

WALID JBILI

Modélisation asymétrique de titres financiers

Mémoire présenté
à la Faculté des études supérieures de l'Université Laval
dans le cadre du programme de maîtrise en statistique
pour l'obtention du grade de Maître ès sciences (M.Sc.)

FACULTÉ DES SCIENCES ET DE GÉNIE
UNIVERSITÉ LAVAL
QUÉBEC

Juillet 2008

Résumé

La théorie de Markowitz a toujours été au centre de la théorie de gestion de portefeuilles. Cependant, elle est l'objet de plusieurs critiques. Dans ce mémoire, on se propose de revoir certains postulats de la théorie de Markowitz.

L'approche que préconise ce mémoire est de considérer le portefeuille dans sa globalité au lieu des titres individuels. Cette approche vise à identifier une loi s'ajustant aux portefeuilles ou à une puissance des portefeuilles. L'identification de la loi s'appuiera sur des portefeuilles simulés et d'autres réels. Plusieurs méthodes seront exploitées pour identifier et vérifier l'adéquation de cette loi.

Abstract

The theory of Markowitz has always been at the center to the theory of portfolio management. However, it undergoes several criticisms. In this thesis, we propose to re-examine certain postulates of the theory of Markowitz.

The approach used in this thesis is to consider the portfolio as a whole instead of the individual titles. This approach aims at identifying a law that fits the distribution of the return (or a power transformation) of the portfolios. The identification of the law will be based on simulated and real portfolios. Several methods will be employed to identify and check the adequacy of the law.

Avant-propos

Tout d'abord, je tiens à remercier mes directeurs de recherche, Thierry Duchesne, professeur au Département de mathématiques et de statistique de l'Université Laval, et Michel Jacques professeur à l'École d'actuariat de l'Université Laval. Ils ont toujours su se montrer disponibles, ce qui fut grandement apprécié. Je dois aussi dire merci à tous les professeur(e)s du Département, à Gaétan Daigle, et à Hélène Crépeau du Service de Consultation Statistique de l'Université, et à toute l'équipe de professionnels du département.

Enfin, je me permets également de remercier mes parents, mes grands-parents pour leur soutien moral et leur encouragement tout au long de mes études. Je ne manquerai de remercier ma femme, mon fils et l'ensemble de ma famille et amis. Vous avez tous contribué, peut-être sans le savoir, à l'avancement de mes travaux.

[À MOHAMED Abd ALLAH!]

Table des matières

| | |
|--|----------|
| Résumé | ii |
| Abstract | iii |
| Avant-Propos | iv |
| Table des matières | ix |
| Liste des tableaux | xi |
| Table des figures | xiii |
| 1 Introduction | 1 |
| 1.1 Motivation financière | 2 |
| 1.2 La structure du mémoire | 5 |
| 2 Familles de distributions | 6 |
| 2.1 Généralités | 6 |
| 2.2 Lois | 8 |
| 2.2.1 La loi lognormale à deux paramètres | 8 |
| 2.2.2 La loi de Weibull à deux paramètres | 9 |
| 2.2.3 La loi gamma | 11 |
| 2.2.4 La loi gamma transformée | 12 |
| 2.2.5 La loi Bêta transformée (GBII) | 14 |
| 2.2.6 Liens entre certaines familles | 16 |
| 2.3 Statistiques descriptives | 16 |
| 2.3.1 Histogramme | 16 |
| 2.3.2 Fonction d'espérance de survie résiduelle | 17 |
| 2.3.3 Ramlau-Hansen | 17 |
| 2.3.4 QQ-plot et PP-plot | 18 |
| 2.4 Aspects inférentiels et numériques | 19 |
| 2.4.1 Approches d'estimation | 20 |
| 2.4.2 Résolution des équations de la méthode des moments | 22 |

| | | |
|----------|---|-----------|
| 3 | Copules | 26 |
| 3.1 | Distributions multivariées | 26 |
| 3.2 | Copules Multivariées | 27 |
| 3.3 | Copules et dépendance | 29 |
| 3.3.1 | Concordance | 29 |
| 3.3.2 | Tau de Kendall | 30 |
| 3.3.3 | Rho de Spearman | 30 |
| 3.4 | Familles de copules | 31 |
| 3.4.1 | Copules de valeurs extrêmes | 31 |
| 3.4.2 | Copules Archimédiennes | 31 |
| 3.4.3 | La distribution multivariée de t et sa copule | 35 |
| 3.5 | Simulation des copules | 36 |
| 3.5.1 | Copules archimédiennes | 36 |
| 3.5.2 | Copule de t | 37 |
| 4 | Modélisation des portefeuilles | 38 |
| 4.1 | Lois des rendements | 39 |
| 4.2 | Rendements des données réelles | 40 |
| 4.2.1 | Berkshire Hathaway | 41 |
| 4.2.2 | Hecla Mining | 42 |
| 4.2.3 | McDonald's | 42 |
| 4.2.4 | Reebok International | 43 |
| 4.2.5 | Scientific Games | 43 |
| 4.3 | Poids des titres et portefeuilles | 44 |
| 4.3.1 | Problématique des poids des titres | 44 |
| 4.3.2 | Conception des portefeuilles | 46 |
| 5 | Étude descriptive | 47 |
| 5.1 | Statistiques descriptives | 47 |
| 5.2 | Histogramme | 55 |
| 5.3 | Espérance de vie résiduelle | 56 |
| 5.4 | Ramlau-Hansen | 57 |
| 5.5 | Conclusion | 58 |
| 6 | Données simulées | 60 |
| 6.1 | Estimation et adéquation | 61 |
| 6.1.1 | La loi de Weibull | 61 |
| 6.1.2 | La loi gamma transformée (GT) | 63 |
| 6.1.3 | La GBII | 66 |
| 6.2 | Conclusion | 68 |

| | | |
|----------|--|------------|
| 7 | Données réelles | 84 |
| 7.1 | Estimation et adéquation | 84 |
| 7.1.1 | La loi de Weibull | 84 |
| 7.1.2 | La loi gamma transformée | 85 |
| 7.1.3 | La loi GBII | 87 |
| 7.2 | Conclusion | 87 |
| 8 | Conclusion | 91 |
| | Bibliographie | 92 |
| A | Approche numérique d'estimation | 94 |
| A.1 | Méthode des moments | 94 |
| A.1.1 | uniroot | 94 |
| A.1.2 | FSolve | 96 |
| A.2 | Méthodes de maximum de vraisemblance et de distance minimale | 98 |
| B | Approche numérique pour la simulation des copules | 102 |
| B.1 | Le package Copula | 102 |
| B.2 | Exemples | 103 |
| C | Simulation | 106 |
| C.1 | Estimation des paramètres | 106 |
| C.2 | Simulation des rendements | 108 |
| C.2.1 | Cas d'indépendance | 108 |
| C.2.2 | Cas de dépendance | 108 |
| C.3 | Simulation des poids et des portefeuilles | 110 |
| D | Programmes R pour l'étude descriptive | 112 |
| D.1 | Histogramme et statistiques descriptives | 112 |
| D.2 | Espérance de survie résiduelle | 113 |
| D.3 | Ramlau-Hansen | 114 |
| E | Histogrammes | 116 |
| F | Définition de nouvelles lois sur R | 129 |
| F.1 | La gamma transformée | 129 |
| F.2 | La GBII | 130 |
| G | Programme R pour l'adéquation | 131 |
| G.1 | PP-plot | 131 |
| G.2 | QQ-plot | 132 |
| G.3 | Superposition | 132 |

| | |
|---|------------|
| G.4 Moments théoriques vs. empiriques | 133 |
| H Adéquation de la GBII | 135 |
| I GBII QQ-plot | 139 |

Liste des tableaux

| | | |
|------|---|----|
| 4.1 | <i>Matrice des coefficients de corrélation de Kendall des données réelles . . .</i> | 41 |
| 5.1 | <i>Indépendance (9 titres)</i> | 49 |
| 5.2 | <i>Copule de Clayton - faible niveau de dépendance (9 titres)</i> | 50 |
| 5.3 | <i>Copule de Clayton - niveau élevé de dépendance (9 titres)</i> | 51 |
| 5.4 | <i>Copule de t - faible niveau de dépendance (9 titres)</i> | 52 |
| 5.5 | <i>Copule de t - niveau élevé de dépendance (9 titres)</i> | 53 |
| 5.6 | <i>Données réelles (9 titres)</i> | 54 |
| 6.1 | <i>Les différents estimateurs de la loi de Weibull pour les neuf portefeuilles simulés sous l'hypothèse de l'indépendance</i> | 69 |
| 6.2 | <i>Les différents estimateurs de la loi de Weibull pour les neuf portefeuilles issus d'une Clayton à faible niveau de dépendance</i> | 70 |
| 6.3 | <i>Les différents estimateurs de la loi de Weibull pour les neuf portefeuilles issus d'une Clayton à niveau élevé de dépendance</i> | 71 |
| 6.4 | <i>Les différents estimateurs de la loi de Weibull pour les neuf portefeuilles issus d'une copule de t à faible niveau de dépendance</i> | 72 |
| 6.5 | <i>Les différents estimateurs de la loi de Weibull pour les neuf portefeuilles issus d'une copule de t à niveau élevé de dépendance</i> | 73 |
| 6.6 | <i>Les différents estimateurs de la gamma transformée pour les neuf portefeuilles simulés sous l'hypothèse de l'indépendance</i> | 74 |
| 6.7 | <i>Les différents estimateurs de la loi gamma transformée pour les neuf portefeuilles issus d'une Clayton à faible niveau de dépendance</i> | 75 |
| 6.8 | <i>Les différents estimateurs de la loi gamma transformée pour les neuf portefeuilles issus d'une Clayton à niveau élevé de dépendance</i> | 76 |
| 6.9 | <i>Les différents estimateurs de la loi gamma transformée pour les neuf portefeuilles issus d'une copule de t à faible niveau de dépendance</i> | 77 |
| 6.10 | <i>Les différents estimateurs de la loi gamma transformée pour les neuf portefeuilles issus d'une copule de t à niveau élevé de dépendance</i> | 78 |
| 6.11 | <i>Les différents estimateurs de la GBII pour les neuf portefeuilles simulés sous l'hypothèse de l'indépendance</i> | 79 |
| 6.12 | <i>Les différents estimateurs de la GBII pour les neuf portefeuilles issus d'une Clayton à faible niveau de dépendance</i> | 80 |

| | | |
|------|---|----|
| 6.13 | <i>Les différents estimateurs de la GBII pour les neuf portefeuilles issus d'une Clayton à niveau élevé de dépendance</i> | 81 |
| 6.14 | <i>Les différents estimateurs de la GBII pour les neuf portefeuilles issus d'une copule de t à faible niveau de dépendance</i> | 82 |
| 6.15 | <i>Les différents estimateurs de la GBII pour les neuf portefeuilles issus d'une copule de t à niveau élevé de dépendance</i> | 83 |
| 7.1 | <i>Les différents estimateurs de la loi de Weibull pour les neuf portefeuilles issus des données réelles</i> | 88 |
| 7.2 | <i>Les différents estimateurs de la loi gamma transformée pour les neuf portefeuilles issus des données réelles</i> | 89 |
| 7.3 | <i>Les différents estimateurs de la loi GBII pour les neuf portefeuilles issus des données réelles</i> | 90 |

Table des figures

| | | |
|-----|---|-----|
| 1.1 | Combinaisons possibles de Espérance de gain - Variance (tiré de Markowitz (1952)) | 4 |
| 5.1 | Exemple de figure d'espérance de vie résiduelle | 57 |
| 5.2 | Exemple de figure de Ramlau-Hansen | 58 |
| 6.1 | Exemples de QQ-plots d'une loi de Weibull pour des portefeuilles simulés sous l'hypothèse de l'indépendance | 62 |
| 6.2 | Exemples de QQ-plots d'une loi de Weibull pour des portefeuilles issus de la copule de t à niveau élevé de dépendance | 63 |
| 6.3 | Exemples de QQ-plots d'une loi gamma transformée ajustée à des portefeuilles simulés sous l'hypothèse de l'indépendance | 65 |
| 6.4 | Exemples de QQ-plots d'une loi gamma transformée ajustée à des portefeuilles issus de la copule de t à niveau élevé de dépendance | 66 |
| 7.1 | Exemples de QQ-plots d'une loi de Weibull pour des portefeuilles issus des données réelles | 85 |
| 7.2 | Exemples de QQ-plots d'une loi gamma transformée pour des portefeuilles issus des données réelles | 86 |
| E.1 | Histogrammes des portefeuilles à neuf titres simulés sous l'hypothèse de l'indépendance | 117 |
| E.2 | Histogrammes des portefeuilles à neuf titres simulés sous l'hypothèse de l'indépendance (Suite) | 118 |
| E.3 | Histogrammes des portefeuilles à neuf titres issus d'une Clayton à un faible niveau de dépendance | 119 |
| E.4 | Histogrammes des portefeuilles à neuf titres issus d'une Clayton à un faible niveau de dépendance (Suite) | 120 |
| E.5 | Histogrammes des portefeuilles à neuf titres issus d'une Clayton à un niveau élevé de dépendance | 121 |
| E.6 | Histogrammes des portefeuilles à neuf titres issus d'une Clayton à un niveau élevé de dépendance (Suite) | 122 |
| E.7 | Histogrammes des portefeuilles à neuf titres issus d'une copule de t à un faible niveau de dépendance | 123 |

| | | |
|------|---|-----|
| E.8 | <i>Histogrammes des portefeuilles à neuf titres issus d'une copule de t à un faible niveau de dépendance (Suite)</i> | 124 |
| E.9 | <i>Histogrammes des portefeuilles à neuf titres issus d'une Clayton à un niveau élevé de dépendance</i> | 125 |
| E.10 | <i>Histogrammes des portefeuilles à neuf titres issus d'une copule de t à un niveau élevé de dépendance (Suite)</i> | 126 |
| E.11 | <i>Histogrammes des portefeuilles à neuf titres simulés à partir des données réelles</i> | 127 |
| E.12 | <i>Histogrammes des portefeuilles à neuf titres simulés à partir des données réelles (Suite)</i> | 128 |
| I.1 | <i>QQ-plots des neuf portefeuilles simulés sous l'hypothèse de l'indépendance</i> | 140 |
| I.2 | <i>QQ-plots des neuf portefeuilles issus d'une Clayton à un faible niveau de dépendance</i> | 141 |
| I.3 | <i>QQ-plots des neuf portefeuilles issus d'une Clayton à un niveau élevé de dépendance</i> | 142 |
| I.4 | <i>QQ-plots des neuf portefeuilles issus d'une copule de t à un faible niveau de dépendance</i> | 143 |
| I.5 | <i>QQ-plots des neuf portefeuilles issus d'une copule de t à un niveau élevé de dépendance</i> | 144 |
| I.6 | <i>QQ-plots des neuf portefeuilles simulés à partir des données réelles</i> . . . | 145 |

Chapitre 1

Introduction

Un des grands problèmes de la finance actuelle est l'allocation de ressources entre plusieurs choix d'investissements. Ce problème est considéré dans ce qu'on appelle la théorie des portefeuilles : fixer la proportion de la richesse totale à investir dans chacun des titres financiers disponibles de sorte à rencontrer un critère d'investissement bien défini. Un des premiers et principaux auteurs ayant traité de cette question est Harry Markowitz dans son article paru dans le *Journal of Finance* en mars 1952. Markowitz cherche à minimiser le risque mesuré par la variance de la variable aléatoire représentant le rendement du portefeuille.

La théorie de Markowitz, quoique révolutionnaire à son époque, a connu plusieurs critiques. L'approche de Markowitz et de ses adeptes porte sur les titres financiers individuels. Implicitement ces titres sont modélisés par une loi normale, donc symétrique. L'objectif principal de ce mémoire est la modélisation asymétrique des titres financiers, hypothèse plus appropriée à la réalité des marchés financiers. Les propriétés de la loi normale permettent de calculer facilement la loi d'un portefeuille, combinaison linéaire de plusieurs titres. Dans le cas d'autres lois, la loi d'une combinaison de titres est plus compliquée, voire non explicite. Pour cette raison, ce mémoire prendra une approche différente, qui est de modéliser directement la loi du portefeuille plutôt que celles des titres individuels. Le choix de la loi n'est cependant pas arbitraire. Nous chercherons une loi, ou plutôt une famille de lois, qui modélise le rendement d'un portefeuille composé de titres individuels dont les rendements sont décrits par des lois aussi variées que possible.

1.1 Motivation financière

Le processus de sélection des poids à accorder à chaque titre financier d'un portefeuille est une approche à deux niveaux. Le premier niveau est celui de l'observation et de la spéculation, alors que le second conduit à la constitution du portefeuille. Au premier niveau, l'observation et l'expérience permettent de spéculer sur le comportement futur du marché financier et d'en identifier les perspectives les plus prometteuses. Au début, l'intervenant observe l'état du marché. Ensuite, il identifie les titres financiers les plus prometteurs dans lesquels il se propose d'investir. Au deuxième niveau, l'intervenant décidera du poids à accorder à chacun de ces titres dans le portefeuille.

Dans un marché financier donné, chaque intervenant vise à maximiser son utilité. Par utilité, on entend la satisfaction suite au placement. Cette satisfaction est étroitement liée au couple profit-risque. Ainsi, l'utilité est une fonction du rendement et du risque, communément mesuré par la variabilité du portefeuille.

En ignorant le risque associé aux titres, on peut accorder davantage intérêt aux titres ayant le plus grand des rendements anticipés. Par contre si on veut tenir compte du risque, on propose souvent qu'un portefeuille diversifié est préférable à tous ceux qui ne le sont pas. En soi, la diversification n'est pas garante d'un portefeuille optimal ; par optimal, on entend un portefeuille qui maximise l'utilité de l'intervenant.

Posons un marché financier ([Markowitz \(1952\)](#)) où s'échangent N titres financiers, et soit r_{it} le rendement du i^e titre au temps t . Soit d_{it} le facteur d'escompte permettant d'actualiser le rendement sur le titre au présent et soit X_i , la proportion du portefeuille investie dans le i^e titre, où $X_i \geq 0$ pour tout i et $\sum X_i = 1$. Ainsi, le revenu anticipé escompté du portefeuille, par dollar d'investissement, est

$$\begin{aligned} S &= \sum_{t=1}^{\infty} \sum_{i=1}^N d_{it} r_{it} X_i \\ &= \sum_{i=1}^N X_i \left(\sum_{t=1}^{\infty} d_{it} r_{it} \right), \end{aligned} \tag{1.1}$$

$$R_i = \sum_{t=1}^{\infty} d_{it} r_{it}, \tag{1.2}$$

où R_i est le rendement escompté du i^e titre. Dès lors, on a $S = \sum_{i=1}^N X_i R_i$. Pour maximiser S , posons $X_i = 1$ pour le rendement R_i maximum et $X_i = 0$ pour tous les autres titres. Si, toutefois, on retrouve plus qu'un maximum, alors toute combinaison des titres de rendement maximum maximisera S .

En général, un investisseur veut maximiser son utilité. Ainsi, un investisseur retiendra tous les titres maximisant son utilité, et répartira ses fonds entre ces derniers. Malheureusement, la diversification ne peut pas éliminer toute la variance, du fait que les rendements des titres sont intercorrélés. Ainsi, un portefeuille maximisant le rendement sur capital n'est pas nécessairement celui de variance minimale.

Markowitz (1952) propose un compromis visant à la fois à maximiser le rendement sur le capital et à minimiser le risque mesuré par la variance. Dès lors, l'investisseur doit choisir entre les diverses combinaisons de Espérance de gain - Variance (risque) celle qui maximise son utilité. On introduira dans ce qui suit la règle «Espérance de gain - Variance», que l'on notera la règle $E-V$.

Soit μ_i , l'espérance du gain R_i ; σ_{ij} la covariance entre R_i et R_j (on notera σ_{ii} la variance de R_i). Soit X_i , le poids accordé par l'investisseur au i^e titre. Nous supposons que X_i n'est pas une variable aléatoire, mais qu'il est fixé par l'investisseur. Par la suite, le rendement du portefeuille $S = \sum_{i=1}^N X_i R_i$ est une combinaison linéaire pondérée d'un ensemble de variables aléatoires. On déduit que S est une variable aléatoire dont l'espérance de gain E est

$$E = \sum_{i=1}^N X_i \mu_i, \quad (1.3)$$

et la variance V est

$$V = \sum_{i=1}^N \sum_{j=1}^N \sigma_{ij} X_i X_j. \quad (1.4)$$

L'investisseur a la tâche de choisir les X_i qui, parmi toutes les combinaisons (E, V) , fournissent la combinaison qui maximise son utilité définie en pénalisant le rendement moyen E par la variance V . Pour un E donné, l'investisseur cherchera le portefeuille qui minimise V , et pour une V donnée, il cherchera le portefeuille qui maximise E . Ces portefeuilles définissent l'ensemble des portefeuilles efficients. Plusieurs techniques ont été élaborées pour retracer cet ensemble et les combinaisons (E, V) associées, pour μ_i et σ_{ij} donnés.

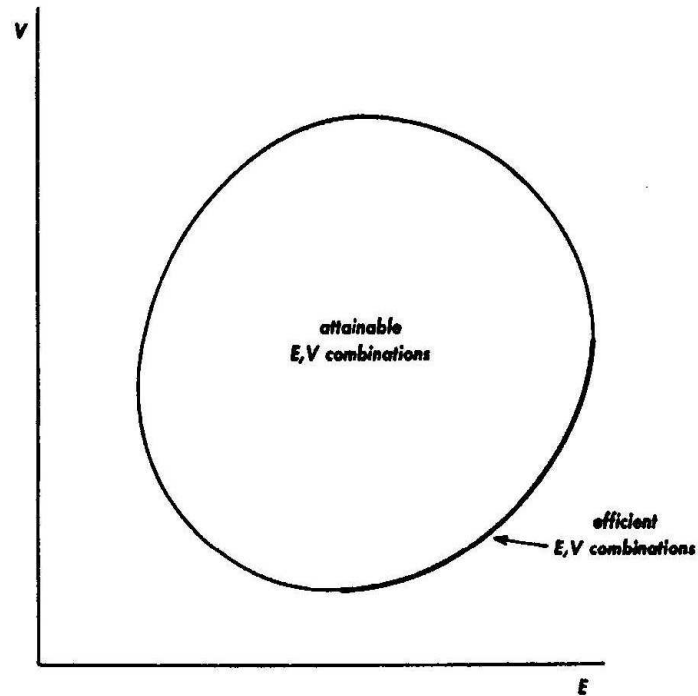


FIG. 1.1 – *Combinaisons possibles de Espérance de gain - Variance (tiré de [Markowitz \(1952\)](#))*

La règle EV de Markowitz a connu ses limites. En effet, mesurer le risque par la variance revient implicitement à supposer que les lois sous-jacentes sont normales, alors que dans la réalité des marchés ([Cencia et Filippinib \(2006\)](#)) les rendements ne suivent pas une distribution normale. L'utilisation de la variance comme mesure de risque pour des distributions asymétriques serait ainsi inappropriée ([Bawa \(1975\)](#)). Plusieurs chercheurs ont abordé des modifications, voire des remaniements structurels à cette règle. Cependant, ces méthodes ne seront pas exposées dans ce mémoire.

La combinaison de lois non normales et corrélées donne des distributions compliquées. Plutôt que de modéliser les titres individuels, nous essaierons plutôt de modéliser directement le rendement du portefeuille, une combinaison linéaire de plusieurs titres individuels. Dans ce mémoire, on considère une approche paramétrique. Nous tenterons donc d'identifier une famille de loi qui s'ajuste aux rendements de portefeuilles.

1.2 La structure du mémoire

Ce mémoire se divise en huit chapitres. Le chapitre 2 se divise en trois parties distinctes. La première partie est l'introduction de certaines lois ainsi que de leurs propriétés. Certaines de ces lois serviront à la simulation de nos données, alors que les autres seront ajustées aux échantillons de rendements de portefeuilles. La deuxième partie expose les différentes approches graphiques utilisées dans ce mémoire afin d'identifier la loi du portefeuille. Enfin, dans la dernière partie, on traite des aspects inférentiels et numériques, tels que les approches d'estimation et la résolution des équations de la méthode des moments. Au chapitre 3, après une introduction aux distributions multivariées et à la théorie des copules, on discute des propriétés et de la simulation des copules retenues pour ce mémoire. Le chapitre 4 traite de la modélisation des portefeuilles. On y aborde la simulation des rendements selon les différents scénarios retenus, la simulation des poids et la théorie sous-jacente, ainsi que la conception des portefeuilles. Le chapitre 5 est le tremplin entre le chapitre 4 et les chapitres 6 et 7. Dans ce chapitre, l'étude descriptive donnera des indices aidant à identifier une éventuelle loi s'ajustant à nos données. Au chapitre 6, on ne s'intéresse qu'aux portefeuilles simulés. Plusieurs approches numériques et graphiques serviront à vérifier l'adéquation de différentes lois à ces données. Au chapitre 7, l'adéquation de la loi retenue sera testée sur les portefeuilles issus de données réelles. À la fin de ce chapitre, on sera en mesure de proposer une avenue de solution à la problématique traitée dans ce mémoire. Le chapitre 8 viendra finalement conclure ce mémoire.

Chapitre 2

Familles de distributions

2.1 Généralités

Une variable aléatoire X est une fonction qui associe à un ensemble d'événements réalisés dans un espace échantillonné, un ensemble de valeurs réelles. Une fonction de probabilité associe à toute réalisation de X une valeur sur le segment $[0, 1]$ qu'on dénotera probabilité associée à cette réalisation. Soit x , la valeur associée à une réalisation de X , et $P[X \leq x]$, la probabilité que la valeur d'une réalisation de X soit inférieure ou égale à x .

Définition 2.1.1. *La fonction $F(x) = P[X \leq x]$ est la fonction de répartition de la variable aléatoire X (CDF).*

Une famille de distributions est un ensemble de distributions dont les éléments sont indexés par un ensemble fixe de paramètres. Une famille est formellement définie par un ensemble de fonctions de répartition

$$\{F(x; \theta) : \theta \in \Theta\},$$

où θ est un scalaire ou un vecteur et où Θ est l'ensemble des valeurs possibles des paramètres. En fixant la famille de lois dont est issue la population, on peut estimer les paramètres et ainsi inférer sur cette population et déterminer toute quantité d'intérêt telle que la moyenne. La précision de ces estimations peut alors également être étudiée. L'estimation des paramètres sera traitée dans la section 2.4.

Par la suite, dans l'intérêt du sujet de ce mémoire, on ne discutera que des distributions continues.

Définition 2.1.2. Une variable aléatoire X avec fonction de répartition $F(x)$ est dite absolument continue si, et seulement si, il existe une fonction non négative $f(x)$, telle que

$$F(x) = \int_{-\infty}^x f(t) dt, \quad (2.1)$$

pour tout réel x . On appelle $f(x)$ fonction de densité de la variable aléatoire X (PDF).

Par ailleurs, si une variable aléatoire X est de distribution absolument continue, alors sa fonction de répartition $F(x)$ est continue.

Soit D , le domaine de définition de X , sous-ensemble de \Re . La fonction de quantile d'une variable aléatoire X de CDF $F(x)$ est définie ainsi :

$$Q(u) = \inf\{x : F(x) \geq u\}; 0 < u < 1. \quad (2.2)$$

Dès lors, pour toute variable aléatoire X de distribution continue, la fonction de quantile $Q(u)$ s'écrit

$$Q(u) = F^{-1}(u); 0 < u < 1. \quad (2.3)$$

La fonction de survie d'une variable aléatoire X est définie par

$$S(x) = P[X > x] = 1 - F(x). \quad (2.4)$$

L'espérance de vie résiduelle, qu'on notera e_X , s'écrit sous la forme suivante :

$$\begin{aligned} e_X(x) &= E[X - x | X \geq x] \\ &= \frac{\int_x^\infty (t - x) f(t) dt}{S(x)} \\ &= \frac{\int_x^\infty S(t) dt}{S(x)}, x \geq 0, \end{aligned} \quad (2.5)$$

la dernière égalité étant vraie si $S(x) \rightarrow 0$ plus rapidement que $1/x$ quand $x \rightarrow \infty$. La fonction de risque d'une variable aléatoire X est définie par

$$h(x) = f(x)/S(x) = -d \ln[S(x)]/dx, \quad (2.6)$$

et la fonction de risque cumulative d'une variable aléatoire X est définie par

$$H(x) = \int_0^x h(t) dt = -\ln[S(x)]. \quad (2.7)$$

2.2 Lois

Dans la section suivante, en fonction du thème à développer dans ce mémoire, on se concentrera sur certaines familles de lois, à savoir, la loi lognormale, la loi de Weibull, la loi gamma, la loi gamma transformée et La loi Bêta transformée (GBII).

Dans ce qui suit, on notera μ_i le i^e moment de X , μ_i le i^e moment centré de X , $V(X)$ la variance de X , β_1 et β_2 respectivement le carré du coefficient d'asymétrie et du coefficient d'aplatissement, et $Med(X)$ la médiane de X .

2.2.1 La loi lognormale à deux paramètres

Une variable aléatoire positive X est dite de distribution lognormale avec paramètres μ et σ si $Y = \ln X$ est normale de moyenne μ et de variance σ^2 , notée $N(\mu, \sigma^2)$. La loi lognormale à deux paramètres sera dénotée $\Lambda(\mu, \sigma)$.

La fonction de densité de X suivant une loi $\Lambda(\mu, \sigma)$ est

$$f(x) = \begin{cases} \frac{1}{\sqrt{2\pi}\sigma x} \exp\left\{-\frac{(\ln(x)-\mu)^2}{2\sigma^2}\right\} & \text{si } x > 0, \\ 0, & \text{si } x \leq 0. \end{cases}$$

On notera que la fonction de densité d'une loi $\Lambda(\mu, \sigma)$ s'écrit

$$f_{\Lambda(\mu, \sigma)}(x) = e^{-\mu} f_{\Lambda(0, \sigma)}(xe^{-\mu}). \quad (2.8)$$

Notons que cette distribution est unimodale et dispose d'un paramètre d'asymétrie positif. Les quantiles de la loi $\Lambda(\mu, \sigma)$ sont donnés par

$$\theta_q = e^{\mu + t_q \sigma}, \quad (2.9)$$

où t_q est tel que

$$q = P(N(0, 1) \leq t_q). \quad (2.10)$$

La distribution lognormale peut représenter des variables aléatoires bornées inférieurement par 0 et n'ayant que peu de grandes valeurs ([Crow et Shimizu \(1988\)](#)). La distribution lognormale est donc appropriée pour la modélisation de phénomènes tels que le poids des adultes, la concentration des minéraux, la durée de vie, la distribution des richesses.

Propriétés 2.2.1. Soit X une variable aléatoire suivant $\Lambda(\mu, \sigma)$, et $Y = \ln X$ une variable aléatoire de distribution $N(\mu, \sigma^2)$. Alors

$$\acute{\mu}_i = \exp(i\mu + 1/2(i\sigma)^2) \quad (2.11)$$

$$E[X] = e^{\mu+1/2\sigma^2} \quad (2.12)$$

$$V[X] = e^{2\mu+\sigma^2}(e^{\sigma^2} - 1) \quad (2.13)$$

$$Med(X) = e^\mu \quad (2.14)$$

$$Mode(X) = e^{\mu-\sigma^2}. \quad (2.15)$$

Corollaire 2.2.1.

$$Mode(X) < Med(X) < E(X)$$

$$\beta_1 = (w - 1)(w + 2)^2, \quad (2.16)$$

$$\beta_2 = w^4 + 2w^3 + 3w^2 - 3, \quad (2.17)$$

où $w = e^{\sigma^2}$.

Communément, σ est dénoté paramètre de forme de la loi lognormale. Plusieurs mesures d'asymétrie sont considérées par Nichols et Gibbons (1979).

Une lognormale standardisée tend vers $N(0, 1)$ à mesure que σ tend vers 0

$$\begin{aligned} (X - E[X])/V[X]^{1/2} &= U + 1/2(U^2 - 1)\sigma \\ &\quad + 1/12(2U^3 - 9U)\sigma^2 + \varepsilon(\sigma^3), \end{aligned} \quad (2.18)$$

où U suit une $N(0, 1)$.

L'asymétrie est seulement fonction du paramètre σ^2 , et à mesure que σ tend vers 0, on perd cette asymétrie.

2.2.2 La loi de Weibull à deux paramètres

Soient X et Y deux variables aléatoires, telles que $X = \alpha Y^{1/\beta}$, où β et α sont positifs. X est dite de distribution loi de Weibull si, et seulement si, Y suit une loi exponentielle avec densité

$$f(y) = e^{-y}, y \geq 0. \quad (2.19)$$

La fonction de densité de X est

$$f(x) = \exp\left\{-(x/\alpha)^\beta\right\} \beta x^{\beta-1}/\alpha^\beta, x \geq 0. \quad (2.20)$$

Ses fonctions de répartition et de survie sont, respectivement,

$$F(x) = 1 - e^{-(x/\alpha)^\beta}, x \geq 0 \quad (2.21)$$

$$S(x) = 1 - F(x) = e^{-(x/\alpha)^\beta}, x \geq 0. \quad (2.22)$$

Sa fonction de risque est

$$h(x) = f(x)/S(x) = \beta x^{\beta-1}/\alpha^\beta, x \geq 0. \quad (2.23)$$

Notons que la fonction de risque est décroissante pour $\beta < 1$, constante pour $\beta = 1$, et croissante pour $\beta > 1$.

L'espérance de vie résiduelle, obtenue à partir de la fonction de survie, est

$$e_X = \int_x^\infty S(u)/S(x) du = e^{(x/\alpha)^\beta} \int_x^\infty e^{-(u/\alpha)^\beta} du. \quad (2.24)$$

Propriétés 2.2.2. On désignera par $W(\alpha, \beta)$ la loi de Weibull de paramètres β et α , et par $Exp(\alpha)$ l'exponentielle de paramètre α .

$$\mu_i = \alpha^i \Gamma(1 + i/\beta) \quad (2.25)$$

$$E[X] = \alpha \Gamma(1 + 1/\beta) \quad (2.26)$$

$$V[X] = \alpha^2 \Gamma(1 + 2/\beta) - (\alpha \Gamma(1 + 1/\beta))^2 \quad (2.27)$$

$$Med(X) = \alpha (\ln 2)^{1/\beta} \quad (2.28)$$

$$Mode(X) = \begin{cases} \alpha(1 - \frac{1}{\beta})^{1/\beta} & , \quad si \beta > 1, \\ 0 & , \quad si \beta \leq 1. \end{cases}$$

On note que $W(\alpha, \beta) \sim \alpha W(1, \beta)$, et $W(\alpha, 1) \sim Exp(\alpha)$.

Propriétés 2.2.3. Pour $0 < \beta \leq 1$:

1. à mesure que $x \rightarrow 0$, $f(x) \rightarrow \infty$.
2. le mode n'existe pas (est à 0).

Propriétés 2.2.4. Pour $c > 1$:

1. $f(x) = 0$, pour $x = 0$.
2. pour $2 < \beta \leq 2.6$, le coefficient d'asymétrie est positif, et ainsi on décèle une asymétrie vers la droite.

3. pour $2.6 < \beta \leq 3.7$, le coefficient d'asymétrie est nul, et donc la loi de Weibull tend vers la Normale.
4. $\beta > 3.7$, le coefficient d'asymétrie est négatif, et on décèle une asymétrie à gauche.

À mesure que α augmente, on observe une asymétrie à droite et une diminution du mode, alors que la diminution de α entraîne une augmentation du mode et une asymétrie prononcée à gauche.

La fonction de survie dépend du paramètre de forme β :

1. elle décroît monotonement et rapidement pour $0 < \beta < 1$,
2. elle décroît monotonement et moins rapidement, dès que $\beta = 1$,
3. elle décroît à mesure que x croît (concave), puis décroît rapidement à partir d'un point d'inflexion (convexe) pour $\beta > 1$.

La fonction de risque dépend du paramètre de forme β :

1. elle décroît pour $\beta < 1$,
2. elle est constante pour $\beta = 1$,
3. elle est croissante et concave pour $1 < \beta < 2$,
4. elle est croissante et convexe pour $\beta > 2$.

2.2.3 La loi gamma

La fonction de densité et la fonction de répartition d'une loi gamma à deux paramètres, $G(\alpha, \beta)$, où α est le paramètre de forme et β est l'inverse du paramètre d'échelle, sont respectivement

$$f(x; \alpha, \beta) = x^{\alpha-1} \frac{\beta^\alpha e^{-\beta x}}{\Gamma(\alpha)}, x > 0, \quad (2.29)$$

$$F(x; \alpha, \beta) = \frac{\Gamma(\alpha, \beta x)}{\Gamma(\alpha)}, x > 0, \quad (2.30)$$

où $\Gamma(\alpha, x) = \int_0^x t^{\alpha-1} e^{-t} dt$ est la fonction gamma incomplète.

Propriétés 2.2.5. Soit une variable aléatoire X suivant une $G(\alpha, \beta)$. Alors

$$\mu_i = E[X^i] = \begin{cases} \frac{\Gamma(\alpha+i)}{\beta^i \Gamma(\alpha)}, & \text{si } i > -\alpha \\ \infty, & \text{sinon} \end{cases}$$

et $kX \sim G(\alpha, k\beta)$.

