une loi de Johnson-SU. On étudie également quelques formes de dépendance que pourraient présenter ces fonds.

### 5.5.1 Construction de fonds synthétiques

On considère un fond synthétique dont la loi mensuelle des rendements suit une loi de Johnson-SU dont les quatres premiers moments sont :

Espérance : 10.89% (annualisée)

Variance : 1.34% (annualisée)

Coefficient d'asymétrie :  $-0.64$

Coefficient d'aplatissement : 6.6382

Ces quatres moments correspondent aux quatres premiers moments d'un fonds de couverture réel.

Ce fonds de couverture présente une certaine particularité. La variance annuelle est très faible, ce qui est plutôt intéressant pour un investissement, lorsque l'analyse se limite à la moyenne et la variance. Cependant, on remarque que les moments d'ordre supérieurs ne sont pas intéressants. La distribution a une asymétrie fortement négative et le coefficient d'aplatissement est assez élevé. Ce fonds de couverture permet de confronter les deux premiers moments aux moments d'ordre supérieurs.

En utilisant la méthodes récursive de détermination des paramètres de la distribution de Johnson-SU, on déduit les paramètres suivants :  $\gamma = 0.3844334233$ ,  $\delta = 1.6054826198$ ,  $\xi = 0.0214691141$  et  $\lambda = 0.0422218865$ .

On considère un investissement initial de 100 pour les trois actifs. Dans un premier temps, on considère un fond indépendant du portefeuille  $P$ . La technique de

réplication évalue le contrat à 100.1449861774. En effet, pour la même distribution de rendements, la réplication coûte plus cher que le fonds de couverture qui, quant à lui, ne coûte que 100. Les statistiques du portefeuille de réplication et de contrat contingent sont présentées dans le tableau ci-dessous.

	Contrat contingent	Portefeuille de réplication
Moyenne	0.008685	0.008658
Variance	0.001397	0.001392
Coefficient d'asymétrie	-0.842305	-0.745763
Coefficient d'aplatissement	4.673280	3.779975

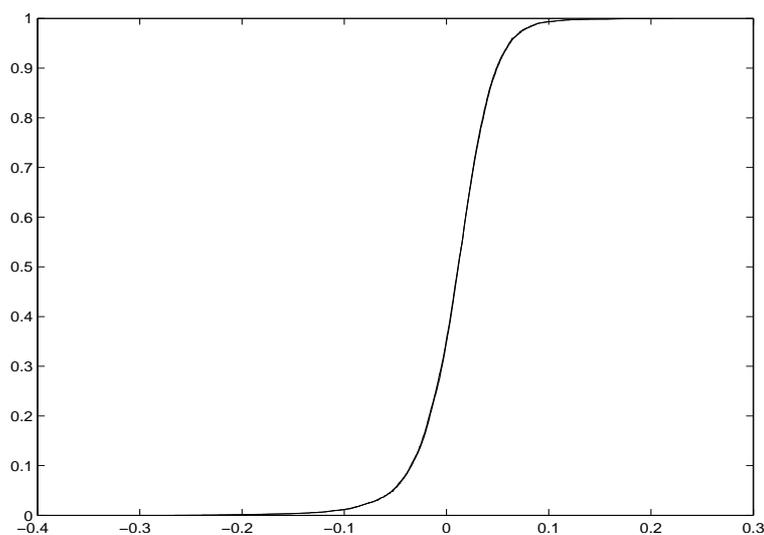
TAB. 5.3 – Statistiques de la réplication : marge Johnson-SU, copule d'indépendance

Il est donc théoriquement sous optimal de construire un tel fond. On remarque tout de même que le rendement du fond synthétique reproduit approximativement la distribution désirée comme l'illustre les graphiques ci-dessous.

