



Mémoire d'actuariat 2009-2010
Mémoire présenté en vue de l'obtention du titre d'Actuaire
et du Master Actuariat cohabilité entre
l'EURIA et TELECOM BRETAGNE

**Structure de dépendance des générateurs de
scénarios économiques**

Modélisation et Simulation

Kamal ARMEL

Elève Ingénieur à TELECOM Bretagne et élève Actuaire à l'EURIA

Jun 2010

Sujet :

Structure de dépendance des générateurs de scénarios économiques : Modélisation et Simulation.

Référents :

- M. Frédéric PLANCHET : Actuaire agrégé, Docteur en Sciences de gestion, Associés chez WINTER & Associés et professeur à l'ISFA
- M. Aymric KAMEGA : Actuaire qualifié, Consultant chez WINTER & Associés et doctorant au sein du laboratoire de science actuarielle et financière à l'ISFA.

Motivation :

Le sujet du mémoire est proposé par M. Frédéric PLANCHET. Plusieurs facteurs ont motivé le choix de ce sujet. Le premier est le goût pour les mathématiques et la volonté de travailler sur un sujet général qui concerne la finance et l'assurance. Les méthodes et les développements mathématiques proposés peuvent être appliqués dans plusieurs situations où l'étude de la structure de dépendance d'un ensemble d'indices est requise.

Remerciements

Je tiens tout d'abord à remercier mes parents pour leur soutien et leur présence permanente tout au long de ma scolarité.

Je tiens également à remercier monsieur Frédéric PLANCHET et monsieur Aymric KAMEGA pour les compétences techniques et relationnelles qu'ils ont mis à ma disposition tout au long de la réalisation de ce mémoire. Je tiens aussi à les remercier de leur disponibilité, de leur confiance et de la qualité professionnelle et académique de mon expérience à WINTER et Associés. Sans leurs apports il m'aurait été impossible de mettre au point ce mémoire.

Enfin, je souhaiterais adresser toute ma gratitude, d'une part aux corps professoral et administratif de l'EURIA et d'autre part à Monsieur Philippe LENCA pour la qualité de son encadrement durant ces deux années de double cursus entre l'EURIA et TELECOM Bretagne.

Table des matières

Table des figures	10
Table des tableaux	12
Introduction générale	14

Partie I

Les générateurs de scénarios économiques, problématique et données historiques	17
---	----

Chapitre I-1 : Générateurs de scénarios économiques	18
--	----

1. Générateurs de scénarios économiques	19
2. Besoin en termes de GSE	20
2.1. La réglementation prudentielle	20
2.2. Evaluation des organismes assureurs	21

Chapitre I-2 : Problématique et structure globale du mémoire	22
---	----

1. Pourquoi le coefficient de corrélation n'est-il pas une bonne mesure de dépendance ?.....	23
2. Problématique	24
2.1. Partie I : générateurs de scénarios économiques, problématique et données historiques	25
2.2. Partie II : la structure de dépendance des indices financiers et macroéconomiques.....	25
2.3. Partie III : modélisation et simulation.....	25

Chapitre I-3 : Présentation des données	27
--	----

1. Le taux d'inflation	28
2. Le taux d'intérêt nominal à long terme.....	29
3. Le taux d'intérêt nominal à court terme.....	29
4. Le taux de rendement des actions	29
5. Le taux de rendement de l'immobilier.....	31
6. Statistiques descriptives.....	32

Partie II

La structure de dépendance des indices financiers et macroéconomiques	35
--	----

Chapitre II-1 : Représentation des lois multi-variées par les copules	36
--	----

1. Rappel sur les fonctions de répartition	37
--	----

1.1.	Fonction de répartition uni-variée.....	37
1.1.1.	Définitions.....	37
1.1.2.	Propriétés.....	37
1.2.	La fonction de répartition multi-variée.....	38
2.	Copules : définition et propriétés.....	39
2.1.	Définitions.....	39
2.2.	Théorème de Sklar.....	39
2.3.	La densité d'une copule.....	40
2.4.	Tau de Kendall et rho de Spearman.....	40
2.4.1.	Tau de Kendall.....	41
2.4.2.	Rho de Spearman.....	41
2.5.	Dépendance sur les queues de distributions.....	42
3.	Exemples de copules paramétriques.....	43
3.1.	Les copules elliptiques.....	43
3.1.1.	La copule gaussienne.....	43
3.1.2.	La copule de Student.....	44
3.2.	Les copules archimédiennes.....	45
3.2.1.	Copule de Cook-Johnson.....	46
3.2.2.	Copule de Franck.....	47
3.2.3.	Copule de Gumbel.....	48
4.	Copules empiriques.....	49
4.1.	Copule de Deheuvels.....	49
4.2.	Estimation empirique par les densités de probabilité.....	50
5.	Calibrage des copules paramétriques.....	50
5.1.	Estimation par le maximum de vraisemblance.....	50
5.2.	Procédure d'estimation de Joe & Xu [1996].....	51
5.3.	Méthode des moments : Calibrage par le tau de Kendall.....	52
6.	Compléments : calcul des densités de la copule de Gumbel et de la copule de Franck.....	52
6.1.	Densité de la copule de Gumbel.....	52
6.2.	Densité de la copule de Franck.....	53
	Chapitre II-2 : Fonction de répartition et densité de probabilité empirique.....	55
1.	La fonction de répartition empirique classique.....	56
2.	Densité de probabilité empirique.....	56
2.1.	Histogramme.....	57
2.2.	L'estimateur naturel de la densité de probabilité.....	58
2.3.	L'estimateur à noyau.....	58

Chapitre II-3 : Procédure du choix de la copule optimale.....	61
1. Estimation et calibrage des densités de probabilité	62
1.1. Estimation de la densité de probabilité empirique	62
1.2. Calibrage des densités marginales et de la densité jointe	62
2. Fonctions de répartition uni-variées et leurs inverses généralisés	63
3. La copule empirique	63
4. Calibrage des copules paramétriques.....	64
4.1. Calibrage de la copule gaussienne	64
4.2. Calibrage de la copule de Student.....	65
4.2.1. Estimation de la matrice de corrélation	65
4.2.2. Estimation du nombre de degrés de liberté de la copule	66
4.3. Calibrage des copules : Cook-Johnson, Franck et Gumbel.....	66
5. Critères de sélection de la copule optimale	66
Chapitre II-4 : Le choix de la copule optimale	68
1. Densité de probabilité et fonction de répartition empiriques.....	69
2. Calibrage des copules	72
2.1. La copule gaussienne	72
2.2. La copule de Student.....	73
2.3. Le calibrage des copules archimédiennes : Cook-Johnson, Gumbel et Franck	73
3. Le choix de la copule optimale.....	74
Partie III	
Modélisation et simulation.....	77
Chapitre III-1 : Présentation du modèle d’Ahlgrim	78
1. Le taux d’inflation	79
2. Les taux d’intérêt	80
3. Le taux de rendement d’un investissement en actions.....	81
3.1. Le rendement hors dividendes	81
3.2. Le taux de dividende.....	82
3.3. Rendement des actions avec dividendes réinvestis : approche alternative	82
4. Le taux de rendement de l’immobilier.....	83
5. Le prix d’une obligation	83
Chapitre III-2 : Calibrage du modèle d’Ahlgrim.....	85

1.	Le taux d'inflation	86
2.	Les taux d'intérêt	86
3.	Le taux de rendement d'un investissement en actions.....	88
4.	Le taux de rendement de l'immobilier.....	88
5.	Matrice de corrélation des résidus	89

Chapitre III-3 : Projection du modèle d'Ahlgrim avec une structure de dépendance gaussienne

1.	Paramètre d'initialisation des projections.....	91
2.	Projection des résidus et de l'excès de rendement.....	91
3.	Projection de l'inflation et des taux nominaux	91
4.	Projection du prix d'une obligation zéro-coupon	94
5.	Projection du taux de rendement d'un investissement en actions.....	95
6.	Projection du taux de rendement de l'immobilier	96

Chapitre III-4 : Projections du modèle d'Ahlgrim en prenant en compte la structure de dépendance optimale : la copule de Cook-Johnson.....

1.	Simulation des copules paramétriques : Généralité.....	98
2.	Projection du modèle d'Ahlgrim et la copule de Cook-Johnson	99
3.	Projection des indices financiers et macroéconomiques.....	99
3.1.	Projection de l'inflation et des taux nominaux	100
3.2.	Projection du prix d'une obligation zéro-coupon en réel et en nominal, du rendement des actions et du rendement de l'immobilier.....	102
4.	Comparaison de la copule gaussienne et la copule de Cook-Johnson dans le cadre du modèle d'Ahlgrim.....	105
4.1.	La Value at Risk.....	105
4.2.	Composition du portefeuille et horizon de comparaison	106
4.3.	Comparaison des deux VaR du portefeuille évalué par les deux approches.....	107

Conclusion générale

Annexe A

	Modèle de Vasicek	112
1.	Modèle de Vasicek à un facteur.....	112
2.	Modèle de Vasicek à deux facteurs	112
2.1.	Premier cas : t et t_0 très proches	112
2.2.	Deuxième cas : cas général	113

Bibliographie..... 115

Table des figures

Figure 1 : Schéma de l'approche globale	24
Figure 2 : Evolution du taux d'inflation, du taux long nominal et du taux court nominal (janvier 1997-juin 2009)	28
Figure 3 : Evolution du rendement annuel et de la volatilité du SBF 250 dividendes réinvestis (janvier 1997-juin 2009)	30
Figure 4 : Evolution de la volatilité (ordonnée) en fonction du rendement (abscisse).....	31
Figure 5 : Evolution du rendement de l'immobilier (janvier 1997-juin 2009)	32
Figure 6: Fonction de répartition et densité de probabilité d'une loi normale centrée réduite	38
Figure 7 : Densité de probabilité jointe de deux variables aléatoires normales centrées réduites indépendantes	38
Figure 8 : Densité de la copule gaussienne pour une corrélation de 0.5	44
Figure 9 : Densité de la copule de Student pour une corrélation de 0.5 et un nombre de degrés de liberté de 14.....	45
Figure 10 : Densité bi-variée de la copule de Cook-Johnson $\alpha=0.5$	47
Figure 11 : Densité de la copule de Franck bi-variée pour $\alpha=2$	48
Figure 12 : Densité de la copule de Gumbel bi-variée pour $\alpha=2$	49
Figure 13 : Fonction de répartition empirique du taux d'inflation.....	56
Figure 14 : Histogramme du taux d'inflation	57
Figure 15 : Densité à noyau gaussien du taux d'inflation	60
Figure 16 : Comparaison de la fonction de répartition empirique à noyau gaussien et la fonction de répartition empirique classique pour le taux d'inflation	60
Figure 17 : Sélection de la copule optimale par une démarche empirique.....	61
Figure 18 : Densités de probabilité empiriques à noyau gaussien du taux d'inflation et des taux d'intérêt nominaux.....	70
Figure 19 : Densités de probabilité empiriques à noyau gaussien du rendement des actions et du rendement de l'immobilier.....	70
Figure 20 : Densité de probabilité empirique à noyau gaussien du rendement des actions sur un intervalle réduit.....	71
Figure 21 : Fonctions de répartition empiriques à noyau gaussien du taux d'inflation et des taux d'intérêt.....	71
Figure 22 : Fonctions de répartition empiriques à noyau gaussien du rendement des actions et de l'immobilier	72
Figure 23 : Chaîne de Markov du modèle de type Hardy à deux régimes	82
Figure 24 : Schéma de la démarche de projection à partir des résidus et de l'excès de rendement.....	90
Figure 25 : Projection du taux d'inflation	92
Figure 26 : Projection du taux d'intérêt réel à long terme	92
Figure 27 : Projection du taux d'intérêt réel à court terme	93

Figure 28 : Projection du taux d'intérêt nominal à long terme	93
Figure 29 : Projection du taux d'intérêt nominal à court terme	94
Figure 30 : Projection du prix d'une obligation zéro-coupon réelle à maturité 5 ans	94
Figure 31 : Projection du prix d'une obligation zéro-coupon nominale à maturité 5 ans	95
Figure 32 : Projection du rendement des actions.....	95
Figure 33 : Projection du taux de rendement de l'immobilier	96
Figure 34 : Projection du taux d'inflation.....	100
Figure 35 : Projection du taux d'intérêt réel à long terme	101
Figure 36 : Projection du taux d'intérêt réel à court terme	101
Figure 37 : Projection du taux d'intérêt nominal à long terme	102
Figure 38 : Projection du taux d'intérêt nominal à court terme	102
Figure 39 : Projection du prix d'une obligation zéro-coupon réelle à maturité 5 ans	103
Figure 40 : Projection du prix d'une obligation zéro-coupon nominale à maturité 5 ans	103
Figure 41 : Projection du rendement des actions.....	104
Figure 42 : Projection du taux de rendement de l'immobilier	104
Figure 43 : Comparaison des VaR à 99,5 % sur un horizon de 3 ans	107

Table des tableaux

Tableau 1 : Statistiques descriptives classiques	33
Tableau 2 : Matrice de corrélation sur l'ensemble de l'historique	33
Tableau 3 : Calibrage des densités empiriques.....	69
Tableau 4 : Matrice de corrélation de la copule gaussienne.....	73
Tableau 5 : Matrice de corrélation de la copule de Student	73
Tableau 6 : Calibrage des copules Cook-Johnson, Gumbel et Franck	73
Tableau 7 : Distances entre les copules paramétriques et la copule empirique.....	74
Tableau 8 : Calibrage du taux d'inflation	86
Tableau 9 : Tests sur le modèle de l'inflation.....	86
Tableau 10 : Calibrage du modèle des taux réels.....	87
Tableau 11 : Tests sur le modèle du taux réel long	87
Tableau 12 : Tests sur le modèle du taux réel court.....	87
Tableau 13 : Calibrage de l'excès de rendement	88
Tableau 14 : Calibrage du modèle sur le rendement de l'immobilier.....	88
Tableau 15 : Tests sur le modèle de l'immobilier.....	89
Tableau 16 : Matrice de corrélation des résidus et de l'excès de rendement.....	89
Tableau 17 : Valeurs initiales	91
Tableau 18 : Valeurs initiales	100

Introduction générale

La récente crise financière et économique à la quelle sont confrontés depuis 2007 les acteurs financiers et économiques conduit à une large mise en cause des modèles mathématiques en finance, notamment en matière de prise en compte des dépendances extrêmes entre les indices financiers. La crise des *subprimes* nous a donné l'exemple de la forte dépendance entre d'une part la capacité des ménages endettés à rembourser leurs emprunts et d'autre part la mise en marché massive des biens immobiliers portés en caution qui a eu pour effet de produire une baisse significative du foncier. Une modélisation mathématique prudente et réaliste doit s'intéresser alors aux structures de dépendances et étudier l'impact des queues de distributions et des événements extrêmes.