Soient N variables aléatoires indépendantes X_1, X_2, \dots, X_N , où $X_i \sim G(\alpha_i, \beta)$, pour $i = 1, 2, \dots, N$, et $\bar{\alpha} = \sum_{i=1}^N \alpha_i$. Alors

$$Y = \sum_{i=1}^N X_i \sim G(\bar{\alpha}, \beta)$$

On note la propriété d'échelle de la famille de lois :

Pour $\alpha > 1$:

1. si $\alpha \leq 2$, alors la PDF a un point d'inflexion à $x = (1/\beta)\sqrt{\alpha-1}(\sqrt{\alpha-1}+1)$.
2. si $\alpha > 2$, alors la PDF a deux points d'inflexion à $x = (1/\beta)\sqrt{\alpha-1}(\sqrt{\alpha-1} \pm 1)$.

Pour $\alpha = 1$:

1. la loi gamma se transforme en une loi Exponentielle.
2. la fonction de risque est constante (β).
3. le mode est à 0.

Pour $0 < \alpha < 1$:

1. à mesure que x tend vers l'infini, la fonction de risque tends vers (β).
2. le mode est à 0.

2.2.4 La loi gamma transformée

La loi gamma transformée a été introduite dans le but d'élargir l'éventail de lois dont disposent les statisticiens, surtout pour l'étude des données industrielles, où l'on recherche de meilleurs ajustements que ceux proposés par l'exponentielle et la loi de Weibull. La loi gamma transformée compte ces lois ainsi que d'autres comme cas particuliers, offrant ainsi plus de maniabilité (Stacy et Mihram (1965)).

Considérons la fonction de densité suivante :

$$f(x; p, a, v) = |p|x^{pv-1}e^{-(x/a)^v} / \{a^{pv}\Gamma(p)\}; x \geq 0, p \neq 0, a, v \geq 0, \quad (2.31)$$

Notons que Stacy et Mihram (1965) imposent la restriction de non-négativité au paramètre p (ici $p > 0$), tout en remarquant que Amoroso, exige que $v + p^{-1} > 0$. Ici, nous supposons que $p > 0$. Notons que si une variable aléatoire X suit une loi gamma transformée, la variable X^m suit également une loi gamma transformée.

Posons $F(x)$, la fonction de répartition de la variable aléatoire X suivant une distribution loi gamma transformée. Alors

$$F(x) = \begin{cases} \frac{\Gamma(v,w)}{\Gamma(v)}, & \text{si } p > 0 \\ 1 - \frac{\Gamma_w(v)}{\Gamma(v)}, & \text{si } p < 0, \end{cases}$$

où $w = (x/a)^p$.

On notera $GT(a, v, p)$ la loi gamma transformée.

Propriétés 2.2.6. *Soit une variable aléatoire X suivant une $GT(a, v, p)$. Alors*

$$\mu_i = E[X^i] = \begin{cases} a^i \frac{\Gamma(\frac{pv+i}{p})}{\Gamma(v)}, & \text{si } i/p > -v \\ \infty, & \text{sinon.} \end{cases}$$

Il est clair que tous les moments de X sont définis, du moment que $p > 0$, ce qui explique que les membres de cette famille ont des queues légères. Cependant, le cas où $p < 0$ est plus courant en pratique et plus délicat à traiter, et sera discuté ultérieurement.

Propriétés 2.2.7.

$$\text{Mode}(X) = \begin{cases} a(\frac{pv-1}{v})^{1/v}, & \text{si } pv > 1 \\ 0, & \text{sinon.} \end{cases}$$

On a $kX \sim GT(p, ka, v)$, $k > 0$ et $X^m \sim GT(p/m, a^m, v)$, $m \neq 0$. Et il s'en suit que $(X/a)^p \sim GT(1, 1, v)$ et $(X/a)^{-p} \sim GT(-1, 1, v)$.

Si $p > 0$, le moindre changement du produit pv affecte la forme de la distribution. Pour $x = 0$, la PDF s'annule, ou devient finie et non nulle, ou encore devient non bornée, selon que le produit pv est inférieur, égal, ou supérieur à 1.

La première dérivée du logarithme de $f(x; a, v, p)$ s'annule pour

$$x = a(v - p^{-1})^{1/p} \geq 0. \quad (2.32)$$

Par conséquent, la PDF est continue en x (pour $x > 0$), et $f(x; a, v, p)$ n'a qu'un seul maximum, quel que soit le produit pv .

2.2.5 La loi Bêta transformée (GBII)

La loi Bêta transformée a été introduite dans un ouvrage traitant de l'assurance par Venter (1983), puis par McDonald (1984) dans un ouvrage traitant d'économie. Ensuite, McDonald et Xu (1995) ont développé cette loi, et au cours de la même année, Johnson et al. (1995) lui ont consacré une partie du chapitre 25 de leur ouvrage. Enfin, Gupta et Nadarajah (2004) et [Nadarajah et Kotz \(2005\)](#) ont longuement étudié les propriétés de cette loi.

Notons *GBII* La loi Bêta transformée. Soit $X \sim GBII(\alpha, \lambda, \gamma, \tau)$. On a

$$f(x; \alpha, \lambda, \gamma, \tau) = \frac{\Gamma(\alpha + \tau)}{\Gamma(\alpha)\Gamma(\tau)} \frac{\gamma(x^\gamma/\lambda)^\tau}{x[1 + (x^\gamma/\lambda)]^{\alpha+\tau}},$$

$$\alpha > 0, \lambda > 0, \gamma > 0, \tau > 0, x > 0. \quad (2.33)$$

Alors

$$F(x) = B(\tau, \alpha; u), \quad (2.34)$$

où $u = \frac{x^\gamma/\lambda}{1+x^\gamma/\lambda}$, et $B(\tau, \alpha; u)$ est la fonction bêta incomplète

$$B(\tau, \alpha; u) = \int_0^u t^{\tau-1}(1-t)^{\alpha-1} dt. \quad (2.35)$$

Propriétés 2.2.8. Soit une variable aléatoire X suivant une $GBII(\alpha, \lambda, \gamma, \tau)$. Alors

$$\mu_i = \frac{\lambda^{i/\gamma} \Gamma(\tau + i/\gamma) \Gamma(\alpha - i/\gamma)}{\Gamma(\alpha) \Gamma(\tau)}, \quad -\tau\gamma < i < \alpha\gamma, \quad (2.36)$$

et

$$Mode(X) = \begin{cases} (\lambda \frac{\tau\gamma-1}{\alpha\gamma+1})^{1/\gamma}, & \text{si } \tau\gamma > 1 \\ 0, & \text{sinon.} \end{cases}$$

Discutons de l'influence des quatre paramètres de La loi Bêta transformée, en analysant les queues à droite et à gauche :

1. Le k^e moment de la *GBII* existe tant que $k < \alpha\gamma$. En conséquence pour tous les membres de la famille, seulement un nombre fini de moments existent. Il en découle que seuls les paramètres α et γ déterminent le comportement de la *GBII* à droite. Dans un premier temps, considérons le ratio suivant :

$$[1 - F(x)]/x^{-\alpha\gamma}. \quad (2.37)$$

À mesure que $x \rightarrow \infty$, le dénominateur ainsi que le numérateur tendent vers 0, ainsi la limite ne peut être déterminée sans le recours à la règle de l'Hôpital :

$$\begin{aligned}
\lim_{x \rightarrow \infty} \frac{[1 - F(x)]}{x^{-\alpha\gamma}} &= \lim_{x \rightarrow \infty} \frac{-f(x)}{-\alpha\gamma x^{-\alpha\gamma-1}} \\
&= \lim_{x \rightarrow \infty} \frac{\Gamma(\alpha + \tau)\gamma x^{\alpha\tau-1}\lambda^{-\tau}}{\Gamma(\alpha)\Gamma(\tau)[1 + \frac{x^\gamma}{\lambda}]^{\alpha+\tau}\alpha\gamma x^{-\alpha\gamma-1}} \\
&= \lim_{x \rightarrow \infty} \frac{\Gamma(\alpha + \tau)}{\Gamma(\alpha)\Gamma(\tau)\lambda^\tau\alpha} \underbrace{\left[\frac{\lambda x^\gamma}{\lambda + x^\gamma} \right]^{\alpha+\tau}}_{\text{}} \quad (2.38)
\end{aligned}$$

À mesure que $x \rightarrow \infty$, le terme souligné tend vers λ . Ainsi,

$$\lim_{x \rightarrow \infty} \frac{[1 - F(x)]}{x^{-\alpha\gamma}} = \frac{\Gamma(\alpha + \tau)\lambda^\alpha}{\Gamma(\alpha)\Gamma(\tau)\alpha}. \quad (2.39)$$

Cette quantité est une constante, et en conséquence, la probabilité dans la queue à droite est approximativement proportionnelle à $x^{-\alpha\gamma}$. Ceci confirme notre hypothèse de départ que le comportement dans la queue de droite est dépendant du produit $\alpha\gamma$. Il s'en suit qu'à mesure que ce produit grandit, la queue devient plus courte et tombe rapidement à 0. À mesure que α tend vers l'infini, on obtient une loi gamma transformée. Tous les moments positifs de la loi gamma transformée existent, ce qui explique que les membres de cette famille disposent de queues plus légères que ceux de la famille Bêta transformée.

2. Les moments négatifs de La loi Bêta transformée existent à mesure que le produit $\gamma\tau$ grandit. Dans un premier temps, considérons le ratio suivant :

$$\begin{aligned}
\lim_{x \rightarrow 0} \frac{F(x)}{x^{\gamma\tau}} &= \lim_{x \rightarrow 0} \frac{f(x)}{-\gamma\tau x^{\gamma\tau-1}} \\
&= \lim_{x \rightarrow 0} \frac{\Gamma(\alpha + \tau)\gamma x^{\gamma\tau-1}\lambda^{-\tau}}{\Gamma(\alpha)\Gamma(\tau)[1 + \frac{x^\gamma}{\lambda}]^{\alpha+\tau}\gamma\tau x^{-\gamma\tau-1}} \\
&= \lim_{x \rightarrow 0} \frac{\Gamma(\alpha + \tau)}{\Gamma(\alpha)\Gamma(\tau)\lambda^\tau\tau} \left[\frac{\lambda}{\lambda + x^\gamma} \right]^{\alpha+\tau} \\
&= \frac{\Gamma(\alpha + \tau)}{\Gamma(\alpha)\Gamma(\tau)\lambda^\tau\tau}. \quad (2.40)
\end{aligned}$$

La probabilité de la queue gauche est approximativement proportionnelle $x^{\gamma\tau}$. Il s'en suit que l'épaisseur de la queue à gauche dépend du produit $\gamma\tau$, et de grandes valeurs de ce produit conduiront à des queues plus minces.

En conclusion, on retiendra que le paramètre γ influe sur les deux queues, celle de droite autant que celle de gauche, et plus ce paramètre augmente, plus les queues deviennent minces. Ainsi, γ contrôle l'aplatissement de la PDF, alors que α et τ contrôlent, respectivement, la forme des queues droite et gauche. Quant au paramètre d'échelle θ , il n'influe pas sur la forme de la distribution.

2.2.6 Liens entre certaines familles

La loi gamma transformée est un cas limite de la *GBII*, tel que

$$GT(x; \alpha, \theta, \tau) = \lim_{\gamma \rightarrow \infty} GBII(x; \alpha, \lambda = \gamma^{1/\alpha} \theta, \gamma, \tau) \quad (2.41)$$

Deux des lois étudiées précédemment, à savoir la loi gamma et la loi de Weibull, sont des cas particuliers de la loi gamma transformée, $GT(\alpha, \theta, \tau)$, lorsque respectivement $\tau = 1$ et $\alpha = 1$.

Enfin, considérons la loi gamma transformée avec la paramétrisation citée plus haut lorsque $\gamma^{-1} \sqrt{\epsilon^\gamma} \rightarrow \sigma$ et $\gamma^{-1}(\epsilon^\gamma \tau - 1) \rightarrow \mu$. Considérons, dans cette nouvelle paramétrisation, la limite $\tau \rightarrow \infty$, $\gamma \rightarrow 0$, $\epsilon \rightarrow 0$, et $\theta = 0$. Le cas limite qui en découle est la loi lognormale $\Lambda(\mu, \sigma)$ (Crow et Shimizu (1988)).

2.3 Statistiques descriptives

Avant de décrire les techniques formelles d'estimation, il est utile de se donner une idée des familles de lois qui peuvent bien s'ajuster aux données. Des tests graphiques simples peuvent parfois aider à ceci.

2.3.1 Histogramme

L'histogramme est une représentation graphique (en tuyaux d'orgue) de la distribution de fréquences d'une variable quantitative. L'histogramme est la fonction de densité correspondant à la fonction de répartition empirique interpolée linéairement. La hauteur du tuyau est proportionnelle à la fréquence de la classe correspondante. Le détail de la méthode est exposée au Johnson *et al.* (1994). Idéalement, l'histogramme permet de distinguer une tendance vers une distribution connue, car sa forme devrait approcher celle de la densité. Il permet ainsi d'orienter la recherche de la loi d'ajustement.

2.3.2 Fonction d'espérance de survie résiduelle

La fonction d'espérance de survie résiduelle permet également, dans certains cas, d'identifier la distribution qui s'ajuste le mieux aux données. La formulation ainsi que l'estimation sont exposées en détail dans [Klein et Moeschberger \(2003\)](#).

L'allure de la courbe de cette fonction peut nous renseigner sur la famille de lois. À titre d'exemple, une forme constante est associée à une exponentielle, alors qu'une courbe croissante peut être associée à la loi lognormale, tandis qu'une courbe décroissante nous suggère une loi gamma ou une loi de Weibull s'il existe une asymptote ([Hogg et Klugman \(1984\)](#)).

2.3.3 Ramlau-Hansen

L'estimateur du maximum de vraisemblance de la fonction de risque cumulé $H(t)$ est l'estimateur de Nelson-Aalen, \tilde{H} :

$$\tilde{H}(x) = \begin{cases} 0 & , \text{ si } x \leq x_{(1)} \\ \sum_{i: x_{(i)} \leq x} \frac{d_i}{Y_i} & , \text{ si } x \geq x_{(1)}, \end{cases}$$

où $x_{(i)}$ est la statistique d'ordre de rang i (i^e plus petite valeur d'un échantillon statistique x_1, \dots, x_n), $d_i = \sum_{l=1}^M I(x_l = x_{(i)})$ est le nombre d'événements à $x_{(i)}$ (en général $d_i = 1$) et $Y_i = \sum_{l=1}^M I(x_l \geq x_{(i)})$. Dans une prochaine étape, on va estimer la fonction de risque $h(s)$ pour s dans une grille $\{s_1, s_2, \dots, s_k\}$.

Ramlau-Hansen (1983) proposa une méthodologie, citée par [Klein et Moeschberger \(2003\)](#), permettant de construire un estimateur de la fonction de risque $\hat{h}(x)$ par lissage. Ce lissage s'opère grâce à des noyaux, en manipulant un paramètre d'ouverture de fenêtre b . Dans notre contexte, cette méthode s'applique à toutes les statistiques d'ordre $x_{(i)} > 0$, où $b \leq x_{(i)} \leq s_k - b$, avec s_k , la borne supérieure d'une séquence de longueur arbitraire k . Ainsi on aboutit à

$$\hat{h}(s_k) = b^{-1} \sum_{i=1}^M K \left(\frac{x_{(i)} - s_k}{b} \right) \Delta \tilde{H}(x_{(i)}), \quad (2.42)$$

où $\Delta \tilde{H}(x_{(i)}) = \tilde{H}(x_{(i)}) - \tilde{H}(x_{(i-1)})$, K est un noyau, $\{s_1, s_2, \dots, s_k\}$ est une séquence de points d'estimation de $h(\cdot)$ de longueur arbitraire k . Dans cette étude, nous utilisons un noyau Epanechnikov :

$$K(x) = \begin{cases} 0.75(1 - x^2) & , \text{ si } -1 \leq x \leq 1 \\ 0 & , \text{ ailleurs.} \end{cases}$$

La taille de l'ouverture de fenêtre sera déterminée de la manière suivante : l'usage juge un paramètre de taille de l'ouverture de fenêtre b adéquat, si ce dernier réduit le bruit sans pour autant l'éliminer. Par contre, un lissage excessif (b trop grand) nous retourne la courbe du noyau, et élimine toute information sur le modèle.

À nouveau, l'allure de la courbe de Ramlau-Hansen est indicatrice d'une famille de lois. Puisque c'est un estimé de la fonction de risque, elle sera par exemple constante dans le cas d'une exponentielle et croissante pour la loi de Weibull.

2.3.4 QQ-plot et PP-plot

Le QQ-plot théorique est un graphe de points dont les abscisses sont les quantiles des données et les ordonnées, les quantiles d'une loi théorique. La fonction quantile empirique d'un échantillon est la fonction quantile de sa distribution empirique.

Définition 2.3.1. Soit $x = (x_1, \dots, x_n)$ un échantillon et $x_{(1)}, \dots, x_{(n)}$ ses statistiques d'ordre. La fonction quantile empirique de l'échantillon est la fonction \hat{Q} qui, pour tout $i = 1, \dots, n$, vaut $x_{(i)}$ sur l'intervalle $]\frac{i-1}{n}, \frac{i}{n}]$:

$$\hat{Q}(u) = x_{(i)}, \forall u \in]\frac{i-1}{n}, \frac{i}{n}]. \quad (2.43)$$

Le QQ-plot est incontestablement un des meilleurs moyens de vérifier l'ajustement et l'adéquation d'une distribution aux données, car il met en évidence l'écart entre le modèle et les données par l'écart entre le graphe et la droite à 45° . En effet, si le graphe se confond avec la première bissectrice, alors on conclut que la distribution s'ajuste bien à nos données.

Le PP-plot est un graphe de points dont les abscisses sont les probabilités associées aux données et les ordonnées, les probabilités d'une loi théorique. La distribution empirique associée à un échantillon est la distribution de probabilité qui affecte à chaque observation un poids $1/n$.

Définition 2.3.2. Soit $x = (x_1, \dots, x_n)$ un échantillon et c_1, \dots, c_k les valeurs distinctes prises par les x_i . Pour $h = 1, \dots, k$, posons

$$n_h = \sum_{i=1}^n 1_{c_h}(x_i), \quad (2.44)$$

le nombre de fois que c_h apparaît dans l'échantillon. La distribution empirique de l'échantillon est la loi de probabilité \hat{P} sur l'ensemble c_1, \dots, c_k , telle que

$$\hat{P}(c_h) = \frac{n_h}{n}. \quad (2.45)$$

Dans certains cas, le PP-plot peut s'avérer plus robuste et plus suggestif que le QQ-plot.

2.4 Aspects inférentiels et numériques

Dans une approche paramétrique, discuter de quantités d'intérêt d'une population s'opère en quatre étapes :

1. Déterminer la famille de distributions qui décrirait le mieux la population à l'étude.
2. Déterminer la valeur des paramètres de cette loi.
3. Valider le choix de la famille.
4. Déterminer la valeur des quantités d'intérêt.
5. Analyser la précision de ces quantités.

L'association d'une valeur à un paramètre d'une loi constitue le processus d'estimation. Ce processus est basé sur un échantillon aléatoire, X_1, \dots, X_n . Dans ce cas, la distribution conjointe de ces n variables est

$$F_{X_1, \dots, X_n}(x_1, \dots, x_n) = F_X(x_1) \cdots F_X(x_n), \quad (2.46)$$

où $F_X(x)$ est la distribution commune à tout l'échantillon.

Les méthodes d'estimation paramétrique peuvent se classer en trois grandes approches. La première est de construire un système avec autant d'équations que d'inconnues (la méthode des moments, par exemple). La deuxième est d'optimiser certaines quantités, qui vont de pair avec notre objectif ultime (méthode de maximum de vraisemblance ou distance minimale, par exemple). Enfin, l'estimation bayésienne est basée sur une autre approche et ne fera pas partie de la discussion qui suivra. Par la suite, on traitera de certaines subtilités relatives aux distributions citées à la section précédente. Dans la première sous-section, on décrira les méthodes des moments, du maximum de vraisemblance et de distance minimale.

2.4.1 Approches d'estimation

Méthode des moments

Soit une famille de distributions à k paramètres. Les équations des moments consistent à imposer que les moments théoriques coïncident avec les moments empiriques :

$$\hat{\mu}_j = E[X^j|\theta] = \frac{1}{n} \sum_{i=1}^n x_i^j, j = 1, \dots, k, \quad (2.47)$$

où θ est le vecteur des k paramètres inconnus. Les estimateurs de la méthode des moments sont les solutions de ces équations pour θ .

Méthode de maximum de vraisemblance

La fonction de vraisemblance de n observations aléatoires d'un échantillon donné est

$$L(\theta|x_1, \dots, x_n) = \prod_{i=1}^n L_i(\theta), \quad (2.48)$$

où $L_i(\theta)$ est la contribution de la i^e observation à la fonction de vraisemblance, et θ est le vecteur des k paramètres inconnus de la famille de distributions retenue. Si la valeur de l'observation provient d'une distribution continue, sa contribution est la fonction de densité à cette valeur, soit

$$L(\theta|x_1, \dots, x_n) = \prod_{i=1}^n f(x_i; \theta). \quad (2.49)$$

De manière imagée, cette fonction donne une probabilité (infinitésimale) de voir l'échantillon se réaliser. L'idée de la méthode d'estimation est de maximiser cette fonction. Notons que rien ne garantit que cette fonction aura un unique maximum. Il faudra donc s'assurer que les méthodes numériques ne nous fournissent pas un extremum local. Ce problème sera davantage traité ultérieurement.

Sous des conditions de régularité ([Hogg et Craig \(1995\)](#)) les estimateurs de maximum de vraisemblance possèdent les propriétés suivantes :

1. Ils sont asymptotiquement non biaisés.
2. De tous les estimateurs asymptotiquement de distribution normale, ils ont la plus petite variance asymptotique, soit

$$Var(\hat{\theta}) = \frac{1}{-E \left[\frac{\partial^2 \ln L(\theta|x)}{\partial \theta^2} \right]}. \quad (2.50)$$

3. L'estimateur de maximum de vraisemblance est invariant sous toute transformation du paramètre. Autrement dit, l'estimateur de maximum de vraisemblance d'une fonction d'un paramètre est ladite fonction évaluée en l'estimateur de maximum de vraisemblance de ce paramètre.
4. La variance asymptotique de l'estimateur de maximum de vraisemblance peut être explicitement obtenue de (2.57). Cependant, l'espérance dans (2.57) peut parfois être difficile à évaluer. On estime alors cette espérance par $-\frac{1}{n} \sum_{i=1}^n \frac{\partial^2 \ln L(\theta|x_i)}{\partial \theta^2}$.

Si la fonction de vraisemblance est dérivable, alors les estimateurs de maximum de vraisemblance $(\theta_1, \dots, \theta_k)$ sont solutions de

$$\frac{\partial}{\partial \theta_i} L(\theta|x) = 0, i = 1, \dots, k. \quad (2.51)$$

Dans la plupart des cas, il s'avère plus simple de maximiser le logarithme népérien de la fonction de vraisemblance,

$$l(\theta|x) = \ln L(\theta|x), \quad (2.52)$$

qui possède les mêmes extrema que $L(\theta|x)$.

Méthodes de distance minimale

Si l'échantillon observé provient d'une certaine loi de probabilité, la distribution empirique de l'échantillon devrait être proche de la distribution théorique de la famille retenue.

La fonction de répartition de l'échantillon (x_1, \dots, x_n) est la fonction en escalier \hat{F} qui vaut 0 avant x_1 , $\frac{i}{n}$ entre x_i et x_{i+1} et ainsi de suite jusqu'à 1 après x_n .

Parmi les moyens de quantifier l'ajustement d'une distribution empirique à une loi théorique, nous retiendrons les mesures suivantes :

Définition 2.4.1. Soit F la fonction de répartition d'une loi de probabilité et \hat{F} la fonction de répartition empirique de l'échantillon (x_1, \dots, x_n) . On appelle distance normée, $L1$:

$$D_{L1} = \sum_{i=1}^n |F(x_i) - \hat{F}(x_i)|. \quad (2.53)$$

Définition 2.4.2. On appelle distance de Cramer-Von Mises, $L2$:

$$D_{L2} = \sum_{i=1}^n (F(x_i) - \hat{F}(x_i))^2. \quad (2.54)$$

Définition 2.4.3. On appelle distance de Kolmogorov-Smirnov, L_∞ :

$$\begin{aligned} D_{L_\infty} &= \sup_x \{|F(x) - \hat{F}(x)|\} \\ &= \max_{i=1}^n \{|F(x_i) - \hat{F}(x_i)|, |F(x_i) - \hat{F}(x_{i-1})|\}, x_0 = -\infty. \end{aligned} \quad (2.55)$$

La méthode consiste à retenir la distribution qui minimise la distance. Elle peut être appliquée avec des poids accordés aux différentes données et cela selon le cas traité et le but poursuivi. En pratique, il est rare que l'on puisse procéder à l'estimation d'un paramètre par ajustement de manière analytique. On doit alors procéder à une minimisation numérique sur les paramètres inconnus. On note, toutefois, que, selon les cas, pour les échantillons de grande taille, la variance de l'estimateur de distance minimale est suffisamment faible et peut se comparer à celle de l'estimateur de maximum de vraisemblance.

2.4.2 Résolution des équations de la méthode des moments

Dans cette section, nous examinons comment les équations de la méthode des moments peuvent être résolues, soit exactement, soit approximativement, pour les familles de lois décrites plus haut. Les autres méthodes doivent toujours être traitées numériquement.

Lois à deux paramètres

La loi gamma

Disposant de la moyenne et de la variance, il est aisé de déduire les paramètres de la loi gamma à deux paramètres par une simple inversion d'équations. Soit θ le paramètre d'échelle et α le paramètre de forme de la loi gamma. Alors

$$\begin{aligned} \theta &= \frac{\mu_2 - \mu_1^2}{\mu_1} \\ &= \frac{Var[X]}{\mu_1}, \end{aligned} \quad (2.56)$$

$$\alpha = \frac{\mu_1}{\theta}. \quad (2.57)$$

On notera aussi $\beta = 1/\theta$.

La loi lognormale

Soient μ et σ les paramètres de la loi lognormale. Alors

$$\begin{aligned}\sigma &= \sqrt{\ln\left(\frac{\mu_2}{\mu_1^2}\right)} \\ &= \sqrt{\ln\left(\frac{\mu_1^2 + \text{Var}[X]}{\mu_1^2}\right)},\end{aligned}\tag{2.58}$$

$$\mu = \ln(\mu_1) - \frac{\sigma^2}{2}.\tag{2.59}$$

La loi de Weibull

Soit α le paramètre d'échelle, et β le paramètre de forme de la loi de Weibull. β est la racine de l'équation

$$\ln\left\{\Gamma\left(1 + \frac{2}{\beta}\right)\right\} - 2\ln\left[\Gamma\left\{\left(1 + \frac{1}{\beta}\right)\right\}\right] + 2\ln(\mu_1) - \ln(\mu_2) = 0,\tag{2.60}$$

alors que $\alpha = C^{-1/\beta}$, où C est la racine de l'équation

$$f(C) = (1/C)^{1/\beta}\Gamma(1 + (1/\beta)) - \mu_1 = 0.\tag{2.61}$$

On doit résoudre numériquement (2.67) pour trouver β . Alors que l'équation (2.68) se résoud facilement de façon analytique une fois (2.67) résolue.

La loi gamma transformée

Soit $X \sim GT(a, v, p)$, $W = (X/a)^p$, et $Z = \ln W$. On a ([Stacy et Mihram \(1965\)](#))

$$\mu_i(Z) = p^i \mu_i(\ln X)\tag{2.62}$$

D'autre part,

$$\begin{aligned}F_Z(z) &= P[Z \leq z] \\ &= P[\ln W \leq z] \\ &= P[W \leq e^z] \\ &= \Gamma(e^z; v)\end{aligned}\tag{2.63}$$

et

$$\begin{aligned}f_Z(z) &= \Gamma'(e^z; \alpha)e^z \\ &= \frac{1}{\Gamma(v)}(e^z)^{v-1}e^{-e^z} \\ &= \exp\{(v-1)z - e^z\}/\Gamma(v).\end{aligned}\tag{2.64}$$

La fonction génératrice des moments de Z est

$$\begin{aligned}
 E[e^{tZ}] &= \int_0^\infty e^{tz} \exp((v-1)z - e^z) / \Gamma(v) dz \\
 &= \int_0^\infty \exp((v+t-1)z - e^z) / \Gamma(v) dz \\
 &= \Gamma(v+t) / \Gamma(v).
 \end{aligned} \tag{2.65}$$

En conséquence, le i^e moment de Z est

$$E[Z^i] = \Gamma^{(i)}(v) / \Gamma(v), \tag{2.66}$$

où $\Gamma^{(i)}(v)$ est la i^e dérivée de $\Gamma(v)$ par rapport à v . En conséquence,

$$\begin{aligned}
 E[Z] &= \Gamma'(v) / \Gamma(v) \\
 &= \Psi(v),
 \end{aligned} \tag{2.67}$$

où $\Psi(v)$ est la dérivée du logarithme de $\Gamma(v)$. Notons $\Psi^{(i)}(v)$ la i^e dérivée de $\Psi(v)$ par rapport à v . On a

$$\begin{aligned}
 \mu_2(Z) &= E[(Z - E[Z])^2] = E[Z^2] - E[Z]^2 \\
 &= \frac{\Gamma^{(2)}(v)}{\Gamma(v)} - \left(\frac{\Gamma'(v)}{\Gamma(v)} \right)^2 \\
 &= \frac{\Gamma^{(2)}(v)\Gamma(v) - \Gamma'(v)^2}{\Gamma(v)^2} \\
 &= \frac{d}{dv} E[Z] = \Psi'(v);
 \end{aligned} \tag{2.68}$$

$$\begin{aligned}
 \mu_3(Z) &= E[(Z - E[Z])^3] \\
 &= E[Z^3] - 3E[Z]E[Z^2] + 2E[Z]^3 \\
 &= \frac{\Gamma^{(3)}(v)}{\Gamma(v)} - 3 \frac{\Gamma'(v)\Gamma^{(2)}(v)}{\Gamma(v)^2} + 2 \left(\frac{\Gamma'(v)}{\Gamma(v)} \right)^3 \\
 &= \frac{\Gamma^{(3)}(v)\Gamma(v)^2 - 3\Gamma^{(2)}(v)\Gamma(v)\Gamma'(v) + 2\Gamma'(v)^3}{\Gamma(v)^3} \\
 &= \frac{d^2}{dv^2} E[Z] = \Psi^{(2)}(v);
 \end{aligned} \tag{2.69}$$

$$\begin{aligned}
 \beta_1 &= \mu_3(Z) / (\mu_2(Z))^{3/2} \\
 &= \mu_3(\ln X) / (\mu_2(\ln X))^{3/2} \\
 &= \Psi^{(2)}(v) / \Psi'(v)^{3/2}.
 \end{aligned} \tag{2.70}$$

Si $\mu_3(Z) < 0$ et n'est pas trop grand comparé à $\mu_2(Z)$, soit \hat{v} la racine de l'équation minimisant la distance entre le coefficient d'asymétrie théorique et le coefficient d'asymétrie

empirique. Il en découle que

$$pE[(\ln X) - \ln a] = \Psi(v), \quad (2.71)$$

$$p^2\mu_2(\ln X) = \Psi'(v), \quad (2.72)$$

$$p^3\mu_3(\ln X) = \Psi^{(2)}(v). \quad (2.73)$$

D'où

$$\hat{p} = (\Psi'(\hat{v})/\mu_2(\ln X))^{1/2} \quad (2.74)$$

$$\hat{a} = 1/\exp\left(\frac{1}{\hat{p}}\Psi(\hat{v}) - E[\ln X]\right). \quad (2.75)$$

La loi Bêta transformée

Soit $X \sim GBII(\alpha, \lambda, \gamma, \tau)$, le i^e moment de X est :

$$E[X^i] = \frac{\lambda^{i/\gamma}\Gamma(\tau + i/\gamma)\Gamma(\alpha - i/\gamma)}{\Gamma(\alpha)\Gamma(\tau)}, \quad -\tau\gamma < i < \alpha\gamma. \quad (2.76)$$

Les équations des moments ne peuvent être résolues analytiquement. Nous devons procéder de manière numérique, par exemple en minimisant simultanément la distance entre les quatre premiers moments empiriques et les quatre moments théoriques. Les approches numériques énumérées seront implantées avec le logiciel R.

Chapitre 3

Copules

Le concept de copule a été introduit par [Sklar \(1959\)](#). D'après [Favre \(2000\)](#), Fréchet et Schweizer y font mention dans le cadre de leurs travaux dénommés «Espaces métriques aléatoires», tel que [Schweizer et Sklar \(1983\)](#). Ce concept sera repris par [Kimeldorf et Sampson \(1975\)](#). Dans les années 1980, Genest et son équipe ([Genest \(1986\)](#), [Genest \(1987\)](#)) se sont penchés sur l'étude du concept.

[Embrechts *et al.* \(1999\)](#) introduisent la notion des copules en finance. Depuis, les copules sont devenues un outil incontournable en modélisation des dépendances pour les données financières. Elles se sont forgé une place en finance et en actuariat, et cela, à cause de la flexibilité et de la robustesse qu'elles offrent lors de modélisation des dépendances entre les risques. Elles sont une des principales approches pour parvenir à l'agrégation des risques individuels en gestion de risque, [Bouyé *et al.* \(2001\)](#).

3.1 Distributions multivariées

Soit un vecteur aléatoire de dimension n , $\mathbf{X} = (X_1, \dots, X_n)$. La fonction $F(x_1, \dots, x_n) = P[X_1 \leq x_1, \dots, X_n \leq x_n]$ est la fonction de répartition conjointe des n variables formant le vecteur aléatoire \mathbf{X} . Notons $\lim_{\{x_j \rightarrow \infty\}, j=1, \dots, n, j \neq i} F(x_1, \dots, x_i, \dots, x_n) = F_{X_i}(x_i) = P[X_i \leq x_i] = F(\infty, \dots, \infty, x_i, \infty, \dots, \infty)$, la fonction de répartition marginale de X_i , $i = 1, \dots, n$.

Définition 3.1.1. *On dit que les variables aléatoires X_1, \dots, X_n formant le vecteur aléatoire \mathbf{X} sont mutuellement indépendantes, si et seulement si pour tout (x_1, \dots, x_n)*

dans \mathfrak{R}^n , on a

$$F(x_1, \dots, x_n) = F_{X_1}(x_1) \cdots F_{X_n}(x_n). \quad (3.1)$$

Définition 3.1.2. Pour toute fonction de répartition multivariée F ayant des marginales F_{X_1}, \dots, F_{X_n} , on appelle bornes de Fréchet-Hoeffding les bornes des inégalités suivantes, qui sont toujours vraies :

$$\max \left\{ \sum_{i=1}^n F_{X_i}(x_i) + 1 - n, 0 \right\} \leq F(x_1, \dots, x_n) \leq \min \{ F_{X_1}(x_1), \dots, F_{X_n}(x_n) \}$$

Afin d'étudier les propriétés des copules, définissons :

– la copule antimonotone, ou copule du maximum

$$W^n(u) = \max \left\{ \sum_{i=1}^n u_i - n + 1, 0 \right\},$$

– la copule produit, ou copule indépendante

$$\Pi^n(u) = \prod_{i=1}^n u_i,$$

– la copule comonotone, ou copule du minimum

$$M^n(u) = \min(u_1, \dots, u_n).$$

La copule comonotone est une copule du vecteur (X_1, \dots, X_n) si et seulement s'il existe des transformations croissantes $g_{i,j}$ telles que $X_i = g_{i,j}(X_j)$. Cette copule correspond à la borne supérieure de Fréchet-Hoeffding. La copule antimonotone, quant à elle, correspond à la borne inférieure de Fréchet-Hoeffding.

3.2 Copules Multivariées

Pour tout entier n , on notera $\bar{\mathfrak{R}}^n$ l'espace $\bar{\mathfrak{R}} \times \dots \times \bar{\mathfrak{R}}$. On utilisera une notation vectorielle pour un point de $\bar{\mathfrak{R}}^n$, tel que $\mathbf{a} = (a_1, \dots, a_n)$. D'autre part, on notera $\mathbf{a} \leq \mathbf{b}$, si, pour tout k , $a_k \leq b_k$. Soit $B = [\mathbf{a}, \mathbf{b}]$ le produit cartésien de n intervalles fermés $[a_1, b_1] \times \dots \times [a_n, b_n]$, qu'on appellera n -boîte. Soit les points $\mathbf{c} = (c_1, \dots, c_n)$ où chaque

c_k est égal soit à a_k , soit à b_k ; on appellera ces points les sommets de la n -boîte B . On définira le n -cube unitaire I^n par le produit $I \times \dots \times I$, avec $I = [0, 1]$.

Soit une fonction multivariée H , définie sur D_H , un sous-ensemble de \mathfrak{R}^n , et d'image $Ran(H)$, un sous-ensemble de R .

Définition 3.2.1. *Une copule est une fonction de répartition multivariée définie de I^n vers I , et dont les lois marginales sont uniformes sur I .*

Une copule C possède les trois propriétés suivantes :

1. $C(u_1, \dots, u_n)$ est croissante en chacune de ses composantes u_k .
2. $C(1, \dots, 1, u_k, 1, \dots, 1) = u_k, \forall k \in \{1, \dots, n\}$.
3. Pour tout (a_1, \dots, a_n) et (b_1, \dots, b_n) de I , où $a_k \leq b_k$, on a :

$$\sum_{i_1=1}^2 \dots \sum_{i_n=1}^2 (-1)^{i_1+\dots+i_n} C(u_{1i_1}, \dots, u_{ni_n}) \geq 0, \quad (3.2)$$

où $u_{k1} = a_k$ et $u_{k2} = b_k, \forall k \in \{1, \dots, n\}$.