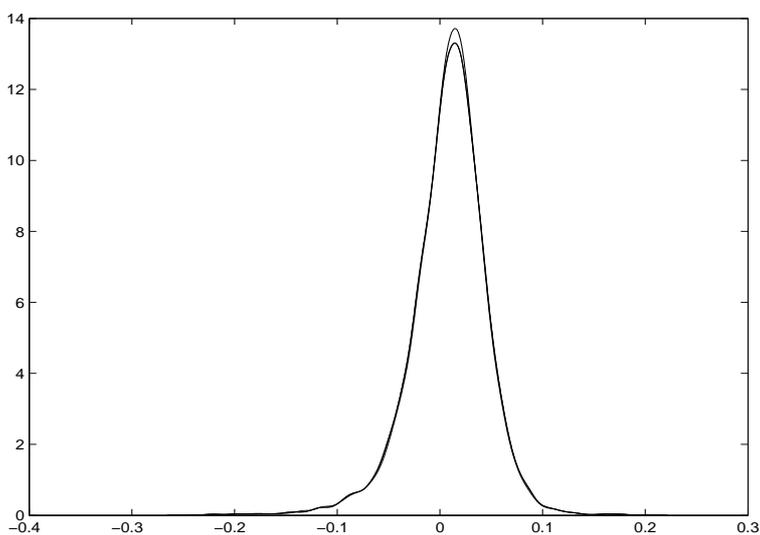
Soit maintenant une dépendance gaussienne entre le portefeuille  $P$  et le fond  $F$ , de paramètre  $\rho = 0.644$ . Le paramètre  $\rho$  correspond au coefficient de van der Waerden d'un fond de couverture d'un portefeuille. Le "KP Measure" est alors de 99.9280260401,. Il est donc optimal de construire un tel fonds synthétique car, pour une même distribution, celui-ci coûte moins cher que le fonds de couverture. Il en est de même lorsque l'on utilise une dépendance de Clayton de paramètre  $\theta = 0.5$ . Le "KP-Measure" est de 99.9259707409. Les statistiques des contrats contingents et des portefeuilles de réplication sont présentées dans les tableaux ci-dessous.

	Contrat contingent	Portefeuille de réplication
Moyenne	0.009020	0.008877
Variance	0.001207	0.001215
Coefficient d'asymétrie	-0.828582	-0.723659
Coefficient d'aplatissement	3.938718	3.243005

TAB. 5.4 – Statistiques de la réplication : marge Johnson-SU, copule gaussienne



(c) Fonction de répartition du contrat et du portefeuille de répliation



(d) Densité du contrat et du portefeuille de répliation

	Contrat contingent	Portefeuille de répliation
Moyenne	0.008254	0.008069
Variance	0.001315	0.001318
Coefficient d'asymétrie	-0.764510	-0.664308
Coefficient d'aplatissement	4.008973	3.215780

TAB. 5.5 – Statistiques de la répliation : marge Johnson-SU, copule de Clayton

Le tableau ci-dessous présente un tableau comparatif de quelques caractéristiques de la stratégie de replication. La composition initiale du portefeuille de replication valide l'intuition financière. Celle-ci suggère une position initiale longue sur l'actif de réserve R et courte sur le portefeuille P.

	Indépendante	Gaussienne	Clayton
KP Measure	100.144986	99.928026	99.925970
Rendement(%)	-0.144881	0.071999	0.074056
Position sur P	-1.410496	-0.581603	-1.002321
Position sur R	3.462202	3.046562	3.400822
Moyenne des erreurs	-0.002926	-0.013843	-0.018354
Moyenne des carrés des erreurs	0.237291	0.202965	0.301163

TAB. 5.6 – Tableau comparatif des performances

En construisant ces fonds, on remarque une certaine influence de la nature de la dépendance entre le fond et le portefeuille. Nous étudierons cette influence dans la section suivante.

### 5.5.2 Sensibilité à la dépendance

La dépendance est importante aussi bien dans l'analyse de la performance des fonds de couverture, que dans la décision de création de fonds synthétiques. Dans cette section, nous étudions le comportement du "KP-Measure", par rapport au paramètre de chacune des deux familles de copule (Gaussienne et Clayton). Les caractéristiques initiales de la distribution marginale du fond sont conservées.