La mesure de la dépendance des indices financiers par le coefficient de corrélation est une pratique largement répandue dans les milieux financiers et académiques et un bon nombre de modèles financiers comme le MEDAF et l'APT se placent dans un cadre gaussien excluant ainsi la dépendance extrême des indices financiers. Dans un univers non gaussien, la mesure de la dépendance par le coefficient de corrélation est non appropriée.

Par ailleurs, les modèles classiques de la théorie financière considèrent que le coefficient de corrélation est constant dans le temps. Cependant, plusieurs travaux, comme Login et Solnik [1995] pour les indices boursiers internationaux, prouvent que la corrélation n'est pas constante dans le temps et qu'elle est d'autant plus forte quand le marché est agité. Or, plusieurs stratégies reposent sur la diversification et donc l'augmentation de la corrélation avec la volatilité est problématique puisque les bénéfices sont réduits au moment où ils sont le plus attendus.

Depuis quelques années, la recherche financière s'oriente vers l'étude d'autres outils permettant de mieux expliquer, appréhender et prévoir les évolutions futures des indices financiers et macroéconomiques. La théorie des copules est un outil puissant permettant de modéliser les dépendances extrêmes.

Nos travaux vont dans ce sens et proposent, dans un premier temps, l'étude de la structure de dépendance d'un ensemble d'indices financiers et macroéconomiques dans un cadre purement empirique qui peut être adapté à plusieurs situations (cf. partie II). L'objectif est de travailler sur un sujet d'actualité et d'adopter une démarche générale qui peut concerner plusieurs secteurs comme la finance ou l'assurance. L'idée est de répondre à la question suivante : disposant des séries historiques d'un ensemble de variables (financières, économiques...) quelle est leur structure de dépendance ?

Pour illustrer l'intérêt de cette étude nous nous sommes intéressés à la génération de scénarios économiques. Notre objectif est de se placer dans un cadre qui touche à la fois l'assurance et la finance et de travailler sur un sujet qui n'est que rarement examiné en assurance. Nous essayons donc de répondre à un réel besoin des assureurs en termes de modélisation mathématique et de structure de dépendance.

En effet, les modèles financiers intéressent de plus en plus les assureurs notamment avec l'avènement de Solvabilité 2. Cette nouvelle réglementation donne une place de choix à la modélisation mathématique dans le but de se placer dans une meilleure représentation du fonctionnement des entreprises et de leurs bilans. Aussi, les assureurs développent une réflexion sur la gestion de leurs actifs. Cette problématique n'est pas nouvelle et la projection sur le long terme des valeurs de marché des actifs financiers et des variables macroéconomiques, constitue un exercice indispensable dans le cadre des modèles ALM en assurance vie, DFA en assurance non vie, ou encore des *embedded value*.

Dans ce cadre, la génération des scénarios économique s'avère incontournable. Planchet, Théron et Kamega [2009] proposent la définition suivante : « un scénario économique correspond à une projection de grandeurs économiques et financières sur un horizon d'intérêt ». Par exemple, les grandeurs de l'économie qui intéressent un assureur sont généralement celles qui ont un impact sur le développement de son activité. Ainsi, les variables classiques des marchés financiers et de la dette, que sont les taux d'intérêt, la performance des actions ou du foncier, sont des grandeurs de l'économie qui intéressent les assureurs ou les gestionnaires de fonds dans la mesure où leur évolution va conditionner le prix futur des actifs financiers mais aussi les conditions de désinvestissement et de réinvestissement futurs. En outre, l'inflation, dont l'évolution influe sur les prix des actifs financiers et sur les prestations futures, apparaît également comme une grandeur économique d'intérêt.

Il est donc particulièrement intéressant d'étudier l'évolution des taux d'inflation, des taux d'intérêt à court terme, des taux d'intérêt à long terme, du rendement des actions et du rendement de l'immobilier. Plusieurs modèles (Wilkie, Brennan et Xia, Ahlgrim : cf. Planchet, Théron et Kamega [2009]) partent de ces indices pour étudier les fonds dont les supports d'investissement sont par exemple : les obligations, les actions et l'immobilier. Ces modèles ne proposent pas, en général, une étude de la structure de dépendance entre les taux d'inflation, les taux d'intérêt à court terme, les taux d'intérêt à long terme, le rendement des actions et le rendement de l'immobilier et se limitent à considérer que le lien entre ces indices est effectué en introduisant les corrélations.

Cependant, les corrélations n'introduisent pas de dépendance de queues et s'avèrent donc limitées en temps de crise. Ainsi avant de faire le choix du ou des modèles les plus pertinents, une réflexion sur la structure de dépendance doit être menée.

Nos travaux s'articulent sur cette problématique et visent dans la deuxième partie à étudier la structure de dépendance des indices présentés ci-dessus en s'appuyant sur la théorie des copules. Nous essayons dans la troisième partie d'exploiter, en utilisant le modèle d'Ahlgrim, les conclusions formulées dans la deuxième partie.

Il est rappelé que les démarches, les calculs et les méthodes présentés dans ce mémoire, notamment lors de l'étude de la structure de dépendance, s'inscrivent dans un cadre plus général et peuvent donc être facilement adaptés à d'autres contextes.

Partie I

Les générateurs de scénarios économiques, problématique et données historiques

L'objet du premier chapitre de cette partie est de revenir sur la notion de scénarios économiques et de présenter leur importance pour un organisme d'assurance. Le second chapitre est consacré à la présentation de la problématique et de la structure globale du mémoire et tentera d'expliquer pourquoi le coefficient de corrélation n'est pas une bonne mesure de dépendance. Enfin, le troisième chapitre est principalement consacré à la présentation des séries de données historiques qui sont utilisées dans notre étude. L'objectif de ce chapitre est de mener une analyse de ces séries afin de se mettre en capacité d'estimer et d'interpréter convenablement les paramètres lors du calibrage des copules et des modèles présentés par la suite.

Chapitre I-1

Générateurs de scénarios économiques

L'objet de la première section de ce premier chapitre est de revenir sur la notion de scénario économique et de présenter quelques aspects caractérisant son contexte d'utilisation et ses perspectives futures. La seconde section donne un aperçu de l'importance de la génération de scénarios économiques pour les organismes assureurs et illustre le besoin en termes de générateurs de scénarios économiques dans le milieu de l'assurance.

1. Générateurs de scénarios économiques

Dans le cadre de son pilotage technique, un organisme assureur ou un gestionnaire de fonds (un fonds de retraite par exemple) doit développer une réflexion sur la gestion de son actif. Cette problématique n'est pas nouvelle et la projection sur le long terme des valeurs du marché des actifs financiers et des variables macroéconomiques, souvent appelée « génération d'un scénario économique » constitue un exercice indispensable dans le cadre des modèles ALM en assurance vie, DFA en assurance non vie, ou encore des *embedded value*.

Dans le livre Planchet, Théron et Kamega [2009] on retrouve la définition suivante d'un scénario économique : « un scénario économique correspond à une projection de grandeurs économiques et financières sur un horizon d'intérêt ». Les générateurs de scénarios économiques, que nous écrirons désormais GSE (en anglais ESG : *Economic Scenario Generator*), peuvent produire des scénarios contenant les différents indicateurs de l'économie, tels que les taux d'intérêt, le taux d'inflation des prix et des salaires, le taux de rendement obligataire, le rendement des actions, le rendement de l'immobilier, les taux de change et le taux de chômage (cf. Ahlgrim et al [2005]). La projection de ces indices financiers et macroéconomiques permettent à un gestionnaire de fonds d'avoir une estimation des valeurs futures des couples rendements/risques des actifs gérés (actions, immobilier, obligations...) et une estimation de la valeur future du passif (provisions techniques, dettes financière...).

Les grandeurs de l'économie qui intéressent un assureur sont généralement celles qui ont un impact sur le développement de son activité. Ainsi, les variables classiques des marchés financiers et de la dette, que sont les taux d'intérêt, la performance des actions ou du foncier sont des grandeurs de l'économie qui intéressent les assureurs dans la mesure où leurs évolutions vont conditionner le prix futur des actifs financiers mais aussi les conditions de désinvestissement et de réinvestissement futurs. En outre, l'inflation, dont l'évolution influe sur les prix des actifs financiers et sur les prestations futures, apparaît également comme une grandeur économique d'intérêt.

Les modèles permettant la construction des GSE sont de plus en plus utilisés par les assureurs. La construction des GSE est proposée dans plusieurs modèles. Ces modèles peuvent être répartis en deux catégories : les modèles composites et les modèles intégrés.

Les modèles composites s'appuient sur une description *ad'hoc* de chaque classe d'actif et les rassemblent ensuite pour avoir une description globale. Typiquement on peut trouver un modèle où les actions et l'immobilier suivent le modèle de Black & Scholes, les taux d'intérêt modélisés par le modèle CIR et l'inflation qui suit le modèle de Vasicek. Bien que facile à concevoir, le seul lien entre les différentes classes d'actifs est la corrélation ce qui fait que la cohérence d'ensemble d'un tel dispositif n'est pas assurée de manière évidente.

Les modèles intégrés proposent une construction structurée de plusieurs classes d'actifs en s'appuyant sur une variable explicative de référence. L'exemple classique de tels modèles est le modèle de Wilkie¹ où les taux d'intérêt, le prix de l'immobilier et le prix des actions sont décrits à partir de l'inflation qui joue ici le rôle de la variable explicative. Les modèles intégrés proposent en général une description de l'inflation dont dérivent les autres indices et donc les dynamiques de l'ensemble des indices décrits dans le modèle sont plus cohérentes que dans les modèles composites. En plus du modèle de Wilkie, d'autres modèles intégrés sont développés. On peut notamment citer le modèle d'Ahlgrim et al [2005] présenté au chapitre III.1 et le modèle de Brennan et Xia².

Cependant, ces modèles (composites et intégrés) ne proposent pas, en général, une étude de la structure de dépendance entre les indices qu'ils modélisent et se limitent à considérer que le lien entre

¹ Le modèle est présenté à Planchet, Théron et Kamega [2009] page 98.

² Le modèle est présenté à Planchet, Théron et Kamega [2009] page 106.

ces indices est simplement effectué en introduisant les corrélations. Bien que la mesure de la dépendance par la corrélation soit une pratique répandue dans les milieux financiers, elle n'introduit pas de dépendance de queues et s'avère donc limitée en temps de crise. Ainsi, avant de faire le choix du ou des modèles les plus pertinents, une réflexion sur la structure de dépendance des indices modélisés doit être menée. Cette problématique est au cœur de nos travaux et sera présentée dans la partie II.

2. Besoin en termes de GSE

Cette section est largement inspirée du livre Planchet, Théron et Kamega [2009] et tente de revenir sur les besoins des assureurs en matière de scénarios économiques.

En effet, les besoins des assureurs en matière de scénarios économiques se sont développés ces dernières années. L'évolution de ces besoins est due principalement à deux facteurs : l'évolution de la réglementation prudentielle et l'évolution de la communication financière des assureurs notamment en matière d'évaluation des entreprises en assurances.

2.1. La réglementation prudentielle

La réglementation prudentielle a connu au fil du temps de nouvelles mesures réglementaires, dont la mise en œuvre, nécessite l'utilisation de scénarios économiques définis par le législateur. Ces générateurs sont généralement adverses et constituent des stress-tests de la solvabilité des entreprises d'assurance.