Le théorème de Sklar en dimension n explique l'intérêt de la notion de copule.

Théorème 3.2.1 (Théorème de Sklar). *Pour toute fonction de répartition multivariée H de marginales F_{X_1}, \dots, F_{X_n} , il existe au moins une copule C telle que*

$$H(x_1, \dots, x_n) = C(F_{X_1}, \dots, F_{X_n}), x \in [-\infty, \infty]^n. \quad (3.3)$$

Si les marginales sont continues, la copule C est unique. Dans les cas contraire, C est uniquement déterminée sur $Ran(F_{X_1}) \times \dots \times Ran(F_{X_n})$.

Soit une copule C et n fonctions de répartition F_{X_1}, \dots, F_{X_n} . Si une fonction H est définie telle que

$$H(x_1, \dots, x_n) = C(F_{X_1}, \dots, F_{X_n}), x \in [-\infty, \infty]^n,$$

alors H est une fonction de répartition multivariée, avec pour marginales F_{X_1}, \dots, F_{X_n} .

Définition 3.2.2. *Posons une fonction de répartition F . Alors F^- est dite une fonction inverse généralisée de F définie sur I , si*

1. $\forall t \in Ran(F), F(F^-(t)) = t$;
2. ailleurs, $F^-(t) = \inf\{x : F(x) \geq t\} = \sup\{x : F(x) \leq t\}$.

Si F est strictement croissante, alors il existe une seule fonction inverse généralisée, qu'on notera F^{-1} .

Corollaire 3.2.1. *Soit une fonction de répartition multivariée H , une copule C , n marginales F_{X_1}, \dots, F_{X_n} , et leurs fonctions inverses généralisées $F_{X_1}^-, \dots, F_{X_n}^-$. Alors, pour tout vecteur $u \in I^n$, on a*

$$C(u_1, \dots, u_n) = H(F_{X_1}^-(u_1), \dots, F_{X_n}^-(u_n)). \quad (3.4)$$

Définition 3.2.3. *La densité de copule, qu'on notera c , est définie par*

$$c(u_1, \dots, u_n) = \frac{\partial C(u_1, \dots, u_n)}{\partial u_1 \dots \partial u_n}. \quad (3.5)$$

Théorème 3.2.2. *Soient n variables aléatoires X_1, \dots, X_n , de fonctions de répartition marginales F_{X_1}, \dots, F_{X_n} , et de fonctions de densité marginales f_{X_1}, \dots, f_{X_n} . Notons f la fonction de densité conjointe. Alors :*

$$f(x_1, \dots, x_n) = c(F_{X_1}(x_1), \dots, F_{X_n}(x_n)) \prod_{i=1}^n f_{X_i}(x_i). \quad (3.6)$$

Théorème 3.2.3. *Pour tout $n \geq 2$, soit n variables aléatoires continues X_1, \dots, X_n . Alors X_1, \dots, X_n sont indépendantes si et seulement si la n -copule reliant ces variables aléatoires est Π^n .*

Théorème 3.2.4 (Propriété d'invariance). *Soit n fonctions strictement croissantes T_1, \dots, T_n . Alors les variables aléatoires X_1, \dots, X_n et les variables aléatoires transformées $T_1(X_1), \dots, T_n(X_n)$ ont la même copule.*

3.3 Copules et dépendance

De par sa propriété d'invariance à toute transformation strictement croissante, la copule capte l'essentiel de la dépendance des distributions conjointes. Les propriétés de dépendance et les mesures d'association sont inter-reliées, et dès lors, il y a plusieurs manières d'approcher le sujet. Le rho de Spearman et le tau de Kendall sont deux mesures d'une forme de dépendance connue sous le nom de «concordance».

3.3.1 Concordance

Deux paires de variables aléatoires sont concordante si de grandes valeurs de l'une sont associées à de grandes valeurs de l'autre, et de même, des petites valeurs de l'une sont associées à de petites valeurs de l'autre.

Considérons deux couples d'observations (x_i, y_i) et (x_j, y_j) issus de deux variables aléatoires X et Y . Si $x_i < x_j$ et $y_i < y_j$, ou $x_i > x_j$ et $y_i > y_j$, alors les deux paires sont dites concordantes. Dans le cas contraire, si $x_i < x_j$ et $y_i > y_j$, ou $x_i > x_j$ et $y_i < y_j$, on parle de discordance. Une formulation alternative de concordance est $(x_i - x_j)(y_i - y_j) > 0$, alors que pour la discordance, on a $(x_i - x_j)(y_i - y_j) < 0$.

3.3.2 Tau de Kendall

Soient les variables aléatoires, (X_1, Y_1) et (X_2, Y_2) indépendantes et identiquement distribuées comme (X, Y) . Le τ de Kendall est la probabilité de concordance moins celle de discordance :

$$\tau(X, Y) = P[(X_1 - X_2)(Y_1 - Y_2) > 0] - P[(X_1 - X_2)(Y_1 - Y_2) < 0] \quad (3.7)$$

Il s'en suit que

1. $-1 \leq \tau \leq 1$;
2. si, et seulement si les variables X et Y sont comonotones (il existe une fonction g strictement croissante telle que $X = g(Y)$), alors $\tau = 1$;
3. si, et seulement si les variables X et Y sont antimonotones (il existe une fonction g strictement décroissante telle que $X = g(Y)$), alors $\tau = -1$;
4. si les variables X et Y sont indépendantes, alors $\tau = 0$. Par contre, si $\tau = 0$, les variables X et Y ne sont pas forcément indépendantes ;
5. si T_1 et T_2 sont deux fonctions strictement croissantes, alors $\tau(T_1(X), T_2(Y)) = \tau(X, Y)$;
6. τ ne dépend que de la copule de (X, Y) . Si (X, Y) est de copule C , alors

$$\tau(X, Y) = 4 \int_0^1 \int_0^1 C(u, v) C(du, dv) - 1 = 4E[C(U, V)] - 1. \quad (3.8)$$

3.3.3 Rho de Spearman

Soient les variables aléatoires, (X_1, Y_1) , (X_2, Y_2) et (X_3, Y_3) indépendantes et identiquement distribuées comme (X, Y) . Le ρ de Spearman se définit comme suit :

$$\rho(X, Y) = 3(P[(X_1 - X_2)(Y_1 - Y_3) > 0] - P[(X_1 - X_2)(Y_1 - Y_3) < 0]). \quad (3.9)$$

Les propriétés du ρ de Spearman sont les mêmes que celle du τ de Kendall :

1. $-1 \leq \rho \leq 1$;
2. si et seulement si les variables X et Y sont comonotones, alors $\rho = 1$;
3. si et seulement si les variables X et Y sont antimonotones, alors $\rho = -1$;
4. si les variables X et Y sont indépendantes, alors $\rho = 0$;
5. si T_1 et T_2 sont deux fonctions strictement croissantes, alors $\rho(T_1(X), T_2(Y)) = \rho(X, Y)$;
6. ρ ne dépend que de la copule de (X, Y) . En effet, soit deux variables aléatoires X et Y , de distributions marginales F_X et F_Y continues, et avec une copule unique C . Dès lors, ρ s'écrit comme suit :

$$\rho(X, Y) = 12 \int_0^1 \int_0^1 uv C(du, dv) - 3, \quad (3.10)$$

que l'on peut ré-écrire :

$$\begin{aligned} \rho(X, Y) &= 12E[UV] - 3 = \frac{E[UV] - 1/4}{1/12} \\ &= \frac{E[UV] - E[U]E[V]}{\sqrt{\text{Var}(U)\text{Var}(V)}}. \end{aligned} \quad (3.11)$$

3.4 Familles de copules

3.4.1 Copules de valeurs extrêmes

Définition 3.4.1. Une copule est dite de valeurs extrêmes, que l'on notera copule EV, si elle vérifie la relation suivante :

$$C(u_1^m, \dots, u_n^m) = C^m(u_1, \dots, u_n), \forall m \in \mathbb{R}^+. \quad (3.12)$$

Dans ce qui suit, lors de la discussion des copules archimédiennes, on abordera la copule de **Gumbel**, qui est un exemple de **copule de valeurs extrêmes**.

3.4.2 Copules Archimédiennes

Définition 3.4.2. Soient un réel u , tel que $0 \leq u \leq 1$, et $\varphi(u)$ une fonction qui vérifie $\varphi(1) = 0$, $\varphi'(u) < 0$, et $\varphi''(u) > 0$ (φ est convexe). Alors, $\varphi(u)$ est appelé la **fonction génératrice** de copule, ou plus simplement **générateur**.

Définition 3.4.3. Soit φ une fonction continue, et strictement décroissante définie de $[0, 1]$ vers $[0, \infty]$, vérifiant les propriétés de la **fonction génératrice** de copule. La fonction pseudo-inverse de φ , notée $\varphi^{[-1]}$, avec un domaine de définition $D_{\varphi^{[-1]}} = [0, \infty]$ et $\text{Ran}(\varphi^{[-1]}) = [0, 1]$, est définie comme suit :

$$\varphi^{[-1]}(t) = \begin{cases} \varphi^{-1}(t), & \text{si } 0 \leq t \leq \varphi(0) \\ 0, & \text{si } \varphi(0) \leq t \leq \infty. \end{cases}$$

On note que $\varphi^{[-1]}$ est continue, non-croissante sur $[0, \infty]$, et strictement décroissante sur $[0, \varphi(0)]$. D'autre part, on a

$$\varphi(\varphi^{[-1]}(t)) = \begin{cases} t, & \text{si } 0 \leq t \leq \varphi(0) \\ \varphi(0), & \text{si } \varphi(0) \leq t \leq \infty \end{cases}$$

Ainsi, si $\varphi(0) = \infty$, alors $\varphi^{-1} = \varphi^{[-1]}$.

Théorème 3.4.1. Soit φ une fonction continue, et strictement croissante définie de I vers $[0, \infty]$, vérifiant les propriétés de la **fonction génératrice** de copule, et $\varphi^{[-1]}$ la fonction pseudo-inverse de φ . Alors

$$C(u_1, \dots, u_n) = \begin{cases} \varphi^{[-1]}(\varphi(u_1), \dots, \varphi(u_n)), & \text{si } \sum_{i=1}^n \varphi(u_i) \leq \varphi(0) \\ 0, & \text{sinon} \end{cases}$$

est une copule.

Ce type de copule est dit **copule archimédienne**, de générateur φ . Cette classe de copule possède de nombreuses propriétés intéressantes, telles que

1. la symétrie

$$\text{Pour tous } x, y \in [0, 1], \quad C(x, y) = C(y, x);$$

2. l'associativité

$$\text{Pour tous } x, y, z \in [0, 1], \quad C(C(x, y), z) = C(x, C(y, z));$$

3. la maniabilité pour le calcul.

Le taux de Kendall de cette classe est de la forme

$$\tau = 1 + 4 \int_0^1 \frac{\varphi(u)}{\varphi'(u)} du. \quad (3.13)$$

Dans ce qui suit, l'attention sera portée sur trois copules de la famille archimédienne : la **Clayton**, la **Frank**, et la **Gumbel**.

Copule de Clayton

Cette copule est générée par

$$\varphi(t) = \frac{1}{\theta}(t^{-\theta} - 1), \theta \in [-1, \infty) \setminus \{0\}. \quad (3.14)$$

Ainsi, on obtient la copule de Clayton :

$$C_{Clayton}(u_1, \dots, u_n) = \left(1 - n + \sum_{i=1}^n u_i^{-\theta}\right)^{-1/\theta}, \forall \theta > 0 \quad (3.15)$$

On peut en tirer comme cas limites :

$$\begin{aligned} \lim_{\theta \rightarrow -1} C_{Clayton} &= W^n; \\ \lim_{\theta \rightarrow 0} C_{Clayton} &= \Pi^n; \\ \lim_{\theta \rightarrow \infty} C_{Clayton} &= M^n. \end{aligned}$$

[Cuvelier \(2003\)](#) ont démontré qu'il est facile de déduire la fonction de densité de cette copule :

$$c(u_1, \dots, u_n) = \left(1 - n + \sum_{i=1}^n u_i^{-\theta}\right)^{-n-\frac{1}{\theta}} \prod_{i=1}^n (u_i^{-\theta-1} \{(i-1)\theta + 1\}). \quad (3.16)$$

Le taux de Kendall de cette copule est

$$\tau_{Clayton}(\theta) = \frac{1}{\frac{2}{\theta} + 1} = \frac{\theta}{2 + \theta}. \quad (3.17)$$

Copule de Frank

Cette copule est générée par

$$\varphi(t) = -\ln\left(\frac{e^{-\theta t} - 1}{e^{-\theta} - 1}\right), \theta \in (-\infty, \infty) \setminus \{0\}. \quad (3.18)$$

et φ^{-1} est donnée par

$$\varphi^{-1}(t) = -\frac{1}{\theta} \ln[1 - (1 - e^{-\theta})e^{-t}]. \quad (3.19)$$

Ainsi, on obtient la copule de Frank :

$$C_{Frank}(u_1, \dots, u_n) = -\frac{1}{\theta} \left\{ 1 + \frac{(e^{-\theta u_1} - 1) \dots (e^{-\theta u_n} - 1)}{(e^{-\theta} - 1)^{n-1}} \right\}. \quad (3.20)$$

On peut en tirer comme cas limites :

$$\begin{aligned} \lim_{\theta \rightarrow -\infty} C_{Frank} &= W^n; \\ \lim_{\theta \rightarrow 0} C_{Frank} &= \Pi^n; \\ \lim_{\theta \rightarrow \infty} C_{Frank} &= M^n. \end{aligned}$$

Le taux de Kendall de cette copule est

$$\tau_{Frank}(\theta) = 1 - \frac{4}{\theta} + \frac{4}{\theta^2} \int_0^\theta \frac{t}{e^t - 1} dt. \quad (3.21)$$

Copule de Gumbel

Cette copule est générée par

$$\varphi(t) = (-\ln(t))^\theta, \theta \in [1, \infty), \quad (3.22)$$

et

$$\varphi^{-1}(t) = \exp(-t^{1/\theta}). \quad (3.23)$$

Ainsi, on obtient la copule de Gumbel :

$$C_{Gumbel}(u_1, \dots, u_n) = \exp(-[\{-\ln(u_1)\}^\theta \dots \{-\ln(u_n)\}^\theta]). \quad (3.24)$$

On peut en tirer comme cas limites :

$$\begin{aligned} \lim_{\theta \rightarrow 1} C_{Gumbel} &= \Pi^n; \\ \lim_{\theta \rightarrow \infty} C_{Gumbel} &= M^n. \end{aligned}$$

Le taux de Kendall de cette copule est

$$\tau_{Gumbel}(\theta) = 1 - \frac{1}{\theta}. \quad (3.25)$$

3.4.3 La distribution multivariée de t et sa copule

Dans cette section, on définira la copule de t et d'introduire, pour cela, la distribution de t multivariée.

La copule de t, ou la **Fang et Fang** (Demarta et McNeil (2004)) peut être introduite à partir de la structure de dépendance sous-jacente à la distribution de t multivariée. Cette copule a reçu énormément d'attention en finance. Plusieurs auteurs, dont Demarta et McNeil (2004), ont démontré que cette copule est souvent appropriée pour la modélisation du phénomène de valeurs extrêmes dépendantes, souvent observé dans les données financières.

Distribution multivariée de t

Un vecteur aléatoire de dimension n , $\mathbf{X} = (X_1, \dots, X_n)'$, est dit de distribution t multivariée de degrés de liberté v , de vecteur-moyenne μ , et de matrice de dispersion définie positive Σ , dénoté $\mathbf{X} \sim t_n(v, \mu, \Sigma)$, si sa fonction de densité est de la forme

$$f(\mathbf{x}) = \frac{\Gamma(\frac{v+n}{2})}{\Gamma(\frac{v}{2})\sqrt{(\pi v)^n |\Sigma|}} \left\{ 1 + \frac{(\mathbf{x} - \mu)' \Sigma^{-1} (\mathbf{x} - \mu)}{v} \right\}^{-\frac{v+n}{2}}, x \in \mathfrak{R}. \quad (3.26)$$

Notons, pour tout $v > 2$, que dans cette paramétrisation standardisée, $cov(\mathbf{X}) = \frac{v}{v-2} \Sigma$. Dès lors, la matrice de covariance n'est pas égale à Σ , et ce tant que $v > 2$.

D'autre part, la distribution de t multivariée fait partie de la classe des mélanges de la loi normale multivariée, car elle peut-être représentée ainsi :

$$\mathbf{X} = \mu + \sqrt{W} \mathbf{Z}, \quad (3.27)$$

où $\mathbf{Z} \sim N_n(\mathbf{0}, \Sigma)$, et W est indépendante de \mathbf{Z} et telle que $v/W \sim \chi_v^2$, c'est-à-dire $W \sim Ig(v/2, v/2)$ où Ig dénote la distribution gamma inverse.

La copule correspondante sera traitée dans la prochaine section.

Copule de t

Soit \mathbf{P} la matrice de corrélation induite par la matrice de dispersion Σ . On notera la distribution de t multivariée équivalente $t_n(v, \mathbf{0}, \mathbf{P})$. Alors, la copule de t s'écrit comme

suit :

$$C_{v,\mathbf{P}}^t(\mathbf{u}) = \int_{-\infty}^{t_v^{-1}(u_1)} \cdots \int_{-\infty}^{t_v^{-1}(u_n)} \frac{\Gamma(\frac{v+n}{2})}{\Gamma(\frac{v}{2})\sqrt{(\pi v)^n |\mathbf{P}|}} \left\{ 1 + \frac{\mathbf{x}/\mathbf{P}^{-1}\mathbf{x}}{v} \right\}^{-\frac{v+n}{2}} dx, \quad u \in I^n, \quad (3.28)$$

où t_v^{-1} la fonction de quantile de la Student univariée standard à v degrés de liberté. Si, toutefois, v tend vers l'infini, alors la copule de t tend vers une copule de distribution gaussienne multivariée. La densité de la copule de t peut être facilement déduite :

$$c_{v,\mathbf{P}}^t(\mathbf{u}) = \frac{f_{v,\mathbf{P}}\{t_v^{-1}(u_1), \dots, t_v^{-1}(u_n)\}}{\prod_{i=1}^n f_v(t_v^{-1}(u_i))}, \quad u \in I^n, \quad (3.29)$$

où $f_{v,\mathbf{P}}$ est la densité conjointe du vecteur aléatoire de distribution $t_n(v, \mathbf{0}, \mathbf{P})$ et f_v est la densité de la t univariée standard à v degrés de liberté.

La relation reliant le coefficient de corrélation de Pearson, que nous noterons r , au τ pour cette classe de copule est

$$\tau(X, Y) = \frac{2}{\pi} \arcsin(r). \quad (3.30)$$

3.5 Simulation des copules

3.5.1 Copules archimédiennes

[Genest \(1986\)](#) proposent un algorithme pour la simulation des copules archimédiennes. Ils prônent l'idée de simuler la distribution conjointe du vecteur aléatoire \mathbf{X} en procédant par des simulations récursives des distributions conditionnelles de X_j sachant X_i . Cet algorithme, tel que résumé par [Lee \(1993\)](#), est

1. Générer un nombre aléatoire uniformément distribué U_1 ;
2. Poser $X_1 = F_1^{-1}(U_1)$;
3. Pour $j = 2, \dots, n$, on calcule récursivement :

$$U_j = F_j(X_j | x_1, \dots, x_{j-1}) = \frac{\varphi^{-1(j-1)}\{c_{j-1} + \varphi[F_j(x_j)]\}}{\varphi^{-1(j-1)}(c_{j-1})},$$

où $c_j = \varphi[F_1(x_1)] + \dots + \varphi[F_j(x_j)]$, et $\varphi^{-1(j)}$ est la j^e dérivée de la fonction inverse de la fonction φ .

4. Poser $X_j = F_j^{-1}(U_j)$;

3.5.2 Copule de t

Afin de simuler des vecteurs aléatoires \mathbf{X} ayant une copule de t avec un certain nombre de degrés de liberté, dl , et de dimension n , on propose l'algorithme suivant (Demarta et McNeil (2004)) :

1. Générer \mathbf{Z} , une réalisation de la loi normale multivariée de moyenne $\mathbf{0}_{n \times 1}$ et de variance $\Sigma_{n \times n}$ avec $\Sigma_{ii} = 1$ $\Sigma_{ij} = r$;
2. Générer $V \sim \chi_{dl}^2$;
3. Dédire $W = dl/V$;
4. Poser $\mathbf{T} = \sqrt{W}\mathbf{Z}$;
5. Soit t_{dl} la distribution de t multivariée avec dl degrés de liberté, de fonction de répartition marginale (Student) F_T . Soit T_j le j^e élément de \mathbf{T} On calcule la variable aléatoire U_j :

$$U_j = F_T(T_j);$$

6. Calculer $X_j = F_j^{-1}(U_j)$, où F_j est la fonction de répartition de X_j .

Chapitre 4

Modélisation des portefeuilles

Ce chapitre traite de la conception des portefeuilles, et fait appel à des techniques et des notions exposées aux chapitres précédents.

Dans un premier temps, on discutera de la simulation des rendements d'un nombre n de titres individuels. M simulations seront effectuées. Le rendement minimal sera -1 puisqu'on ne permettra pas les ventes à découvert. On modélise donc des variables aléatoires à support $[-1, \infty)$. Dans un deuxième temps, les rendements des n titres, mis côte à côte, constitueront une matrice $(M \times n)$ de rendement. Enfin, on définira un vecteur de poids, qui correspond à la proportion des différents titres dans le portefeuille. Ainsi, M simulations du rendement d'un portefeuille correspondent au produit de la matrice des rendements $(M \times n)$ et du vecteur de poids $(n \times 1)$.

Notre but premier est de trouver la loi s'ajustant aux rendements du portefeuille dans le plus de cas possibles. Afin de diversifier notre éventail de portefeuilles, nous avons choisi de simuler neuf jeux différents de poids et de construire une matrice de poids $(n \times 9)$ en joignant côte à côte ces neuf vecteurs de poids. Le produit matriciel de la matrice des rendements $(M \times n)$ et de celle des poids $(n \times 9)$ est une matrice de neuf colonnes, dont chacune représente M simulations d'un portefeuille.

Les détails des procédures de simulation des rendements des titres individuels et de leurs poids dans les portefeuilles sont donnés dans la suite de ce chapitre.

4.1 Lois des rendements

Le rendement d'un titre financier sur un intervalle de temps donné est

(perte ou gain durant l'intervalle) / valeur au début de l'intervalle.

L'idée de ce mémoire est de modéliser le rendement du portefeuille et non des titres individuels. On cherche donc quelle loi s'ajuste au mieux à un vaste éventail de données simulées. Cet ajustement sera validé par diverses techniques dont l'exposé suivra dans les prochains chapitres. Pour tenter d'identifier cette loi, on simulera des titres individuels, puis on ajustera la loi au portefeuille. Afin d'avoir suffisamment de diversité, on considérera les cas suivants :

- différentes lois pour les titres individuels,
- différentes dépendances ou indépendance entre les titres individuels

Pour les titres individuels, on s'est fixé trois différents couples de moyenne - écart-type : (0.05, 0.01), (0.10, 0.03), et (0.15, 0.08). Les couples moyenne - écart-type vont être associés à trois lois à deux paramètres : la lognormale, la Weibull, la gamma. Puisque ces lois sont à support positif, nous considérons donc les couples suivants : (1.05, 0.01), (1.10, 0.03), et (1.15, 0.08). Ainsi, on simule une matrice $R = \text{rendement} + 1$.

Il faut donc déduire les paramètres de la distribution à simuler à partir du couple (moyenne , écart-type) associé, de façon analogue à la méthode des moments. Sous le postulat de dépendance, on simule à partir d'une copule. À cet effet, on a retenu quatre copules pour la cause : trois copules archimédiennes (la Clayton, la Frank, et la Gumbel) et la copule de t. On a aussi choisi quatre différents niveaux de dépendance exprimés en fonction du τ de Kendall : $\tau = 0.1, 0.3, 0.5$ et 0.8 , dont on déduira les paramètres respectifs des copules.

Enfin, pour chaque combinaison possible, nous engendrerons $M = 10\ 000$ observations.

De multiples combinaisons des trois lois (la lognormale, la Weibull, la gamma) et des trois couples de (moyenne , écart-type) sont possibles. Nous avons choisi de retenir trois combinaisons possibles de trois titres et une unique combinaison de neuf titres. Ces quatre combinaisons sont les suivantes :

1. $R_1 = 1 + \text{rendement}_1$ gamma et paire (moyenne , écart-type) = (1.05, 0.01),
 $R_2 = 1 + \text{rendement}_2$ lognormale et paire (moyenne , écart-type) = (1.10, 0.03),
 et $R_3 = 1 + \text{rendement}_3$ Weibull et paire (moyenne , écart-type) = (1.15, 0.08).

2. $R_1 = 1 + \text{rendement}_1$ gamma et paire (moyenne , écart-type) = (1.15, 0.08),
 $R_2 = 1 + \text{rendement}_2$ lognormale et paire (moyenne , écart-type) = (1.05, 0.01),
 et $R_3 = 1 + \text{rendement}_3$ Weibull et paire (moyenne , écart-type) = (1.10, 0.03).
3. $R_1 = 1 + \text{rendement}_1$ gamma et paire (moyenne , écart-type) = (1.10, 0.03),
 $R_2 = 1 + \text{rendement}_2$ lognormale et paire (moyenne , écart-type) = (1.15, 0.08),
 et $R_3 = 1 + \text{rendement}_3$ Weibull et paire (moyenne , écart-type) = (1.05, 0.01).
4. $R_1 = 1 + \text{rendement}_1$ gamma et paire (moyenne , écart-type) = (1.05, 0.01),
 $R_2 = 1 + \text{rendement}_2$ lognormale et paire (moyenne , écart-type) = (1.10, 0.03),
 $R_3 = 1 + \text{rendement}_3$ Weibull et paire (moyenne , écart-type) = (1.15, 0.08),
 $R_4 = 1 + \text{rendement}_4$ gamma et paire (moyenne , écart-type) = (1.15, 0.08),
 $R_5 = 1 + \text{rendement}_5$ lognormale et paire (moyenne , écart-type) = (1.05, 0.01),
 $R_6 = 1 + \text{rendement}_6$ Weibull et paire (moyenne , écart-type) = (1.10, 0.03),
 $R_7 = 1 + \text{rendement}_7$ gamma et paire (moyenne , écart-type) = (1.10, 0.03),
 $R_8 = 1 + \text{rendement}_8$ lognormale et paire (moyenne , écart-type) = (1.15, 0.08),
 et $R_9 = 1 + \text{rendement}_9$ Weibull et paire (moyenne , écart-type) = (1.05, 0.01).

Sous le postulat de dépendance, les combinaisons de trois titres seront simulées selon quatre copules différentes ; la Clayton, la Frank, la Gumbel, et la copule de t. Quant à la combinaison de neuf titres, elle sera simulée avec la Clayton et la copule de t (simuler les rendements de neuf titres dépendants nécessite énormément de temps et de mémoire, et cette simulation n'a pas abouti pour la Frank et la Gumbel). Par ailleurs, chacune des copules sera simulée selon quatre niveaux de dépendance. On aboutit, ainsi, à quarante-huit combinaisons pour trois titres (3 combinaisons de lois \times 4 copules \times 4 niveaux de dépendance), huit combinaisons pour neuf titres (2 copules \times 4 niveaux de dépendance) et une combinaison de neuf titres sous l'hypothèse d'indépendance. Pour de raisons de clarté de la présentation, nous ne présenterons que cinq combinaisons parmi ces 57. Ces cinq combinaisons sont le cas d'indépendance à neuf titres, la copule de Clayton à neuf titres et la copule de t à neuf titres aux niveaux de dépendance 0.3 et 0.8.

4.2 Rendements des données réelles

Nous considérons également un jeu de données réelles provenant de cinq titres. On dispose des rendements quotidiens (jours ouvrables) sur la période du 1 juin 2001 au 31 mai 2004 d'un marché comportant les titres de Berkshire Hathaway, Hecla Mining, McDonald's, Reebok International et Scientific Games.

Un exposé des secteurs d'activité de chacune de ces entreprises nous permettra de mieux cerner ce marché. L'exposé et les tableaux de corrélation qui suivent nous montrent la diversité des secteurs économiques de ce marché ainsi que la faible dépendance entre ceux-ci.

| | Titre 1 | Titre 2 | Titre 3 | Titre 4 | Titre 5 |
|----------------|----------------|----------------|----------------|----------------|----------------|
| Titre 1 | 1.0000 | 0.0346 | 0.1111 | 0.1596 | 0.1309 |
| Titre 2 | 0.0346 | 1.0000 | 0.0036 | -0.0704 | 0.0067 |
| Titre 3 | 0.1111 | 0.0036 | 1.0000 | 0.2354 | 0.1380 |
| Titre 4 | 0.1596 | -0.0704 | 0.2354 | 1.0000 | 0.2080 |
| Titre 5 | 0.1309 | 0.0067 | 0.1380 | 0.2080 | 1.0000 |

TAB. 4.1 – Matrice des coefficients de corrélation de Kendall des données réelles

4.2.1 Berkshire Hathaway

L'année 1950 marque la naissance de Berkshire Hathaway suite à la fusion des sociétés textiles Hathaway Manufacturing Company et Berkshire Fine Spinning Associates. En 1962, elle est rachetée par Warren Buffett. Aujourd'hui, ce dernier partage la direction de la société avec Charlie Munger et Bill Gates. Au cours de cette même année, il entreprend de la transformer en société d'investissement et de relocaliser son siège social à Omaha dans le Nebraska (USA). Aujourd'hui, elle est présente dans les secteurs de

- l'assurance et de la finance,
- l'agro-alimentaire,
- le textile,
- la distribution d'énergie et de services collectifs,
- les matériaux de construction,
- les médias,
- la logistique,
- le luxe
- etc...

Pour l'anecdote, 10 000 \$ investis dans l'action BRK en 1965 valaient 30 000 000 \$ en 2005.

4.2.2 Hecla Mining

Depuis 1891, date de sa création, Hecla Mining gère plusieurs exploitations minières en Amérique du Nord et en Amérique centrale. Elle siège à Alene au coeur de l'Idaho. En 2006, elle devient le premier producteur d'argent à bas coût en Amérique du Nord, produisant 5,5 millions d'onces d'argent à un coût total moyen de 24 cents par once. Elle a également produit 179 276 onces d'or à un coût total moyen de 345 \$ par once.

Hecla extrait de l'argent et de l'or aux États-Unis, au Mexique et au Vénézuéla. Elle exploite trois mines d'argent ; Greens Creek à Admiralty Island en Alaska, Lucky Friday à Mullan en Idaho et San Sebastian Property à Durango au Mexique, ainsi que la mine d'or La Camorra à Bolivar au Vénézuéla. Actuellement, elle est sur le point de lancer l'exploitation de la mine d'or de Hollister Block à Midas, Nevada.

Les actions de Hecla Mining sont cotées sur la bourse de New York depuis plus de 40 ans sous le symbole «HL».

4.2.3 McDonald's

McDonald's est la plus grande chaîne de restauration rapide du monde. En géant de la restauration rapide, McDonald's est présent dans les plus grandes métropoles mondiales. Elle possède plus de 30 000 franchises de restaurants rapides sous la marque McDonald's, dans 121 pays. Elle dirige également d'autres marques de restaurants, comme Aroma Café, Boston Market, Chipotle Mexican Grill, Donatos Pizza (depuis décembre 2003) et Prêt À Manger. Ses revenus pour 2001 étaient de 14,87 milliards de dollars, avec un revenu net de 1,64 milliard de dollars.

En 1986, le magazine The Economist a proposé l'utilisation de l'indice Big Mac pour estimer le niveau de vie des pays dans lesquels McDonald's est présent.

Les actions de McDonald's sont cotées à la bourse de New York. Le titre est aussi membre de l'indice Dow Jones. Ces principaux actionnaires sont Jim Skinner et ZZ Investment Manag.

4.2.4 Reebok International

Reebok, fondé en 1890, est un équipementier sportif anglais siégeant à Bolton (Angleterre). À l'époque, Joseph William Foster était l'un des premiers à fabriquer des chaussures d'athlétisme à pointe (par exemple, souliers pour golf). La société J.W. Foster and Sons fabrique tous ses souliers à la main et acquiert une clientèle d'athlètes internationaux, en fournissant notamment les chaussures portées aux Jeux olympiques de Paris (1924).

En 1958, les deux petits-fils de J.W. Foster renomment la société Reebok, nom qui rappelle celui d'une variété d'antilope sud-africaine, le rhebok. En 1979, la société se lance à l'assaut du marché nord-américain en vendant des chaussures de course à un prix jamais vu : 60 dollars US. En 1981, la filiale états-unienne vend pour plus de 1,5 milliard de dollars US et sa croissance semble assurée pour plusieurs années. En 1985, Reebok devient cotée à la Bourse de New York. En 2005, elle détenait les marques Rockport (chaussures), Greg Norman (golf) et Ralph Lauren Footwear. Toutefois, la marque Reebok représente 84% des ventes de l'entreprise.

4.2.5 Scientific Games

Scientific Games est un chef de fil des industries de la loterie, du pari-mutuel, et des télécommunications. Son expertise porte sur

- la conception et la production des billets instantanés de loterie,
- la gestion des systèmes de loterie en ligne,
- le développement et la commercialisation de divertissement,
- la fourniture de pari-mutuel tel que les courses de chevaux,
- la production des cartes prépayées de téléphone cellulaire.

Elle affiche un taux de croissance des plus élevés sur le marché des loteries en ligne. Ses actions sont cotées sur la bourse du NASDAQ sous le symbole de GS. En avril 2008, elles s'échangeaient à 23,87 dollars US.

4.3 Poids des titres et portefeuilles

Un portefeuille est une combinaison linéaire des titres, les coefficients de la combinaison représentant la proportion des différents titres formant le portefeuille. Ces coefficients somment à 1. Puisque nous excluons la possibilité de ventes à découvert, nous supposons tous ces coefficients positifs.

4.3.1 Problématique des poids des titres

Simuler uniformément des poids non négatifs qui somment à 1 n'est pas chose aisée. En fait, on cherche à simuler uniformément sur l'ensemble $\{x_1, \dots, x_n \in [0, 1]^n : \sum_{i=1}^n x_i = 1, 0 \leq x_i \leq 1\}$. Il ne s'agit donc pas de chercher simplement des poids qui somment à 1 en ayant de moins en moins de poids au fur et à mesure qu'on avance dans le processus. En effet, on cherche une diversité dans les combinaisons, de manière à ce que chaque titre ait la chance de figurer dans le portefeuille, et que les combinaisons ne favorisent pas toujours un titre par rapport aux autres.

Commençons par étudier le cas à 4 titres. Soient X_1, X_2, X_3 et X_4 quatre variables aléatoires sommant à 1. La contrainte $X_4 = 1 - X_1 - X_2 - X_3$ laisse trois variables libres. Une distribution uniforme sur l'hyperplan $\{0 \leq x_i \leq 1, \sum_{i=1}^4 x_i = 1\}$ est

$$f(x_1, x_2, x_3) = \begin{cases} 6 & , \quad 0 \leq x_1 \leq 1, 0 \leq x_2 \leq 1 - x_1, 0 \leq x_3 \leq 1 - x_1 - x_2 \\ 0 & , \quad \text{sinon.} \end{cases}$$

La loi marginale de X_1 est

$$\begin{aligned} f_{X_1}(x_1) &= \int_0^{1-x_1} \int_0^{1-x_1-x_2} 6 \, dx_2 dx_3, \\ &= \int_0^{1-x_1} 6(1-x_1-x_2) \, dx_2, \\ &= 3(1-x_1)^2, \quad 0 < x_1 < 1, \end{aligned} \tag{4.1}$$

et on reconnaît une loi $Beta(1, 3)$. La loi de X_2 , une fois X_1 connue, est obtenue ainsi :

$$\begin{aligned} f_{X_1, X_2}(x_1, x_2) &= \int_0^{1-x_1-x_2} 6 \, dx_3, \\ &= 6(1-x_1-x_2), \quad 0 \leq x_1 \leq 1, 0 \leq x_2 \leq 1-x_1. \end{aligned} \tag{4.2}$$

On obtient

$$\begin{aligned}
 f_{X_2|X_1=x_1}(x_2) &= \frac{6(1-x_1-x_2)}{3(1-x_1)^2}, \\
 &= \frac{2(1-x_1-x_2)}{(1-x_1)^2}, \\
 &= \frac{2(1-\frac{x_2}{1-x_1})}{1-x_1}, \quad 0 < x_2 < 1-x_1.
 \end{aligned} \tag{4.3}$$

Posons $Y_2 = \frac{X_2}{1-x_1}$. On a

$$F_{Y_2}(y) = F_{X_2|X_1=x_1}((1-x_1)y), \tag{4.4}$$

$$\begin{aligned}
 f_{Y_2}(y) &= f_{X_2|X_1=x_1}((1-x_1)y)(1-x_1), \\
 &= 2(1-y), \quad 0 < y < 1.
 \end{aligned} \tag{4.5}$$

et on reconnaît une loi $Beta(1, 2)$. Finalement,

$$\begin{aligned}
 f_{X_3|X_1=x_1, X_2=x_2}(x_3) &= \frac{6}{6(1-x_1-x_2)}, \\
 &= \frac{1}{1-x_1-x_2}.
 \end{aligned} \tag{4.6}$$

Posons $Y_3 = \frac{X_3}{1-x_1-x_2}$. On a

$$F_{Y_3}(y) = 1, \quad 0 < y < 1, \tag{4.7}$$

et on reconnaît la loi $Unif(0, 1) \sim Beta(1, 1)$.