Considérons, tout d'abord le cas de la famille des copules normales. Le tableau ci-après présente le comportement de la replication en fonction du paramètre de la copule. "MHE" et "SHE" représentent les moyennes et carrés des erreurs de replication.  $\Delta P$  et  $\Delta R$  représentent les positions initiales respectives sur les actifs P et R.

$\rho$	KP-Measure	Performance	Delta P	Delta R	MHE	SHE
-1.000000	101.000978	-0.009960	-1.332390	0.001945	-0.004595	0.104135
-0.900000	100.770242	-0.007673	-1.799145	1.447177	-0.005681	0.109843
-0.800000	100.655854	-0.006537	-1.893214	1.999485	-0.009874	0.110586
-0.700000	100.564938	-0.005633	-1.916834	2.382592	-0.011980	0.110905
-0.600000	100.487494	-0.004863	-1.905526	2.680473	-0.012719	0.110781
-0.500000	100.414954	-0.004141	-1.864179	2.907288	-0.016245	0.110914
-0.400000	100.352794	-0.003522	-1.798782	3.080208	-0.015389	0.110192
-0.300000	100.292206	-0.002918	-1.714524	3.208060	-0.017290	0.109551
-0.200000	100.237357	-0.002371	-1.621314	3.310301	-0.017097	0.108435
-0.100000	100.185492	-0.001853	-1.506970	3.367496	-0.016870	0.106556
0.000000	100.135060	-0.001350	-1.378934	3.389103	-0.017905	0.104917
0.100000	100.086851	-0.000868	-1.238276	3.385586	-0.019046	0.103091
0.200000	100.040708	-0.000407	-1.078765	3.330846	-0.020493	0.101130
0.300000	99.999222	0.000008	-0.908424	3.253924	-0.018853	0.098869
0.400000	99.960327	0.000397	-0.717296	3.125895	-0.017245	0.097264
0.500000	99.922997	0.000770	-0.509372	2.962399	-0.015961	0.096893
0.600000	99.888231	0.001118	-0.279293	2.737035	-0.015035	0.098582
0.700000	99.858939	0.001412	-0.019503	2.442971	-0.012585	0.102922
0.800000	99.836152	0.001640	0.275637	2.057925	-0.009673	0.111319
0.900000	99.825814	0.001743	0.645951	1.499834	-0.005883	0.125447
1.000000	99.888799	0.001113	1.389128	0.006340	-0.004660	0.158230

TAB. 5.7 – Performance de la réplication en fonction du paramètre de la copule gaussienne

Le fond devient avantageux lorsque  $\rho \geq 0.3$ . L'analyse du rendement de l'investissement permet de constater qu'il a une croissance plus importante lorsque la dépendance est de moins en moins négative puis devient plus stable lorsque  $\rho$  est proche de 1.

Dans le cas la famille de copule de Clayton, l'investissement est sous optimal pour des valeurs de  $\theta$  très proches de 0 (voir Tableau 5.8). Ce constat est tout à fait cohérent à la nature même de cette dépendance : lorsque  $\theta$  est proche de 0 par valeurs positives (la copule est proche la copule d'indépendance), la dépendance des queues de distribution est faible. Lorsque cette valeur est de plus en plus importante, cette dépendance devient importante. Dans une logique d'analyse de la performance des fonds de couverture, ce dernier devient non optimal lorsque la dépendance des pertes extrêmes est élevée.

$\theta$	KP Measure	Performance	Delta P	Delta R	MHE	SHE
0.100000	100.063386	-0.000634	-1.250585	3.400907	-0.019399	0.119462
0.200000	99.992739	0.000073	-1.123022	3.387535	-0.022459	0.157977
0.300000	99.936104	0.000639	-0.990935	3.330580	-0.018137	0.198336
0.400000	99.883316	0.001168	-0.876063	3.273993	-0.018927	0.233877
0.500000	99.843662	0.001565	-0.761649	3.187475	-0.013148	0.263269
0.600000	99.806956	0.001932	-0.651647	3.104181	-0.011912	0.287009
0.700000	99.772314	0.002279	-0.560992	3.025703	-0.013817	0.306897
0.800000	99.745806	0.002545	-0.478392	2.946978	-0.013108	0.323271
0.900000	99.721414	0.002790	-0.395909	2.857275	-0.014550	0.336199
1.000000	99.708799	0.002916	-0.324189	2.773295	-0.008789	0.346484