Nous allons nous contenter dans la suite de présenter quelques scénarios économiques fournis par le législateur sans entrer dans le détail des articles du code des assurances. Pour plus d'approfondissement se référer à ce code ou au livre Planchet, Théron et Kamega [2009].

Les deux premiers scénarios économiques que nous présentons concernent la provision pour aléas financiers. Définie à l'article R331-3 du code des assurances comme étant destinée à compenser la baisse de rendement de l'actif, son calcul est détaillé à l'article A331-2. Cet article explicite deux scénarios. Le premier consiste à considérer que les actifs financiers ne produiront des revenus qu'à hauteur de 60 % du TME à la date du calcul. Le deuxième consiste à distinguer les titres obligataires, pour lesquels un rendement à hauteur du TME est à considérer, des autres titres et emplois pour lesquels un rendement à hauteur de 75 % du TME est à considérer s'il s'agit d'investissement à court terme et à hauteur de 60 % s'il s'agit d'investissement à plus long terme.

En ce qui concerne la gestion actif-passif, les assureurs doivent remettre trimestriellement un rapport, l'état T3, à l'Autorité de Contrôle des Assurances et des Mutuelles (ACAM). Cet état décrit les impacts d'une variation des taux d'intérêt ou d'une baisse des cours boursiers sur les actifs et les passifs des compagnies d'assurances et « s'inscrit dans le cadre d'une démarche préventive, dont le but est de faire que les acteurs du marché, en prenant conscience de l'impact potentiel de perturbations brutales de leur environnement financier, se dotent des outils de gestion adéquats »³. En revanche « il doit permettre d'apprécier l'efficacité de l'outil de gestion actif-passif existant au sein de l'entreprise ». Dans ce cadre des chocs sur les prix des actions (jusqu'à -40 %), de l'immobilier (jusqu'à -40 %) et sur les taux (de -300 pb à +400 pb par rapport à la partie entière de la référence TEC 10⁴) devraient être envisagés (Planchet, Théron et Kamega [2009]).

Aussi, le législateur explicite à l'article A344-15 les modalités de calcul du test d'exigibilité qui consiste en un stress-test de la capacité de l'assureur à honorer ses engagements à un horizon de 5 ans.

³ Commission de Contrôle des Assurances : Rapport d'Activité 2000/2001.

⁴ Le TEC 10 est un indice quotidien des rendements des emprunts d'État à long terme.

Ce test « est pratiqué à partir d'hypothèses financières standardisées. Ces hypothèses consistent, par rapport à leur moyenne respective constatée sur les trois dernières années :

- en une baisse de l'indice boursier de référence de 30 % ;
- en une hausse de deux points des taux d'intérêt de l'obligation de référence ;
- en une baisse de 20 % des prix des transactions immobilières. » (cf. A344-15)

Nous constatons que les chocs décrits sont statiques et s'appliquent à une assiette déterminée de manière prospective en intégrant les flux de trésorerie attendus des actifs à échéance au cours de cette période et des prestations correspondantes.

Il apparaît donc que, jusqu'à solvabilité 2, les utilisations de scénarios économiques demandées par la réglementation avaient en commun un caractère statique : elles apparaissaient d'avantage comme des stress-tests prudentiels que comme des prévisions de l'évolution des indices financiers et des variables macroéconomiques.

Par ailleurs, dans le référentiel Solvabilité II, la génération de scénarios économiques intervient au moins à deux niveaux lors de l'évaluation de la solvabilité globale d'un assureur. Le premier niveau concerne le calcul des provisions techniques relatives aux engagements d'assurance contenant des options et des garanties financières. Le second niveau concerne l'estimation du capital de solvabilité par le biais d'un modèle interne.

2.2. Evaluation des organismes assureurs

Devenue un standard dans la valorisation des entreprises d'assurance, l'*embedded value*, ou valeur intrinsèque, d'une compagnie d'assurance-vie se calcule comme la somme de l'actif net réévalué, de la valeur actualisée des profits futurs sur les polices d'assurances actuellement en cours sous déduction du coût du capital. Que ce soit à des fins de transactions (fusions acquisitions) ou de communication, les calculs de l'*embedded value* se sont généralisés et nécessitent la génération de scénarios économiques.

Globalement, la génération de scénarios économiques est aujourd'hui une problématique incontournable pour les organismes assureurs et également un passage obligé en matière de gestion d'organismes d'assurances et plus particulièrement dans la recherche d'allocation stratégique d'actif.

La grande majorité des générateurs de scénarios économiques modélisent la dépendance entre les indices financiers par des corrélations. Cette modélisation semble inadaptée à des situations extrêmes et peut conduire à sous estimer de manière significative les besoins en capitaux notamment lorsqu'il s'agit de déterminer un quantile extrême d'une distribution engendrée par un grand nombre de risques qui interagissent entre eux. Cette problématique est au cœur des travaux présentés par la suite et est détaillée au chapitre suivant.

Problématique et structure globale du mémoire

Face à l'importance et la nécessité de générer des scénarios économiques, les assureurs doivent développer une réflexion en matière de modélisation leur permettant de générer des scénarios réalistes et estimant au mieux la tendance globale des indices financiers et des variables macroéconomiques. Ces estimations permettront, entre autres, à l'assureur d'estimer sa marge de solvabilité, et donc de mobiliser les fonds nécessaires pour éviter le risque de ruine (les assureurs devront à partir de 2012 estimer le niveau de fonds propre pour ne pas être en ruine sur un horizon d'un an avec une probabilité 99,5 %), et permet aussi une gestion plus évoluée de l'organisme assureur notamment en ce qui concerne la détermination de l'allocation optimale et les politiques de réinvestissement et de désinvestissement futures.

Comme nous l'avons évoqué précédemment, plusieurs modèles permettent de générer des scénarios économiques (Wilkie, Brennan et Xia, Ahlgrim). Toutefois, la grande majorité de ces générateurs de scénarios économiques modélisent la dépendance entre les indices financiers et les variables macroéconomiques par des corrélations. Cela tient à plusieurs facteurs. Le plus classique est la simplicité des modèles à dépendance linéaire (corrélation) que ce soit au niveau de l'estimation des paramètres qu'au niveau de la mise en œuvre opérationnelle. Aussi, l'utilisation de modèles plus développés (utilisant par exemple les copules) peut s'avérer complexe notamment lors du calibrage ou de la mise en œuvre.

Les corrélations n'introduisent pas de dépendance de queues et s'avèrent donc limitées notamment en temps de crise. La première section de ce chapitre tentera d'expliquer pourquoi le coefficient de corrélation n'est pas une bonne mesure de dépendance. La deuxième section est consacrée à la présentation de la problématique et la structure globale du mémoire.

1. Pourquoi le coefficient de corrélation n'est-il pas une bonne mesure de dépendance ?

La mesure de la dépendance des indices financiers par le coefficient de corrélation est une pratique largement répandue dans les milieux financiers et académiques. Par exemple, les modèles du MEDAF (Modèle d'évaluation des actifs financiers) et de l'APT (*Arbitrage Pricing Theory*) utilisent la corrélation comme mesure de dépendance entre plusieurs instruments financiers et développent une théorie essentiellement basée sur la normalité multi-variée des rendements afin de construire un portefeuille optimal.

Bien que la corrélation linéaire soit l'un des concepts les plus présents, elle est aussi l'un des plus mal utilisée. En effet, la corrélation est seulement une mesure parmi d'autres (cf. chapitre II-1 : tau de Kendall, rho de Spearman...) de la dépendance stochastique entre plusieurs variables. Le coefficient de corrélation est la mesure canonique de la dépendance dans le cadre d'un modèle multi-varié gaussien et plus généralement pour les distributions elliptiques (cf. chapitre II-1). Donc utiliser le coefficient de corrélation pour mesurer la dépendance en dehors de ce cadre peut être un générateur d'erreurs fréquentes.

D'après Embrechts et al. [1999] la corrélation est naturellement utilisée pour décrire la structure de dépendance lorsque la distribution jointe des indices financiers étudiés appartient à la famille des distributions elliptiques (cf. chapitre II-1). Il est à noter que la famille des distributions elliptiques contient la distribution normale et la distribution de Student. Aussi, Embrechts et al. [1999] listent les principaux défauts du coefficient de corrélation. On retrouve notamment les points suivants :

- La corrélation linéaire n'est pas définie si l'une des deux variances des deux variables aléatoires considérées n'est pas finie. Il s'agit d'un réel problème lorsque l'on considère des distributions ayant des queues épaisses. Cette situation peut être rencontrée lors de la modélisation des rendements des actions.
- Le coefficient de corrélation de deux variables aléatoires peut s'annuler sans que les deux variables soient indépendantes. L'exemple classique de ce genre de situation, consiste à considérer une variable normale : $X \approx N(0,1)$, et de calculer ensuite sa corrélation avec la variable $Y = X^2$. La corrélation s'annule bien que X et X^2 soient parfaitement dépendants.
- La corrélation linéaire n'est pas invariante sous transformations croissantes non linéaires. Cela signifie que d'un point de vue économique et informationnel, le coefficient de corrélation linéaire n'est pas une mesure de dépendance cohérente. Par exemple, Embrechts et al. [1999] affirme que la corrélation entre un couple de variables (deux indices financiers) X et Y est en général différente de la corrélation entre le couple de variables $\ln(X)$ et $\ln(Y)$. Cela veut dire que les coefficients de corrélation sont différents bien que le contenu informationnel dans les deux couples soit le même.
- Le coefficient de corrélation ne peut être calculé qu'à partir des deux volatilités du couple de variables aléatoires étudié. Se limiter aux coefficients de corrélation pour décrire la dépendance signifie que l'on ne peut étudier celle-ci indépendamment de la volatilité. Or un bon nombre d'études et de modèles (Heston [1993]) affirme que la volatilité suit un processus stochastique et dans ce cadre l'étude de la structure de dépendance devient complexe.

Ainsi le coefficient de corrélation linéaire ne permet pas d'étudier, de manière réaliste, la structure de dépendance d'un ensemble d'indices financiers et macroéconomiques. Il existe toutefois une alternative à la fois plus consistante et plus réaliste permettant une étude plus développée de la dépendance. Il s'agit de la théorie des copules. Notre problématique s'inscrit dans ce cadre et s'appuie sur cette théorie pour proposer une méthode plus réaliste et plus générale permettant une étude correcte de la structure de dépendance entre un ensemble de variables financières et macroéconomiques.

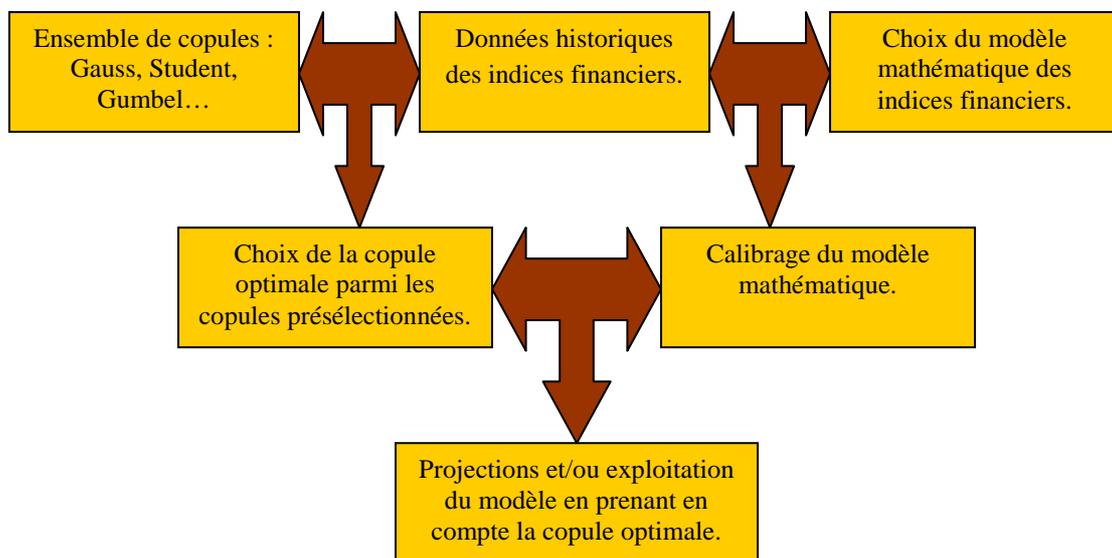
2. Problématique

L'étude de la structure de dépendance des indices financiers et macroéconomiques d'un générateur de scénarios économiques constitue le cœur de nos travaux. Notre objectif est de présenter un ensemble de méthodes et d'outils permettant d'estimer de manière réaliste et affinée les performances à long terme d'un fonds d'actifs géré par un organisme assureur (classiquement composé d'obligations, d'actions et d'immobilier).

Dans ce cadre, il est particulièrement intéressant d'étudier la structure de dépendance et l'évolution des taux d'inflation, des taux d'intérêt à long terme, des taux d'intérêt à court terme, du rendement d'un investissement en actions et du rendement d'un investissement en immobilier. Les projections de ces indices peuvent servir pour étudier les fonds dont les supports d'investissement sont, entre autres : les obligations, les actions et l'immobilier.