Si on généralise cette méthode en dimension 4 à la dimension n , on peut établir l'algorithme suivant :

1. On simule le premier poids, $poids_1$, d'une $Beta(1, n-1)$;
2. on simule une $Beta(1, n-2)$ et le deuxième poids, $poids_2$, est obtenu en multipliant ce tir aléatoire par $(1 - poids_1)$;
3. on simule une $Beta(1, n-j)$, et le j^e poids, $poids_j$, est obtenu en multipliant ce tir aléatoire par $(1 - \sum_{i=1}^{j-1} poids_i)$;
4. on simule une $Beta(1, 1)$, où $Beta(1, 1)$ est l'uniforme $[0, 1]$ et l'avant dernier poids est obtenu en multipliant ce tir aléatoire par $(1 - \sum_{i=1}^{n-2} poids_i)$;
5. le dernier poids est obtenu en soustrayant les $n-1$ autres poids de 1 :

$$x_n = 1 - \sum_{i=1}^{n-1} poids_i.$$

4.3.2 Conception des portefeuilles

Soient $R_1 - 1, \dots, R_n - 1$ les rendements simulés et x_1, \dots, x_n les poids simulés tel que décrit plus haut. On se propose de créer la matrice \mathbf{R} de taille $(M \times n)$ en joignant les n vecteurs les uns à côté des autres ($M = 10\ 000, n = 3$ ou 9). La matrice \mathbf{R} s'écrit donc

$$\mathbf{R} = \begin{pmatrix} R_{1,1} & \dots & R_{1,n} \\ R_{2,1} & \dots & R_{2,n} \\ \vdots & & \vdots \\ R_{10\ 000,1} & \dots & R_{10\ 000,n} \end{pmatrix}$$

On crée ensuite une matrice X de dimension $(n \times m)$ obtenue en simulant m jeux différents de poids $\{x_i\}$ placés côte à côte. Dans ce mémoire, on a retenu $m = 9$. La matrice X , obtenue ainsi, est

$$\mathbf{X} = \begin{pmatrix} X_{1,1} & X_{1,2} & \dots & X_{1,9} \\ X_{2,1} & X_{2,2} & \dots & X_{2,9} \\ \vdots & \vdots & & \vdots \\ X_{n,1} & X_{n,2} & \dots & X_{n,9} \end{pmatrix}$$

Le produit matriciel $\mathbf{R} \times \mathbf{X}$ fournit $M = 10\ 000$ simulations de *rendement* + 1 de neuf portefeuilles différents S_1, \dots, S_9 , présentés dans une matrice S

$$\mathbf{S} = \begin{pmatrix} S_{1,1} & \dots & S_{1,9} \\ S_{2,1} & \dots & S_{2,9} \\ \vdots & & \vdots \\ S_{10\ 000,1} & \dots & S_{10\ 000,9} \end{pmatrix}$$

Les rendements de portefeuilles sont représentés par $S - 1$. Comme expliqué plus haut, cette procédure est répétée 48 fois lorsque $n = 3$ et 9 fois lorsque $n = 9$.

Chapitre 5

Étude descriptive

Dans le chapitre précédent, on a abordé la conception des portefeuilles. Dans ce chapitre, on étudiera les principales caractéristiques empiriques des portefeuilles simulés afin de trouver une famille de lois qui permet de modéliser la distribution de leurs rendements.

En raison du nombre très élevé de portefeuilles simulés, seulement les résultats de certains d'entre eux seront étudiés. Suite à l'observation des histogrammes et des statistiques descriptives, on a retenu les portefeuilles qui semblent qualitativement les plus représentatifs de l'ensemble de nos données. Ces portefeuilles correspondent au cas d'indépendance à neuf titres, à la copule de Clayton à neuf titres et à la copule de t à neuf titres aux niveaux de dépendance 0.3 et 0.8. Il s'ajoute à ces derniers les neuf portefeuilles à neuf titres issus de données réelles.

Dans les sections suivantes, on présentera une analyse descriptive des résultats de nos simulations, ainsi que certains graphiques tels que les histogrammes, les courbes d'espérance de vie résiduelle et de Ramlaou-Hansen. Cette analyse nous permettra de réduire l'ensemble des familles de lois pouvant potentiellement bien modéliser la distribution de nos données.

5.1 Statistiques descriptives

Dans ce qui suit, nous présenterons les statistiques descriptives des portefeuilles retenus. Ces statistiques sont le minimum et le maximum, les premier et troisième

quartiles, la moyenne et la médiane. Le but est d'étudier les spécificités de chaque portefeuille afin de trouver une famille de lois qui permet de modéliser la distribution du rendement.

| portefeuille | Min | 1er Quartile | Médiane | Moyenne | 3ème Quartile | Max |
|--------------|----------|--------------|---------|---------|---------------|--------|
| 1 | -0.02781 | 0.07783 | 0.09956 | 0.1002 | 0.1219 | 0.2305 |
| 2 | 0.01735 | 0.0827 | 0.09537 | 0.09563 | 0.1086 | 0.1754 |
| 3 | -0.04695 | 0.09009 | 0.1168 | 0.1164 | 0.1438 | 0.2739 |
| 4 | -0.03163 | 0.08907 | 0.1093 | 0.1085 | 0.129 | 0.2086 |
| 5 | 0.01571 | 0.07641 | 0.08754 | 0.08743 | 0.09843 | 0.1543 |
| 6 | -0.01871 | 0.09423 | 0.114 | 0.1132 | 0.1326 | 0.2291 |
| 7 | 0.003496 | 0.07965 | 0.09773 | 0.09831 | 0.1169 | 0.2112 |
| 8 | -0.04201 | 0.05422 | 0.07381 | 0.07431 | 0.09392 | 0.1865 |
| 9 | 0.003302 | 0.07696 | 0.0905 | 0.09026 | 0.1039 | 0.17 |

TAB. 5.1 – *Indépendance (9 titres)*

Au tableau 6.1, on retrouve les caractéristiques des portefeuilles dont les rendements des neuf titres individuels sont indépendants. Le rendement moyen observé de ces portefeuilles se situe entre 7% et 12%. Toutefois, le comportement de nos données est relativement stable, peu importe les poids simulés. En premier lieu, on observe une quasi-égalité entre la moyenne et la médiane. En second lieu, on décèle une relative symétrie. En effet, les écarts entre le premier quartile et la médiane, et entre cette dernière et le troisième quartile sont relativement égaux. Ces deux observations sont caractéristiques de la loi normale. Par ailleurs, le rendement maximal se situe entre 17% et 28%, alors que le rendement minimal se situe au dessus de -5% .

| portefeuille | Min | 1er Quartile | Médiane | Moyenne | 3ème Quartile | Max |
|--------------|----------|--------------|---------|---------|---------------|--------|
| 1 | -0.1349 | 0.07207 | 0.1008 | 0.09611 | 0.1252 | 0.2035 |
| 2 | -0.09977 | 0.07502 | 0.09737 | 0.09318 | 0.1155 | 0.1827 |
| 3 | -0.1238 | 0.0772 | 0.1056 | 0.1015 | 0.1301 | 0.2251 |
| 4 | -0.1869 | 0.07845 | 0.1148 | 0.1076 | 0.1443 | 0.2512 |
| 5 | -0.1697 | 0.06419 | 0.1016 | 0.09678 | 0.1341 | 0.274 |
| 6 | -0.1724 | 0.07346 | 0.1106 | 0.105 | 0.1412 | 0.2626 |
| 7 | -0.1762 | 0.08629 | 0.1226 | 0.1168 | 0.1541 | 0.2686 |
| 8 | -0.1917 | 0.06911 | 0.1133 | 0.1099 | 0.1535 | 0.324 |
| 9 | -0.114 | 0.08049 | 0.1074 | 0.103 | 0.1303 | 0.218 |

TAB. 5.2 – Copule de Clayton - faible niveau de dépendance (9 titres)

Contrairement à l'indépendance, quand la dépendance entre les rendements des neuf titres individuels est régie par une copule de Clayton à faible niveau de dépendance, on observe, au tableau 6.2, d'une part une différence entre le rendement moyen et la médiane des rendements du portefeuille, et d'autre part une asymétrie à gauche. Par ailleurs, l'étendue a doublé comparativement à l'indépendance. Quant au rendement minimal, il atteint -19% , alors que le rendement maximal se situe entre 18% et 33% .

À partir de ces observations, on peut éliminer certaines lois telles que la loi normale, la loi log normale et la loi gamma. En effet, ces deux dernières présentent une asymétrie à droite, alors que la loi normale est une distribution symétrique.

| portefeuille | Min | 1er Quartile | Médiane | Moyenne | 3ème Quartile | Max |
|--------------|----------|--------------|---------|---------|---------------|--------|
| 1 | -0.0934 | 0.06415 | 0.1004 | 0.09928 | 0.137 | 0.2512 |
| 2 | -0.1576 | 0.06377 | 0.1124 | 0.1105 | 0.1609 | 0.3 |
| 3 | -0.1171 | 0.06393 | 0.1005 | 0.09799 | 0.1356 | 0.2446 |
| 4 | -0.1798 | 0.0522 | 0.1003 | 0.0981 | 0.1478 | 0.3008 |
| 5 | -0.1595 | 0.06893 | 0.1139 | 0.1101 | 0.156 | 0.2797 |
| 6 | -0.1426 | 0.06273 | 0.1026 | 0.09939 | 0.1403 | 0.2406 |
| 7 | -0.1632 | 0.04975 | 0.0951 | 0.09256 | 0.1391 | 0.268 |
| 8 | -0.09336 | 0.0541 | 0.08713 | 0.08593 | 0.1198 | 0.2192 |
| 9 | -0.1277 | 0.07554 | 0.112 | 0.108 | 0.1456 | 0.2283 |

TAB. 5.3 – Copule de Clayton - niveau élevé de dépendance (9 titres)

Au tableau 6.3, la dépendance entre les rendements des neuf titres individuels est modélisée par une copule de Clayton à niveau élevé de dépendance. On observe alors que, d'une part, l'écart entre le rendement moyen et la médiane, bien que réduit en passant d'un faible niveau de dépendance à un niveau élevé de dépendance, est encore significatif et que, d'autre part, l'étendue et l'écart interquartile ont doublé comparativement à l'indépendance. Par ailleurs, le rendement minimal et le rendement maximal demeurent les mêmes que pour la Clayton à faible niveau de dépendance. En effet, le minimum atteint -18% ; quant au maximum, il se situe entre 22% et 30% .

En résumé, pour la copule de Clayton, quelque soit le niveau de dépendance, le rendement moyen est différent de la médiane. Quant à l'écart interquartile et l'étendue, ils ont doublé comparativement à l'indépendance. Enfin, on y observe des distributions asymétriques à gauche. On peut donc éliminer la loi normale, la loi lognormale et la loi gamma.

| portefeuille | Min | 1er Quartile | Médiane | Moyenne | 3ème Quartile | Max |
|--------------|----------|--------------|---------|---------|---------------|--------|
| 1 | -0.1275 | 0.06619 | 0.1087 | 0.1091 | 0.1502 | 0.3674 |
| 2 | -0.1205 | 0.07645 | 0.1099 | 0.1071 | 0.1403 | 0.288 |
| 3 | -0.1075 | 0.07608 | 0.1127 | 0.1134 | 0.1501 | 0.3507 |
| 4 | -0.05581 | 0.08548 | 0.1128 | 0.1119 | 0.1377 | 0.2892 |
| 5 | -0.0541 | 0.07543 | 0.1008 | 0.1003 | 0.1249 | 0.2599 |
| 6 | -0.07365 | 0.06685 | 0.09488 | 0.0955 | 0.1232 | 0.2965 |
| 7 | -0.08164 | 0.07227 | 0.1025 | 0.1024 | 0.1317 | 0.2977 |
| 8 | -0.07666 | 0.08018 | 0.1113 | 0.1105 | 0.1395 | 0.2996 |
| 9 | -0.0687 | 0.07658 | 0.1017 | 0.09975 | 0.1241 | 0.2527 |

TAB. 5.4 – Copule de t - faible niveau de dépendance (9 titres)

Pour la copule de t à faible niveau de dépendance, on observe au tableau 6.4 une quasi-égalité entre le rendement moyen et la médiane. Quant à l'asymétrie, elle est tantôt à gauche tantôt à droite. De plus, l'étendue et l'écart interquartile ont augmenté comparativement à l'indépendance et à la Clayton à faible niveau de dépendance. Par ailleurs, le rendement minimal n'atteint que -13% , tandis que le rendement maximal augmente et se situe entre 25% et 37% .

| portefeuille | Min | 1er Quartile | Médiane | Moyenne | 3ème Quartile | Max |
|--------------|----------|--------------|---------|---------|---------------|--------|
| 1 | -0.1741 | 0.05341 | 0.1076 | 0.1087 | 0.1616 | 0.4427 |
| 2 | -0.1751 | 0.06776 | 0.1106 | 0.1069 | 0.1504 | 0.3183 |
| 3 | -0.1436 | 0.0657 | 0.113 | 0.1132 | 0.1601 | 0.3994 |
| 4 | -0.09346 | 0.07615 | 0.112 | 0.1117 | 0.1469 | 0.3219 |
| 5 | -0.09521 | 0.06791 | 0.1008 | 0.1 | 0.1325 | 0.2866 |
| 6 | -0.1024 | 0.05635 | 0.09425 | 0.09531 | 0.1328 | 0.3341 |
| 7 | -0.118 | 0.0627 | 0.1022 | 0.1022 | 0.1412 | 0.3389 |
| 8 | -0.1112 | 0.07174 | 0.1107 | 0.1101 | 0.1482 | 0.3367 |
| 9 | -0.1084 | 0.06902 | 0.1016 | 0.09956 | 0.1321 | 0.2694 |

TAB. 5.5 – Copule de t - niveau élevé de dépendance (9 titres)

Au tableau 6.5, un niveau élevé de dépendance pour la copule de t s'est traduit par l'augmentation de l'étendue et de l'écart interquartile, et des queues plus épaisses. De plus, le maximum et le minimum atteignent respectivement 44% et -18%.

La copule de t se caractérise par un grand écart interquartile et une grande étendue comparativement à l'indépendance et la Clayton. De plus, les distributions sont asymétriques tantôt à gauche, tantôt à droite. Par ailleurs, l'augmentation du niveau de dépendance de la copule de t s'accompagne de queues plus épaisses et d'une hausse du rendement maximal. En conséquence, on doit chercher parmi les familles de lois assez flexibles au niveau de l'asymétrie et de l'épaisseur des queues.

De la confrontation de tous ces résultats, il ressort quelques observations majeures. En premier lieu, l'augmentation du niveau de dépendance entraîne une augmentation de l'étendue, de l'écart interquartile et du rendement maximal et, ce, tant au niveau de la copule de Clayton que de la copule de t . On notera aussi que lors du passage de l'indépendance à la dépendance, le rendement minimal devient plus petit. Quant aux queues, elles deviennent plus épaisses pour les niveaux élevés de dépendance. En second lieu, on distingue trois comportements distincts d'asymétrie. En effet, les distributions sont relativement symétriques à l'indépendance, asymétrique à gauche pour la Clayton, et asymétrique tantôt à gauche tantôt à droite pour la copule de t . Enfin, pour un même niveau de dépendance, les étendues et les écarts interquartile observés pour la copule de t sont plus importants que ceux observés pour la Clayton. Toutes ces observations nous amènent à chercher parmi les familles de lois flexibles au niveau de l'asymétrie. On élimine ainsi des lois telles que la normale, la log normale et la gamma.

| portefeuille | Min | 1er Quartile | Médiane | Moyenne | 3ème Quartile | Max |
|--------------|-----------|--------------|---------|---------|---------------|--------|
| 1 | -0.01188 | 0.04235 | 0.05097 | 0.05143 | 0.0596 | 0.1394 |
| 2 | -0.002436 | 0.04303 | 0.0505 | 0.05123 | 0.05872 | 0.1313 |
| 3 | -0.01332 | 0.04315 | 0.05005 | 0.05079 | 0.05791 | 0.1175 |
| 4 | -0.03991 | 0.03999 | 0.05104 | 0.05198 | 0.06346 | 0.1677 |
| 5 | -0.02962 | 0.03762 | 0.05021 | 0.05202 | 0.06521 | 0.2147 |
| 6 | 0.004025 | 0.04319 | 0.0505 | 0.05109 | 0.05889 | 0.1243 |
| 7 | -0.002885 | 0.04438 | 0.05112 | 0.05094 | 0.05708 | 0.1328 |
| 8 | 0.004622 | 0.04279 | 0.05053 | 0.05084 | 0.05858 | 0.1259 |
| 9 | -0.04068 | 0.04128 | 0.05176 | 0.05168 | 0.06135 | 0.1471 |

TAB. 5.6 – *Données réelles (9 titres)*

Dans le cas des rendements des titres réels décrits au chapitre 5, on observe au tableau 6.6 la quasi-égalité entre le rendement moyen et la médiane. De plus, les distributions présentent une asymétrie à droite. Par ailleurs, l'écart interquartile observé pour ces données réelles ne dépasse pas 0.02, comparativement à un minimum de 0.045 pour les données simulées.

Par ailleurs, le rendement maximal se situe en deçà de 15%, sauf pour les portefeuilles 4 et 5 où il est respectivement de 16.77% et 21.47%, tandis que le rendement minimal observé est comparable à celui observé pour les données simulées indépendantes.

De ces résultats ressort un nombre d'observations. En premier lieu, le rendement maximal est moins important pour les données réelles. Toutefois, notons que nous avons simulé avec des rendements moyens beaucoup plus élevés que ceux des données réelles. En second lieu, l'écart interquartile observé pour les données réelles est plus faible que celui observé pour les données simulées. Enfin, on observe des distributions asymétriques à droite.

En conclusion, les données réelles ont un comportement plus semblable aux données simulées sous l'hypothèse de l'indépendance que celles simulées sous l'hypothèse de la dépendance. En effet, cette observation est en cohérence avec le tableau 5.1 à la page 47. Dans ce tableau, on note la faible corrélation entre certains titres (ne dépassant pas les 25%) et la quasi-indépendance entre certains. Ce comportement amène à la question : sur des marchés financiers, observe-t-on réellement de la dépendance ? Une hypothèse peut être formulée à ce sujet : notre sélection n'est pas représentative de la réalité de

la totalité des marchés financiers. Puisque les titres financiers considérés pour le cas des données réelles proviennent de secteurs économiques à très faible dépendance (voir quasi-indépendants), les titres financiers seront appelés à être moins dépendants. À ce moment, qu'en serait-il si on avait retenu des titres financiers provenant du même secteur, de la même filière, ou de secteurs interdépendants (à titre d'exemple, la métallurgie et l'industrie des polymères) ? Est-ce que la dépendance serait plus forte ? Tout ceci reste pure spéculation, et on ne peut être sûr à moins de considérer plus de titres financiers provenant de différentes combinaisons et suivis de façon conjointe sur une longue période de temps.

5.2 Histogramme

Avant d'entamer cette section, rappelons que pour cette étude, on a utilisé cinq routines de simulation de portefeuilles ainsi que les données réelles. Chacune des cinq routines correspond à une combinaison de copule - niveau de dépendance, à l'exception de la dernière qui correspond aux neuf portefeuilles issus des données réelles. La première de ces routines correspond à l'indépendance avec neuf titres. La deuxième et la troisième correspondent à la copule de Clayton avec respectivement des niveaux de dépendance de 30% et 80%. Enfin, les deux dernières correspondent à la copule de t avec respectivement des niveaux de dépendance de 30% et 80%. Chacune de ces routines a abouti à une matrice $S - 1$, la matrice de rendements de neuf portefeuilles. On se propose, dès maintenant, d'étudier les histogrammes des rendements de chaque portefeuille (voir l'annexe E). Cette étude nous permet de relever certains points.

1. L'asymétrie

On observe une asymétrie à gauche pour la copule de Clayton, à droite pour les données réelles, et tantôt à gauche et tantôt à droite pour la copule de t . On observe des distributions relativement symétriques à l'indépendance.

2. Les queues

On relève que les queues sont de plus en plus épaisses à mesure que le niveau de dépendance augmente.

3. Rendement maximal et rendement minimal

Le maximum augmente avec l'augmentation du niveau de dépendance. Quant au minimum, il diminue en passant de l'indépendance à la dépendance.

Ces observations rejoignent celles de la section précédente. Ainsi, nous aboutissons aux mêmes conclusions. Afin de raffiner cette étude, on se propose donc d'examiner d'autres approches graphiques.

5.3 Espérance de vie résiduelle

Certains ouvrages ont clairement identifié les courbes d'espérance de vie résiduelle relatives à certaines lois. Ainsi, l'allure de la courbe peut parfois nous renseigner sur la famille de lois. À titre d'exemple, une forme constante horizontale est associée à une loi exponentielle, une ligne droite croissante est associée à une loi de Pareto, alors qu'une courbe croissante peut être associée à la loi lognormale, tandis qu'une courbe décroissante nous suggère une loi gamma ou une loi de Weibull s'il existe une asymptote ([Hogg et Klugman \(1984\)](#)). Dans cette perspective, on a essayé de confronter les courbes engendrées par nos données à celles de cet ouvrage. L'allure de nos courbes s'apparente à la loi de Weibull avec paramètre de forme c supérieur à 1 et dont la courbe décroît vers une asymptote (voir figure 6.1). Toutefois, on observe pour nos courbes des points d'inflexion. Une question se pose dès lors : peut-on envisager la loi de Weibull ?

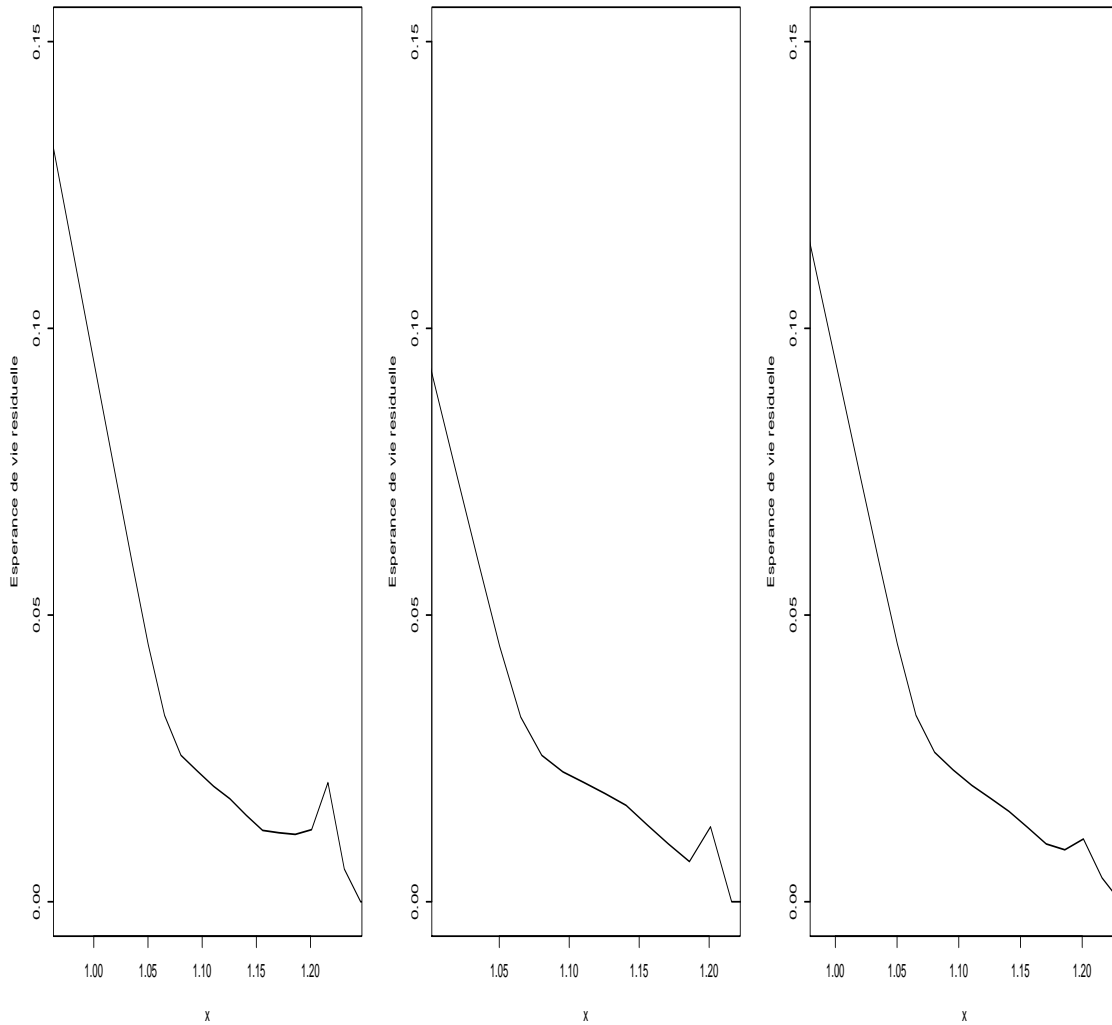


FIG. 5.1 – Exemple de figure d’espérance de vie résiduelle

5.4 Ramlau-Hansen

L’allure des courbes de Ramlau-Hansen peut également nous renseigner sur la distribution de nos données. Pour conclure que la loi de Weibull à paramètre de forme supérieur à 1 est appropriée, ces courbes doivent croître de façon exponentielle. Suite à l’analyse des graphiques de Ramlau-Hansen, on s’aperçoit que seulement au début ils ont la même allure croissante que la loi de Weibull puis il deviennent rapidement décroissantes puis constantes (voir figure 6.2). Ce comportement est du moins intrigant et nous pousse à réviser l’hypothèse selon laquelle la loi de Weibull modéliserait

les rendements de nos portefeuilles.

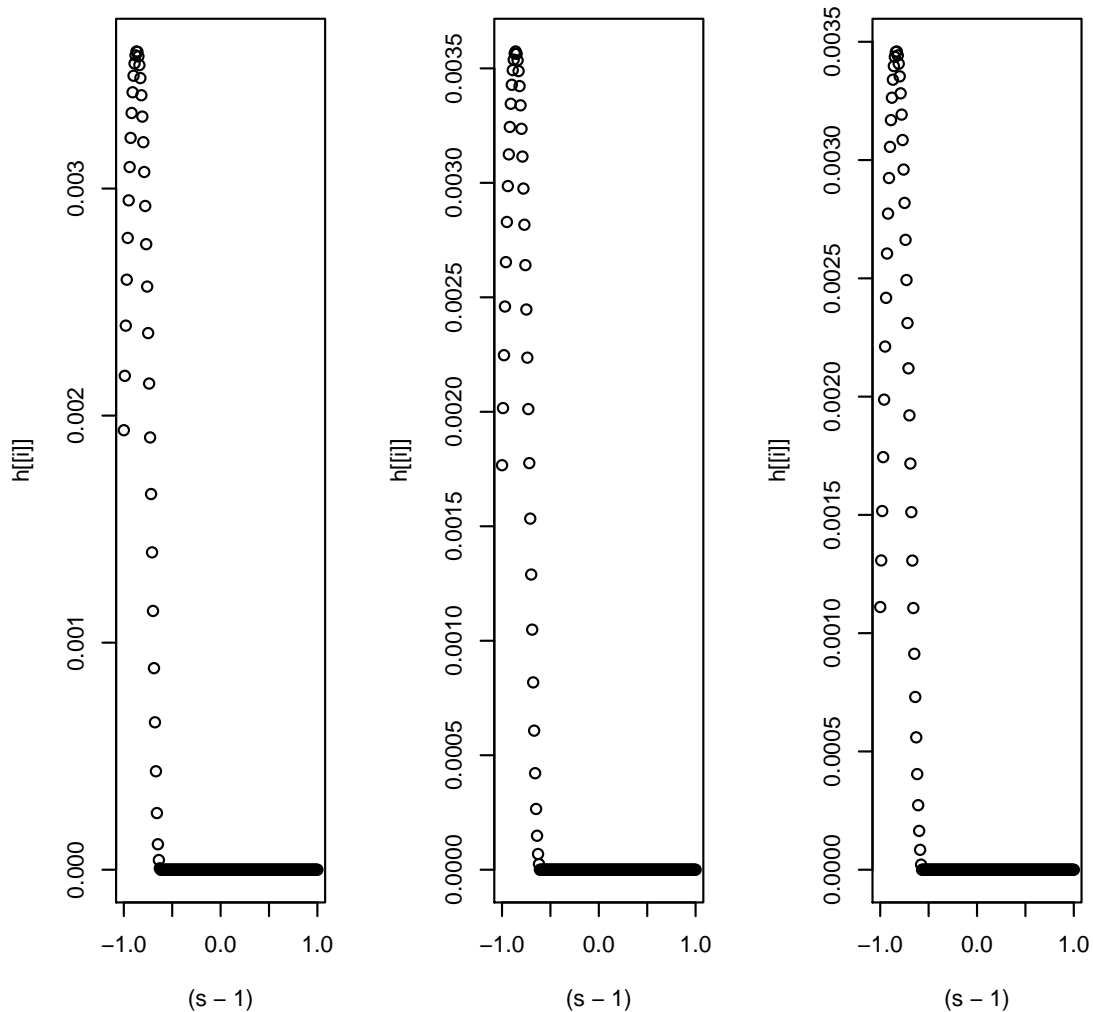


FIG. 5.2 – Exemple de figure de Ramlau-Hansen

5.5 Conclusion

Les courbes d'espérance de vie résiduelle nous ont suggéré de partir avec la loi de Weibull avec un paramètre de forme supérieur à 1. Toutefois, les courbes de Ramlau-Hansen nous poussent à explorer des lois offrant plus de souplesse. Par souplesse, on désigne des lois ayant plus de paramètres pouvant ainsi modéliser davantage les ren-

dements de nos portefeuilles. Dans le chapitre suivant, on essayera donc de vérifier l'adéquation de nos données à un nombre de lois, dont la loi de Weibull.

Chapitre 6

Données simulées

Lors du chapitre précédent, l'analyse graphique nous a suggéré de commencer avec une loi de Weibull, tout en s'attendant à avoir recours à des lois offrant plus de souplesse. Dans ce chapitre, on considérera l'estimation des paramètres des différentes lois retenues, ainsi que leur adéquation à nos données simulées. L'estimation est opérée numériquement pour les différentes méthodes vues dans la section 2.4. Quant à l'adéquation, elle est vérifiée en comparant les moyennes empiriques aux estimés de maximum de vraisemblance des moyennes, ainsi qu'à l'aide de certains graphiques tels que les QQ-plots, les PP-plots et la superposition de l'estimé de maximum de vraisemblance des courbes de densité avec les histogrammes.

En raison du nombre élevé de graphiques et de résultats, on présentera seulement certains d'entre eux. En particulier, on ne considèrera que les estimateurs de maximum de vraisemblance avec points de départ quelconques. En effet, pour certaines lois, les méthodes des moments et de distance minimale n'ont pas convergé. Ainsi, on ne pouvait présenter ni les estimateurs de la méthode des moments, ni ceux de maximum de vraisemblance ayant les estimateurs de la méthode des moments comme points de départ, ni ceux de la distance minimale. Quant à l'adéquation, elle sera seulement vérifiée à l'aide des QQ-plots (les autres méthodes ne seront pas exposées dans ce mémoire) et de la comparaison des moyennes empiriques aux estimés de maximum de vraisemblance des moyennes.

Dans la section suivante, on discutera l'estimation des paramètres et l'adéquation de la loi de Weibull avec un paramètre de forme supérieur à 1.

6.1 Estimation et adéquation

6.1.1 La loi de Weibull

On a estimé les paramètres de la loi de Weibull avec le logiciel **R** en utilisant les méthodes des moments, de maximum de vraisemblance et de distance minimale. Dans les tableaux 6.1 à 6.5 aux pages 69 à 73, on expose, cas par cas, les estimateurs obtenus par la méthode de maximum de vraisemblance avec points de départ quelconques, la valeur de la fonction objectif ($-l(\hat{\Theta})$ où l est la fonction log-vraisemblance et $\hat{\Theta}$ l'estimateur de maximum de vraisemblance du vecteur des paramètres Θ), ainsi que l'estimé de maximum de vraisemblance de la moyenne et la moyenne empirique.

Dans tous les cas, on a obtenu des lois de Weibull avec un paramètre de forme supérieur à 1. Toutefois, les QQ-plots révèlent un énorme manque d'ajustement au niveau des queues. En effet, au milieu, quelques points collent à la première bissectrice contrairement aux queues où aucun point ne colle. Les deux figures qui suivent nous montrent des ajustements de la loi de Weibull aux données simulées sous l'hypothèse d'indépendance et de dépendance élevée.

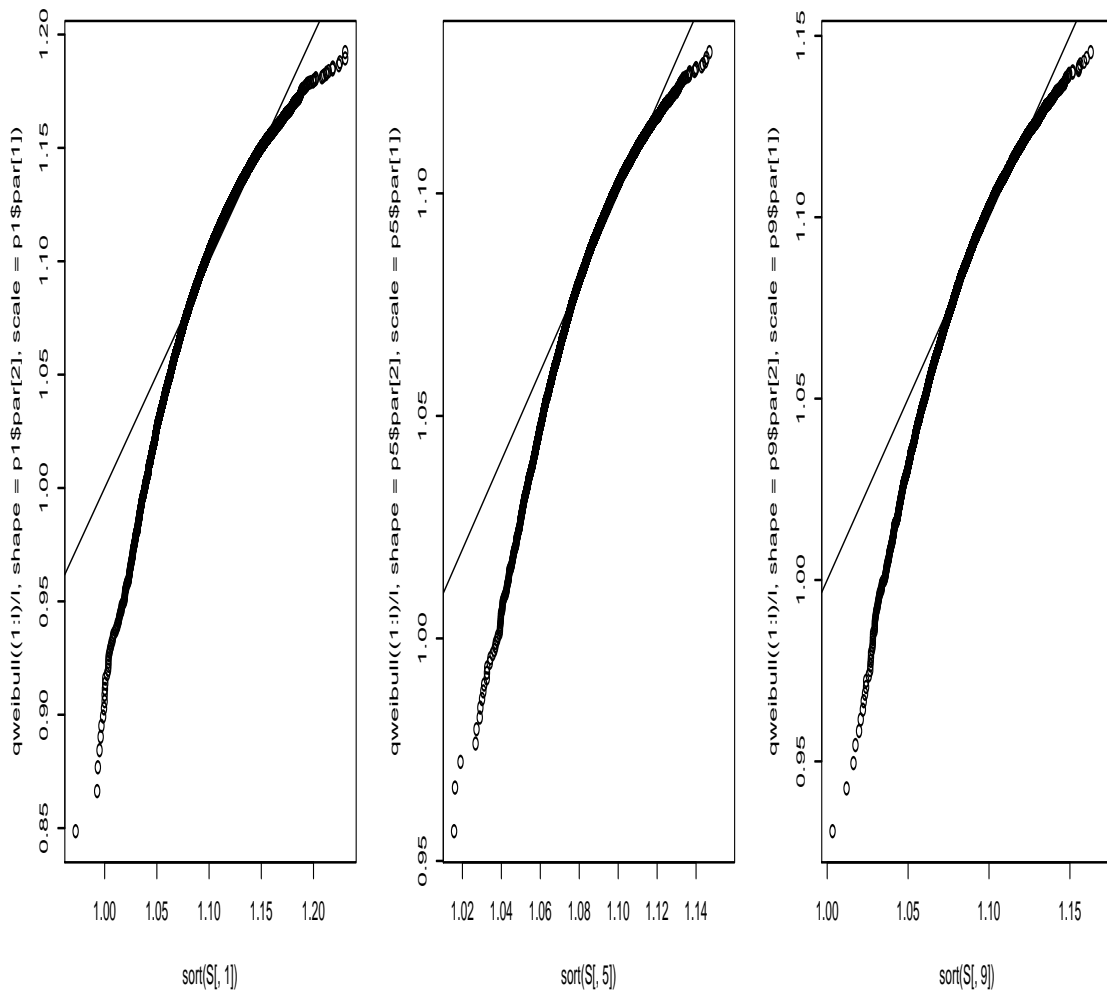


FIG. 6.1 – Exemples de QQ-plots d’une loi de Weibull pour des portefeuilles simulés sous l’hypothèse de l’indépendance

Sachant que la loi de Weibull est un sous cas de la loi gamma transformée à trois paramètres, on a décidé de vérifier si cette dernière modélise mieux les rendements de nos portefeuilles. L’idée est qu’une loi à trois paramètres sera peut-être en mesure d’offrir une plus grande souplesse ainsi qu’une meilleure adéquation au niveau des queues de la distribution.