TAB. 5.8 – Performance de la réplication en fonction du paramètre de la copule de Clayton

## 5.6 Conclusion du chapitre

De ce chapitre, il ressort que la procédure de réplication permet de construire des stratégies d'investissement à partir d'un ou deux actifs. Il est donc possible de

construire une distribution de rendements correspondant aux attentes de l'investisseur. La décision d'investissement dans ces fonds dépend de la valeur du "KP-Measure". Toutefois, cette mesure peut dépendre fortement du comportement de la dépendance entre le fond et le portefeuille.

# Chapitre 6

## Conclusion générale

L'approche révolutionnaire de Dybvig (1988a), puis de Kat and Palaro (2005) a permis de construire un nouveau cadre d'évaluation de la performance des investissements. Le présent mémoire utilise une méthode rigoureuse proposée par Papa-georgiou et al. (2007) permettant d'élaborer une stratégie optimale de répliation, minimisant le carré des erreurs de la procédure. Elle ne nécessite pas un changement de mesure de probabilité.

Cette procédure a ainsi été validée dans la pratique en utilisant une distribution de valeurs extrêmes. Cette étude a permis de prouver l'efficacité de la procédure de répliation elle-même. Le carré des erreurs, dans ce contexte, est quasiment nul.

Fort de la validation de la procédure, il a été élaboré un modèle de marché réaliste, construit à partir de rendements de deux actifs. Des contrats synthétiques ont ainsi été créés pour caractériser certaines particularités des fonds de couverture. De cette étude, il ressort que le comportement de la dépendance influence fortement la structure du portefeuille de répliation. La décision d'investissement pour des fins de diversification devrait tenir compte de la structure de la dépendance. Le carré des erreurs de répliation est de l'ordre de 0.20% pour la classe de fonds synthétiques

qui a été étudiée.

L'erreur de réplication est due principalement à trois facteurs qui ne sont pas forcément indépendants : les incohérences du modèle de marché, la non linéarité des contrats conditionnels et la procédure de réplication elle-même (qui inclut l'erreur due à la représentativité des distributions). Des études effectuées précédemment, il ressort que l'erreur de la procédure elle-même est négligeable. Ainsi, le principal enjeu dans la procédure de réplication est le contrôle l'erreur intrinsèque au modèle et celle due à la non linéarité des contrats conditionnels. Des études futures pourraient permettre de mieux cerner l'erreur liée à la nature des contrats. Cette erreur constituerait une règle de décision importante dans la sélection des familles de dépendance et des lois marginales.

# Bibliographie

- Agarwal, V. and Naik, N. (2003). Performance evaluation of hedge funds with option based and buy-and-hold strategies. *Review of Financial Studies*.
- Ané, T. and Geman, H. (2000). Order flow, transaction clock, and normality of asset returns. *Journal of Finance*, 55 :2259–2284.
- Berger, J. and Mandelbrot, B. (1963). A new model for error clustering in telephone circuits. *IBM Journal*, pages 224–236.
- Berrada, T., Dupuis, D., Jacquier, E., Papageorgiou, N., and Rémillard, B. (2005). Credit migration and derivatives pricing using copulas. *Journal of Computational Finance*, 10 :43–68.
- Bochner, S. (1955). Harmonic Analysis and the Theory of Probability. *University of California Press*.
- Brown, S., Goetzman, N., and Liang, B. (2004). Fees on fees on fund of funds. Technical report, Yale International Center for Finance.
- Charlier, C. and Wicksell, S. (1924). On the dissection of frequency functions. *Arkiv for Matematik Astronomi och Fysik*.
- Cherubini, U. and Luciano, E. (2002). Bivariate option pricing with copulas. *Applied Mathematical Finance*, 9 :69–85.