Il est à noter que les démarches, les calculs et les méthodes présentés dans ce mémoire, notamment lors de l'étude de la structure de dépendance (cf. partie II), s'inscrivent dans un cadre plus général et peuvent donc être facilement adaptés à d'autres contextes comme l'étude de la structure de dépendance des indices boursiers (nationaux et internationaux) ou l'étude de la structure de dépendance des prix des matières premières... L'approche globale consiste à considérer un ensemble d'indices financiers et macroéconomique dont l'étude de la structure de dépendance permet de choisir, parmi un ensemble de copules, la copule adaptée aux données historiques. Cette copule peut être utilisée ensuite pour projeter ces indices en passant par un modèle décrivant leurs dynamiques. La figure suivante schématise cette approche :

Figure 1 : Schéma de l'approche globale



En ce qui concerne les travaux présentés dans ce mémoire :

- Les indices financiers et macroéconomiques étudiés sont : le taux d'inflation, le taux d'intérêt à long terme, le taux d'intérêt à court terme, le rendement d'un investissement en actions et le rendement d'un investissement en immobilier.
- Les copules retenues sont des copules classiquement étudiées en finance (en général dans le cadre d'un modèle bi-varié) à savoir : la copule gaussienne, la copule de Student, la copule de Cook-Johnson, la copule de Gumbel et la copule de Franck.

- Le modèle d'Ahlgrim et al. [2005] est utilisé pour la projection des indices financiers et macroéconomiques (cf. partie III).

Le mémoire se compose donc de trois parties présentées ci-dessous.

2.1. Partie I : générateurs de scénarios économiques, problématique et données historiques

La première partie reprend la notion de générateurs de scénarios économiques et décrit de manière succincte la place de la génération des scénarios économiques en assurance. Le second chapitre de cette partie présente la problématique du mémoire et le troisième présente les données utilisées dans le cadre de cette problématique.

2.2. Partie II : la structure de dépendance des indices financiers et macroéconomiques

Nous présentons dans cette partie les démarches et les méthodes qui permettent d'extraire la structure de dépendance d'un ensemble d'indices en identifiant la copule adaptée à leurs historiques.

L'identification de la copule adaptée aux données étudiées est une problématique aussi importante que son utilisation : quelle est la copule qui reflète le mieux la structure de dépendance des données considérées?

La réponse que nous apportons à cette question se base sur un ensemble de copules paramétriques prédéfinies et sur la série de données que l'on veut étudier⁵. Plus précisément, si nous disposons d'un historique de d indices financiers, alors la démarche proposée permet de sélectionner la copule paramétrique la plus adaptée à l'historique parmi les copules paramétriques présélectionnées.

L'ensemble que nous avons retenu contient cinq copules présentant cinq structures de dépendance différentes. Il s'agit de la copule gaussienne, la copule de Student, la copule de Franck, la copule de Gumbel et la copule de Cook-Johnson.

Genest et Rivest [1993] proposent une procédure qui permet de comparer un ensemble de copules archimédiennes. Cette procédure exclue donc la copule gaussienne et la copule de Student et n'est donc pas adaptée à notre étude.

Notons finalement que les conclusions que l'on peut faire à l'issue de cette partie sont indépendantes du modèle mathématique qui peut décrire les dynamiques des indices financiers et ne dépendent que des séries historiques retenues.

2.3. Partie III : modélisation et simulation

Cette partie consiste en une réflexion sur la modélisation des indices présentés ci-dessus en intégrant les conclusions formulées à l'issue de la deuxième partie. Le modèle qui est à la base de cette réflexion est le modèle d'Ahlgrim. L'un des objectifs de ce volet est de projeter les indices en question ce qui peut servir ensuite à évaluer les performances d'un fonds composé d'obligations, d'actions et de l'immobilier.

Avant de se lancer dans la recherche de la copule optimale, il est particulièrement intéressant d'étudier les données historiques des indices financiers et macroéconomiques retenus, et ce afin de se mettre en capacité de mieux interpréter et exploiter les résultats observés lors de l'étude de la structure

⁵ Cette méthode étend celle présentée par Kharoubi-Rakotomalala [2008] où la mise en œuvre ne concerne que les copules bi variées.

de dépendance ou lors du calibrage et la projection du modèle d'Ahlgren. Le chapitre suivant présente donc les séries de données historiques utilisées dans l'ensemble des travaux.

Chapitre I-3

Présentation des données

Comme nous l'avons évoqué au chapitre I-2, nous nous intéressons à la génération de scénarios économiques dont les indices financiers et macroéconomiques sont le taux d'inflation, le taux d'intérêt à long terme, le taux d'intérêt à court terme, le rendement d'un investissement en actions et le rendement d'un investissement en immobilier. Pour un organisme assureur, la génération de tels scénarios doit intégrer une perspective de moyen ou long terme dans laquelle s'inscrit la gestion de son actif. Cette perspective doit être prise en compte lors de l'estimation des paramètres du modèle retenu. De manière plus précise, la profondeur de l'historique et l'horizon de projection doivent être cohérents.

Ce chapitre est principalement consacré à la présentation des séries de données historiques qui sont utilisées dans notre étude. L'objectif est de mener une analyse de ces séries afin de se mettre en capacité d'estimer et d'interpréter convenablement les paramètres lors du calibrage des copules et des modèles présentés par la suite. Cette estimation doit être effectuée sur une plage temporelle large avec une fréquence significative. Cela permet, dans un premier temps, une meilleure estimation des paramètres et permet, dans un second temps, de capter la structure de dépendance des indices financiers étudiés.

Les séries de données retenues datent de janvier 1997 à juin 2009, et couvrent donc une période d'environ 13 ans, ce qui est cohérent avec un horizon de projection à moyen terme. Aussi, la construction de ces séries de données s'est appuyée sur les sources utilisées par Friggit [2007] et mis à part la série historique du rendement de l'immobilier, qui est à fréquence trimestrielle, toutes les autres séries sont à fréquences mensuelles.

Les statistiques descriptives classiques de chaque indice financier sont présentées dans la suite. Le test de Jarque-Bera⁶ est retenu pour tester la normalité. Pour ce test, l'hypothèse nulle H_0 est celle de la normalité de l'indice. Ce test vaut 1 si on peut refuser H_0 avec un niveau de significativité de 5 % et vaut 0 sinon.

⁶ Pour les lecteurs intéressés, se référer à Bera et Jarque [1980] ou consulter l'aide en ligne Matlab sur le site de Mathworks : <http://www.mathworks.com/access/helpdesk/help/toolbox/stats/jbtest.html>.

1. Le taux d'inflation

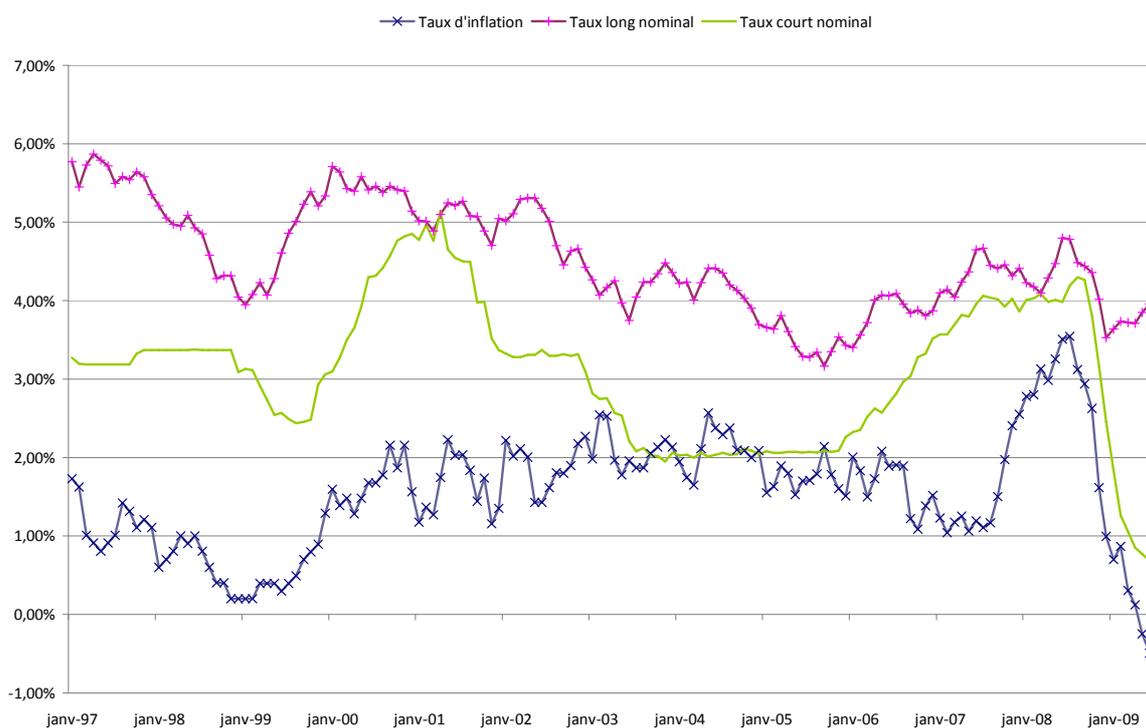
L'historique du taux d'inflation est construit à partir de l'indice des prix à la consommation (IPC) publié sur le site de l'INSEE⁷. Cet historique correspond à l'IPC de la France entière à périodicité mensuelle et est régulièrement mis à jour depuis janvier 1990. Cet indice est utilisé pour calculer le taux d'inflation annuel à fréquence mensuelle (taux sur 12 mois glissants) entre janvier 1997 et juin 2009. Plus précisément, le taux d'inflation annuel au mois m s'écrit :

$$q_m = \ln\left(\frac{IPC_m}{IPC_{m-12\text{mois}}}\right)$$

où IPC_m est l'indice des prix à la consommation au mois m .

Le tableau 1 (page 31) présente les statistiques descriptives classiques du taux d'inflation annuel (entre janvier 1997 et juin 2009). La figure suivante illustre l'évolution du taux d'inflation entre janvier 1997 et juin 2009. On observe en particulier le fort niveau de l'inflation en 2008 et la chute brutale, jusqu'à des niveaux négatifs, en 2009 suite à la crise financière.

Figure 2 : Evolution du taux d'inflation, du taux long nominal et du taux court nominal (janvier 1997-juin 2009)



⁷ www.indices.insee.fr/bsweb/servlet/bsweb.

2. Le taux d'intérêt nominal à long terme

Le taux nominal à long terme retenu correspond au taux moyen des emprunts d'état (TME) publié par la Caisse des Dépôt et des Consignations sur le site de la Banque de France⁸. Cet historique est régulièrement mis à jour et commence au mois de janvier 1970.

L'historique des taux d'intérêt réels à long terme est construit à partir des valeurs historiques des taux d'intérêt nominaux à long terme auxquels on retranche les valeurs historiques du taux d'inflation.

Donc le taux réel à long terme au mois m s'écrit :

$${}_m l_{réel} = {}_m l_{nominal} - q_m$$

Le graphe ci-dessus illustre l'évolution du taux d'intérêt nominal à long terme et le tableau 1 (page 31) présente ces statistiques descriptives classiques (entre janvier 1997 et juin 2009).

3. Le taux d'intérêt nominal à court terme

Le taux nominal à court terme retenu correspond au taux moyen du marché monétaire (TMM ou T4M) publié par la Caisse des Dépôt et des Consignations sur le site de la Banque de France⁹. Cet historique est régulièrement mis à jour et commence au mois de janvier 1952.

L'historique des taux d'intérêt réels à court terme est construit à partir des valeurs historiques des taux d'intérêt nominaux à court terme auxquels on retranche les valeurs historiques du taux d'inflation. Donc le taux réel au mois m s'écrit :

$${}_m r_{réel} = {}_m r_{nominal} - q_m$$

Le graphe ci-dessus illustre l'évolution du taux d'intérêt nominal à court terme et le tableau 1 présente les statistiques descriptives classiques du taux d'intérêt nominal à court terme (entre janvier 1997 et juin 2009).

4. Le taux de rendement des actions

Afin de construire l'historique des rendements des actions, l'indice SBF 250 avec dividendes réinvestis (*total return*) a été retenu. Le choix de cet indice répond à deux contraintes :

Une contrainte de cohérence : un placement en action possède deux sources de rendement : le gain en capital (plus value) et les dividendes. Ainsi, afin que le rendement des actions soit comparable et cohérent avec le rendement de l'immobilier, à l'inflation et aux taux d'intérêt, l'indice que nous avons retenu suppose que les dividendes sont réinvestis.

Une contrainte de diversification : les assureurs possèdent des portefeuilles en actions largement diversifiés. L'indice SBF 250, regroupant les 250 premières capitalisations de l'indice CAC All Shares (CAC All Shares regroupe toutes les sociétés cotées en actions à la Bourse de Paris), semble le plus adapté lors du calcul du rendement d'un portefeuille largement diversifié sur le marché boursier français.