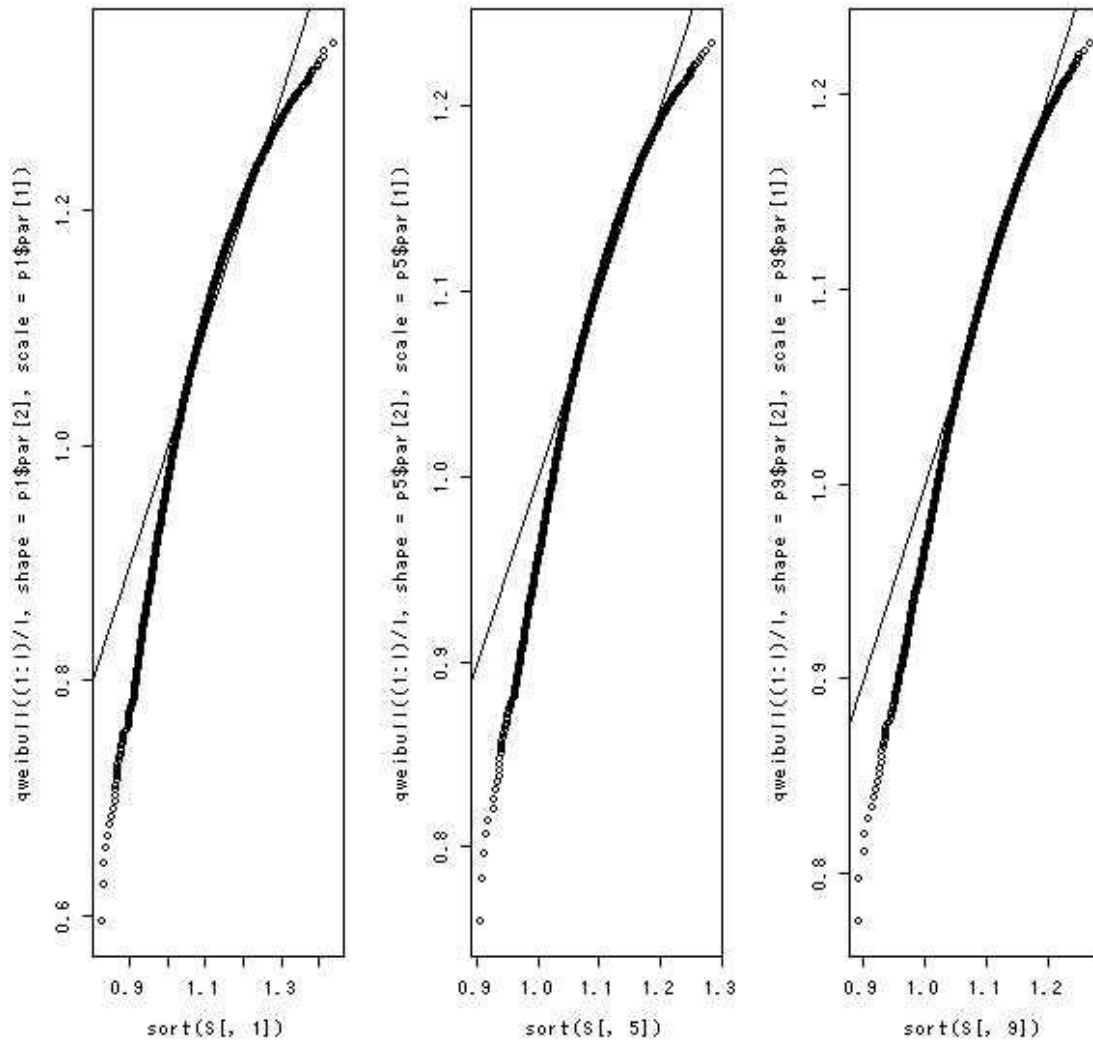


FIG. 6.2 – Exemples de QQ-plots d’une loi de Weibull pour des portefeuilles issus de la copule de t à niveau élevé de dépendance

6.1.2 La loi gamma transformée (GT)

On retrouve la loi loi gamma transformée sur \mathbf{R} dans les bibliothèques «actuar» et «VGAM». Cependant, on a choisi de programmer ses fonctions de densité, de répartition et de quantile. Quant à l’estimation de ses paramètres, on a utilisé les méthodes de moments et de maximum de vraisemblance discutées au chapitre 3. Cependant, la tâche n’a pas toujours été aisée. En effet, la résolution numérique des équations de la méthode des moments n’a pas toujours abouti. Les problèmes rencontrés s’expliquent en partie

par le non respect des conditions énoncées dans la sous-section 2.4.2. Rappelons-nous que les conditions nécessaires à la convergence de la méthode sont « $\mu_3(Z) < 0$ et pas trop grand comparé à $\mu_2(Z)$ ». Dans les tableaux 6.6 à 6.10 aux pages 74 à 78, on expose, cas par cas, les estimateurs obtenus, la valeur de la fonction objectif ainsi que l'estimé de maximum de vraisemblance de la moyenne et la moyenne empirique.

La comparaison des fonctions objectifs et des QQ-plots obtenus pour la loi de Weibull et la loi gamma transformée montrent une nette amélioration au niveau de l'ajustement. Cependant, ce résultat est attendu puisque la loi de Weibull est un cas particulier de la loi gamma transformée. En particulier, on s'attend à ce que la fonction objectif diminue (la fonction de vraisemblance augmente).

Les figures 6.3 et 6.4. montrent, respectivement, un bon et un mauvais ajustement de la loi gamma transformée. Dans le cas de la figure 6.4, on a décelé au niveau des QQ-plots, un petit manque d'ajustement aux niveaux des queues. On observe des queues à droite plus légères de la loi gamma transformée. En conséquence, on doit chercher une famille de lois dont d'une part les queues sont plus lourdes et d'autre part la loi gamma transformée est une sous-loi ou un cas limite.

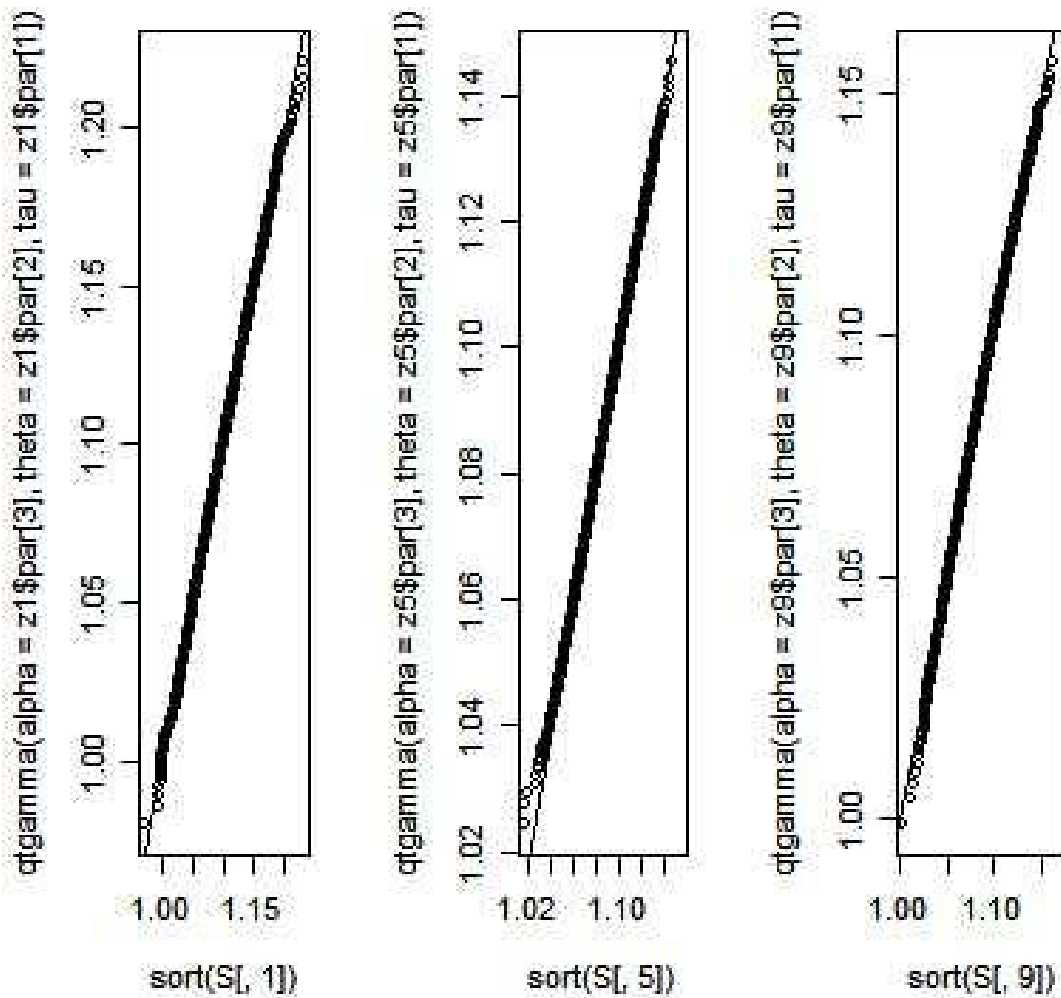


FIG. 6.3 – Exemples de QQ-plots d’une loi gamma transformée ajustée à des portefeuilles simulés sous l’hypothèse de l’indépendance

Il se trouve que la loi gamma transformée est un cas limite de la GBII, tel que discuté dans la sous-section 2.2.7. Par ailleurs, tous les moments positifs de la loi gamma transformée existent, ce qui explique que les membres de cette famille disposent de queues plus légères que ceux de la famille Bêta transformée. On en conclut donc que la GBII pourrait éventuellement mieux modéliser les rendements de nos portefeuilles.

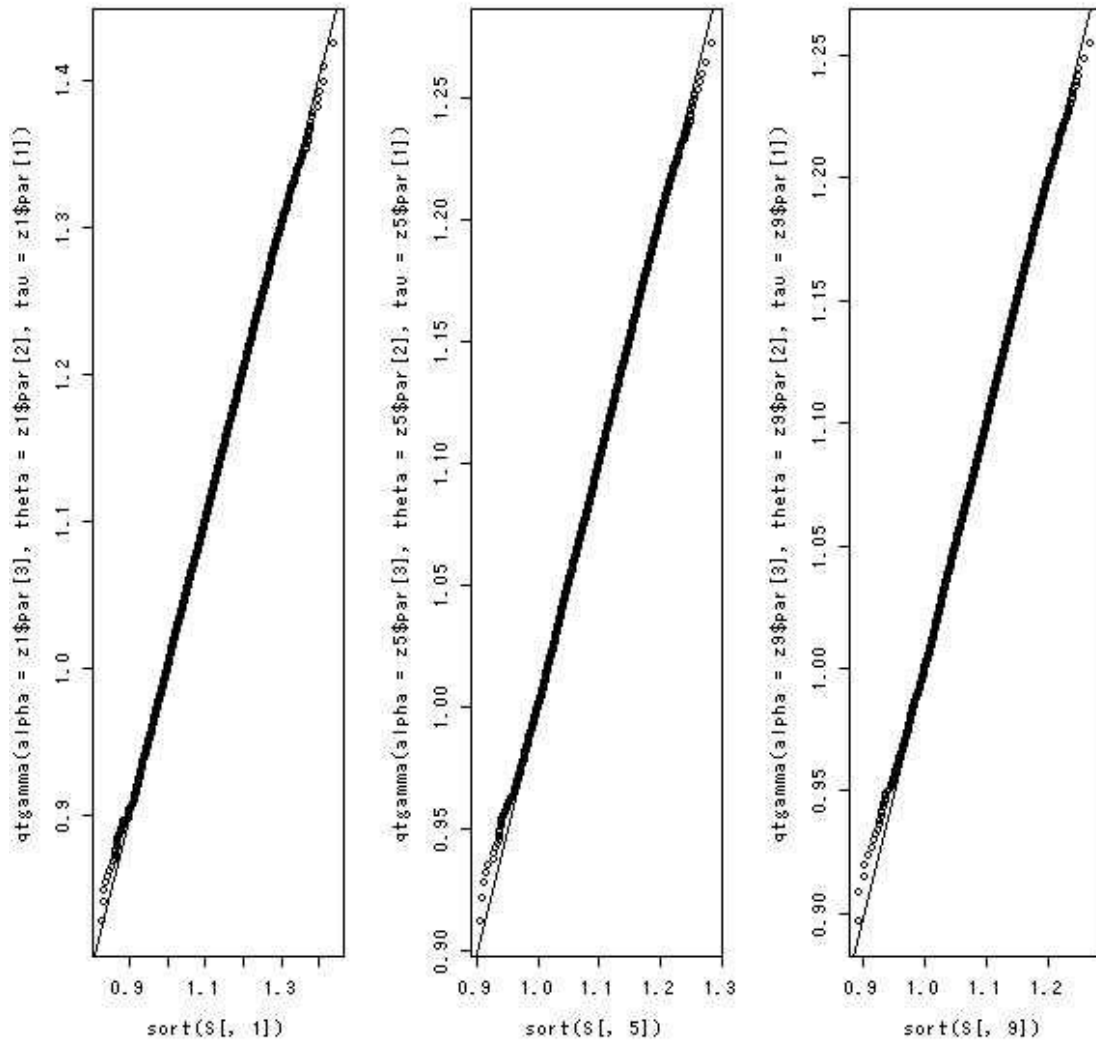


FIG. 6.4 – Exemples de QQ-plots d’une loi gamma transformée ajustée à des portefeuilles issus de la copule de t à niveau élevé de dépendance

6.1.3 La GBII

Après avoir programmé les fonctions de densité, de répartition, et de quantile de la loi GBII sur \mathbf{R} , on a estimé, pour chacun des portefeuilles, les paramètres de la GBII avec la méthode des moments, et celle du maximum de vraisemblance. L’ajustement offert par la GBII est meilleur que celui de la GT. Toutefois, il existait toujours un petit manque d’ajustement au niveau des queues et tout particulièrement de la queue gauche. Dans cette sous-section, on considèrera plutôt des puissances de portefeuilles (voir plus

bas). Le choix de cette transformation vise en premier lieu à améliorer l'ajustement. En second lieu, le choix d'une telle transformation plutôt qu'une autre réside d'une part dans l'éventail de scénarios qu'elle offre et d'autre part dans la facilité de son implantation.

On obtient une puissance de portefeuille $(S_i - 1)$ en élevant à une puissance donnée u_j de l'intervalle $[-2; 2]$ (choix arbitraire) les rendements de ce dernier.

$$(\mathbf{S}_i - \mathbf{1})^{u_j} = \begin{pmatrix} (S_{1,i} - 1)^{u_j} \\ (S_{2,i} - 1)^{u_j} \\ \vdots \\ (S_{10\ 000,i} - 1)^{u_j} \end{pmatrix}$$

Soit $G = \{-2, -2 + h, -2 + 2h, \dots, 2 - 2h, 2 - h, 2\}$, une fine grille de valeurs incluses dans $[-2, 2]$ (nous utilisons $h = 0.1$ dans nos analyses). Pour tout $u_j \in G$, nous ajustons la loi GBII au portefeuille $(S_i - 1)^{u_j}$ et nous mesurons la distance de Kolmogorov-Smirnov (L_∞) entre la loi empirique de $(S_i - 1)^{u_j}$ et la loi ajustée. La puissance retenue sera celle qui fournit la plus petite distance parmi ces valeurs. Le code **R** pour implanter cet algorithme est donné à l'annexe H.

Cette démarche est plus facile à implanter que d'avoir à estimer la puissance. En effet, estimer quatre paramètres est déjà un processus délicat, alors il est raisonnable d'éviter l'ajout d'un cinquième paramètre dans le problème de maximisation de la vraisemblance. Dans les tableaux 6.10 à 6.15 aux pages 79 à 83, on expose, cas par cas, les estimateurs obtenus, la valeur de la fonction objectif, la puissance ainsi que l'estimé de maximum de vraisemblance de la moyenne et la moyenne empirique.

En observant les tableaux, on note quelques puissances qui se répètent, telles que ± 0.1 , ± 0.2 et 0.3 . Quant à la fonction objectif, elle n'est plus un critère de comparaison entre les lois car les données sous-jacentes ont été transformées. En effet, les moments sont dans des échelles différentes car on considère des rendements transformés. Cependant, la loi de Weibull est une sous-loi de la GT qui est un cas limite de la GBII qui, à son tour, est un sous-cas d'une GBII avec puissance (où $u_j = 1$). Ainsi, si on avait estimé la puissance comme un cinquième paramètre, on s'attendrait à ce que la fonction objectif diminue davantage.

Par ailleurs, les QQ-plots de la loi GBII (voir annexe I) nous offrent des ajustements satisfaisants comparativement aux QQ-plots de la loi Weibull et de la loi gamma

transformée. La loi GBII ajustée à la puissance modélise donc mieux la plupart des rendements de portefeuilles simulées.

6.2 Conclusion

Au long de ce chapitre, on a abordé l'estimation des paramètres des différentes lois retenues, ainsi que leur adéquation à nos données simulées. La première de ces lois est la loi de Weibull. Pour celle-ci, les QQ-plots ont relevé un énorme manque d'ajustement. Pour cette raison, on a considéré la loi gamma transformée, dont la loi de Weibull est une sous loi. Malgré la nette amélioration de la modélisation des rendements, on a dû écarter cette loi. En effet, la loi gamma transformée a des queues trop légères. Dans notre recherche d'une loi aux queues plus épaisses, on a vérifié l'adéquation de la loi GBII. Toutefois, pour un meilleur ajustement, on s'est proposé de traiter des puissances de portefeuilles au lieu des portefeuilles. Cette démarche se montre fructueuse. Effectivement, la loi GBII s'est mieux ajustée à la plupart de nos données simulées après en avoir pris une puissance. Dans le chapitre suivant, nous verrons ce qu'il en est pour les données réelles.

| Portefeuille | Shape | Scale | Fonction objectif | Théorique | Empirique |
|--------------|--------|---------|-------------------|------------|-----------|
| 1 | 1.1161 | 33.6250 | -19241.3220 | 0.09784662 | 0.1002 |
| 2 | 1.1051 | 56.7450 | -24497.9779 | 0.09420406 | 0.09563 |
| 3 | 1.1357 | 28.9421 | -17419.7632 | 0.11433458 | 0.1164 |
| 4 | 1.1230 | 38.8060 | -20397.6048 | 0.10697413 | 0.1085 |
| 5 | 1.0954 | 68.0255 | -26316.6422 | 0.08633513 | 0.08743 |
| 6 | 1.1271 | 39.9899 | -20718.8104 | 0.11151292 | 0.1132 |
| 7 | 1.1119 | 39.4779 | -20842.7374 | 0.09633629 | 0.09831 |
| 8 | 1.0886 | 37.1709 | -20401.1953 | 0.07243469 | 0.07431 |
| 9 | 1.1002 | 55.0604 | -24109.7149 | 0.08905936 | 0.09026 |

TAB. 6.1 – Les différents estimateurs de la loi de Weibull pour les neuf portefeuilles simulés sous l'hypothèse de l'indépendance

| Portefeuille | Shape | Scale | Fonction objectif | théorique | empirique |
|--------------|--------|---------|-------------------|------------|-----------|
| 1 | 1.1149 | 31.4172 | -17993.3852 | 0.09551013 | 0.09611 |
| 2 | 1.1077 | 41.3157 | -20715.5462 | 0.09280407 | 0.09318 |
| 3 | 1.1201 | 31.5833 | -18068.5640 | 0.10067377 | 0.1015 |
| 4 | 1.1308 | 25.9997 | -15892.7638 | 0.10729409 | 0.1076 |
| 5 | 1.1212 | 23.4404 | -15189.9110 | 0.09552401 | 0.09678 |
| 6 | 1.1288 | 24.4904 | -15511.8594 | 0.10395759 | 0.105 |
| 7 | 1.1407 | 24.9884 | -15505.9142 | 0.11608197 | 0.1168 |
| 8 | 1.1393 | 18.9356 | -13123.6898 | 0.10755654 | 0.1099 |
| 9 | 1.1207 | 33.6038 | -18616.0737 | 0.10236951 | 0.103 |

TAB. 6.2 – Les différents estimateurs de la loi de Weibull pour les neuf portefeuilles issus d'une Clayton à faible niveau de dépendance

| Portefeuille | Shape | Scale | Fonction objectif | théorique | empirique |
|--------------|---------|---------|-------------------|------------|-----------|
| 1 | 1.1233 | 23.8179 | -15370.34120600 | 0.09792556 | 0.09928 |
| 2 | 1.1421 | 18.2090 | -12544.6327 | 0.10911542 | 0.1105 |
| 3 | 1.1217 | 24.4163 | -15546.7985 | 0.09696738 | 0.09799 |
| 4 | 1.1296 | 18.0707 | -12571.2843 | 0.09672744 | 0.0981 |
| 5 | 1.1388 | 20.3684 | -13587.4149 | 0.10910019 | 0.1101 |
| 6 | 1.1250 | 22.8030 | -14785.7520 | 0.09858482 | 0.09939 |
| 7 | 1.1220 | 19.3018 | -13269.8171 | 0.09129641 | 0.09256 |
| 8 | 1.10766 | 26.0623 | -16389.3584 | 0.08468452 | 0.08593 |
| 9 | 1.1313 | 25.6609 | -15823.7404 | 0.10750197 | 0.108 |

TAB. 6.3 – Les différents estimateurs de la loi de Weibull pour les neuf portefeuilles issus d'une Clayton à niveau élevé de dépendance

| Portefeuille | Shape | Scale | Fonction objectif | Théorique | Empirique |
|--------------|--------|---------|-------------------|------------|-----------|
| 1 | 1.1392 | 17.3444 | -12575.0518 | 0.10489223 | 0.1091 |
| 2 | 1.1301 | 24.3293 | -15660.3515 | 0.10510307 | 0.1071 |
| 3 | 1.1399 | 19.9899 | -13925.1755 | 0.10972480 | 0.1134 |
| 4 | 1.1313 | 27.5346 | -17167.2089 | 0.10898283 | 0.1119 |
| 5 | 1.1186 | 29.1776 | -17808.4424 | 0.09773149 | 0.1003 |
| 6 | 1.1165 | 24.7465 | -16287.7735 | 0.09220291 | 0.0955 |
| 7 | 1.1241 | 24.2296 | -15974.8772 | 0.09918547 | 0.1024 |
| 8 | 1.1320 | 24.7108 | -16081.1267 | 0.10733972 | 0.1105 |
| 9 | 1.1172 | 31.4921 | -18416.8064 | 0.09780202 | 0.09975 |

TAB. 6.4 – Les différents estimateurs de la loi de Weibull pour les neuf portefeuilles issus d'une copule de t à faible niveau de dépendance

| Portefeuille | Shape | Scale | Fonction objectif | Théorique | Empirique |
|--------------|--------|---------|-------------------|------------|-----------|
| 1 | 1.1462 | 14.0655 | -10379.9330 | 0.10455713 | 0.1087 |
| 2 | 1.1357 | 19.5593 | -13397.6849 | 0.10494552 | 0.1069 |
| 3 | 1.1460 | 16.3293 | -11808.2488 | 0.10952755 | 0.1132 |
| 4 | 1.1367 | 21.7218 | -14658.1509 | 0.10877681 | 0.1117 |
| 5 | 1.1230 | 23.5969 | -15553.4615 | 0.09749128 | 0.1 |
| 6 | 1.1224 | 19.3788 | -13750.4908 | 0.09180679 | 0.09531 |
| 7 | 1.1299 | 19.3192 | -13591.8123 | 0.09895958 | 0.1022 |
| 8 | 1.1371 | 20.0937 | -13885.6009 | 0.10710906 | 0.1101 |
| 9 | 1.1219 | 24.9301 | -15976.6791 | 0.09761621 | 0.09956 |

TAB. 6.5 – Les différents estimateurs de la loi de Weibull pour les neuf portefeuilles issus d'une copule de t à niveau élevé de dépendance

| Portefeuille | p | a | v | Fonction objectif | Théorique | Empirique |
|--------------|---------|--------|----------|-------------------|------------|-----------|
| 1 | 2.9149 | 0.2044 | 135.4444 | -20096.3396 | 0.10018651 | 0.1002 |
| 2 | 4.9089 | 0.4036 | 134.9675 | -25329.2036 | 0.09563349 | 0.09563 |
| 3 | 5.2415 | 0.5910 | 28.4474 | -17962.8099 | 0.11643030 | 0.1164 |
| 4 | 9.6492 | 0.8400 | 14.9791 | -20875.0002 | 0.10853325 | 0.1085 |
| 5 | 8.1318 | 0.6477 | 67.9939 | -26996.0022 | 0.08743701 | 0.08743 |
| 6 | 8.0208 | 0.7527 | 23.5249 | -21282.2889 | 0.11320374 | 0.1132 |
| 7 | 3.4174 | 0.2615 | 135.2481 | -21696.3869 | 0.09830924 | 0.09831 |
| 8 | 5.2287 | 0.5105 | 49.3187 | -21172.3675 | 0.07426423 | 0.07431 |
| 9 | 18.2486 | 0.9711 | 8.7315 | -24666.4997 | 0.09024533 | 0.09026 |

TAB. 6.6 – Les différents estimateurs de la gamma transformée pour les neuf portefeuilles simulés sous l'hypothèse de l'indépendance

| Portefeuille | p | a | v | Fonction objectif | Théorique | Empirique |
|--------------|---------|--------|--------|-------------------|------------|-----------|
| 1 | 22.1821 | 1.0800 | 1.8387 | -18048.3400 | 0.09611064 | 0.09611 |
| 2 | 31.7146 | 1.0890 | 1.5797 | -20749.3119 | 0.09317661 | 0.09318 |
| 3 | 19.4010 | 1.0646 | 2.3888 | -18169.1619 | 0.10147811 | 0.1015 |
| 4 | 22.3605 | 1.1153 | 1.2916 | -15904.3107 | 0.10762575 | 0.1076 |
| 5 | 13.3842 | 1.0300 | 2.7647 | -15326.7987 | 0.09677711 | 0.09678 |
| 6 | 9.8895 | 0.9444 | 5.1684 | -15576.8612 | 0.10506994 | 0.105 |
| 7 | 18.2321 | 1.1013 | 1.7315 | -15551.9160 | 0.11674801 | 0.1168 |
| 8 | 7.12709 | 0.8637 | 6.3992 | -13446.1857 | 0.10988301 | 0.1099 |
| 9 | 23.1802 | 1.0848 | 1.9202 | -18678.1139 | 0.10297519 | 0.103 |

TAB. 6.7 – Les différents estimateurs de la loi gamma transformée pour les neuf portefeuilles issus d'une Clayton à faible niveau de dépendance

| Portefeuille | p | a | v | Fonction objectif | Théorique | Empirique |
|--------------|---------|--------|--------|-------------------|------------|-----------|
| 1 | 8.1474 | 0.8665 | 7.3831 | -15622.3600 | 0.09929089 | 0.09928 |
| 2 | 7.7665 | 0.9194 | 4.7614 | -12732.8736 | 0.11054737 | 0.1105 |
| 3 | 11.9483 | 0.9964 | 3.6374 | -15692.8343 | 0.09799424 | 0.09799 |
| 4 | 8.2547 | 0.9359 | 4.1693 | -12741.8198 | 0.09810181 | 0.0981 |
| 5 | 11.4238 | 1.0298 | 2.7977 | -13692.4057 | 0.11009227 | 0.1101 |
| 6 | 13.2846 | 1.0378 | 2.5954 | -14875.1201 | 0.09940713 | 0.09939 |
| 7 | 8.6239 | 0.9330 | 4.3342 | -13435.6609 | 0.09257953 | 0.09256 |
| 8 | 9.2119 | 0.8865 | 6.9253 | -16635.2506 | 0.08593032 | 0.08593 |
| 9 | 17.6783 | 1.0847 | 1.9017 | -15870.1041 | 0.10805612 | 0.108 |

TAB. 6.8 – Les différents estimateurs de la loi gamma transformée pour les neuf portefeuilles issus d'une Clayton à niveau élevé de dépendance

| Portefeuille | p | a | v | Fonction objectif | Théorique | Empirique |
|--------------|---------|--------|----------|-------------------|------------|-----------|
| 1 | 1.7461 | 0.0792 | 100.4756 | -13390.8773 | 0.10906922 | 0.1091 |
| 2 | 10.1373 | 0.9477 | 5.2794 | -15971.0448 | 0.10709109 | 0.1071 |
| 3 | 1.7434 | 0.0678 | 131.7904 | -14707.9762 | 0.11335081 | 0.1134 |
| 4 | 2.7756 | 0.2120 | 99.8350 | -17966.9412 | 0.11190423 | 0.1119 |
| 5 | 4.3498 | 0.4604 | 44.6444 | -18545.4471 | 0.10030061 | 0.1003 |
| 6 | 2.1769 | 0.1153 | 134.6100 | -17193.2523 | 0.09549205 | 0.0955 |
| 7 | 2.4878 | 0.1758 | 96.5560 | -16790.1412 | 0.10239318 | 0.1024 |
| 8 | 2.5899 | 0.1943 | 91.6365 | -16858.6819 | 0.11045428 | 0.1105 |
| 9 | 9.6973 | 0.8724 | 9.8882 | -18871.3074 | 0.09975308 | 0.09975 |

TAB. 6.9 – Les différents estimateurs de la loi gamma transformée pour les neuf portefeuilles issus d'une copule de t à faible niveau de dépendance

| Portfeuille | p | a | v | Fonction objectif | Théorique | Empirique |
|-------------|--------|--------|----------|-------------------|------------|-----------|
| 1 | 1.5439 | 0.0648 | 80.2604 | -11054.5392 | 0.10862457 | 0.1087 |
| 2 | 8.9000 | 0.9496 | 4.3455 | -13638.1216 | 0.10687753 | 0.1069 |
| 3 | 2.1432 | 0.1710 | 55.7121 | -12450.1941 | 0.11314837 | 0.1132 |
| 4 | 3.3366 | 0.3677 | 40.4363 | -15273.5160 | 0.11166184 | 0.1117 |
| 5 | 4.7136 | 0.5647 | 23.5742 | -16102.0388 | 0.10003541 | 0.1 |
| 6 | 1.6642 | 0.0574 | 135.4729 | -14543.2326 | 0.09531929 | 0.09531 |
| 7 | 2.4073 | 0.1989 | 61.9548 | -14243.1013 | 0.10217908 | 0.1022 |
| 8 | 3.2810 | 0.3746 | 35.6655 | -14484.3077 | 0.11014225 | 0.1101 |
| 9 | 8.5896 | 0.8736 | 7.6468 | -16329.9851 | 0.09955410 | 0.09956 |

TAB. 6.10 – Les différents estimateurs de la loi gamma transformée pour les neuf portefeuilles issus d'une copule de t à niveau élevé de dépendance

| Portefeuille | Alpha | Lambda | Gamma | Tau | Objectif | Puissance | Théorique | Empirique |
|--------------|--------|---------|----------|--------|-------------|-----------|------------|------------|
| 1 | 1.7697 | 32.6111 | 409.2370 | 2.2941 | -44616.4733 | 0.1 | 1.00937966 | 1.00938014 |
| 2 | 2.2993 | 0.0317 | 206.0711 | 1.7056 | -37914.2359 | -0.2 | 0.98155461 | 0.98155606 |
| 3 | 2.0222 | 20.9905 | 185.5348 | 2.8472 | -37642.8380 | 0.2 | 1.01888129 | 1.01888329 |
| 4 | 2.0583 | 0.0239 | 434.8862 | 1.6055 | -44778.5662 | -0.1 | 0.99070550 | 0.99070515 |
| 5 | 1.2495 | 77.2907 | 168.1047 | 1.6112 | -33300.0549 | 0.3 | 1.02845937 | 1.02846030 |
| 6 | 2.1275 | 0.0254 | 211.8726 | 1.7153 | -38005.5714 | -0.2 | 0.98152757 | 0.98152793 |
| 7 | 1.2627 | 78.2591 | 509.1079 | 1.6579 | -44684.9886 | 0.1 | 1.00937163 | 1.00937109 |
| 8 | 3.2951 | 6.2070 | 262.6918 | 5.7448 | -44507.7376 | 0.1 | 1.00937923 | 1.00938194 |
| 9 | 2.7232 | 11.0455 | 306.0192 | 4.0070 | -44598.8114 | 0.1 | 1.00937058 | 1.00937810 |

TAB. 6.11 – Les différents estimateurs de la GBII pour les neuf portefeuilles simulés sous l'hypothèse de l'indépendance

| Portefeuille | Alpha | Lambda | Gamma | Tau | Objectif | Puissance | Théorique | Empirique |
|--------------|--------|----------|----------|--------|-------------|-----------|------------|------------|
| 1 | 2.4554 | 54.4083 | 305.5267 | 1.1564 | -40149.8920 | 0.1 | 1.00977716 | 1.00978384 |
| 2 | 2.0007 | 109.5036 | 119.7250 | 0.8916 | -28950.4965 | 0.3 | 1.02977109 | 1.02976930 |
| 3 | 2.4675 | 64.7889 | 313.8475 | 1.0961 | -40173.4513 | 0.1 | 1.00979063 | 1.00979513 |
| 4 | 2.0911 | 100.4376 | 350.6751 | 0.9201 | -40136.4218 | 0.1 | 1.00978928 | 1.00979395 |
| 5 | 2.4464 | 67.0920 | 316.2516 | 1.0773 | -40147.8086 | 0.1 | 1.00977346 | 1.00977446 |
| 6 | 1.7887 | 115.4152 | 373.1268 | 0.8594 | -40109.0778 | 0.1 | 1.00980317 | 1.00980397 |
| 7 | 2.3796 | 59.8075 | 311.8183 | 1.1056 | -40073.6787 | 0.1 | 1.00978431 | 1.00978591 |
| 8 | 2.2773 | 90.4590 | 337.4387 | 0.9673 | -40144.3413 | 0.1 | 1.00979273 | 1.00979146 |
| 9 | 1.6003 | 137.4366 | 395.3058 | 0.8051 | -40123.4395 | 0.1 | 1.00978032 | 1.00977959 |

TAB. 6.12 – Les différents estimateurs de la GBII pour les neuf portefeuilles issus d'une Clayton à faible niveau de dépendance

| Portefeuille | Alpha | Lambda | Gamma | Tau | Objectif | Puissance | Théorique | Empirique |
|--------------|----------|----------|----------|---------|-------------|-----------|-------------|------------|
| 1 | 9.2526 | 18.9313 | 161.9953 | 2.5878 | -39459.7939 | 0.1 | 1.00941502 | 1.00941235 |
| 2 | 154.9292 | 14.7733 | 21.2506 | 16.6853 | -29605.3334 | 0.2 | 1.020870571 | 1.02087930 |
| 3 | 161.6842 | 33.3328 | 36.1629 | 9.9273 | -32485.7731 | 0.2 | 1.01868933 | 1.01869404 |
| 4 | 21.7520 | 20.2180 | 102.4665 | 3.1761 | -36580.7229 | 0.1 | 1.00923685 | 1.00922402 |
| 5 | 149.9887 | 89.7610 | 44.1351 | 4.6175 | -30591.1911 | 0.2 | 1.02084188 | 1.02084419 |
| 6 | 124.5369 | 247.4432 | 141.4497 | 2.3587 | -38739.6227 | 0.1 | 1.00940318 | 1.00940265 |
| 7 | 31.4601 | 28.3770 | 106.0266 | 3.2239 | -37234.2916 | 0.1 | 1.00873642 | 1.00873590 |
| 8 | 144.0808 | 7.9470 | 24.7162 | 27.5323 | -33341.0985 | 0.2 | 1.01647446 | 1.01647579 |
| 9 | 125.2324 | 236.3293 | 72.1954 | 2.7677 | -32760.7579 | 0.2 | 1.02055225 | 1.02055408 |

TAB. 6.13 – Les différents estimateurs de la GBII pour les neuf portefeuilles issus d'une Clayton à niveau élevé de dépendance

| Portefeuille | Alpha | Lambda | Gamma | Tau | Objectif | Puissance | Théorique | Empirique |
|--------------|--------|---------|----------|--------|-------------|-----------|------------|------------|
| 1 | 6.0263 | 3.9575 | 111.7628 | 4.8590 | -37349.8365 | 0.1 | 1.01026138 | 1.01025801 |
| 2 | 1.5582 | 0.0270 | 238.6676 | 4.1481 | -40129.1454 | -0.1 | 0.98998815 | 0.98998756 |
| 3 | 5.4259 | 5.3945 | 135.6489 | 4.3495 | -38699.1766 | 0.1 | 1.01068405 | 1.01068290 |
| 4 | 2.6188 | 0.0549 | 251.3323 | 3.2386 | -42176.9133 | -0.1 | 0.98951934 | 0.98951918 |
| 5 | 2.8173 | 0.0675 | 247.3723 | 3.7863 | -42638.4997 | -0.1 | 0.99054833 | 0.99054853 |
| 6 | 4.2866 | 0.1866 | 181.4612 | 4.4346 | -41234.4632 | -0.1 | 0.99100641 | 0.99100585 |
| 7 | 3.8045 | 8.5765 | 200.0131 | 3.1575 | -40704.5721 | 0.1 | 1.00972034 | 1.00972100 |
| 8 | 3.7883 | 11.1027 | 206.0878 | 3.0175 | -40840.0436 | 0.1 | 1.01045952 | 1.01045870 |
| 9 | 1.9378 | 0.0266 | 289.3476 | 4.0697 | -42962.0824 | -0.1 | 0.99059870 | 0.99059838 |

TAB. 6.14 – Les différents estimateurs de la GBII pour les neuf portefeuilles issus d'une copule de t à faible niveau de dépendance

| Portefeuille | Alpha | Lambda | Gamma | Tau | Objectif | Puissance | Théorique | Empirique |
|--------------|----------|---------|----------|---------|-------------|-----------|-------------|------------|
| 1 | 2.2621 | 0.1992 | 136.2901 | 2.7565 | -35170.8634 | -0.1 | 0.99001774 | 0.99002379 |
| 2 | 156.8240 | 17.8560 | 24.5366 | 14.7895 | -30476.2674 | 0.2 | 1.02024103 | 1.02025220 |
| 3 | 2.4134 | 0.1502 | 148.7798 | 3.1740 | -36622.5305 | -0.1 | 0.98954616 | 0.98954952 |
| 4 | 3.3716 | 0.1181 | 157.0731 | 5.2078 | -39452.3764 | -0.1 | 0.98959097 | 0.98959108 |
| 5 | 2.5141 | 0.0872 | 196.9578 | 4.1247 | -40165.5336 | -0.1 | 0.99061668 | 0.99061729 |
| 6 | 5.6428 | 3.2569 | 122.8612 | 5.2522 | -38377.0220 | 0.1 | 1.009023772 | 1.00902415 |
| 7 | 4.3483 | 5.0895 | 144.5214 | 3.5178 | -38127.9343 | 0.1 | 1.009648963 | 1.00964947 |
| 8 | 6.1800 | 6.0018 | 130.3315 | 4.1190 | -38439.4210 | 0.1 | 1.010372241 | 1.01038363 |
| 9 | 1.9231 | 0.0454 | 219.9299 | 4.5386 | -40389.1929 | -0.1 | 0.990655745 | 0.99065689 |

TAB. 6.15 – Les différents estimateurs de la GBII pour les neuf portefeuilles issus d'une copule de t à niveau élevé de dépendance

Chapitre 7

Données réelles

Dans ce chapitre, on se propose d'appliquer le travail fait au chapitre précédent aux données réelles. Ainsi, la procédure suivie est similaire à celle suivie au chapitre 6. Dans le chapitre précédent, on a procédé à l'estimation des paramètres des différentes lois retenues, puis à la vérification de leur adéquation à nos données simulées. Cette vérification a été faite en comparant les moyennes empiriques aux estimés de maximum de vraisemblance des moyennes, ainsi qu'à l'aide des QQ-plots. Dans cette perspective, on considère certains des résultats des neuf portefeuilles issus des données réelles.