- Clark, P. (1973). A subordinated stochastic process with finite variance for speculative prices. *Econometrica*, pages 135–156.
- Dave, R., Kat, H., and Lu, S. (2004). Fund of Hedge Funds Portfolio Selection : A Multiple objective approach. *International Securities Market Association Centre*.
- Deheuvels, P. (1979). La fonction de dépendance empirique et ses propriétés : Un test non paramétrique d'indépendance. *Acad. Royal Bel., Bull. Class. Sci., 5e série*, 65 :274–292.
- Dempster, A., Lair, N., and Rubin, D. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, 39 :1–38.
- Dybvig, P. (1988a). Distributional Analysis of Portfolio Choice. *Journal of Business*, 61 :369–393.
- Dybvig, P. (1988b). Inefficient Dynamic Portfolio Strategies or How to Throw Away a Million Dollars in the Stock Market. *Review of Financial Studies*, 1 :67–88.
- Dybvig, P. and Ingersoll, J. (1982). Mean-variance theory in complete markets. *Journal of Business*, pages 233–251.
- Elderton, W. and Johnson, N. (1969). *Systems of frequency curves*.
- Elton, E., Gruber, M., and Rubin, C. (1995). Fundamental Economic Variables, Expected Returns, and Bond Fund Performance. *The Journal of Finance*, 50 :1229–1256.
- Embrechts, P., McNeil, A., and Straumann, D. (1999). Correlation : Pitfalls and alternatives. *RISK Magazine*, pages 69–71.
- Fama, E. and French, K. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*.

- Fryer, J. and Robertson, C. (1972). A comparison of some methods for estimating mixed normal distributions. *Biometrika*, pages 639–648.
- Genest, C., Ghoudi, K., and Rivest, L. (1995). A semi parametric estimation procedure of dependence parameter in multivariate family distributions. *Biometrika*, 82 :543–552.
- Genest, C. and MacKay, J. (1986). The joy of copulas : bivariate distributions with uniform marginals. *The American Statistician*, 40 :280–283.
- Genest, C. and Rémillard, B. (2008). Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models. *Annales de l'Institut Henri-Poincaré : Probabilités et Statistique*, (sous presse).
- Harrison, J. and Pliska, S. (1981). Martingales and Stochastic Integrals in the Theory of Continuous Trading. *Stochastic Processes and their Applications*, 11 :215–260.
- Harrison, J. M. and Kreps, D. M. (1979). Martingales and arbitrage in multiperiod securities markets. *Journal of Economic Theory*, 20 :381–408.
- Jones, C., Kaul, G., and Lipson, M. (1994). Transactions, volume and volatility. *Review of financial Studies*, pages 631–651.
- Kat, H. and Palaro, H. (2005). Who needs hedge funds? A copula-based approach to hedge fund return replication. Technical report, Cass Business School, City University.
- Liang, B. (1999). On the Performance of Hedge Funds. *Financial Analysts Journal*.
- Lo, A. (2001). Risk Management of Hedge Funds :Introduction and Overview. *Financial Analysts Journal*.
- Nelsen, R. (2005). *An Introduction to Copulas*. Springer Verlag.

- Papageorgiou, N., Rémillard, B., and Hocquard, A. (2007). Replicating the properties of hedge fund returns. Technical report, DGAM-HEC Research on Alternative Investments.
- Pearson, K. (1894). Contributions of the theory of mathematical evolution. *Philosophical Transactions of the Royal Society of London*, pages 343–414.
- Roncalli, T. (2004). *La Gestion des Risques Financiers*. Gestion. Economica.
- Rosenberg, J. V. (1999). Semiparametric pricing of multivariate contingent claims. Technical report, Stern school of business.
- Rosenberg, J. V. (2003). Non-parametric pricing of multivariate contingent claims. *The Journal of Derivatives*, 10(3) :9–26.
- Rosenblatt, M. (1952). Remarks on a multivariate transformation. *Ann. Math. Stat.*, 23 :470–472.
- Schweizer, M. (1995). Variance-Optimal Hedging in Discrete Time. *Mathematics of Operations Research*, 20 :10–32.
- Sklar, M. (1959). Fonctions de Répartition à N dimensions et Leurs Marges. *Publ. Inst. Statist. Univ. Paris*, 8 :229–231.
- Tan, W. and Chang, W. (1972). Some comparisons of the method of moments and the method of maximum likelihood in estimating parameters of a mixture of two normal densities. *Journal of the American Statistical Association*, pages 702–708.