L'historique de l'indice SBF 250 avec dividendes réinvestis est en téléchargement libre sur le site d'Euronext¹⁰ et est régulièrement mis à jour depuis le mois de décembre 1990.

⁸ www.banque-france.fr/fr/statistiques/base/statmon/html/tmf_mens_france_fr_tauxmarfinan.htm.

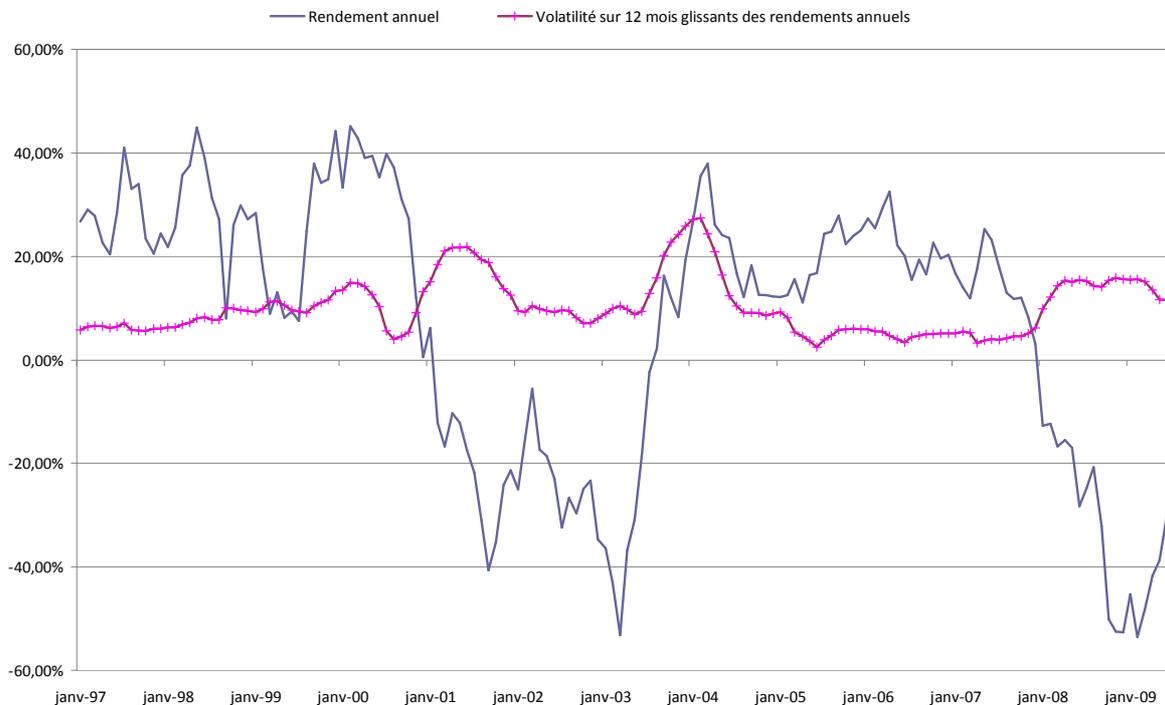
⁹ www.banque-france.fr/fr/statistiques/base/statmon/html/tmf_mens_zeuro_fr_tauxmarmonet.htm.

Le rendement annuel au mois m est calculé avec la formule suivante :

$$s_m = \ln\left(\frac{ind_m}{ind_{m-12\text{mois}}}\right)$$

où ind_m représente l'indice du SBF 250 avec dividendes réinvestis au mois m . La figure suivante illustre l'évolution du rendement annuel et de la volatilité sur 12 mois glissants du rendement annuel entre janvier 1997 et juin 2009. On observe en particulier les chutes brutales du rendement lors des deux dernières crises 2003 et 2008. Il est à noter que la volatilité est calculée chaque mois comme la volatilité des rendements annuels des 12 mois précédant le mois du calcul. Par exemple, la volatilité annuelle au 01 janvier 2009 est la volatilité des rendements annuels au 01 de chaque mois, de février 2008 jusqu'au janvier 2009.

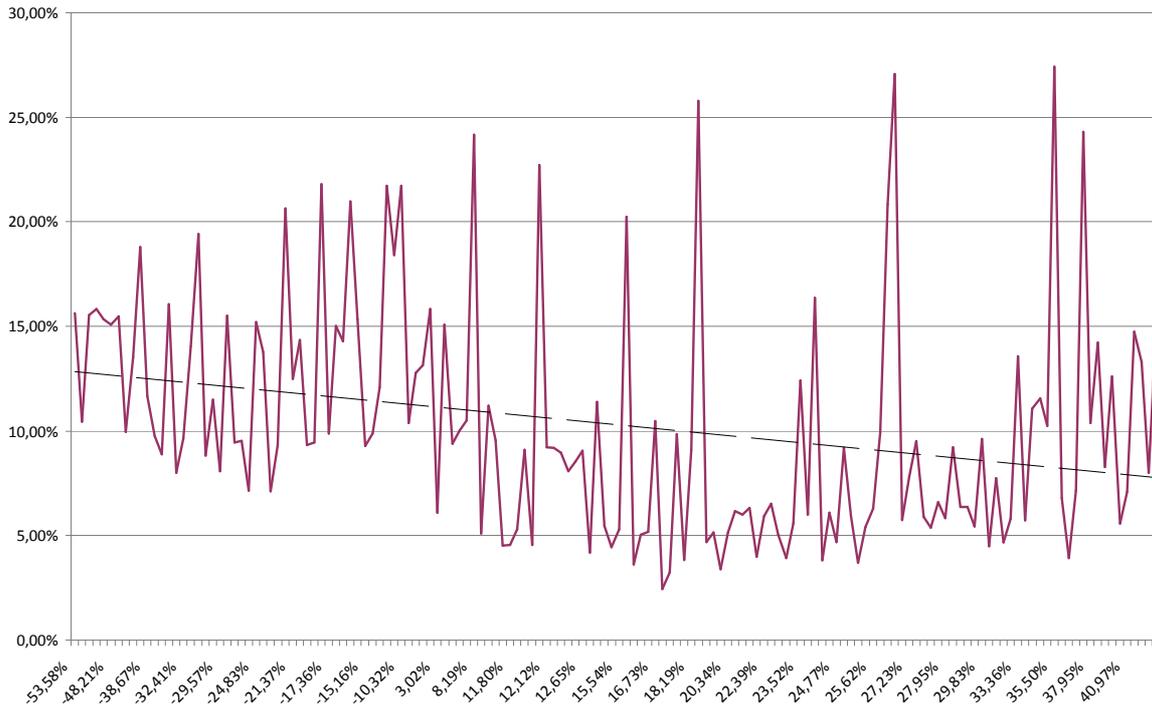
Figure 3 : Evolution du rendement annuel et de la volatilité du SBF 250 dividendes réinvestis (janvier 1997-juin 2009)



On observe que la volatilité est plus forte quand le rendement annuel atteint des niveaux bas. Cela s'illustre plus clairement dans le graphe suivant (une droite en tiret donnant la tendance linéaire est ajoutée) où l'on observe que la volatilité a une allure décroissante en fonction du rendement et donc les périodes de fortes volatilités peuvent être associées aux périodes de faibles rendements.

¹⁰ www.euronext.com/editorial/wide/editorial-7350-FR.html.

Figure 4 : Evolution de la volatilité (ordonnée) en fonction du rendement (abscisse)



5. Le taux de rendement de l'immobilier

L'historique du rendement de l'immobilier est construit à partir de l'indice des prix des logements en France (IPL) disponible sur le site de l'INSEE¹¹. Cet indice est à fréquence trimestrielle et est régulièrement mis à jour depuis le premier trimestre de l'année 1996. L'IPL est utilisé pour calculer le rendement annuel de l'immobilier à fréquence trimestrielle (rendement sur 4 trimestres glissants) entre le premier trimestre de 1997 et le second trimestre de 2009. Plus précisément, le rendement annuel de l'immobilier au trimestre m s'écrit :

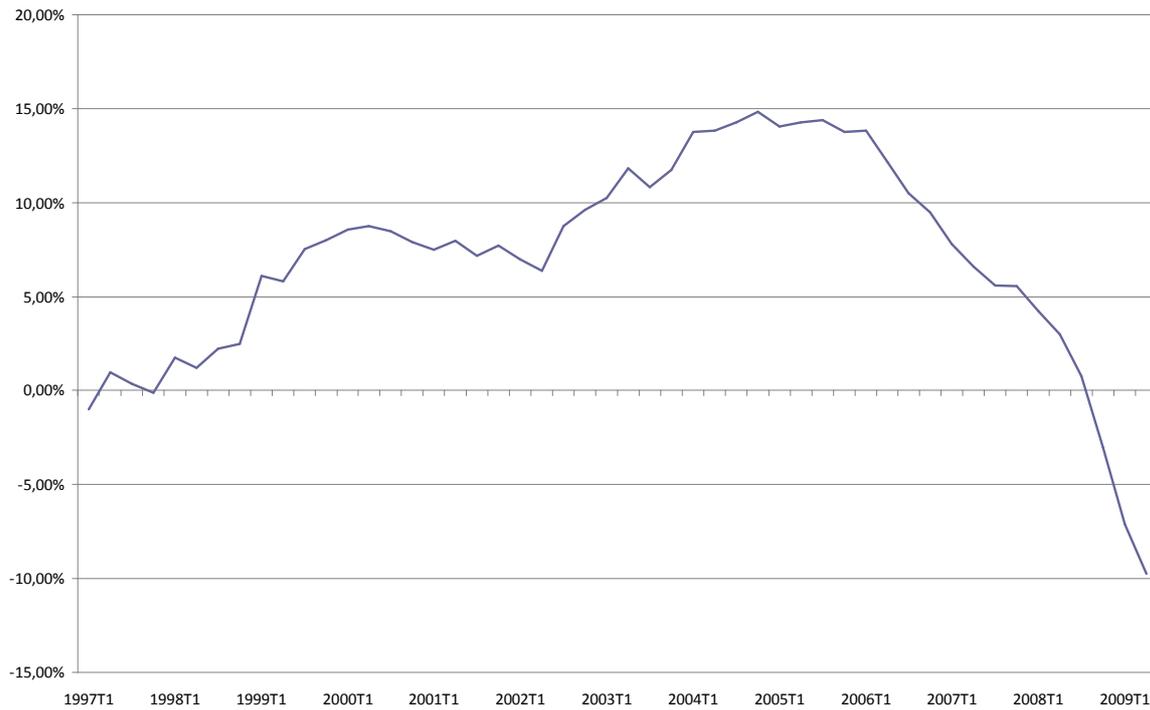
$$Im_m = \ln\left(\frac{IPL_m}{IPL_{m-4\text{trimestre}}}\right)$$

où IPL_m est l'indice des prix des logements au trimestre m .

La figure suivante illustre l'évolution du rendement annuel de l'immobilier entre janvier 1997 et juin 2009. On observe en particulier la chute brutale en 2009 à cause de la crise financière.

¹¹ www.indices.insee.fr/bsweb/servlet/bsweb?action=BS_RECHGUIDEE&BS_IDARBO=05000000000000.

Figure 5 : Evolution du rendement de l'immobilier (janvier 1997-juin 2009)



6. Statistiques descriptives

Les deux tableaux suivants présentent les statistiques descriptives de base et la matrice de corrélation de l'ensemble des séries objet de l'étude.

Tableau 1 : Statistiques descriptives classiques

	Taux d'inflation	Taux long nominal	Taux court nominal	Action	Immobilier
Moyenne	1,56 %	4,53 %	3,10 %	6,70 %	6,97 %
Volatilité	0,74 %	0,67 %	0,91 %	26,36 %	5,61 %
Médiane	1,62 %	4,41 %	3,19 %	15,53 %	7,76 %
Maximum	3,55 %	5,87 %	5,14 %	45,17 %	14,85 %
Minimum	-0,49 %	3,17 %	0,68 %	-53,58 %	-9,75 %
Skewness	-0,03	0,12	-0,11	-0,70	-0,80
Kurtosis	3,16	2,01	2,71	2,32	3,53
Test Jarque-Bera	0	1	0	1	0

Tableau 2 : Matrice de corrélation sur l'ensemble de l'historique

	Taux d'inflation	Taux long nominal	Taux court nominal	Actions	Immobilier
Taux d'inflation	100,00 %	-76,88 %	-50,80 %	-17,32 %	48,73 %
Taux long nominal	-76,88 %	100,00 %	72,59 %	25,08 %	-54,86 %
Taux court nominal	-50,80 %	72,59 %	100,00 %	11,20 %	-42,45 %
Actions	-17,32 %	25,08 %	11,20 %	100,00 %	30,27 %
Immobilier	48,73 %	-54,86 %	-42,45 %	30,27 %	100,00 %

On observe que les séries présentées dans cette étude présentent un coefficient d'asymétrie négatif (ce coefficient est égal à 0 pour une distribution normal) et certaines séries (le taux d'intérêt nominal à long terme et le rendement des actions) présentent des coefficients d'aplatissement significativement différents de 3 (valeur du coefficient pour une distribution normale). Aussi, mis à part le taux d'intérêt nominal à long terme et le rendement des actions, les autres séries répondent positivement au test de normalité de Jarque-Bera.