7.1 Estimation et adéquation

7.1.1 La loi de Weibull

À l'aide des différentes méthodes d'estimation programmées avec le logiciel **R**, on se propose d'estimer les paramètres de la loi de Weibull pour les neuf portefeuilles issus des données réelles. Dans le tableau 7.1 à la page 88, on expose les estimateurs obtenus, la valeur de la fonction objectif, ainsi que l'estimé de maximum de vraisemblance de la moyenne et la moyenne empirique.

On observe toujours la loi de Weibull avec un paramètre de forme supérieur à 1. Toutefois, l'écart entre la moyenne théorique et la moyenne empirique grandit. Par ailleurs, les QQ-plots montrent un énorme manque d'ajustement au niveau de la queue gauche, comme le graphique de la figure 7.1 en témoigne.

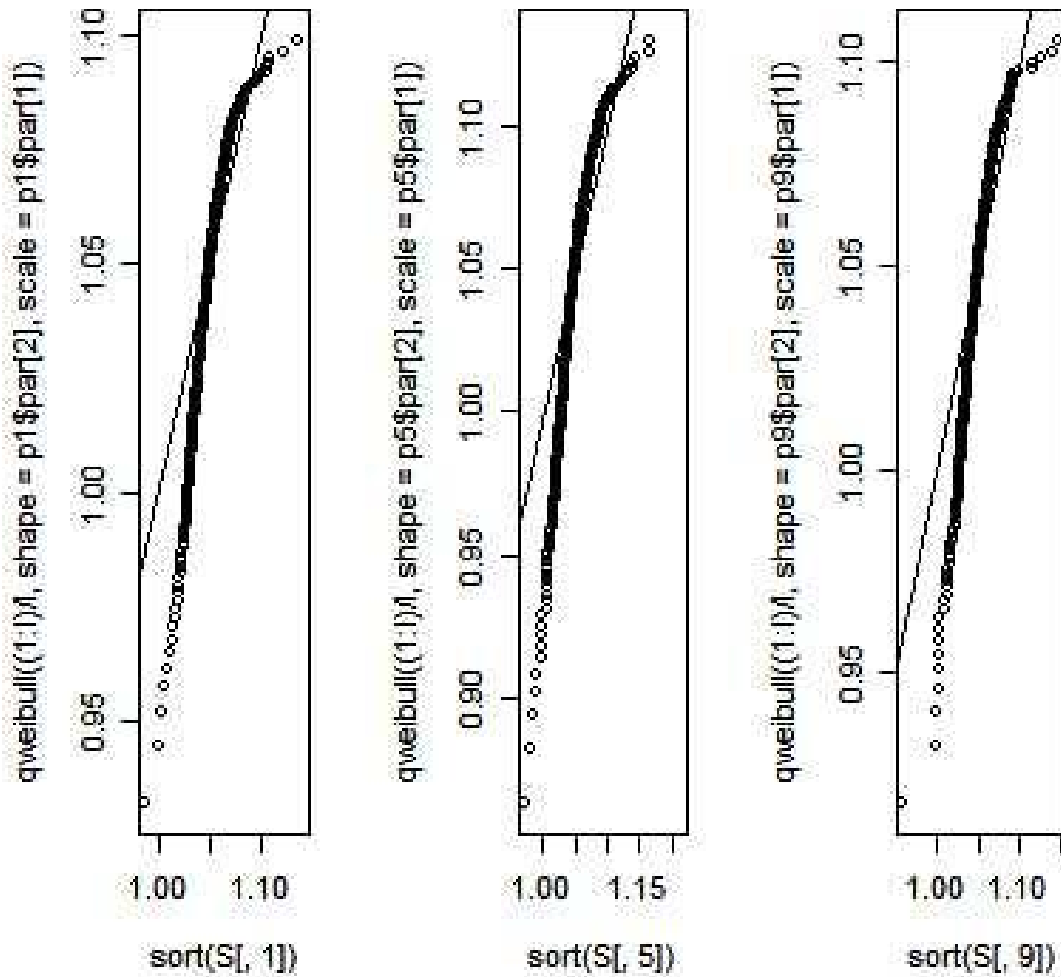


FIG. 7.1 – Exemples de QQ-plots d’une loi de Weibull pour des portefeuilles issus des données réelles

Dans une démarche similaire au chapitre précédent, on considère la loi gamma transformée dont la loi de Weibull est un sous-cas.

7.1.2 La loi gamma transformée

On se propose de vérifier l’adéquation de la loi gamma transformée aux neuf portefeuilles issus des données réelles. Dans le tableau 7.2 à la page 89, on expose les

estimateurs obtenus, la valeur de la fonction objectif, ainsi que l'estimé de maximum de vraisemblance de la moyenne et la moyenne empirique.

Les QQ-plots font ressortir un manque d'ajustement aux niveaux des queues. Comme le montre la figure 7.2, ce manque d'ajustement est plus important que celui observé pour les données simulées, mais beaucoup moins flagrant qu'avec la loi de Weibull.

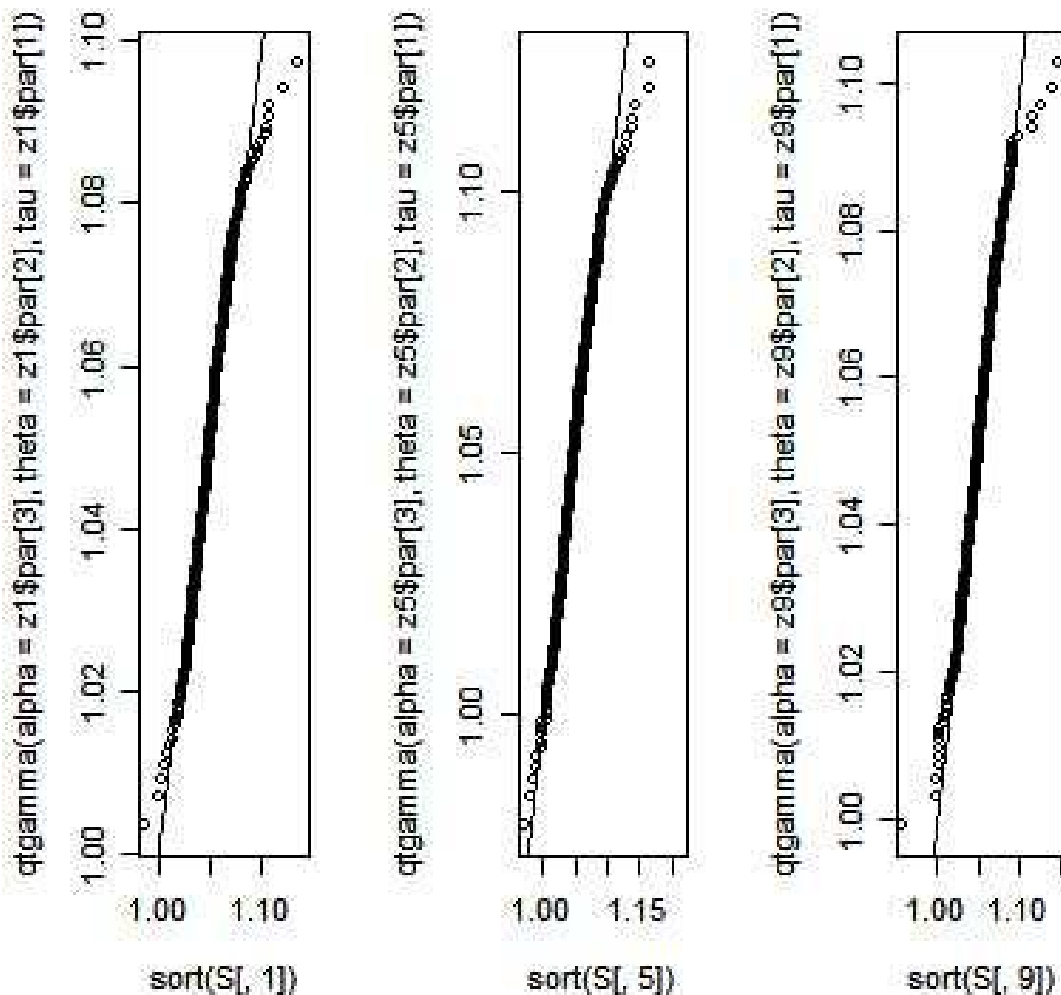


FIG. 7.2 – Exemples de QQ-plots d'une loi gamma transformée pour des portefeuilles issus des données réelles

La démarche qu'on s'est proposée pour ce chapitre implique de considérer la loi GBII, dont la loi gamma transformée est un cas limite.

7.1.3 La loi GBII

On se propose de vérifier l'adéquation de la loi GBII aux neuf portefeuilles issus des données réelles. Dans le tableau 7.3 à la page 90, on expose les estimateurs obtenus, la valeur de la fonction objectif, la puissance, ainsi que l'estimé de maximum de vraisemblance de la moyenne et la moyenne empirique.

On observe dans ce tableau à six reprises une puissance de -0.1 , à deux reprises une puissance de 0.1 , et enfin une puissance de 0.2 . Aucune famille de lois n'offre un ajustement parfait pour tous les portefeuilles considérés, surtout dans la queue gauche. Cependant, la loi GBII présente en général une meilleure adéquation que les autres lois.

7.2 Conclusion

Dans le chapitre 6, la loi GBII nous a offert le meilleur ajustement pour une puissance de nos données simulées. En appliquant, dans ce chapitre, le travail fait au chapitre précédent aux données réelles, on trouve que la loi GBII modélise raisonnablement les rendements d'une puissance de portefeuille sauf au niveau de la queue gauche.

| Portefeuille | Shape | Scale | Fonction objectif | Théorique | Empirique |
|--------------|--------|---------|-------------------|------------|-----------|
| 1 | 1.0593 | 51.7599 | -1874.4290 | 0.04785942 | 0.05143 |
| 2 | 1.0585 | 56.0955 | -1932.2021 | 0.04797770 | 0.05123 |
| 3 | 1.0576 | 65.0499 | -2019.6557 | 0.04846551 | 0.05079 |
| 4 | 1.0627 | 38.8826 | -1652.6775 | 0.04757160 | 0.05198 |
| 5 | 1.0644 | 31.4727 | -1512.5690 | 0.04591973 | 0.05202 |
| 6 | 1.0582 | 60.9324 | -1978.8744 | 0.04842639 | 0.05109 |
| 7 | 1.0570 | 63.2899 | -2041.7529 | 0.04762047 | 0.05094 |
| 8 | 1.0573 | 65.4611 | -2040.2078 | 0.04818498 | 0.05084 |
| 9 | 1.0605 | 45.8874 | -1783.6715 | 0.04769191 | 0.05168 |

TAB. 7.1 – Les différents estimateurs de la loi de Weibull pour les neuf portefeuilles issus des données réelles

| Portefeuille | p | a | v | Fonction objectif | Théorique | Empirique |
|--------------|---------|--------|----------|-------------------|------------|-----------|
| 1 | 6.3032 | 0.4960 | 114.3362 | -2065.2748 | 0.05139485 | 0.05143 |
| 2 | 6.4621 | 0.4939 | 132.1898 | -2119.1448 | 0.05126357 | 0.05123 |
| 3 | 10.2439 | 0.7111 | 55.0244 | -2160.0675 | 0.05071734 | 0.05079 |
| 4 | 4.3179 | 0.3393 | 132.7361 | -1825.8413 | 0.05196716 | 0.05198 |
| 5 | 3.7951 | 0.2915 | 130.7908 | -1725.4578 | 0.05199124 | 0.05202 |
| 6 | 6.4988 | 0.4958 | 132.4909 | -2129.8986 | 0.05110396 | 0.05109 |
| 7 | 7.5691 | 0.5526 | 130.2395 | -2238.9080 | 0.05091882 | 0.05094 |
| 8 | 7.0729 | 0.5274 | 131.6155 | -2191.1231 | 0.05083239 | 0.05084 |
| 9 | 5.3045 | 0.4189 | 132.5080 | -1964.9324 | 0.05172481 | 0.05168 |

TAB. 7.2 – Les différents estimateurs de la loi gamma transformée pour les neuf portefeuilles issus des données réelles

| Portefeuille | Alpha | Lambda | Gamma | Tau | Objectif | Puissance | Théorique | Empirique |
|--------------|------------|-------------|-----------|--------|------------|-----------|------------|------------|
| 1 | 0.45724255 | 7.16 e-06 | 2440.7650 | 0.3971 | -3844.1237 | -0.1 | 0.99499936 | 0.99500168 |
| 2 | 0.4801 | 1.62 e-05 | 2272.0897 | 0.4307 | -3831.1015 | -0.1 | 0.99503217 | 0.99503273 |
| 3 | 0.3833 | 110681.5082 | 2433.0274 | 0.4635 | -3825.8553 | 0.1 | 1.00500577 | 1.00500453 |
| 4 | 0.5716 | 6.61 e-05 | 2038.9966 | 0.4531 | -3816.6324 | -0.1 | 0.99501546 | 0.99501556 |
| 5 | 0.3878 | 95192.6480 | 2374.2637 | 0.4652 | -3812.3931 | 0.1 | 1.00505551 | 1.00505570 |
| 6 | 0.3958 | 3.31 e-06 | 2572.4695 | 0.3664 | -3812.1580 | -0.1 | 0.99501795 | 0.99501787 |
| 7 | 0.2471 | 2.56 e-09 | 2012.3734 | 0.2315 | -3321.5890 | -0.2 | 0.99007800 | 0.99007915 |
| 8 | 0.3632 | 9.40 e-07 | 2892.9232 | 0.3173 | -3823.4188 | -0.1 | 0.99506168 | 0.99506191 |
| 9 | 0.4056 | 2.64 e-06 | 2665.1152 | 0.3509 | -3831.1260 | -0.1 | 0.99502826 | 0.99502837 |

TAB. 7.3 – Les différents estimateurs de la loi GBII pour les neuf portefeuilles issus des données réelles

Chapitre 8

Conclusion

L'objectif de ce mémoire était d'identifier une famille de lois qui modélise au mieux la distribution des rendements de portefeuilles et ce sous une vaste gamme de conditions quant à la distribution conjointe des rendements des titres individuels constituant ce portefeuille. Au début des chapitres 6 et 7, on a vu que l'adéquation de la loi de Weibull a été réfutée. Par la suite, les queues légères de la loi gamma transformée nous ont suggéré la loi GBII. On a toutefois considéré des puissances de portefeuilles. En effet, l'étude des QQ-plots de la loi GBII nous ont mené à conclure que la loi GBII manquait légèrement d'ajustement. Cependant, elle pourrait peut-être mieux modéliser les rendements d'une puissance de portefeuille.

La loi GBII serait un choix raisonnable mais peut-être pourrait-on améliorer ce résultat. En effet, il existe des lois plus générales et plus flexibles telles que la loi bêta généralisée à cinq paramètres (GB) introduite par [McDonalda et Xu \(1995\)](#), et récemment, la loi Compound Confluent Hypergeometric (CCH) à six paramètres introduite par [Gordy \(1998\)](#), qui l'a utilisée pour modéliser la répartition des revenus familiaux aux États Unis. Ces lois pourrait possiblement offrir une meilleure modélisation des rendements de portefeuilles. Toutefois, l'utilisation de ces lois dans le cadre de l'approche préconisée dans ce mémoire conduit à de sérieuses difficultés numériques (par exemple, la maximisation de fonctions irrégulières en six dimensions). Il sera intéressant de poursuivre cette étude lorsqu'une approche numérique automatisée permettant de résoudre ces difficultés aura été identifiée.

Bibliographie

- BAWA, V. S. (1975). Optimal rules for ordering uncertain prospects. *J. of Financial Economics*, (No. 2):95–121.
- BOUYÉ, E., DURRLEMAN, V., NIKEGHBALI, A., RIBOULET, G. et RONCALLI, T. (2001). Copulas : an open field for risk management.
- CENCIA, M. et FILIPPINIB, F. (2006). Portfolio selection : A linear approach with dual expected utility. *Applied Mathematics and Computation*, Volume 179(Issue 2):523–534.
- CROW, E. L. et SHIMIZU, K. (1988). *Lognormal Distributions : Theory and Applications, Statistics : textbooks and monographs, Volume 88*. Marcel Dekker, Inc., New York.
- CUVELIER, Etienne et Noirhomme-Fraiture, M. (2003). Mélange de distributions de distributions, décomposition de mélange avec la copule de clayton. *XXXV èmes Journées de Statistiques*.
- DEMARTA, S. et MCNEIL, A. J. (2004). The t copulas and related copulas.
- EMBRECHTS, P., MCNEIL, A., STRAUMANN, D. et KAUFMANN, R. (1999). Modeling extremal events for insurance and finance. *SAV Summer School, Université de Lausanne, one-week course*.
- FAVRE, A.-C. (2000). Copules. *Chaire en Hydrologie Statistique, INRS, Université du Québec*.
- GENEST, C. (1987). Frank's family of bivariate distributions. *Biometrika*, No. 47:549–555.
- GENEST, C. et Mackay, R. (1986). Copules archimédiennes et familles de lois bidimensionnelles dont les marges sont données. *The Canadian journal of statistics*, No. 14:145–159.
- GORDY, M. B. (1998). A generalization of generalized beta distributions. *Board of Governors of the Federal Reserve System*.

- HOGG, R. V. et CRAIG, A. T. (1995). *Introduction to Mathematical Statistics*. Prentice Hall, New Jersey.
- HOGG, R. V. et KLUGMAN, S. A. (1984). *Loss Distributions*. John Wiley and Sons, Inc., New York.
- JOHNSON, N. L., KOTZ, S. et BALAKRISHNAN, N. (1994). *Continuous Univariate Distributions, Volume 1, Second Edition*. John Wiley and Sons, Inc., New York.
- KIMELDORF, G. et SAMPSON, A. (1975). Uniform representations of bivariate distributions. *Com. Statist.*, No. 4:617–627.
- KLEIN, J. P. et MOESCHBERGER, M. L. (2003). *Survival Analysis : Techniques for Censored and Truncated Data, Second Edition*. Springer-Verlag New York, Inc. New York.
- LEE, A. J. (1993). Generating random binary deviates having fixed marginal distributions and specified degrees of association. *Am. Statistician*, No. 47 Part 3:209–215.
- MARKOWITZ, H. (1952). Portfolio selection. *The Journal of Finance*, Volume 7:77–91.
- MCDONALDA, J. B. et XU, Y. J. (1995). A generalization of the beta distribution with applications. *Journal of Econometrics*, No. 69(Issue 2):427–428.
- NADARAJAH, S. et KOTZ, S. (2005). A generalized beta distribution II.
- SCHWEIZER, B. et SKLAR, A. (1983). Probabilistic metric spaces. *North-Holland, Amsterdam*.
- SKLAR, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut de Statistique de L'Université de Paris*, No. 8:229–231.
- STACY, E. W. et MIHRAM, G. A. (1965). Parameter estimation for a generalized gamma distribution. *Technometrics*, Volume 7(No. 3):349–358.

Annexe A

Approche numérique d'estimation

A.1 Méthode des moments

A.1.1 `uniroot`

Définition

La procédure `uniroot` nous permet de résoudre numériquement les équations de la méthode des moments pour les lois à deux paramètres. Lors de la programmation, nous utiliserons la nomenclature suivante :

1. Soit `scalegamma` le paramètre d'échelle, `shapegamma` le paramètre de forme, et `rategamma` le taux de la Gamma ;
2. Soit `meanlog` et `sdlog` les paramètres de la Lognormale ;
3. Soit `scale` le paramètre d'échelle, `shape` le paramètre de forme de la Weibull, et une inconnue `C`.

Exemples

Exemple 1 : Estimation des paramètres de la Weibull

```
Paramweibull <-function(moy,sigma)
```

```
{
f1<-function(shape)
{
log(gamma(1+ 2/shape))- 2*log(gamma(1 + 1/shape))
+ 2*log(1 + moy) - log((1 + moy)^2 + sigma^2)
}
```

```
shape<-uniroot (f=f1,lower=1,upper=1000)$root
```

```
f2<-function (C)
{
(1/C)^(1/shape)*gamma(1 + 1/shape) - (1 + moy)
}
```

```
C<-uniroot (f=f2,lower=0,upper=1000)$root
```

```
scale<-C^(-1/shape)
}
```

Exemple 2 : Estimation des paramètres de la gamma transformée par la méthode des moments

```
estim.moments.GT<-function (alpha)
{
((mean(x^3)-3*mean(x^2)*mean(xx)+2*(mean(x)^3))/
((mean(x^2)-mean(x)^2)^(3/2)))-psigamma(alpha,deriv=2)
}
```

```
alphaMMGT<-uniroot (f=estim.moments.GT,lower=1,upper=1000)$root
```

```
tauMMGT<-sqrt(trigamma(alpha1)/(mean(x^2)-mean(x)^2))
```

```
thetaMMGT<-1/exp((digamma(alpha1)/tau1)-mean(x))
```

* Avec

```
x <- log(S[,i]) ; i = 1,2,...,9
```

A.1.2 FSolve

Pour les lois les plus complexes, le principe est de construire autant d'équations que de paramètres. La commande `FSolve`, du package `RMinpack` de R, nous permet de résoudre simultanément les équations des moments. Comme toute méthode numérique, `FSolve` exige des points de départ comme argument ; un vecteur de points de départ de longueur égale au nombre des paramètres. En option, on peut se fixer une tolérance, mais, autrement, il est impossible de se fixer un nombre d'itérations (la commande `FSolve` dispose d'un nombre maximal d'itérations, déjà implanté dans son code avant compilation).

$$FSolve(f, x, tol = 1e - 12, \dots)$$

où

`f` : la fonction à optimiser, et dont les racines constituent les estimateurs des paramètres. La solution porte sur le premier paramètre qui doit être défini comme un vecteur numérique (de longueur arbitraire). Cette fonction doit retourner un vecteur numérique de même longueur que le vecteur des paramètres.

`x` : première approximation (points de départ).

`tol` : un critère de tolérance qui, une fois atteint, arrête les itérations. Il peut être défini comme l'erreur relative de l'estimation.

`...` : tout autre paramètre à assigner à la fonction 'f'

La valeur retournée par `FSolve` est une liste dont les composantes sont

`root` : les racines recherchées.

`value` : la valeur de la fonction - on note un succès plus cette valeur s'approche de 0.

`icode` : un entier signalant la raison de l'arrêt des itérations (1 signifie le succès des itérations, les autres valeurs indiquent une erreur)

Exemples

Exemple 1 : Minimisation de l'écart entre les moments théoriques et empiriques de la Weibull

```
estim.moments.Weibull<-function(p)
{
diff<-p[2]^(1:2)*gamma(1 + (1:2)/p[1])
- c(moment(y,1),moment(y,2))

diff
}
```

* charger les packages RMinpack et moments

```
pMMW<-FSolve(estim.moments.Weibull,p0,tol=1e-20)$root
```

Exemple 2 : Minimisation de l'écart entre les moments théoriques et empiriques de la GBII

```
estim.moments.GBII<-function(p)
{
diff<-(p[2]^(1:4/p[3])*gamma(p[4] + 1:4/p[3])*
gamma(p[1] - 1:4/p[3])/(gamma(p[1])*gamma(p[4])))
-c(moment(y,1),moment(y,2),moment(y,3),moment(y,4))

diff
}
```

```
pMMGB <- FSolve(estim.moments.GBII,p0,tol=1e-20)$root
```

* Avec

```
y<-S[,i] ; i = 1,2,...,9
```

A.2 Méthodes de maximum de vraisemblance et de distance minimale

La procédure `optim` de R permet d'optimiser la fonction de vraisemblance (ou l'une des distances citées à la section 3.4). Par défaut, `optim` minimise. Nous minimisons l'opposé de la fonction de log-vraisemblance (ou l'une des distances de la section 3.4).

```
optim(par, fn, gr = NULL, method = c("Nelder - Mead", "BFGS", "CG", "L - BFGS - B", "SANN"), lower = -Inf, upper = Inf, control = list(), hessian = FALSE, ...)
```

où

`par` : une première approximation (points de départ).

`fn` : la fonction à minimiser (ou maximiser), avec pour premier argument un vecteur de paramètres, sur lesquels la minimisation va être opérée. La fonction doit retourner un scalaire.

`gr` : une fonction qui retournera le gradient pour les méthodes `"BFGS"`, `"CG"` et `"L-BFGS-B"`. Si elle vaut `'NULL'`, une différence finie approximée sera utilisée. Pour la méthode `"SANN"`, elle spécifie une fonction qui génère de nouveaux candidats. Si elle vaut `'NULL'`, un noyau (Gaussian Markov kernel) sera utilisé.

`method` : la méthode à utiliser.

`lower, upper` : les bornes des paramètres pour la méthode `"L-BFGS-B"`.

`control` : la liste des paramètres de contrôle.

`hessian` : de type logique. Une indicatrice Vrai/Faux décidant si une matrice hessienne, obtenue numériquement, doit être retournée.

`...` : autres arguments des fonctions `'fn'` et `'gr'`.

Notons que la fonction 'fn' peut retourner 'NA' ou 'Inf', lorsqu'elle ne peut pas être évaluée. Certaines options de la liste 'control' méritent davantage d'attention, comme 'maxit' pour le nombre d'itérations, et 'reitol' pour la tolérance.

La commande `optim` retourne une liste dont les éléments sont

`par` : les valeurs des paramètres optimisant la fonction objectif.

`value` : la valeur de 'fn' correspondant à 'par'.

`counts` : un vecteur de deux entiers retournant, respectivement le nombre d'appels de 'fn' et 'gr', excluant ceux utilisés pour l'optimisation du hessien, si requis, et tout autre appel de 'fn' pour optimiser une approximation du gradient.

`convergence` : un entier signalant la raison de l'arrêt des itérations (0 signifie le succès des itérations, les autres valeurs indiquent une erreur)

`message` : donne de plus amples informations concernant l'optimisation, ou retourne 'NULL'.

`hessian` : seulement si l'argument 'hessian' est vrai. Il retourne une matrice symétrique estimant le hessien correspondant à la solution retournée.

Exemples

Exemple 1 : Maximum de vraisemblance de la Weibull

```
logvrais.Weibull<- fonction(p)
{
e<- dweibull(y,p[1],p[2])
-sum(log(e))
}
```

```
pMVW <- optim(p0,logvrais.ind, control=list(maxit=1e6))$par
```

Exemple 2 : Maximum de vraisemblance de la gamma transformée


```

logvrais.GT<-function(p)
{
e<-(p[3]*((y/p[2])^(p[1]*p[3]))*(exp(-((y/p[2])^(p[3])))))/(y*gamma(p[1])))

if(any(e <= 0)) Inf else -sum(log(e))
}

```

```

pEMVGT<-optim(p0,logvrais.GT,control=list(maxit=1e6))$par

```

Exemple 3: Distance minimale de la gamma transformée

```

ptgamma<- function(x,alpha,teta,tau)
{
u<-(x/teta)^tau

pgamma(u,shape=alpha)
}

```

```

ecdf2<-function(t,k)
{
  len <- length(k)
  sum(k<=t)/len
}

```

```

DM.GT<-function(p)
{
e<-(1-sapply(x,ecdf2,k=x))*((ptgamma

```

```
(x=sort(y),alpha=p[1],teta=p[2],tau=p[3]) -
sapply(x,ecdf2,k=x))^2)
```

```
sum(e)
}
```

```
pDMGT<-optim(p0,DM.GT,control=list(maxit=1e6,tol=1e-20))$par
```

Exemple 4 : Maximum de vraisemblance de la GBII

```
logvrais.GBII<- fonction(p)
{
e<-(log((gamma(p[1] + p[4])*p[3]*(y/(p[2]^(1/p[3]))))^(p[3]*p[4])/y/
gamma(p[1])/gamma(p[4]))/(p[2] + y^p[3])^(p[1] + p[4])))
-sum(e)
}
```

```
pEMVGB<-optim(p1,logvrais.GBII,control=list(maxit=1e6))$par
```

Annexe B

Approche numérique pour la simulation des copules

B.1 Le package Copula

La simulation des copules archimédiennes s'opère facilement avec le package **Copula** de **R**. Cependant, le nombre des variables aléatoires est limité pour certaines copules telles que la Frank et la Gumbel. En fait, la commande **mvdc** peine à partir de six variables. Toutefois, la Clayton est assez rapide d'exécution.

mvdc(**copula** = *nom de la copule* (param = *valeur du paramètre de la copule* , dim = *n*) , **margins** = *c(nom des n lois marginales retenues)* , **paranMargins** = *list(list(liste des paramètres de la 1ère loi) , ... , list(liste des paramètres de la nème loi))*)

Avec les copules données comme suit :

claytonCopula (param = , dim =)

frankCopula (param = , dim =)

gumbelCopula (param = , dim =)

La simulation de la copule de t est relativement plus simple, et il suffisait de se conformer à l'algorithme énoncé dans la section 3.5.. Dans la librairie **MASS** du logiciel

R, on retrouve :

mvrnorm(**n** , **mu**= *vecteur de la moyenne* , **Sigma** = *matrice de covariance*)

La commande **mvrnorm** simule une normale multivariée, alors que **n** est le nombre de vecteurs de v variables aléatoires à simuler, **mu** est le vecteur des v moyennes, **Sigma** est la matrice de covariance $v \times v$. Par ailleurs, la Khi-deux est simulée comme suit :

rchisq(**n** = *taille de l'échantillon simulé*, **dl** = *degrés de liberté*)

La commande **pt** nous retourne des uniformes. Enfin, les commandes **qgamma**, **qlnorm**, et **qweibull** nous permettent d'avoir des échantillons issus des lois marginales désirées.

Toutefois, le package **Copula** de **R** offre une alternative intéressante pour la simulation d'une copule de t . Il suffit d'utiliser la commande **tCopula** pour définir les différents paramètres de la copule, puis la commande **rcopula** pour la simuler.

B.2 Exemples

Exemple 1 : Copules archimédiennes et la simulation des rendements

```
Cop <- c ("claytonCopula","frankCopula","gumbelCopula")
```

```
Copule<- fonction(i,theta,n,L)
```

```
{
```

```
C<-match.fun (paste(Cop[i],sep=""))
```

```
yy<- mvdc(copula = C , (param=theta , dim=n), margins = dist,
paramMargins = list(list(shape=shapegamma1,rate=rategamma1),
list(meanlog=meanlog1,sdlog=sdlog1),
list(shape=shape1,scale=scale1),list(shape=shapegamma2,rate=rategamma2),
list(meanlog=meanlog2,sdlog=sdlog2),
list(shape=shape2,scale=scale2),list(shape=shapegamma3,rate=rategamma3),
list(meanlog=meanlog3,sdlog=sdlog3),
```

```
list(shape=shape3,scale=scale3)))  
  
R <- rmvdc(yy,L)  
  
}
```

Exemple 2 : Copule de t et simulation des rendements

```
t.Cop<- function(rho.mult,n,L)  
{  
  
yy<- tCopula(param = rho.mult, dim  
= n, dispstr = "toep", df = n - 1)  
  
TT<-rcopula(yy,L)  
  
}
```

* TT est une matrice de variables Student multivariées simulées avec une copule t. On se définit deux matrices U et R :

```
U<-matrix(,10000,9)
```

```
R<-matrix(,10000,9)
```

* U est une matrice de variables uniformes [0,1] correspondant aux probabilités des variables de la matrice TT

```
for (i in 1:9)  
{  
  
U[,i]<-pt(TT[,i],8)  
  
}
```

* R est la matrice des rendements, obtenue en appliquant les lois marginales retenues

```
R[,1]<-qgamma(u[,1],shapegamma1,rategamma1)
R[,2]<-qlnorm(u[,2],meanlog1,sdlog1)
R[,3]<-qweibull(u[,3],shape1,scale1)
R[,4]<-qgamma(u[,4],shapegamma2,rategamma2)
R[,5]<-qlnorm(u[,5],meanlog2,sdlog2)
R[,6]<-qweibull(u[,6],shape2,scale2)
R[,7]<-qgamma(u[,7],shapegamma3,rategamma3)
R[,8]<-qlnorm(u[,8],meanlog3,sdlog3)
R[,9]<-qweibull(u[,9],shape3,scale3)
```

Annexe C

Simulation

C.1 Estimation des paramètres

```
Paramgamma <-function(moy,sigma)
{
scalegamma<-sigma^2/(1 + moy)
shapegamma<-(1 + moy)/rategamma
rategamma<-1/scalegamma
}
```

```
Paramlnorm <-function(moy,sigma)
{
sdlog<-sqrt ( log (((1 + moy)^2+sigma^2)/(1 + moy)^2))
meanlog<-log (1 + moy)-((sdlog^2)/2)
}
```

```
Paramweibull <-function(moy,sigma)
{
f1<-function(shape)
{ log (gamma (1+ (2/shape)))- 2*log (gamma ((1+
(1/shape)))) + 2*log (1 + moy) - log ((1 + moy)^2 + sigma^2) }
```

```
shape<-uniroot (f=f1,lower=1,upper=1000)$root
```

```
f2<-function (C)
```

```
{((1/C)^(1/shape))*gamma (1+(1/shape)) - (1 + moy)}
```

```
C<-uniroot (f=f2,lower=0,upper=1000)$root
```

```
scale<-C^(-1/shape)
```

```
}
```

* On obtient les estimés des différents paramètres en fixant la valeur du couple moyenne-variance (moy,sigma).

```
dist<-c("gamma","lnorm","weibull","gamma","lnorm","weibull",
"gamma","lnorm","weibull")
```

```
par<-list(c("shapegamma","rategamma"),c("meanlog","sdlog"),
c("shape","scale"),c("shapegamma","rategamma"),
c("meanlog","sdlog"),c("shape","scale"),
c("shapegamma","rategamma"),c("meanlog","sdlog"),
c("shape","scale"))
```

```
p<-list(c(shapegamma1,rategamma1),c(meanlog1,sdlog1),
c(shape1,scale1),c(shapegamma2,rategamma2),
c(meanlog2,sdlog2),c(shape2,scale2),
c(shapegamma3,rategamma3),c(meanlog3,sdlog3),
c(shape3,scale3))
```


C.2 Simulation des rendements

C.2.1 Cas d'indépendance

```
loi<-list (dist,par,p)

M<-length (loi[[1]])

R<-matrix (,10000,M)

simul.rdt<-function (loi)
{

i<-1:M for (i in 1:M)

    {
    f<-match.fun (paste ("r",loi$dist[i],sep=""))

    R[,i]<-f (10000,loi$p[[i]])
    }

return (R)

}

R<- simul.rdt (loi=list (dist,par,p))
```

C.2.2 Cas de dépendance

```
Cop <- c ("claytonCopula","frankCopula","gumbelCopula")

Copule<- function(i,theta,n,L)

{
```

```
C<-match.fun (paste(Cop[i],sep=""))

yy<- mvdc(copula = C , (param=theta , dim=n), margins = dist,
paramMargins = list(list(shape=shapegamma1,rate=rategamma1),
list(meanlog=meanlog1,sdlog=sdlog1),
list(shape=shape1,scale=scale1),list(shape=shapegamma2,rate=rategamma2),
list(meanlog=meanlog2,sdlog=sdlog2),
list(shape=shape2,scale=scale2),list(shape=shapegamma3,rate=rategamma3),
list(meanlog=meanlog3,sdlog=sdlog3),
list(shape=shape3,scale=scale3)))

R <- rmvdc(yy,L)

}

t.Cop<- function(rho.mult,n,L)
{

yy<- tCopula(param = rho.mult, dim
= n, dispstr = "toep", df = n - 1)

TT<-rcopula(yy,L)

}

U<-matrix(,10000,9)

R<-matrix(,10000,9)

for (i in 1:9)
{

U[,i]<-pt(TT[,i],8)

}
```

```

R[,1]<-qgamma(u[,1],shapegamma1,rategamma1)
R[,2]<-qlnorm(u[,2],meanlog1,sdlog1)
R[,3]<-qweibull(u[,3],shape1,scale1)
R[,4]<-qgamma(u[,4],shapegamma2,rategamma2)
R[,5]<-qlnorm(u[,5],meanlog2,sdlog2)
R[,6]<-qweibull(u[,6],shape2,scale2)
R[,7]<-qgamma(u[,7],shapegamma3,rategamma3)
R[,8]<-qlnorm(u[,8],meanlog3,sdlog3)
R[,9]<-qweibull(u[,9],shape3,scale3)