# Annexe A

## Preuve de la proposition 1

Considérons la fonction caractéristique des rendements  $R_t, t = 1, \dots, m$ . Nous avons, pour tout  $\lambda$ ,

$$\begin{aligned}\phi_{R_t}(\lambda) &= E[e^{i\lambda R_t}] \\ &= E[e^{i\lambda\sqrt{2\sigma}Z\sqrt{X}}] \\ &= E[e^{-0.5(\lambda\sqrt{2\sigma}\sqrt{X})^2}] \\ &= E[e^{-\lambda^2\sigma^2X}] \\ &= (1 + \sigma^2\lambda^2)^{-\frac{1}{m}}.\end{aligned}$$

Par la propriété de la fonction caractéristique de la somme des variables aléatoires indépendantes, la fonction caractéristique de la somme de variable aléatoires indépendantes est égale au produit des fonctions caractéristiques de chacune d'elles, on déduit que

$$\phi_{\sum_{t=1}^m R_t}(\lambda) = (1 + \sigma^2\lambda^2)^{-1},$$

et  $\phi_{R_T}$  est bien la fonction caractéristique d'une distribution de Laplace centrée. Par l'unicité de cette fonction, on déduit que  $R_T$  suit une loi de Laplace centrée. Q.E.D.

## Annexe B

# Détermination des paramètres de la distribution de Johnson-SU

Il s'agit d'une procédure récursive permettant de déterminer les paramètres d'une distribution de Johnson-SU à partir de ses quatre premiers moments. Cette méthode est extraite de Elderton and Johnson (1969).

Elle est décrite de la manière suivante :

Tout d'abord, on note  $\beta_1$  et  $\beta_2$  respectivement le carré du coefficient d'asymétrie et le coefficient d'aplatissement de la distribution.

1. Fixer des paramètres initiaux pour  $\delta = \delta_{(1)}$  à partir de la table de Johnson présentée dans *Biometrika, LII(1965)*.
2. Poser  $\omega_{(1)} = \exp(-\delta_{(1)}^{-2})$ .

3. Déterminer  $\gamma_{(1)}$ , en résolvant l'équation suivante :

$$\beta_2 - 3 = \frac{(\omega_{(1)} - 1)[A_2(\omega_{(1)})m_{(1)}^2 + A_1(\omega_{(1)})m_{(1)} + A_0(\omega_{(1)})]}{2(2m_{(1)} + \omega_{(1)} + 1)^2},$$

avec

$$\begin{aligned} m_{(1)} &= \omega_{(1)} \sinh^2(\gamma_{(1)}/\delta_{(1)}), \\ A_2(\omega) &= 8(\omega^3 + 3\omega^2 + 6\omega + 6), \\ A_1(\omega) &= 8(\omega^4 + 3\omega^3 + 6\omega^2 + 7\omega + 3), \\ A_0(\omega) &= \omega^5 + 3\omega^4 + 6\omega^3 + 10\omega^2 + 9\omega + 3. \end{aligned}$$

4. Avec les valeurs de  $m_{(1)}$  et  $\omega_{(1)}$ , calculer  $\beta_1$  en utilisant la formule suivante :

$$\beta_1 = \frac{m(\omega - 1)[4(\omega + 2)m + 3(\omega + 1)^2]^2}{2(2m + \omega + 1)^3}.$$

On dénote cette valeur par  $\beta_1(\delta_{(1)}, \beta_2)$ .