De la même façon que la différence des coefficients d'asymétrie et d'aplatissement révèle des différences distributionnelles importantes entre les indices, les différences observées au niveau des coefficients de corrélation suggèrent une grande diversité des structures de dépendance des séries bi-variées.

La mesure de la dépendance par la corrélation repose fondamentalement sur l'hypothèse de normalité multi-variée (et donc sur l'hypothèse que les lois marginales sont normales). Or nous observons que deux des cinq séries étudiées ne répondent pas au test de normalité de Jarque-Bera. Cela signifie que la distribution multi-variée des cinq indices objet d'étude n'est pas normale. Ainsi, se contenter de décrire la structure de dépendance par la simple mesure des coefficients de corrélation s'avère insuffisant et peut être une source d'erreur. Les chapitres de la partie II tentent de proposer une méthode pour décrire de manière plus réaliste la structure de dépendance d'un ensemble de variables en partant de leurs historiques. La structure de dépendance du taux d'inflation, du taux d'intérêt nominal à long terme, du taux d'intérêt nominal à court terme, du rendement des actions et du rendement de l'immobilier sera examinée au chapitre 4 de la partie II.

Partie II

La structure de dépendance des indices financiers et macroéconomiques

Nous essayons à travers cette partie de proposer une démarche purement empirique permettant d'extraire la structure de dépendance d'un ensemble de variables (financières, économiques...) en partant de leurs historiques. Cette démarche consiste à choisir la copule paramétrique la plus adaptée parmi la copule gaussienne, la copule de Student, la copule de Frank, la copule de Gumbel et la copule de Cook-Johnson. Ce choix se base essentiellement sur le calcul de la distance entre ces copules, après les avoir calibrées, et la copule empirique.

Pour ce faire, il est nécessaire de disposer de la densité de probabilité de chaque variable étudiée ainsi que la densité de probabilité jointe. Nous envisageons de déterminer ces densités empiriquement par un estimateur à noyau gaussien (cf. chapitre II-2). Ces densités empiriques permettent de déterminer la copule empirique.

Le premier chapitre comprend un rappel sur les distributions multi-variées et les copules et présente des méthodes de calibrage concernant notamment les copules retenues dans le cadre de notre étude. Le second chapitre présente une démarche de construction des densités de probabilité marginales et la densité jointe d'un ensemble de variables aléatoires en partant de leurs historiques. Cette construction se base sur les travaux de Silverman [1986]. Enfin, le troisième chapitre présente la démarche à suivre pour trouver la copule paramétrique optimale parmi les cinq copules étudiées. Le quatrième chapitre illustre une application de la démarche sur les indices financiers et macroéconomiques présentés au chapitre I-3.

Chapitre II-1

Représentation des lois multi-variées par les copules

La copule est devenue un outil de base dans la modélisation des distributions multi-variées. La connaissance de cet outil probabiliste est essentielle à l'appréhension de nombreux domaines d'application de la finance quantitative et de l'assurance : mesure du risque multiple de crédit, évaluation de produits de crédits structurés, réplique de la performance des *hedge funds*, mesure du risque multiple de marché, gestion de portefeuille à l'aide de simulations Monte Carlo, *pricing* d'options à plusieurs sous-jacents, etc.

Une mesure de dépendance régulièrement utilisée est la corrélation linéaire. Cet indicateur est performant lorsque la relation de dépendance est linéaire et l'univers considéré gaussien, cadre d'analyse rare en général. Pour remédier à cela, nous avons recours à d'autres indicateurs de dépendance comme le tau de Kendall ou encore le rho de Spearman (ces deux mesures sont définies dans la suite). En outre, contrairement au coefficient de corrélation, un intérêt pratique à l'utilisation des copules consiste en la possibilité de modéliser la dépendance d'événements extrêmes.

Dans la suite de ce chapitre, nous étudierons en premier lieu les définitions et propriétés fondamentales des copules en explicitant quelques exemples de copules paramétriques, pour ensuite, analyser leur mise en œuvre pratique via des méthodes de simulation et d'inférence statistique. Une attention particulière est accordée aux cinq copules qui nous intéressent dans le cadre de notre étude : la copule gaussienne, la copule de Student, la copule de Frank, la copule de Gumbel et la copule de Cook-Johnson.

1. Rappel sur les fonctions de répartition

La bibliographie étant abondante, nous nous limitons dans cette partie à définir et lister quelques propriétés de la fonction de répartition sans entrer dans le détail des démonstrations. Pour plus de détails le lecteur pourra se référer à Jacod [2001-2002] ou à ENPC [2008]. On suppose dans la suite que l'on dispose d'un espace probabilisé (Ω, A, P) sur \mathfrak{R} .

1.1. Fonction de répartition uni-variée

1.1.1. Définitions

Soit Y une variable aléatoire réelle. On appelle fonction de répartition de Y la fonction $F : \mathfrak{R} \rightarrow [0,1]$, parfois notée F_Y , définie par $F(y) = P(Y \leq y)$ pour tout $y \in \mathfrak{R}$. F est càd-làg.

L'inverse généralisé de F , noté F^{-1} est la fonction croissante continue à gauche, définie sur $]0,1]$ par :

$$F^{-1}(\alpha) = \inf\{y \in \mathfrak{R}, F(y) \geq \alpha\}$$

Si la fonction F est bijective de \mathfrak{R} dans $]0,1[$, l'inverse généralisé et l'inverse coïncident.

L'inverse généralisé de la fonction de répartition permet d'introduire la notion de quantile très utilisé en statistique. En effet, pour $\alpha \in]0,1]$, la quantité $F^{-1}(\alpha)$ s'appelle le quantile ou fractile d'ordre α de la loi de Y .

1.1.2. Propriétés

La fonction de répartition F est croissante, continue à droite et admet les limites 0 en $-\infty$ et 1 en $+\infty$. Elle vérifie en outre les propriétés suivantes :

a. La fonction F^{-1} est croissante et continue à gauche et on a l'équivalence suivante pour tout $y \in \mathfrak{R}$ et tout $\alpha \in]0,1]$:

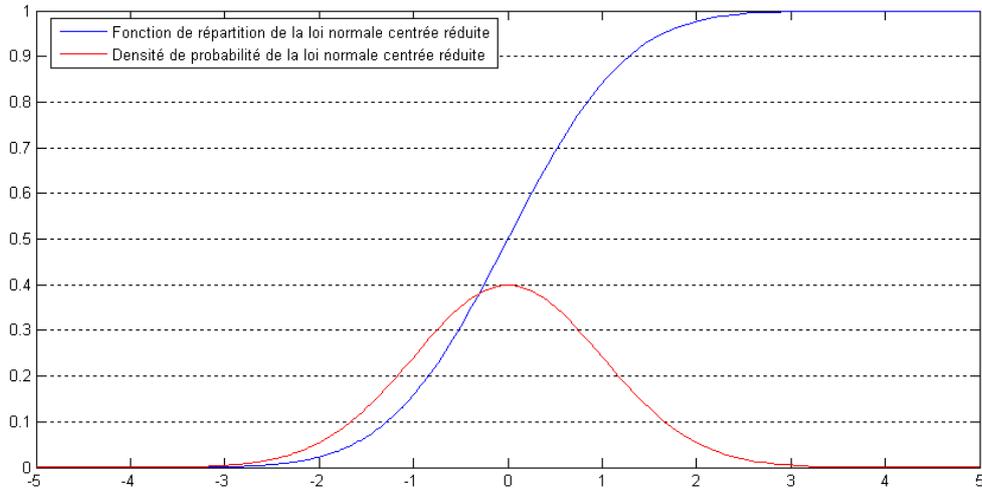
$$F(y) \geq \alpha \Leftrightarrow y \geq F^{-1}(\alpha)$$

b. Pour tout $\alpha \in]0,1]$, on a $F(F^{-1}(\alpha)) \geq \alpha$ avec égalité si $F^{-1}(\alpha) > -\infty$ et si F est continue en $F^{-1}(\alpha)$.

c. Soit U une variable aléatoire de loi uniforme sur $[0,1]$. La fonction F est la fonction de répartition de la variable aléatoire $F^{-1}(U)$.

d. Soit Y une variable aléatoire à valeurs dans \mathfrak{R} de fonction de répartition F . Si F est continue, alors $F(Y)$ est de loi uniforme sur $[0,1]$.

Figure 6: Fonction de répartition et densité de probabilité d'une loi normale centrée réduite



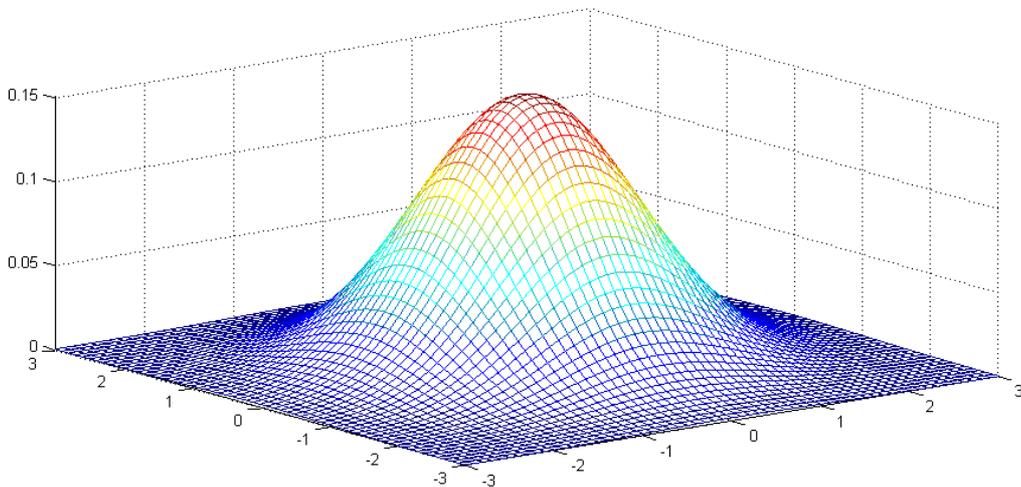
1.2. La fonction de répartition multi-variée

On peut caractériser la loi d'une variable aléatoire vectorielle à l'aide de sa fonction de répartition multidimensionnelle. Soit $d \geq 1$, $a = (a_1, \dots, a_d)$ et $b = (b_1, \dots, b_d)$ dans \mathfrak{R}^d , on note dans la suite $a \leq b$ si $a_i \leq b_i$ pour tout $1 \leq i \leq d$.

Soit $d \geq 1$ et X une variable aléatoire à valeurs dans \mathfrak{R}^d . On appelle fonction de répartition de X la fonction $F : \mathfrak{R}^d \rightarrow [0,1]$, parfois notée F_X définie par $F(x) = P(X \leq x)$ pour tout $x \in \mathfrak{R}^d$. En plus si elle admet une densité alors celle-ci s'écrit :

$$f = \frac{\partial^d F}{\partial x_1 \dots \partial x_d}$$

Figure 7 : Densité de probabilité jointe de deux variables aléatoires normales centrées réduites indépendantes



Par ailleurs, la connaissance de la fonction de répartition jointe permet de déterminer les lois marginales des variables aléatoires aussi bien que leur structure de dépendance. Cependant, il est rare de disposer de la fonction de répartition jointe dans les modèles mathématiques où seules les marginales sont décrites. La fonction copule a, dans ce contexte, un réel intérêt car elle permet, en partant des lois marginales, de déterminer la fonction de répartition jointe tout en ayant un contrôle sur la structure de dépendance.

2. Copules : définition et propriétés

Les propriétés des copules sont nombreuses et variées. L'objet de cette section n'est pas de toutes les recenser mais de présenter les définitions et les propriétés utiles pour la suite. Pour plus de détail et d'approfondissement, le lecteur pourra se référer au livre de Nelson [1999] ou à celui de Joe [1997].

2.1. Définitions

Une fonction $C : [0,1]^d \rightarrow [0,1]$ est appelée copule de dimension d , ou simplement copule, si C est (la restriction à $[0,1]^d$ de) la fonction de répartition d'une variable aléatoire $U = (U_1, \dots, U_d)$ à valeurs dans \mathfrak{R}^d , où les variables aléatoires U_1, \dots, U_d sont de lois uniformes sur $[0,1]$: $C(u) = P(U \leq u)$ pour tout $u \in [0,1]^d$.

Soit $X = (X_1, \dots, X_d)$ une variable aléatoire à valeurs dans \mathfrak{R}^d de fonction de répartition F . Pour $1 \leq k \leq d$, soit F_k la fonction de répartition de X_k . Une copule pour X est une copule vérifiant $F(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d))$ pour tout $x = (x_1, \dots, x_d) \in \mathfrak{R}^d$.

L'exemple le plus simple est la copule $C(u_1, \dots, u_d) = \prod_{k=1}^d u_k$ qui est la fonction de répartition des variables aléatoires indépendantes U_1, \dots, U_d de lois uniformes sur $[0,1]$.