```

C.3 Simulation des poids et des portefeuilles

```

xx<-matrix (,M,1)

simul.inv<-function(xx)
{

xx[1]<-rbeta (1,1,(M-1))

i<-1:M for (i in 2:(M-1))

{
xx[i]<-(rbeta (1,1,(M-i)))*(1-sum (xx[1:(i-1)]))
}

xx[M]<-1-sum (xx[1:(M-1)])

return (xx)

```

```
}
```

```
X<-replicate (9 ,simul.inv(xx))
```

```
S <-R %*% X
```

```
Portefeuilles<-S - 1
```

```
invPortefeuilles <- - Portefeuilles
```

```
lS<-log(S)
```

Annexe D

Programmes R pour l'étude descriptive

D.1 Histogramme et statistiques descriptives

```
h<-matrix(,9,1)
```

```
d<-list()
```

```
for (i in 1:9)
```

```
{
```

```
  d[[i]]<-summary (Portefeuilles[,i])
```

```
  hist(Portefeuilles[,i],
```

```
  xlim=c(min(Portefeuilles[,i])-0.1 , max(Portefeuilles[,i])+0.1),
```

```
  prob=T, breaks=100)
```

```
}
```

D.2 Espérance de survie résiduelle

```
x<-seq(length=200,-1,2)

Y<-list()

y<-matrix(,200,1)

x<-matrix(x,200,1)

L<-length(Portefeuilles[1,])

fct <- function (Portefeuilles,x,y)
{

i<-1:L

for (i in 1:L)
{

j<-1:200 for
(j in 1:200)
{

y[j,]<-mean((Portefeuilles[,i]-x[j,])[Portefeuilles[,i]>x[j,]])

}

Y[[i]]<-y

Y[[i]][is.na (Y[[i]])]<-0

}
```

```
return (Y)

}

Y<-fct (Portefeuilles,x,y)

i<-1:L

for (i in 1:L)
{

a<-min (Portefeuilles[,i])

b<-max(Portefeuilles[,i])

plot (x,Y[[i]],xlim=c(a,b),ylab="Esperance
de vie residuelle",xlab="x",type="l")

}
```

D.3 Ramlau-Hansen

```
c<-rep(1,10000)

s<-seq(length=200,0,2)

h<-list()

Ramlau.hansen<- function (S,s,b)
{
```

```
i<-1:L

for (i in 1:L)
{

k<-1:200

for (k in 1:200)
{

y[k]<- (1/10000)*sum ((3/(4*b))*(1-(pmin (rep
(1,1e4),(((sort (S[,i])-s[k])/b)^2))))*1/(10000:1))

}

h[[i]]<-y

plot((s-1),h[[i]])

}

}

Ramlau.hansen (S,s,b=0.2)
```


Annexe E

Histogrammes

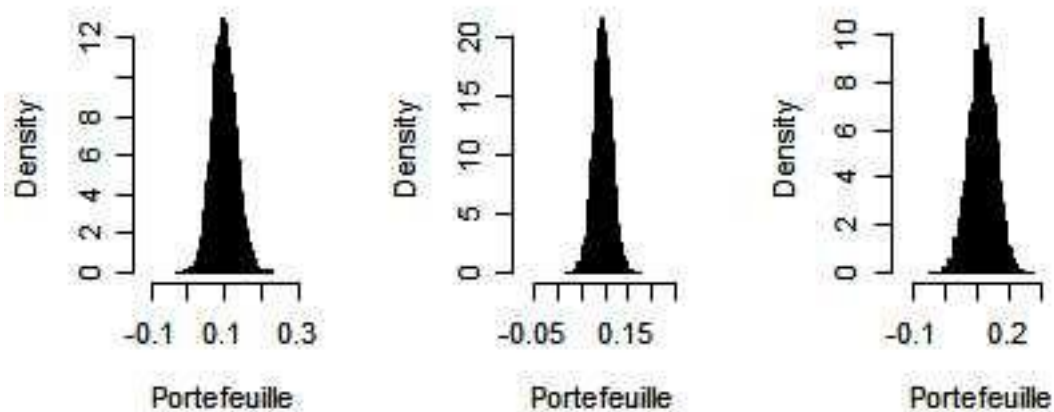
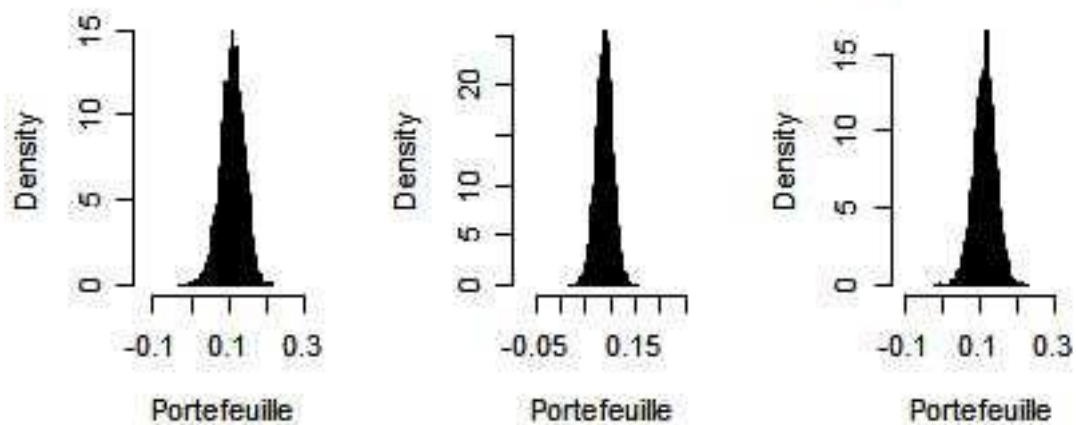
Histogram of Portefeu Histogram of Portefeu Histogram of Portefeu**Histogram of Portefeu Histogram of Portefeu Histogram of Portefeu**

FIG. E.1 – *Histogrammes des portefeuilles à neuf titres simulés sous l'hypothèse de l'indépendance*

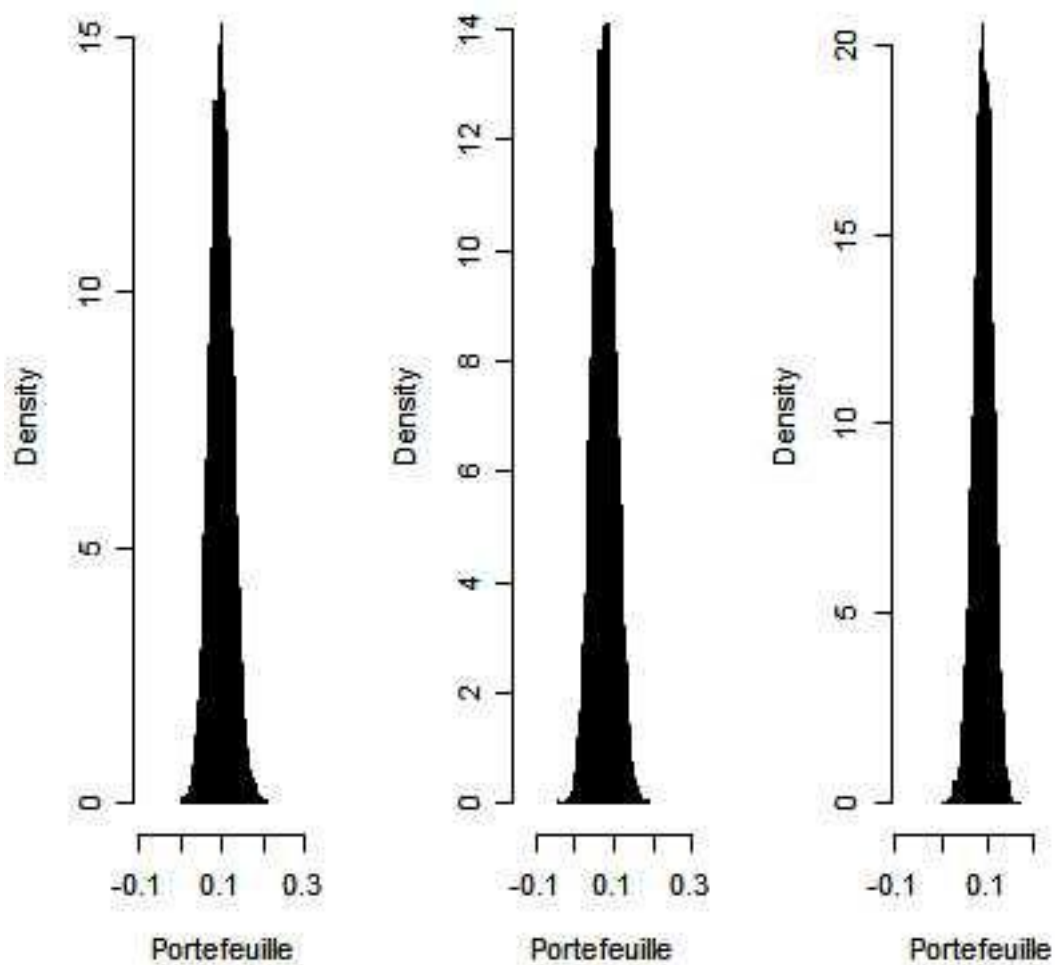
Histogram of Portefeu Histogram of Portefeu Histogram of Portefeu

FIG. E.2 – Histogrammes des portefeuilles à neuf titres simulés sous l'hypothèse de l'indépendance (Suite)

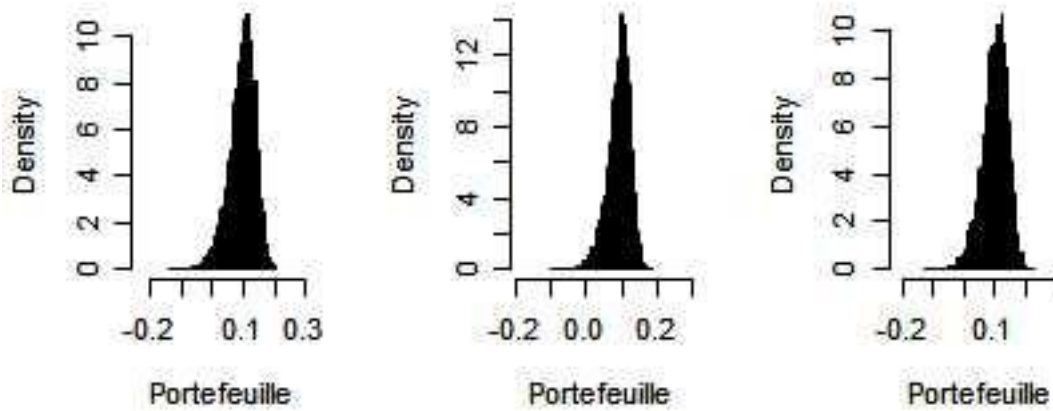
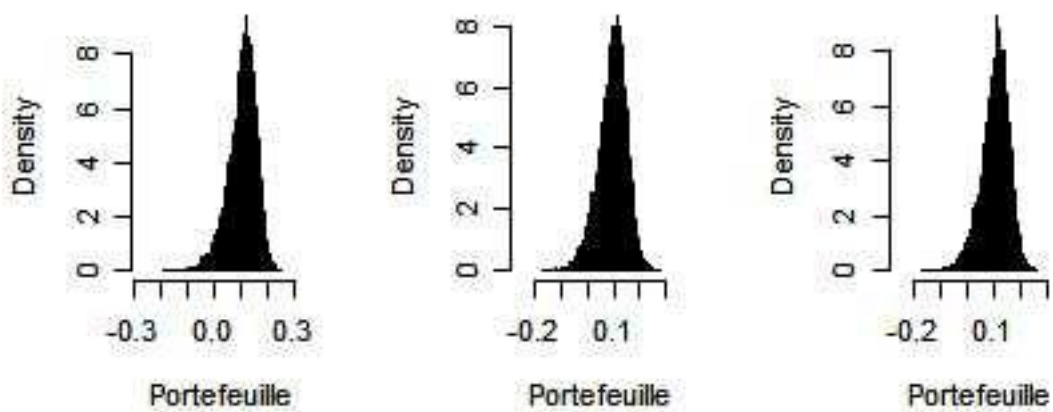
Histogram of Portefeu Histogram of Portefeu Histogram of Portefeu**Histogram of Portefeu Histogram of Portefeu Histogram of Portefeu**

FIG. E.3 – Histogrammes des portefeuilles à neuf titres issus d'une Clayton à un faible niveau de dépendance

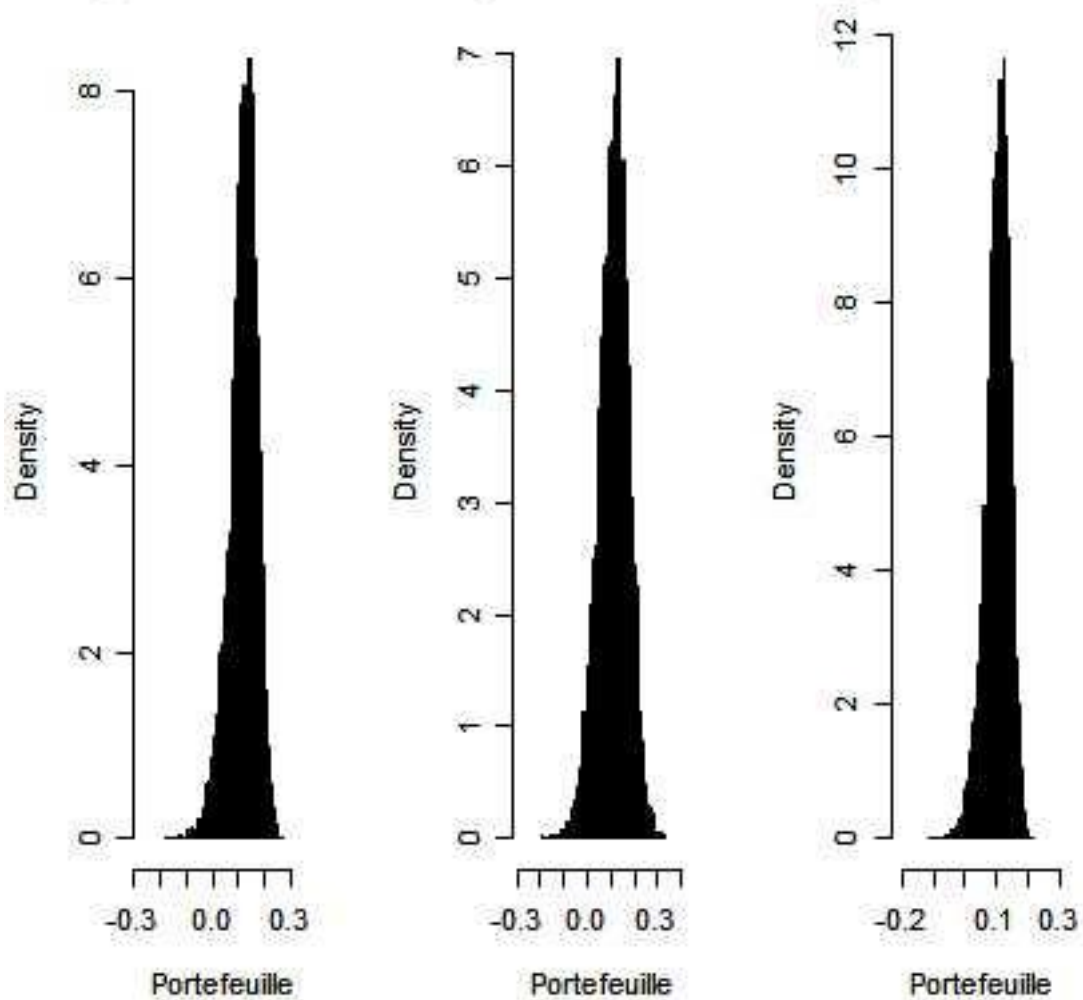
Histogram of Portefeu Histogram of Portefeu Histogram of Portefeu

FIG. E.4 – Histogrammes des portefeuilles à neuf titres issus d'une Clayton à un faible niveau de dépendance (Suite)

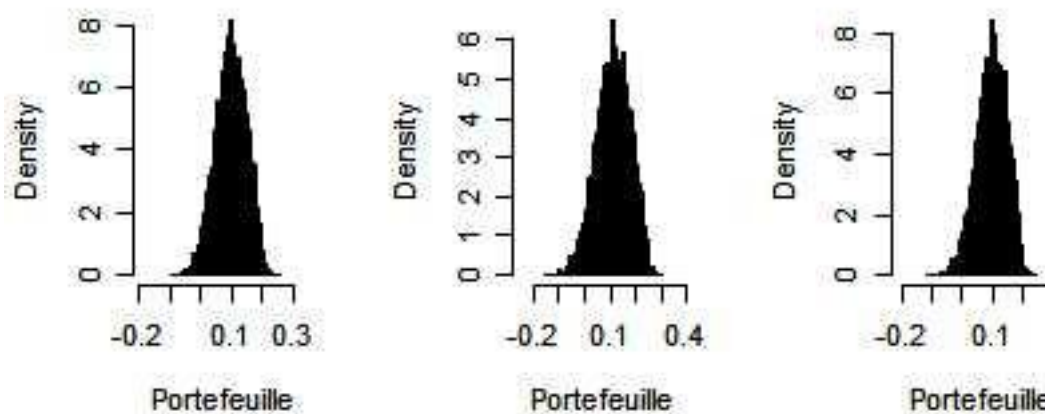
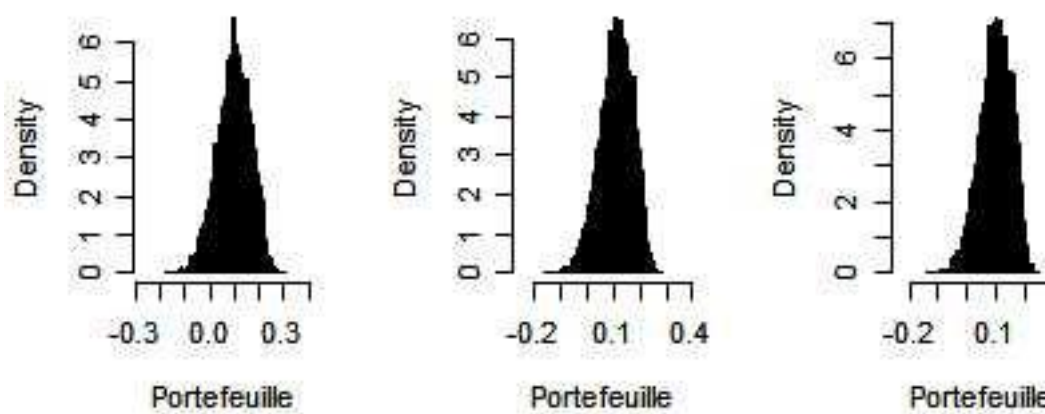
Histogram of Portefeu Histogram of Portefeu Histogram of Portefeu**Histogram of Portefeu Histogram of Portefeu Histogram of Portefeu**

FIG. E.5 – *Histogrammes des portefeuilles à neuf titres issus d'une Clayton à un niveau élevé de dépendance*

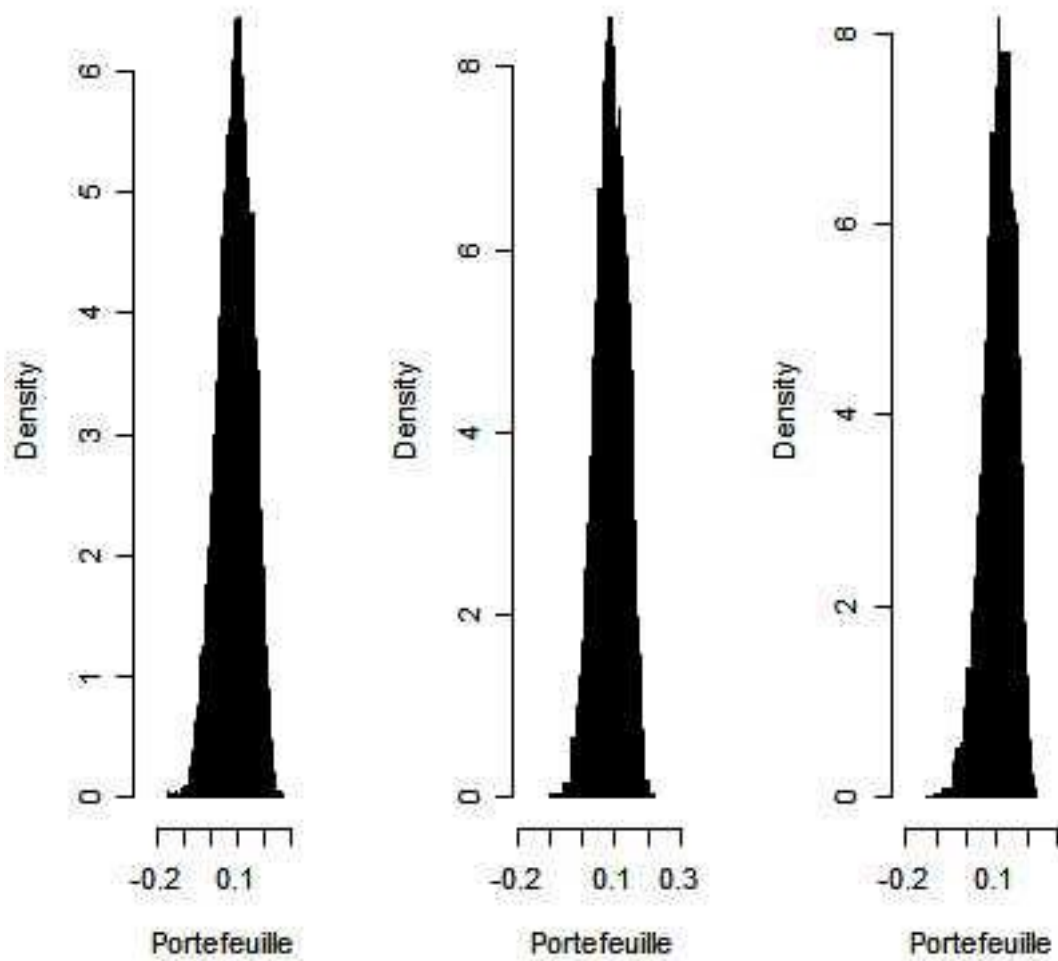
Histogram of Portefeu Histogram of Portefeu Histogram of Portefeu

FIG. E.6 – Histogrammes des portefeuilles à neuf titres issus d'une Clayton à un niveau élevé de dépendance (Suite)

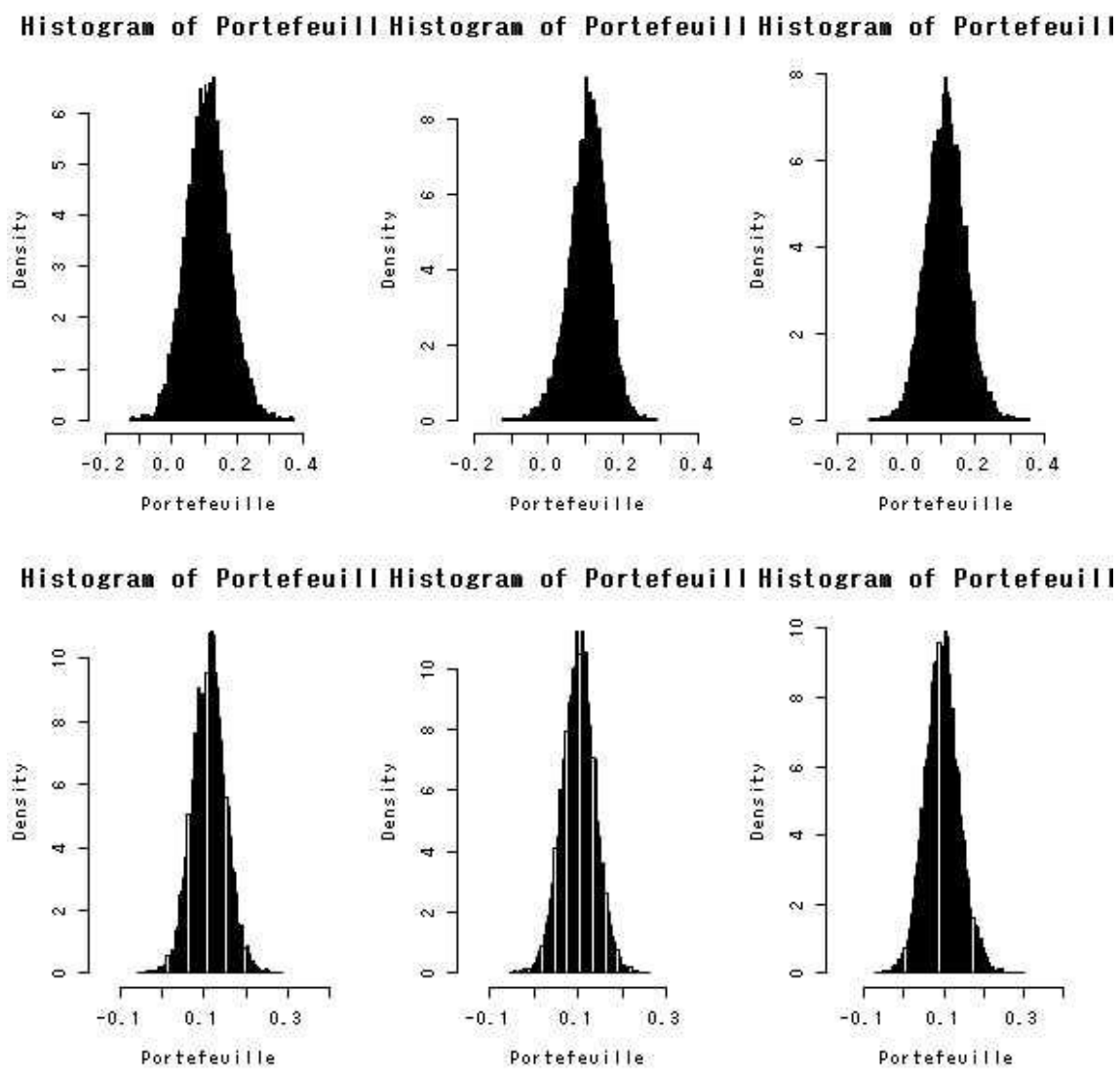


FIG. E.7 – *Histogrammes des portefeuilles à neuf titres issus d'une copule de t à un faible niveau de dépendance*

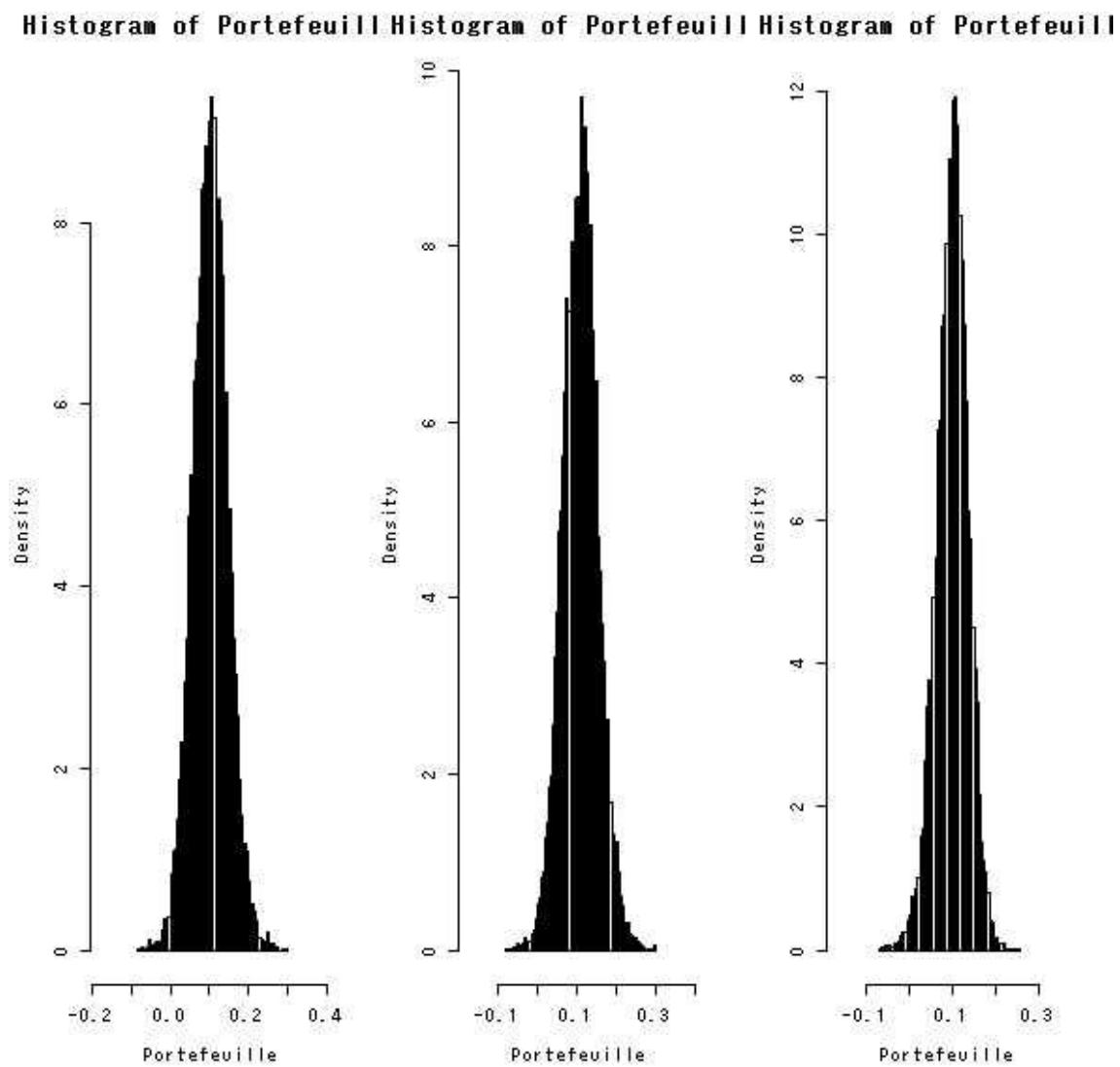


FIG. E.8 – *Histogrammes des portefeuilles à neuf titres issus d’une copule de t à un faible niveau de dépendance (Suite)*

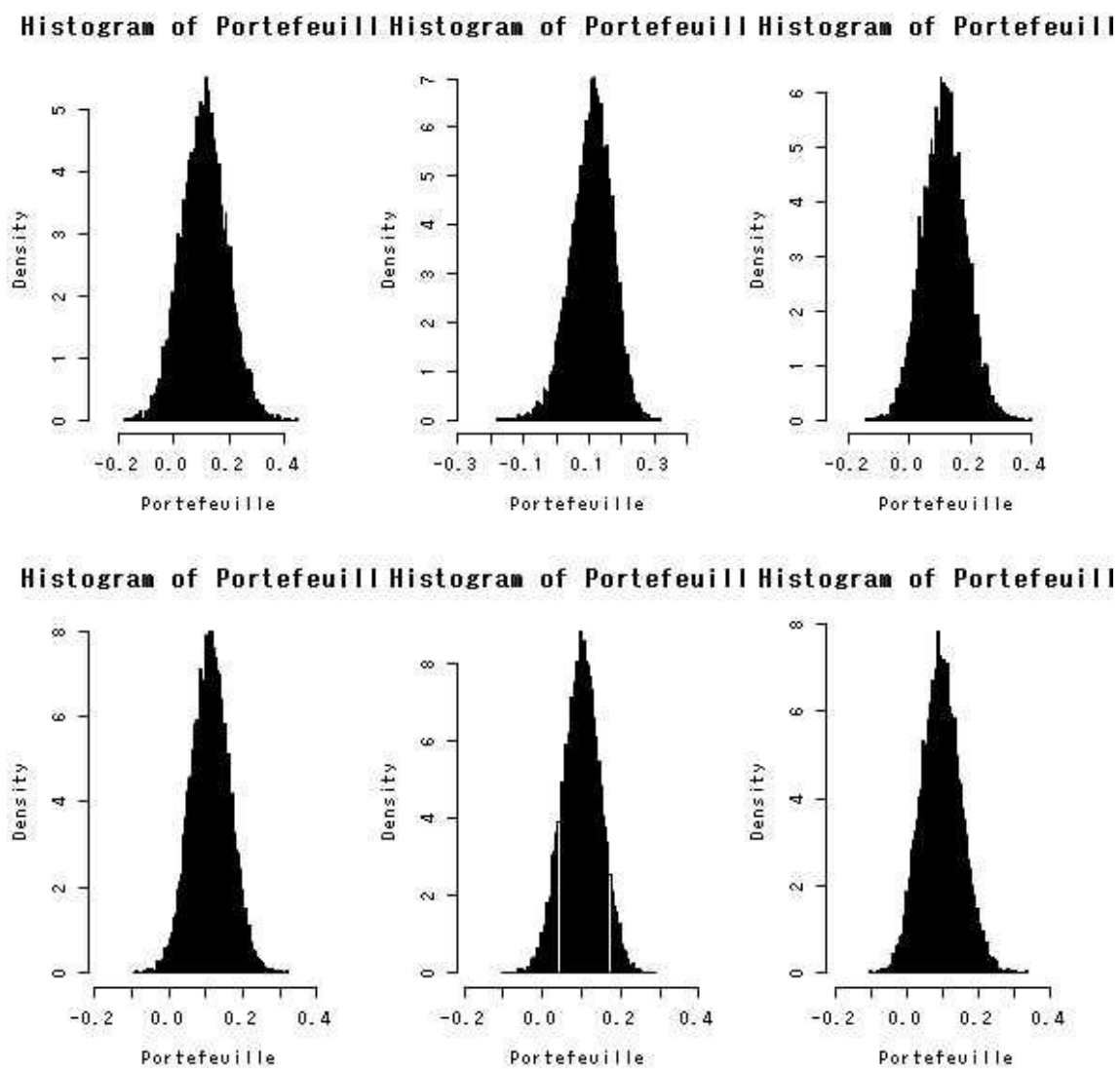


FIG. E.9 – Histogrammes des portefeuilles à neuf titres issus d'une Clayton à un niveau élevé de dépendance

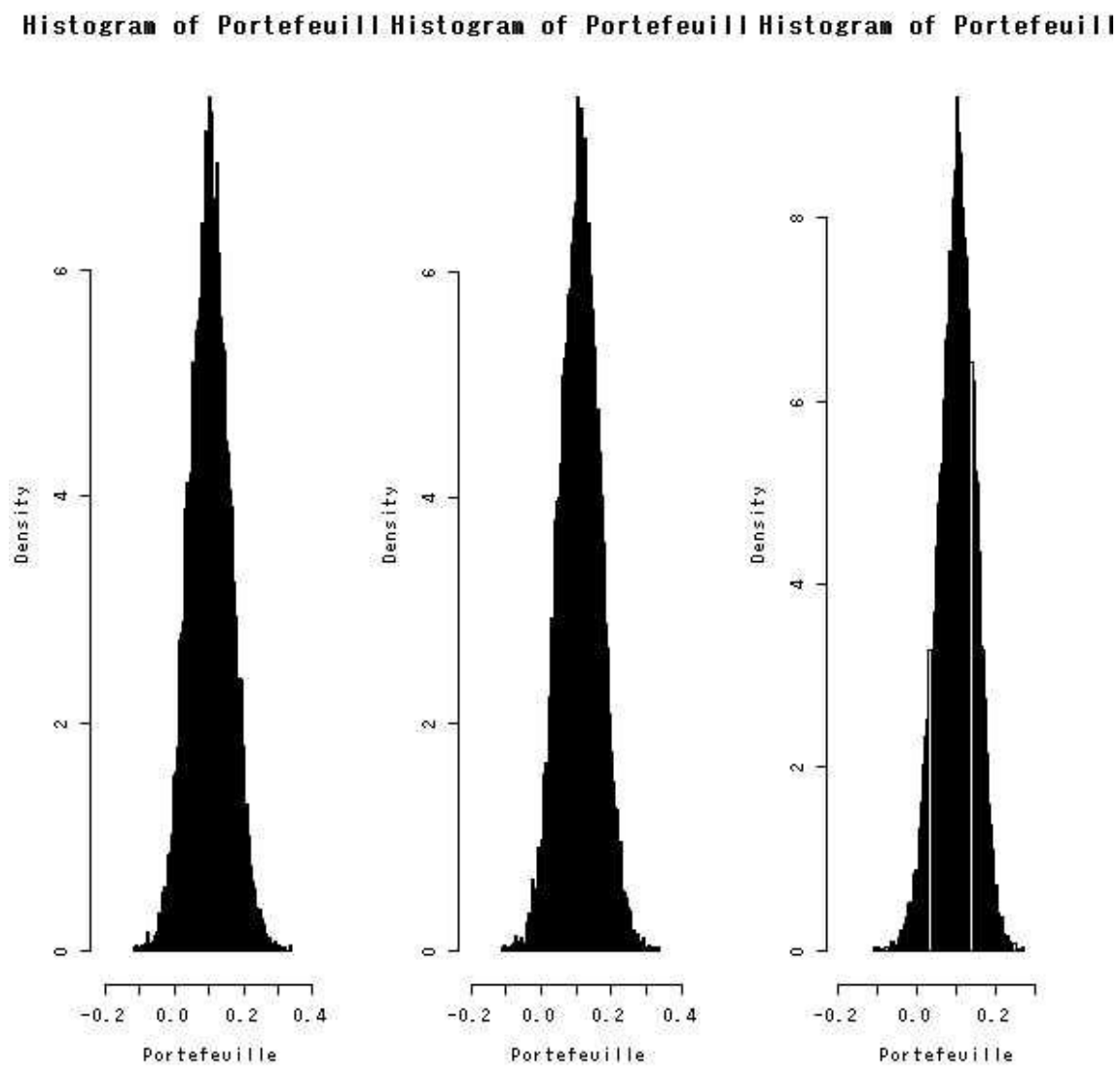


FIG. E.10 – *Histogrammes des portefeuilles à neuf titres issus d'une copule de t à un niveau élevé de dépendance (Suite)*