5. Déterminer une nouvelle valeur de  $\omega_2$  de  $\omega$  en résolvant l'équation suivante :

$$\frac{\beta_1}{\beta_1(\delta_{(1)}, \beta_2)} = \frac{\beta_2 - \frac{1}{2}(\omega_{(2)}^4 + 2\omega_{(2)}^2 + 3)}{\beta_2 - \frac{1}{2}(\omega_{(1)}^4 + 2\omega_{(1)}^2 + 3)}.$$

6. Recommencer les étapes précédentes jusqu'à ce que  $\beta_1(\delta_{(n)}, \beta_2)$  soit proche de  $\beta_1$ .

7. Déduire les paramètres  $\lambda$  et  $\xi$  en utilisant les formules de l'espérance et de la variance.

# Annexe C

## Copules

### C.1 Introduction aux copules

Les copules sont de plus en plus utilisées en ingénierie financière. Elles ont été introduites par Sklar (1959). Une copule est une fonction de répartition multivariée dont les marges sont uniformes. En particulier, pour une copule bivariée  $C$ , on a :

1.  $\text{Dom } C = [0, 1] \times [0, 1]$ .
2.  $C(0, u) = C(u, 0) = 0$  et  $C(1, u) = C(u, 1) = u$  pour tout  $u$  dans  $[0, 1]$ .
3.  $C$  est 2-increasing :

$$C(v_1, v_2) - C(v_1, u_2) - C(u_1, v_2) + C(u_1, u_2) \geq 0$$

pour tous points  $(u_1, u_2), (v_1, v_2) \in [0, 1]^2$ , tels que  $0 \leq u_1 \leq v_1 \leq 1$  et  $0 \leq u_2 \leq v_2 \leq 1$ .

Le théorème de Sklar (1959) fait la jonction entre la fonction copule et les variables aléatoires.

**Théorème 2** Soit  $F$  une fonction de distribution bidimensionnelle dont les marges sont  $F_1$  et  $F_2$ . Alors il existe une copule  $C$  telle que  $F$  admette la représentation copule :

$$F(x_1, x_2) = C(F_1(x_1), F_2(x_2)), \quad \forall x_1, x_2.$$

La copule  $C$  est unique si les marges sont continues.

Pour de plus amples détails, voir Roncalli (2004).

## C.2 Quelques mesures de dépendance

Une mesure de dépendance est une mesure exprimée exclusivement à partir de la structure de dépendance entre les variables aléatoires. Il existe plusieurs mesures de dépendance des variables aléatoires. Nous nous intéressons au *tau de Kendall* et au *coefficient de van der Waerden*.

Le  $\tau$  de Kendall est une mesure de concordance entre deux variables aléatoires. On dit que deux points  $(x_1, x_2)$  et  $(x'_1, x'_2)$  sont discordants si et seulement si  $x_1 \leq x'_1$  et  $x_2 > x'_2$  ou  $x_2 \leq x'_2$  et  $x_2 > x'_2$ . Dans le cas contraire, ils sont concordants.

Le  $\tau$  de Kendall est défini par

$$\tau = 4 \int \int_{[0,1]^2} C(u_1, u_2) d(C(u_1, u_2)) - 1 = 2P_C - 1,$$

où  $P_C$  est la probabilité que deux points choisis au hasard soient concordants.

$\tau$  est estimé, à partir d'un échantillon de  $n$  points, par  $(C - D)/\binom{n}{2}$ , où  $C$  est le nombre de paires de points concordants et  $D$  est le nombre de paires de points discordants.

On considère un couple de variables aléatoires  $(X_1, X_2)$  de fonction de répartition respectives  $F_1$  et  $F_2$ . Le *coefficient de van der Waerden* est le coefficient de corrélation de Spearman (classique) des variables aléatoires  $\Phi^{-1}\{F_1(X_1)\}$  et  $\Phi^{-1}\{F_2(X_2)\}$  où  $\Phi^{-1}$  est la fonction de répartition inverse de la loi normale centrée réduite. À noter que dans le cas d'un vecteur gaussien, ce coefficient correspond exactement au coefficient de corrélation classique.