Le théorème suivant (Sklar [1959]), précise le lien défini par la copule C , entre les fonctions de répartition marginales uni-variées F_1, \dots, F_d et la distribution complète multi-variée F . On peut trouver la démonstration de ce théorème dans Nelson [1999], Sklar [1959] ou ENPC [2008].

2.2. Théorème de Sklar

- a. Soit C une copule et F_1, \dots, F_d des fonctions de répartition de variables aléatoires à valeurs dans \mathfrak{R} . La fonction F définie par $F(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d))$ est la fonction de répartition d'une variable aléatoire à valeurs dans \mathfrak{R}^d .
- b. Soit $X = (X_1, \dots, X_d)$ une variable aléatoire à valeurs dans \mathfrak{R}^d . Il existe une copule pour X . Si en plus les fonctions de répartition des variables aléatoires X_1, \dots, X_d sont continues, alors la copule est unique.

Nous remarquons donc qu'une copule permet d'exprimer une fonction de répartition multi-variée en fonction des fonctions de distributions marginales et que celle-ci contient toute l'information sur la structure de dépendance.

Par définition, une copule est une fonction de répartition d'un vecteur de variables aléatoires uniformes sur $[0,1]$. Donc sa densité, lorsqu'elle existe, correspond à la différentiation de cette copule (cf. la section 1.2 de ce chapitre). L'intérêt de travailler avec la densité de la copule est de pouvoir séparer la structure de dépendance des densités marginales comme le montre la section suivante.

2.3. La densité d'une copule

Si nous supposons que les distributions marginales F_1, \dots, F_d et C sont différentiables, alors la densité jointe de la variable aléatoire $X = (X_1, \dots, X_d)$ prend la forme suivante :

$$f(x_1, \dots, x_d) = f_1(x_1) \dots f_d(x_d) c(F_1(x_1), \dots, F_d(x_d))$$

où, pour $1 \leq k \leq d$, f_k est la densité de probabilité dérivée de F_k , f est la densité jointe issue de F et c est la densité de la copule C définie par :

$$c = \frac{\partial^d C}{\partial u_1 \dots \partial u_d}$$

Nous constatons ainsi que l'on peut séparer la densité jointe en deux blocs. Le premier : $c(f_1(x_1), \dots, f_d(x_d))$, contient l'information sur la structure de dépendance des variables aléatoires X_1, \dots, X_d . Le second est le produit des densités marginales. Cela montre que les copules représentent un moyen d'extraire la structure de dépendance de la distribution jointe et de la séparer des comportements marginaux.

Il est possible donc de caractériser la densité d'une copule entièrement par la densité jointe et les densités marginales. En effet, si on suppose que F est une fonction de répartition admettant la densité continue f , alors les lois marginales F_1, \dots, F_d sont continues et admettent des densités notées respectivement f_1, \dots, f_d et la copule de F admet la densité c définie par :

$$c(u_1, \dots, u_d) = \frac{f(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d))}{f_1(F_1^{-1}(u_1)) \dots f_d(F_d^{-1}(u_d))}$$

Cette écriture va permettre d'estimer la densité d'une copule à partir de l'estimation des lois marginales et de la loi jointe.

Comme nous l'avons souligné au chapitre I-2, la corrélation est une mesure adaptée aux distributions elliptiques. En dehors de cet univers, nous avons recours à d'autres indicateurs de dépendance comme le tau de Kendall ou encore le rho de Spearman. L'idée est de généraliser la notion de corrélation et de synthétiser, dans un indicateur numérique, l'intensité du lien entre deux variables aléatoires. La section suivante présente ces deux indicateurs.

2.4. Tau de Kendall et rho de Spearman

Le tau de Kendall et le rho de Spearman sont deux mesures de concordance bien connues en statistiques. Elles donnent une mesure de la corrélation entre les rangs des observations, à la différence du coefficient de corrélation linéaire qui lui apprécie la corrélation entre les valeurs des observations. Elles offrent par ailleurs l'avantage de s'exprimer simplement en fonction de la copule associée au couple de variables aléatoires. Ces notions figurent dans plusieurs travaux comme Cadoux et al. [2006], Embrechts et al. [2002], Demarta et McNeil [2004] et Lindskog et al. [2003].

2.4.1. Tau de Kendall

Le tau de Kendall joue le même rôle pour les fonctions de dépendance que le coefficient de corrélation linéaire pour la distribution gaussienne. L'estimation de ce paramètre permet de calibrer certaines copules notamment la copule de Student et la copule gaussienne.

Soient (X_1, X_2) un couple de vecteurs aléatoires et (X_1', X_2') un couple de vecteurs identique à (X_1, X_2) (une copie en tout point identique à (X_1, X_2)). Le tau de Kendall s'écrit alors :

$$\tau(X_1, X_2) = P\{(X_1 - X_1')(X_2 - X_2') > 0\} - P\{(X_1 - X_1')(X_2 - X_2') < 0\}$$

Le tau de Kendall n'est autre que la différence entre la probabilité de concordance et celle de discordance. Si C est la copule associée au couple (X_1, X_2) l'expression du tau de Kendall devient :

$$\tau(X_1, X_2) = 4 \iint_{[0,1]^2} C(u_1, u_2) dC(u_1, u_2) - 1 = 1 - 4 \iint_{[0,1]^2} \partial_1 C(u_1, u_2) \partial_2 C(u_1, u_2) du_1 du_2$$

Ce qui peut s'écrire encore (U_1, U_2 sont de loi uniforme sur $[0,1]$) :

$$\tau(X_1, X_2) = 4E(C(U_1, U_2)) - 1$$

Si on dispose d'un échantillon d'observations de taille T de (X_1, X_2) , $(x_1^i, x_2^i)_{1 \leq i \leq T}$, on peut construire un estimateur empirique du tau de Kendall. Celui-ci est donné par la formule suivante :

$$\hat{\tau}(X_1, X_2) = \frac{2}{T(T-1)} \sum_{j=2}^T \sum_{i=1}^{j-1} \text{sign}\{(x_1^j - x_1^i)(x_2^j - x_2^i)\}$$

La fonction $\text{sign}(z)$ est égale à 1 si z est positif et à -1 si z est strictement négatif.

2.4.2. Rho de Spearman

Le rho de Spearman est la corrélation entre les fonctions de répartition marginales des deux variables aléatoires du couple (X_1, X_2) . On peut écrire le rho de Spearman en fonction du coefficient ρ de la corrélation linéaire de Pearson comme suit :

$$\rho_s(X_1, X_2) = \rho(F_1(X_1), F_2(X_2))$$

où F_1 et F_2 sont les fonctions de répartition respectives de X_1 et X_2 . Si C est la copule associée au couple (X_1, X_2) l'expression du rho de Spearman devient :

$$\rho_s(X_1, X_2) = 12 \iint_{[0,1]^2} u_1 u_2 dC(u_1, u_2) - 3 = 12 \iint_{[0,1]^2} C(u_1, u_2) du_1 du_2 - 3$$

Si on dispose d'un échantillon d'observation de taille T de (X_1, X_2) , $(x_1^i, x_2^i)_{1 \leq i \leq T}$, on peut construire un estimateur empirique du rho de Spearman. Celui-ci est donné par la formule suivante :

$$\hat{\rho}(X_1, X_2) = \frac{\sum_{i=1}^T (a_i - \bar{a})(b_i - \bar{b})}{\sqrt{\sum_{i=1}^T (a_i - \bar{a})^2} \sqrt{\sum_{i=1}^T (b_i - \bar{b})^2}}$$

où, pour tout $1 \leq i \leq T$, a_i est le rang de x_1^i , b_i celui de x_2^i et on note : $\bar{z} = \frac{1}{T} \sum_{i=1}^T z_i$.

Comme le coefficient de corrélation, le tau de Kendall et le rho de Spearman sont des mesures de dépendance globales. Il est toutefois intéressant d'examiner la dépendance sur les queues de distributions. Ce point fait l'objet de la section suivante.

2.5. Dépendance sur les queues de distributions

Le concept de dépendance de queue fournit une description de la dépendance au niveau des queues de distribution, très intéressante pour étudier la survenance simultanée de valeurs extrêmes. C'est une mesure locale contrairement au tau de Kendall et au rho de Spearman qui mesurent la dépendance sur l'ensemble de la distribution. On peut retrouver les définitions et les propriétés qui suivent dans Cadoux et al. [2006] ou dans le livre Kharoubi-Rakotomalala [2008].

Le coefficient de dépendance de queue inférieure ou *lower tail dependence coefficient* de deux variables aléatoires X_1 et X_2 , de fonctions de répartition respectives F_1 et F_2 , est défini par (lorsque la limite existe) :

$$\lambda_L(X_1, X_2) = \lim_{\alpha \rightarrow 0^+} P(X_1 \leq F_1^{-1}(\alpha) / X_2 \leq F_2^{-1}(\alpha))$$

Le coefficient de dépendance de queue supérieure ou *upper tail dependence coefficient* de deux variables aléatoires X_1 et X_2 , de fonctions de répartition respectives F_1 et F_2 , est défini par (lorsque la limite existe) :

$$\lambda_U(X_1, X_2) = \lim_{\alpha \rightarrow 1^-} P(X_1 > F_1^{-1}(\alpha) / X_2 > F_2^{-1}(\alpha))$$

Soit C la copule associé à X_1 et X_2 . Alors la dépendance de queue inférieure et la dépendance de queue supérieure s'écrivent :

$$\lambda_L(X_1, X_2) = \lim_{u \rightarrow 0^+} \frac{C(u, u)}{u}$$

$$\lambda_U(X_1, X_2) = \lim_{u \rightarrow 1^-} \frac{1 - 2u + C(u, u)}{1 - u}$$

On dit que les variables aléatoires X_1 et X_2 sont asymptotiquement dépendantes au niveau supérieur de la queue de distribution (ou dépendantes sur les queues à droite) si $\lambda_U \in]0,1]$ et asymptotiquement indépendantes au niveau supérieur de la queue de distribution si $\lambda_U = 0$. Lorsque $\lambda_U > 0$, cela signifie que les valeurs extrêmes positives sont corrélées.

De manière symétrique, les variables X_1 et X_2 sont dépendantes au niveau inférieur de la queue de distribution (ou dépendantes sur les queues à gauche) si $\lambda_L \in]0,1]$ et elles sont indépendantes si $\lambda_L = 0$. Cela signifie que lorsque $\lambda_L > 0$, alors les valeurs extrêmes négatives sont corrélées.

Il existe des copules ne présentant pas de dépendance de queues à l'image de la copule de Franck et de la copule gaussienne présentées par la suite. Ainsi, pour la copule gaussienne, même avec une forte corrélation les rendements extrêmes sont asymptotiquement indépendants. Cela indique que la corrélation linéaire est mal adaptée aux situations extrêmes, notamment les périodes de crise. La section suivante présente quelques exemples de deux familles de copules : les copules elliptiques et les copules archimédiennes.

3. Exemples de copules paramétriques

L'objectif de notre étude est de trouver une copule paramétrique adaptée à la structure de dépendance d'un ensemble d'indices financiers et macroéconomiques. La recherche de la copule optimale est réalisée sur un ensemble de cinq copules sélectionnées *a priori* et présentées dans cette section. La flexibilité, la simplicité analytique et la diversité des formes de dépendance sont, entre autres, les critères du choix de ces copules. Cet ensemble présélectionné comprend deux familles de copules. On retrouve ainsi deux copules de la famille des copules elliptiques et trois copules de la famille des copules archimédiennes.

3.1. Les copules elliptiques

Soit $M_d(\mathfrak{R})$ l'ensemble des matrices carrées réelles de taille d^2 . Une loi continue est dite elliptique de paramètre de position $\mu = (\mu_1, \dots, \mu_d) \in \mathfrak{R}^d$ et de matrice de forme symétrique définie positive $\Sigma \in M_d(\mathfrak{R})$ si sa densité f peut s'écrire pour tout $x = (x_1, \dots, x_d) \in \mathfrak{R}^d$:

$$f(x) = (\det \Sigma)^{-\frac{1}{2}} g((x - \mu)\Sigma^{-1}(x - \mu)')$$

où l'on désigne par z' la transposée de z et g est une fonction à valeurs positives vérifiant $\int_{\mathfrak{R}^d} g(xx') dx = 1$.

On note $\xi(\mu, \Sigma, g)$ cette famille de lois (dites elliptiques car les courbes de niveaux de la densité sont des ellipses). La loi est dite sphérique si $\Sigma = kI_d$ où $k > 0$ et I_d est la matrice unité de $M_d(\mathfrak{R})$.

Les lois elliptiques associées à la même fonction g font partie de la même famille elliptique, dans laquelle on distingue le représentant standard (centré réduit) pour lequel $\mu = 0$ et $\Sigma = I_d$. On appellera par la suite la matrice Σ : matrice de corrélation.

La loi d'un vecteur gaussien est un exemple classique de lois elliptiques, associé au choix de $g(y) = (2\pi)^{-d/2} \exp(-y/2)$. Ces lois vérifient, comme les vecteurs gaussiens, des propriétés algébriques intéressantes. La propriété la plus importante réside dans le fait que les lois elliptiques forment une classe stable par transformation affine. Pour une présentation détaillée de ces lois se référer à Fang et al. [1990].