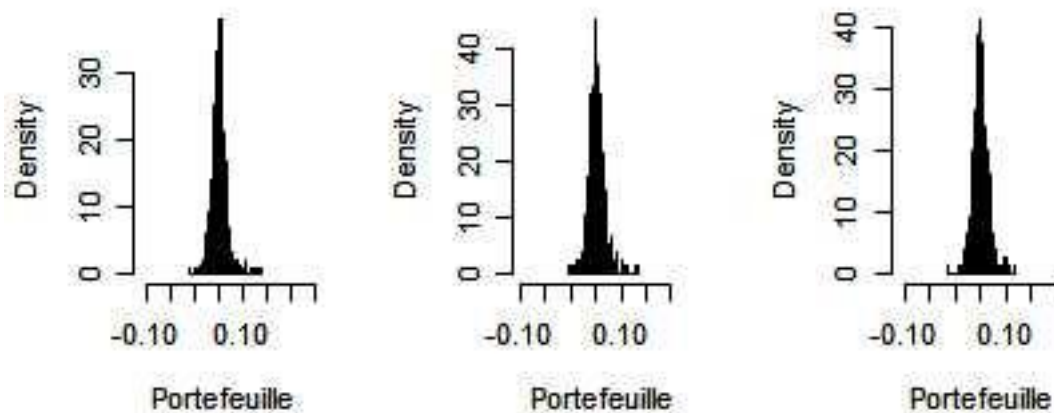
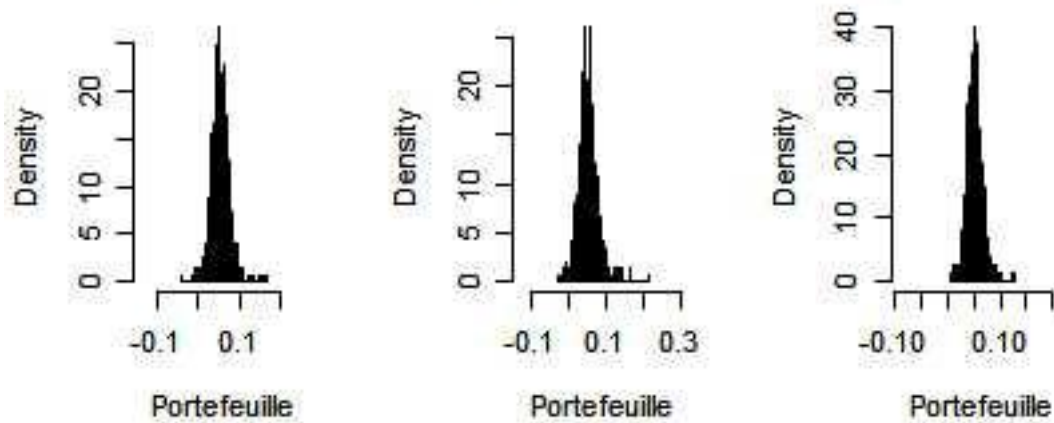
Histogram of Portefeu Histogram of Portefeu Histogram of Portefeu**Histogram of Portefeu Histogram of Portefeu Histogram of Portefeu**

FIG. E.11 – *Histogrammes des portefeuilles à neuf titres simulés à partir des données réelles*

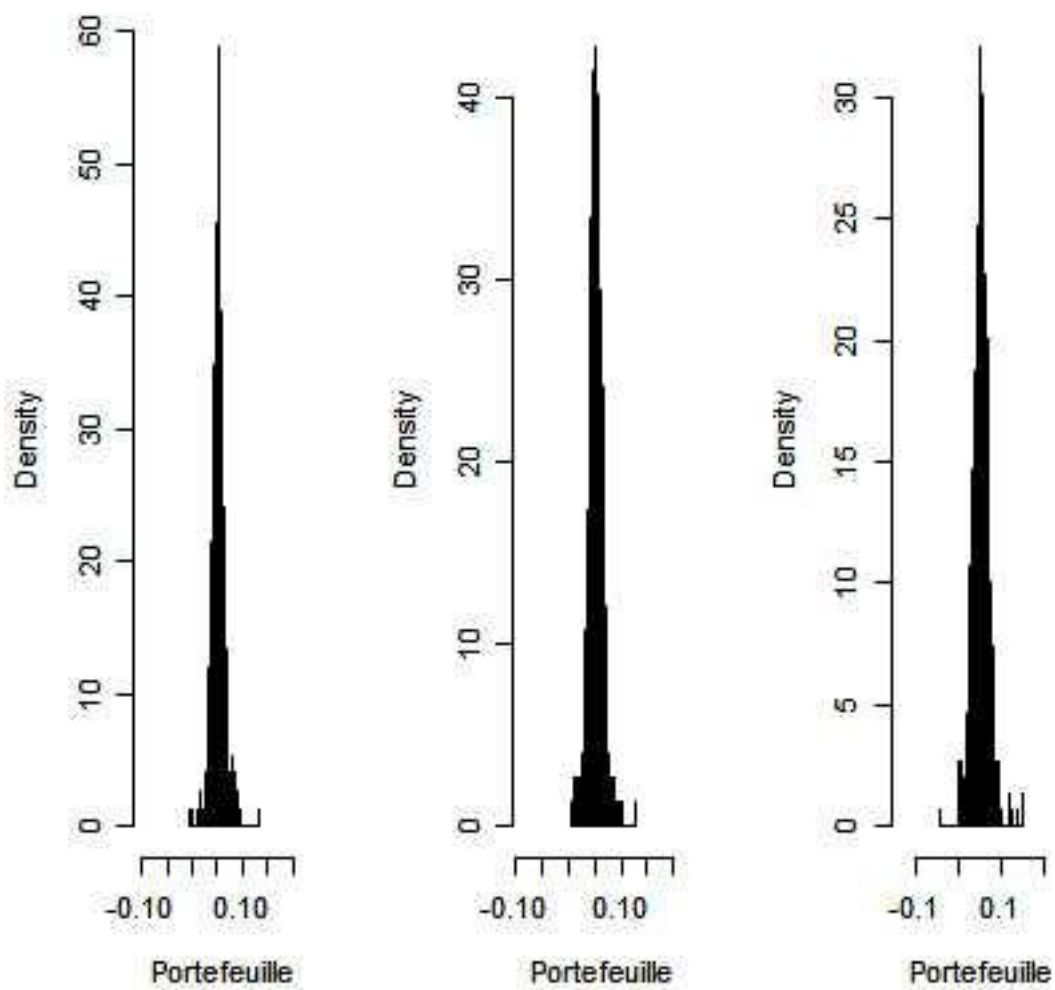
Histogram of Portefeu Histogram of Portefeu Histogram of Portefeu

FIG. E.12 – *Histogrammes des portefeuilles à neuf titres simulés à partir des données réelles (Suite)*

Annexe F

Définition de nouvelles lois sur \mathbb{R}

F.1 La gamma transformée

```
ptgamma<- fonction(x,alpha,theta,tau)
{

u<-(x/theta)^tau

pgamma(u,shape=alpha)

}
```

```
qtgamma<- fonction(alpha,theta,tau)
{

theta*(qgamma((1:1)/1,shape=alpha)^(1/tau))

}
```

F.2 La GBII

```

ptbeta<- fonction(x,alpha,lambda,gam1,tau)
{

u<-(x^gam)/((lambda)+(x^gam))

pbeta(u,shape1=tau,shape2=alpha)

}

qtbeta<- fonction(alpha,lambda,gam,tau)
{

(lambda^(1/gam))*((qbeta((1:1)/1,shape1=tau,
shape2=alpha)/(1-qbeta((1:1)/1,shape1=tau,
shape2=alpha)))^(1/gam))

}

dtbeta <- fonction(x,alpha,lambda,gam,tau)
{

gamma(alpha + tau)*gam*(x/(lambda^(1/gam))^(gam*tau)/
x/gamma(alpha)/gamma(tau)/(lambda + x^gam)^(alpha + tau)

}

```

Annexe G

Programme R pour l'adéquation

G.1 PP-plot

Exemple 1 : La Weibull

```
plot((1:1)/1,pweibull(sort(y),shape=p[1],scale=p[2]))  
  
abline(0,1)
```

Exemple 2 : La gamma transformée

```
plot((1:1)/1,ptgamma(x=sort(y),alpha=p[1],theta=p[2],tau=p[3]))  
  
abline(0,1)
```

Exemple 3 : La GBII

```
plot((1:1)/1,ptbeta(x=sort(y),alpha=p[1],lambda=p[2],gam=p[3],tau=p[4]))  
  
abline(0,1)
```


G.2 QQ-plot

Exemple 1 : La Weibull

```
plot(sort(y),qweibull((1:1)/1,shape=p[1],scale=p[2]))  
  
abline(0,1)
```

Exemple 2 : La gamma transformée

```
plot(sort(y),qtgamma(alpha=p[1],theta=p[2],tau=p[3]))  
  
abline(0,1)
```

Exemple 3 : La GBII

```
plot(sort(y),qtbeta(alpha=p[1],lambda=p[2],gam=p[3],tau=p[4]))  
  
abline(0,1)
```

G.3 Superposition

```
# Courbe bleue: estimateur des moments  
  
# Courbe rouge: maximum de vraisemblance avec moments au départ  
  
# Courbe verte: maximum de vraisemblance avec départ "quelconque"  
  
hist(y,prob="T",breaks=50)  
  
curve(dtbeta(x,p1[1],p1[2],p1[3],p1[4]),add=T,col="blue")
```

```

curve(dtbeta (x,p2[1],p2[2],p2[3],p2[4]),add=T,col="red")

curve(dtbeta (x,p3[1],p3[2],p3[3],p3[4]),add=T,col="green")

* Autre méthode

hist (y,xlim = c(min (y),max (y)),
ylim = c(0,1),freq = F)

lines(portefeuille1,ptbeta(y,alpha=p[1],lambda=p[2],tau=p[3],gam=p[4]))

y.dns<-density (y, n = 200, width = "SJ")

lines(y.dns$x,y.dns$y,lty = 2)

```

G.4 Moments théoriques vs. empiriques

```

TK<- function (k)
{
mean(y^k)
}

```

* On peut plutôt utiliser le package moments de R

```
moment(yy,1); moment(yy,2); moment(yy,3); moment(yy,4)
```

```

SK<-function (k,alpha,lambda,gam,tau)
{
(lambda^(k/gam))*gamma*(tau + (k/gam))*
gamma(alpha - (k/gam))/(gamma(alpha)*gamma(tau))
}

```

* Asymétrie empirique

```
Asymétrie.empirique<-(mean(y^3) - 3*mean(y^2)*mean(y) +
3*mean(y)*(mean(y)^2) - (mean(y)^3))/(var(y)^(3/2))
```

* Aplatissement empirique

```
Applatissement.empirique<-(mean(y^4) - 4*mean(y^3)*mean(y) +
6*mean(y^2)*(mean(y)^2) - 4*mean(y)*(mean(y)^3) +
(mean(y)^4))/(var(y)^2)
```

* Asymétrie théorique

```
Asymétrie.théorique<-(SK(3,p[1],p[2],p[3],p[4]) -
3*SK(2,p[1],p[2],p[3],p[4])*SK(1,p[1],p[2],p[3],p[4]) +
3*SK(1,p[1],p[2],p[3],p[4])*((SK(1,p[1],p[2],p[3],p[4]))^2) -
((SK(1,p[1],p[2],p[3],p[4]))^3))/(((SK(2,p[1],p[2],p[3],p[4]) -
(SK(1,p[1],p[2],p[3],p[4]))^2))^(3/2))
```

* Aplatissement théorique

```
Applatissement.théorique<-(SK(4,p[1],p[2],p[3],p[4]) -
4*SK(3,p[1],p[2],p[3],p[4])*SK(1,p[1],p[2],p[3],p[4]) +
6*SK(2,p[1],p[2],p[3],p[4])*((SK(1,p[1],p[2],p[3],p[4]))^2) -
4*SK(1,p[1],p[2],p[3],p[4])*((SK(1,p[1],p[2],p[3],p[4]))^3) +
((SK(1,p[1],p[2],p[3],p[4]))^4))/(((SK(2,p[1],p[2],p[3],p[4]) -
(SK(1,p[1],p[2],p[3],p[4]))^2))^(3/2))
```

Annexe H

Adéquation de la GBII

```
ptbeta<- function(x,p)
{
u<-(x^p[3])/((p[2])+(x^p[3]))

pbeta(u,shape1=p[4],shape2=p[1])
}
```

```
qtbeta<- function(p)
{
(p[2]^(1/p[3]))*((qbeta((1:1)/1,shape1=p[4],
shape2=p[1])/(1-qbeta((1:1)/1,shape1=p[4],
shape2=p[1])))^(1/p[3]))
}
```

```
dtbeta <- function(x,p)
{
((gamma(p[1] + p[4])*p[3]*(x/(p[2]^(1/p[3])))^(p[3]*p[4])/x/gamma(p[1])/
gamma(p[4]))/(p[2] + x^p[3])^(p[1] + p[4]))
}
```

```
#Minimisation de l'écart entre les moments théoriques
et empiriques de la GBII
```

```
estim.moments.GBII<-function(p)
{
diff<-(p[2]^(1:4/p[3])*gamma(p[4] + 1:4/p[3])*gamma(p[1] - 1:4/p[3])/
(gamma(p[1])*gamma(p[4]))) - c(moment(y,1),moment(y,2),moment(y,3),moment(y,4))

diff
}
```

```
#Maximum de vraisemblance de la GBII
```

```
logvrais.GBII<- fonction(p)
{
e<-(log((gamma(p[1] + p[4])*p[3]*(y/(p[2]^(1/p[3]))))^(p[3]*p[4])/y/
gamma(p[1])/gamma(p[4])/(p[2] + y^p[3])^(p[1] + p[4]))))

-sum(e)
}
```

```
#####
```

```
# Charger les packages RMinpack, moments, stats.
```

```
#####
```

```
y<-S[,i]
```

```
p0<-c(5,1,1,1)
D<-numeric(0)
C<-list()
MM<-list()
OP1<-list()
```

```
OP2<-list()
s<-seq(-2,2,0.1)

j<-1:length(s)

for (j in 1:length(s))
{

YY<-y^s[j]

pp<-numeric(0)

MM<-FSolve(estim.moments.GBII,p0,tol=1e-20)

OP1<-optim(MM$root,logvrais.GBII,control=list(maxit=1e6))

OP2<-optim(p0,logvrais.GBII,control=list(maxit=1e6))

if(OP1$convergence == 0)
{
C[[j]]<-list(MM$value,MM$root,OP1$value,OP1$par)

pp<-OP1$par
}
else
{
C[[j]]<-list(MM$value,MM$root,OP2$value,OP2$par)

pp<-OP2$par
}

ecdf2<-function(t,k)
{
  len <- length(k)
  sum(k<=t)/len
}
```

```
D[j]<- abs(max( sapply(yy,ecdf2,k=yy) - dtbeta(sort(yy),pp)))
```

```
}
```

```
s[D==min(D)]
```

```
C[[which(D == min(D))]]
```

Annexe I

GBII QQ-plot

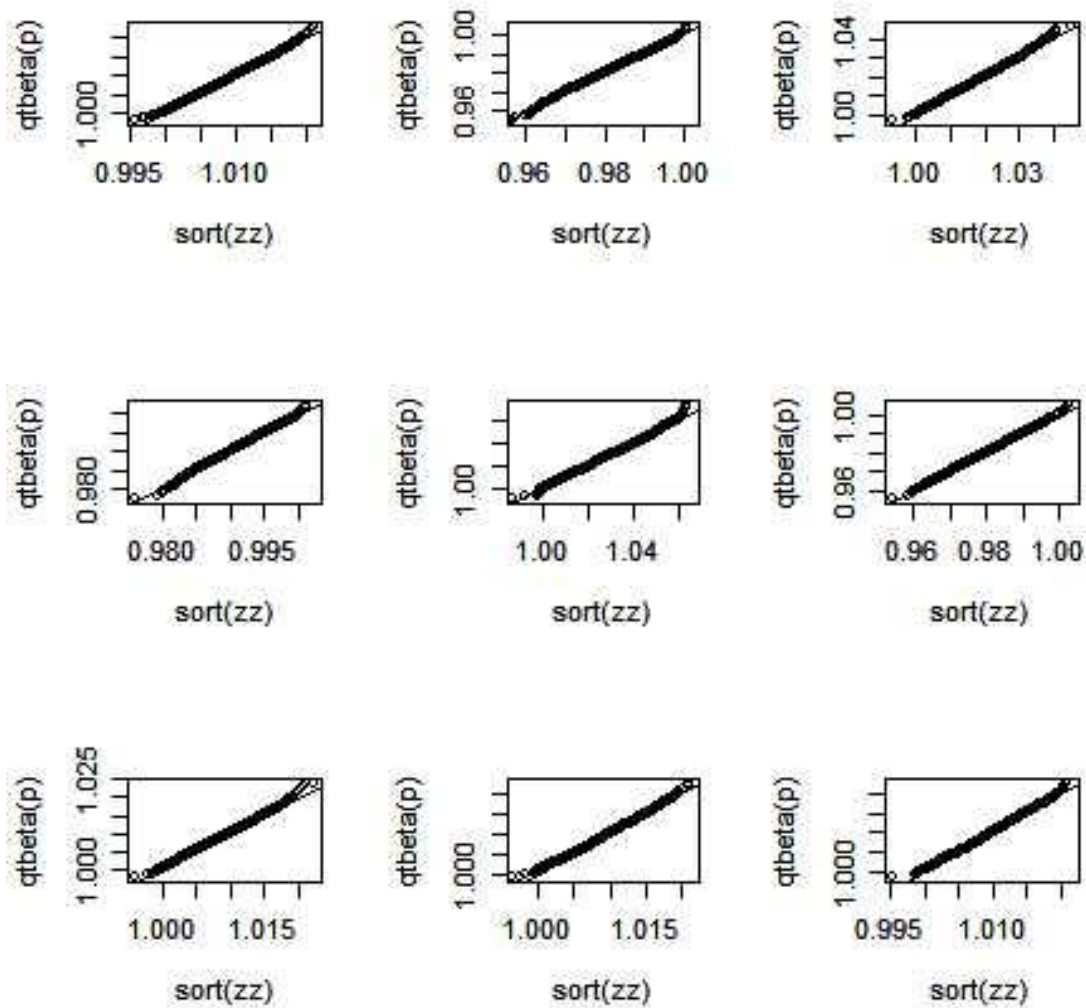


FIG. I.1 – QQ-plots des neuf portefeuilles simulés sous l'hypothèse de l'indépendance

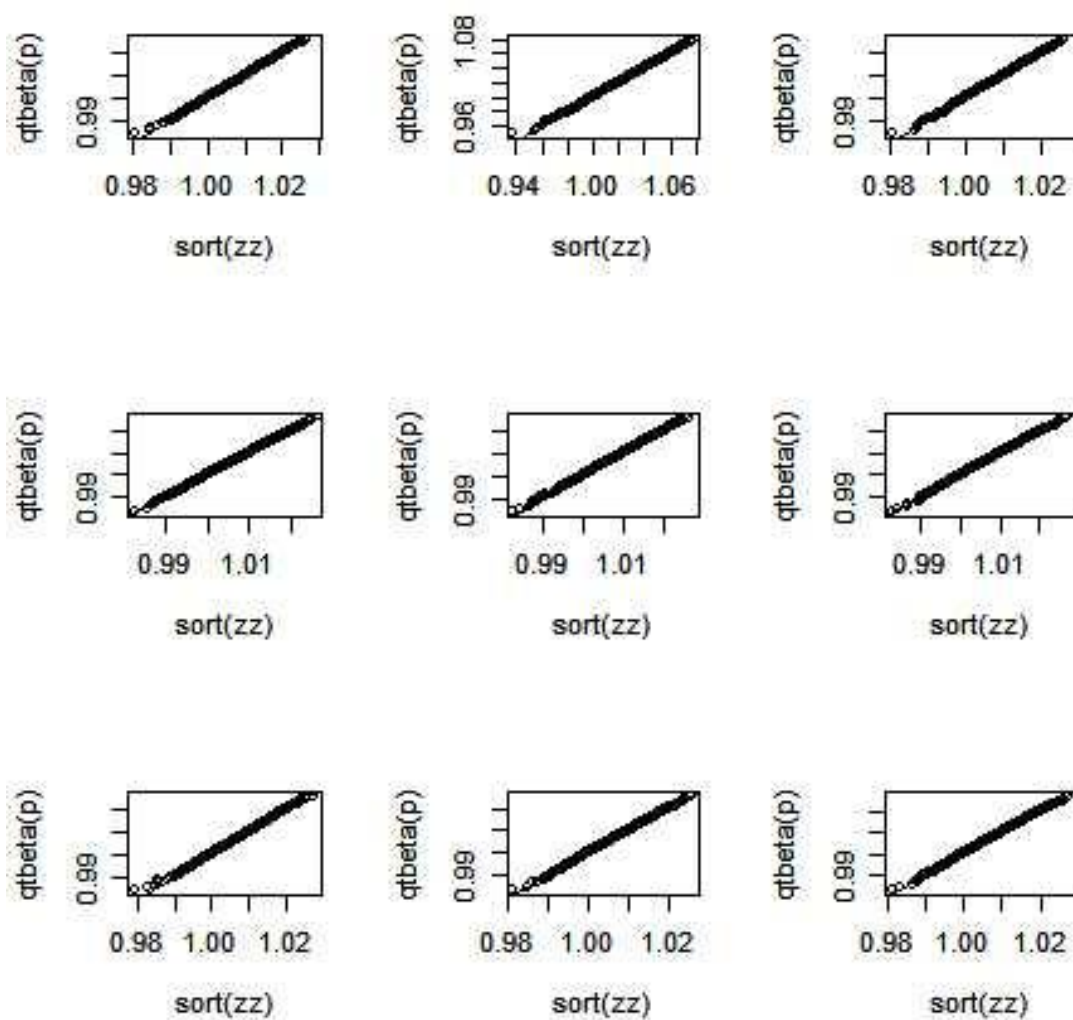


FIG. I.2 – QQ-plots des neuf portefeuilles issus d'une Clayton à un faible niveau de dépendance

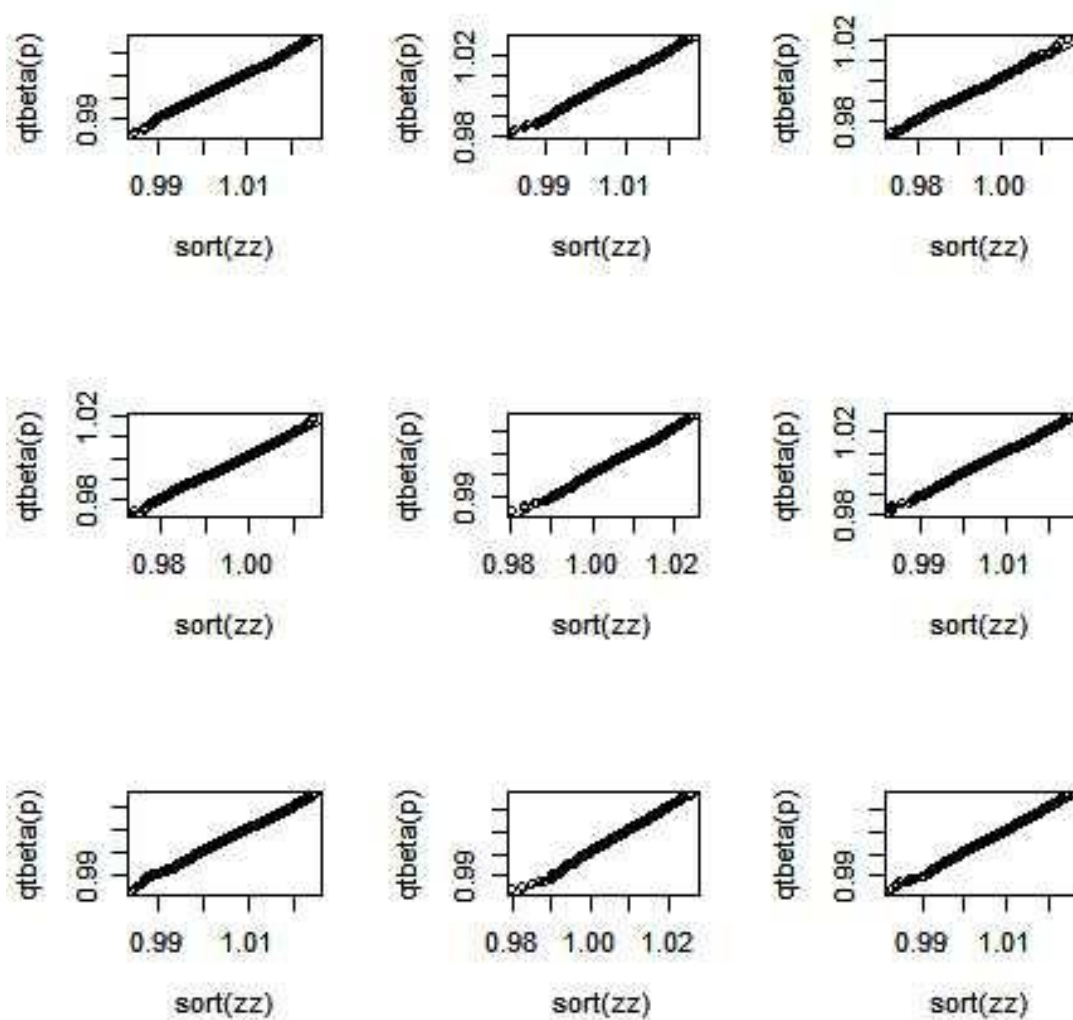


FIG. I.3 – QQ-plots des neuf portefeuilles issus d'une Clayton à un niveau élevé de dépendance

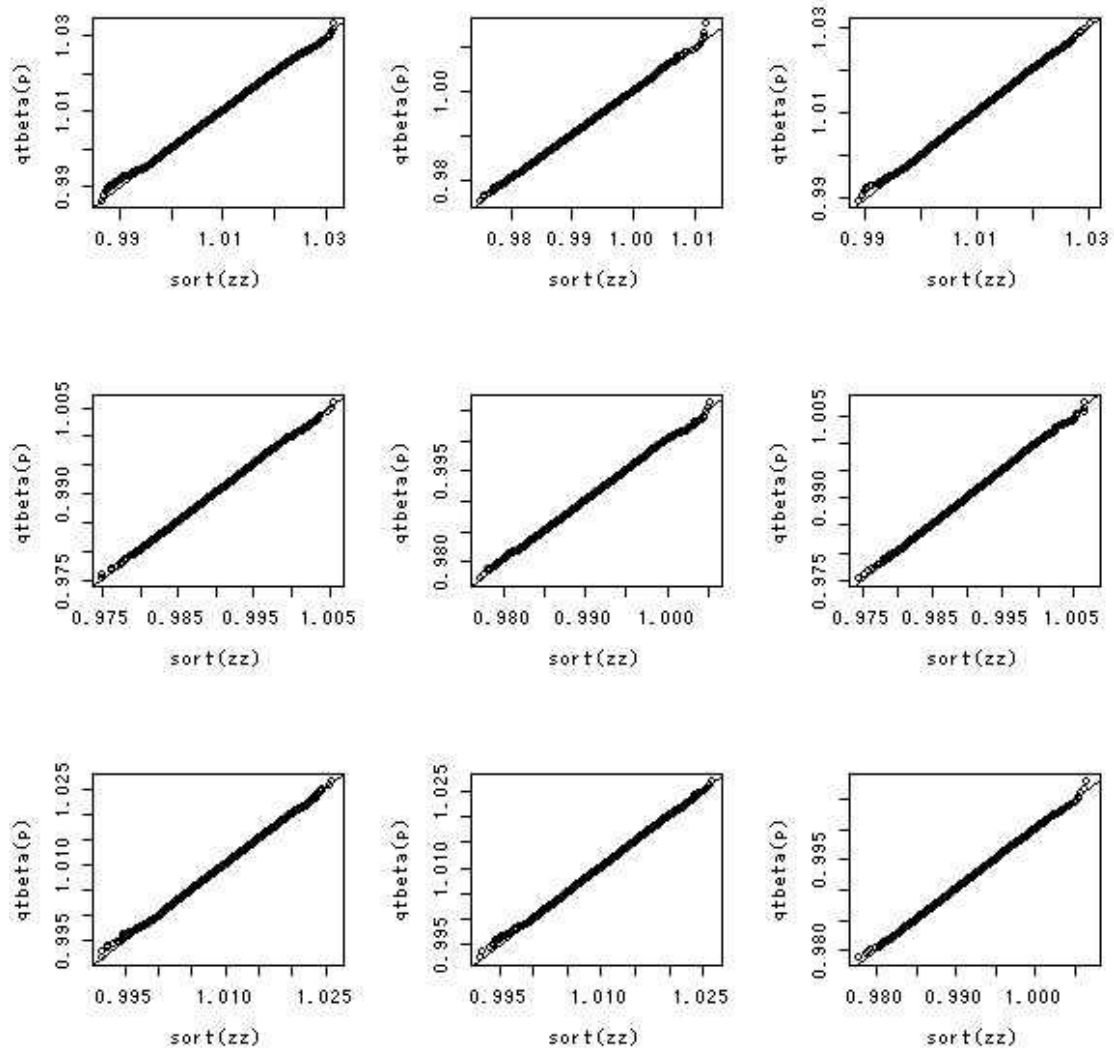


FIG. I.4 – QQ-plots des neuf portefeuilles issus d'une copule de t à un faible niveau de dépendance

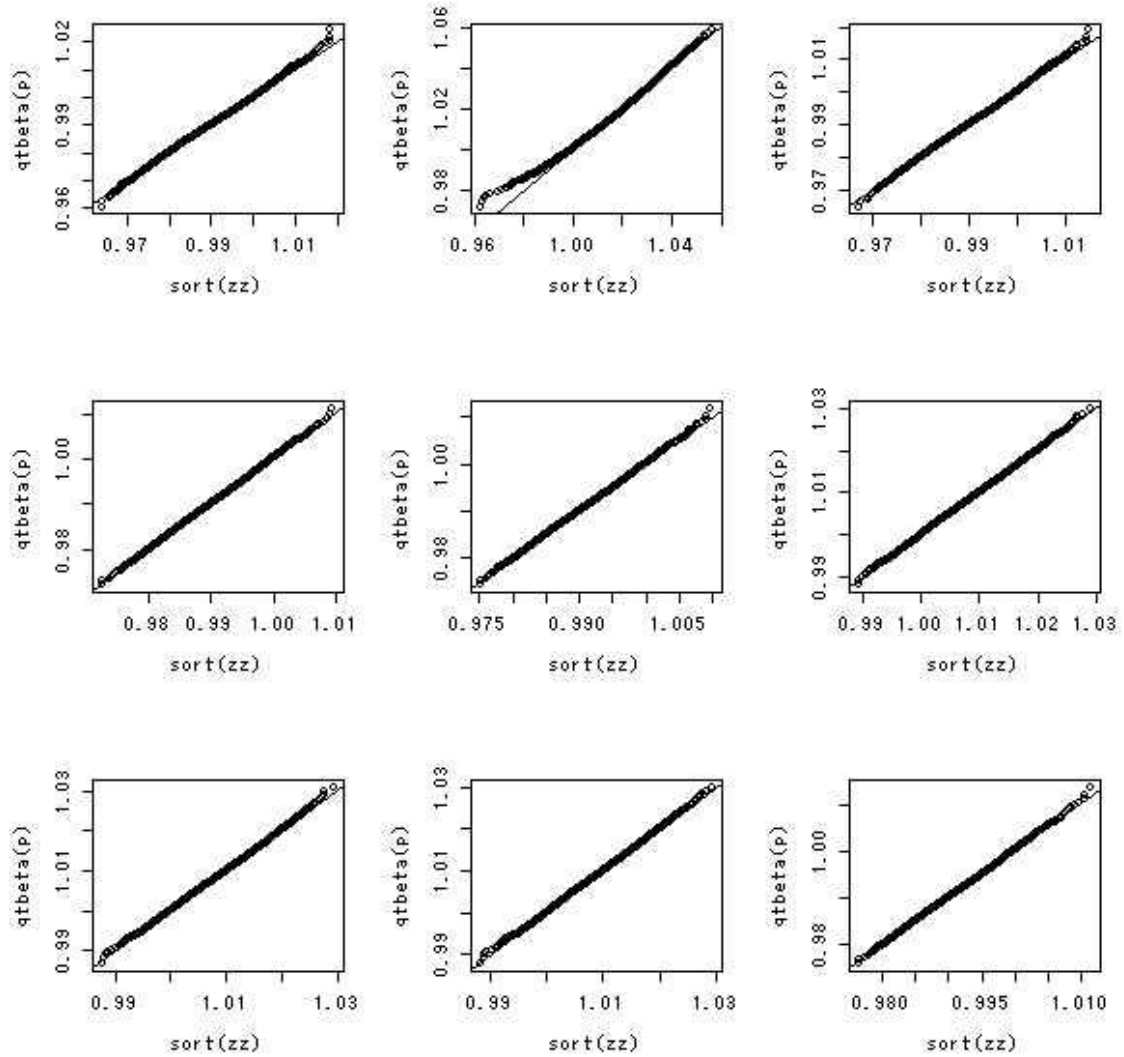


FIG. I.5 – QQ-plots des neuf portefeuilles issus d'une copule de t à un niveau élevé de dépendance

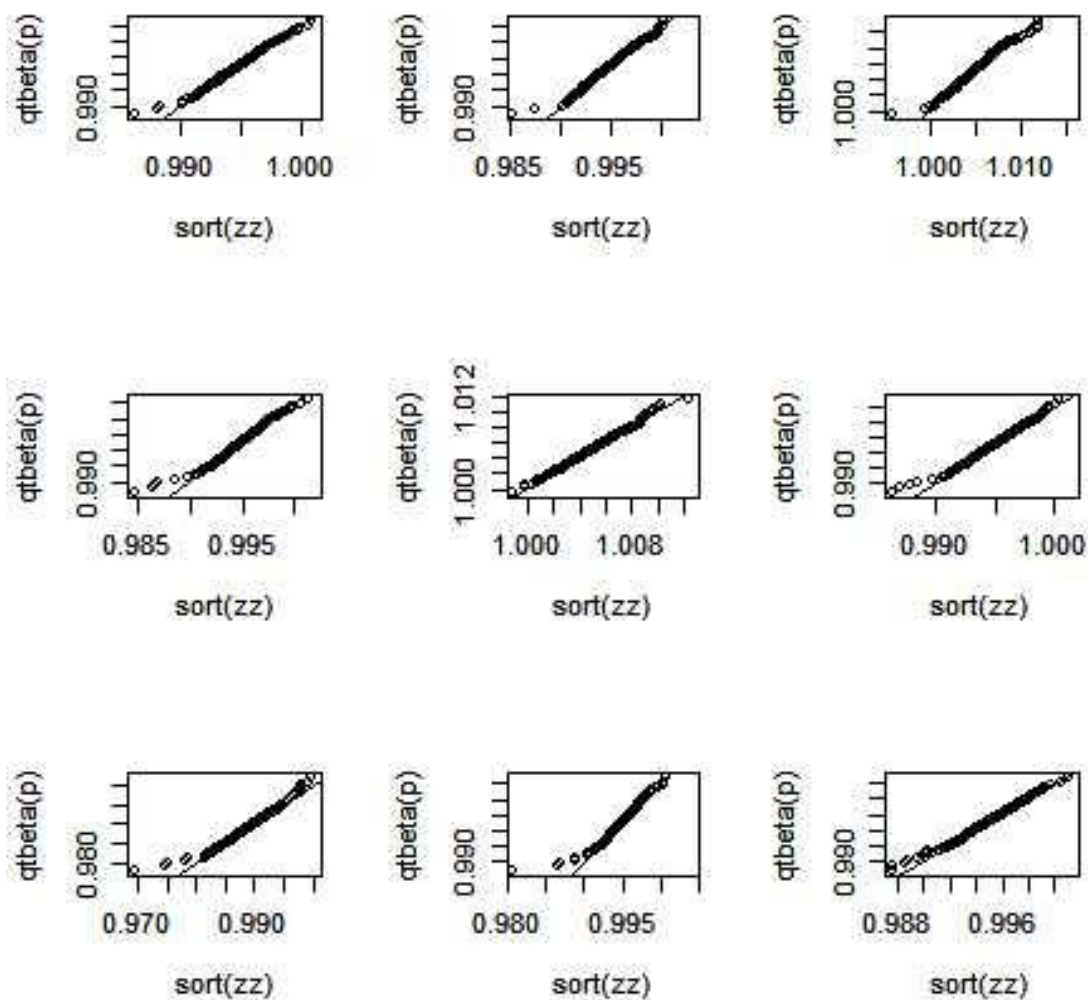


FIG. I.6 – QQ-plots des neuf portefeuilles simulés à partir des données réelles