Les copules elliptiques sont définies à partir des familles des lois elliptiques. En effet, une copule est dite elliptique si elle est la copule d'une loi elliptique. Ci-après deux exemples classiques de copules elliptiques. Il s'agit de la copule gaussienne et la copule de Student.

3.1.1. La copule gaussienne

La copule gaussienne ne présente pas de dépendance de queue et n'est donc pas adaptée à des valeurs extrêmes. L'importance de cette copule réside dans le fait qu'elle est sous-jacente à la distribution normale multi-variée. En effet, modéliser la structure de dépendance d'un échantillon par une copule gaussienne est cohérent avec la mesure de cette dépendance par le coefficient de corrélation linéaire. Cependant lorsque l'on souhaite modéliser une dépendance non linéaire ou entre événements extrêmes, on va faire appel à d'autres copules dont des exemples sont explicités dans la suite. La fonction de distribution de la copule gaussienne d -dimensionnelle, s'écrit pour tout $(u_1, \dots, u_d) \in [0,1]^d$:

$$C(u_1, \dots, u_d) = \Phi_{\Sigma}(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_d))$$

La fonction de distribution Φ^{-1} est l'inverse généralisé de la distribution normale centrée réduite uni-variée. La fonction Φ_{Σ} est la fonction de répartition de la loi normale centrée réduite et Σ sa matrice de variance covariance (égale à la matrice de corrélation dans ce cas). A titre de rappel la densité de Φ_{Σ} s'écrit pour tout $x = (x_1, \dots, x_d) \in \mathfrak{R}^d$:

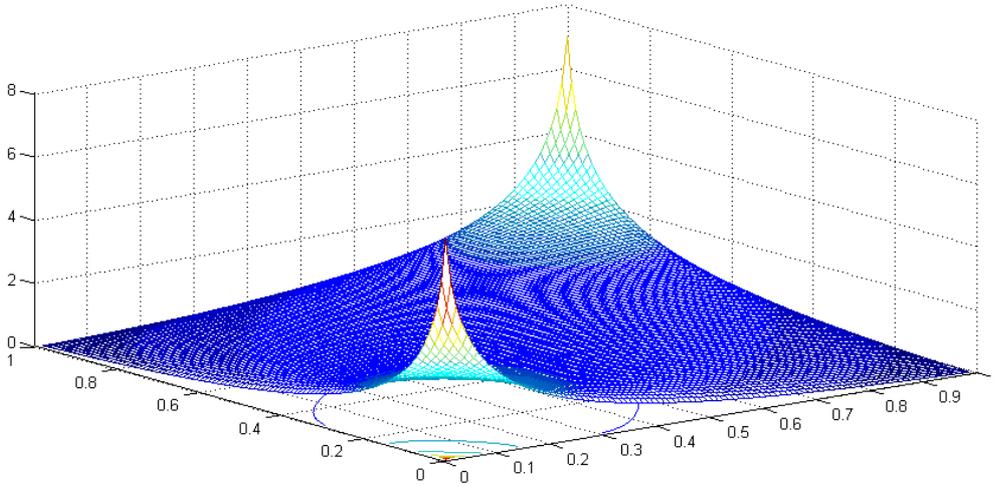
$$\varphi_{\Sigma}(x) = \frac{\exp(-\frac{1}{2} x \Sigma^{-1} x')}{(2\pi)^{d/2} \det(\Sigma)^{1/2}}$$

En dérivant la formule définissant la copule gaussienne on peut facilement extraire la densité de la copule gaussienne d-variée qui s'écrit :

$$c(u_1, \dots, u_d) = \frac{1}{\det(\Sigma)^{1/2}} \exp(-\frac{1}{2} \beta(\Sigma^{-1} - I_d) \beta')$$

où I_d est la matrice unité de $M_d(\mathfrak{R})$ et $\beta = (\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_d))$. La figure suivante illustre la densité de la copule gaussienne bi-variée pour une corrélation de $\Sigma = 0.5$.

Figure 8 : Densité de la copule gaussienne pour une corrélation de 0.5



3.1.2. La copule de Student

La copule de Student (*t copula*) est la copule sous-jacente à une distribution multi-variée de Student. Cette structure de dépendance a reçu dernièrement une attention particulière notamment en finance quantitative. Un bon nombre de documents (Marshall et Zeevi [2002], Breymann et al. [2003]) ont montré que la copule de Student reflète mieux la structure de dépendance d'un ensemble d'indices financiers que la copule gaussienne. L'une des raisons de ce constat est la capacité de la copule de Student à capter les dépendances extrêmes qui est un phénomène souvent observé sur les marchés financiers.

La copule de Student est construite de la même manière que la copule gaussienne mais cette fois-ci à partir de la distribution de Student centrée réduite. La fonction de densité de la copule de Student d-variée, s'écrit pour tout $(u_1, \dots, u_d) \in [0,1]^d$:

$$c(u_1, \dots, u_d) = \frac{f_{v, \Sigma}(t_v^{-1}(u_1), \dots, t_v^{-1}(u_d))}{\prod_{i=1}^d f_v(t_v^{-1}(u_i))}$$

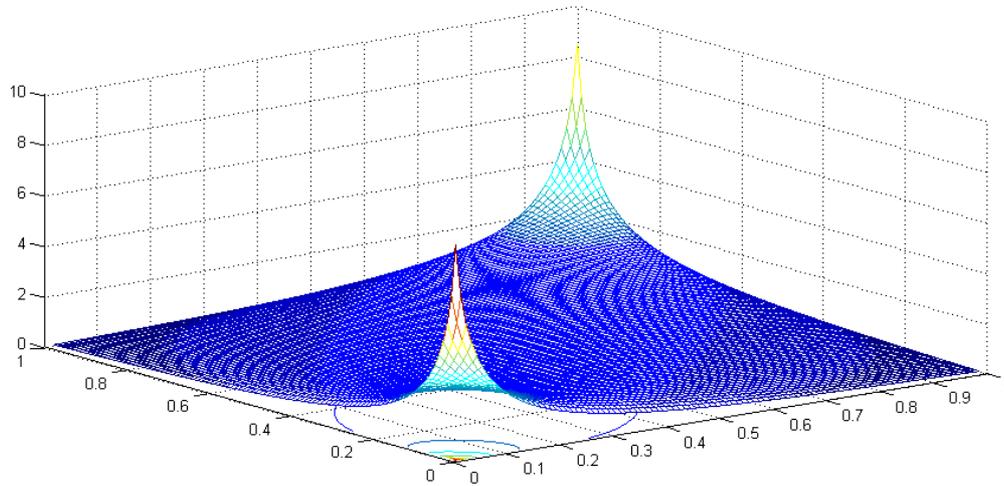
La fonction de distribution t_v^{-1} est l'inverse généralisé de la distribution de Student centrée réduite uni-variée à ν degrés de liberté. La fonction $f_{v, \Sigma}$ est la densité de probabilité de la loi de Student centrée réduite, Σ sa matrice de corrélation et f_v est la densité uni-variée de la loi de Student centrée réduite ($\Sigma = 1$). A titre de rappel sa densité s'écrit pour tout $x = (x_1, \dots, x_d) \in \mathfrak{R}^d$:

$$f_{v, \Sigma} = \frac{\Gamma(\frac{\nu+d}{2})}{\Gamma(\frac{\nu}{2}) \sqrt{(\pi\nu)^d \det(\Sigma)}} \left(1 + \frac{x \Sigma^{-1} x'}{\nu}\right)^{-(\nu+d)/2}$$

où Γ est la fonction gamma.

Par ailleurs, une étude intéressante reprenant plusieurs propriétés et proposant une méthode de calibrage de la copule de Student est proposée dans Demarta et McNeil [2004]. La figure suivante illustre la densité de la copule de Student bi-variée pour une corrélation $\Sigma = 0.5$ et un nombre de degrés de liberté $\nu = 14$.

Figure 9 : Densité de la copule de Student pour une corrélation de 0.5 et un nombre de degrés de liberté de 14



3.2. Les copules archimédiennes

La classe des copules archimédiennes, définies par Genest et Mackay [1986], joue un rôle important. En effet, contrairement aux copules gaussiennes et aux copules de Student, les copules archimédiennes ont le grand avantage de décrire des structures de dépendance très diverses dont notamment les dépendances dites asymétriques, où les coefficients de queue inférieure et de queue supérieure diffèrent.

Soit φ une fonction strictement décroissante et continue sur $]0,1]$. La fonction définie pour tout $(u_1, \dots, u_d) \in [0,1]^d$ par : $C(u_1, \dots, u_d) = \varphi^{-1}(\varphi(u_1) + \dots + \varphi(u_d))$ est une copule si φ est d fois dérivable sur $]0,1[$ et si pour tout $1 \leq i \leq d$, $\varphi^i > 0$ pour i paire et $\varphi^i < 0$ sinon. Les fonctions s'écrivant de la sorte s'appellent les copules archimédiennes dont le générateur est φ .

De la définition précédente on déduit que la densité d'une copule archimédienne d -variée s'écrit :

$$c(u_1, \dots, u_d) = (\varphi^{-1})^d (\varphi(u_1) + \dots + \varphi(u_d)) \prod_{i=1}^d \varphi'(u_i)$$

Les copules archimédienne présentent un double intérêt. D'une part, elles permettent de construire une grande variété de familles de copules et donc représenter une grande variété de structures de dépendance. D'autre part, les copules ainsi générées ont des formes analytiques fermées et sont faciles à simuler. Pour plus d'éléments sur cette famille de copules le lecteur pourra se référer à Nelsen [1998].

Les trois familles de copules archimédiennes qui nous intéressent dans le cadre de notre étude sont présentées dans la section suivante. Cela concerne les copules de Cook-Johnson, de Franck et de Gumbel.

3.2.1. Copule de Cook-Johnson

La copule de Cook-Johnson, connu aussi sous les noms : copule de Clayton ou copule de Kimeldorf-Sampson, est une copule archimédienne dont le générateur est défini, pour $\alpha > 0$ et pour $u \in]0,1]$, par :

$$\varphi(u) = \alpha^{-1} (u^{-\alpha} - 1)$$

La copule d -variée de Cook-Johnson s'écrit donc :

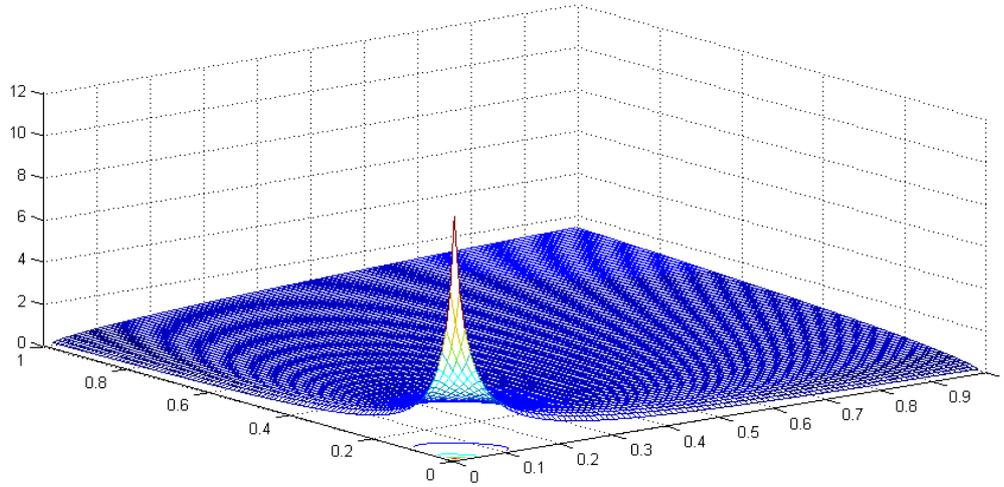
$$C(u_1, \dots, u_d) = \left(1 - d + \sum_{i=1}^d u_i^{-\alpha} \right)^{-1/\alpha}$$

Cette copule est différentiable et sa densité se calcule facilement :

$$c(u_1, \dots, u_d) = \left(1 - d + \sum_{i=1}^d u_i^{-\alpha} \right)^{-d - \frac{1}{\alpha}} \prod_{j=1}^d (u_j^{-\alpha-1} (j\alpha - \alpha + 1))$$

La copule de Cook-Johnson présente une dépendance asymptotique à gauche (sur les valeurs négatives) ce qui n'est pas le cas de la copule gaussienne. La figure suivante illustre la densité bi-variée de la copule de Cook-Johnson pour $\alpha = 0.5$.

Figure 10 : Densité bi-variée de la copule de Cook-Johnson $\alpha=0.5$



3.2.2. Copule de Franck

La copule de Franck permet de modéliser les dépendances aussi bien positives que négatives. Cependant, elle ne présente pas de dépendance de queue. Le générateur de cette copule archimédienne est :

$$\varphi(u) = -\ln\left(\frac{e^{-\alpha u} - 1}{e^{-\alpha} - 1}\right)$$

où $\alpha \neq 0$ et $u \in]0,1]$.

La copule de Franck d-variée s'écrit donc :

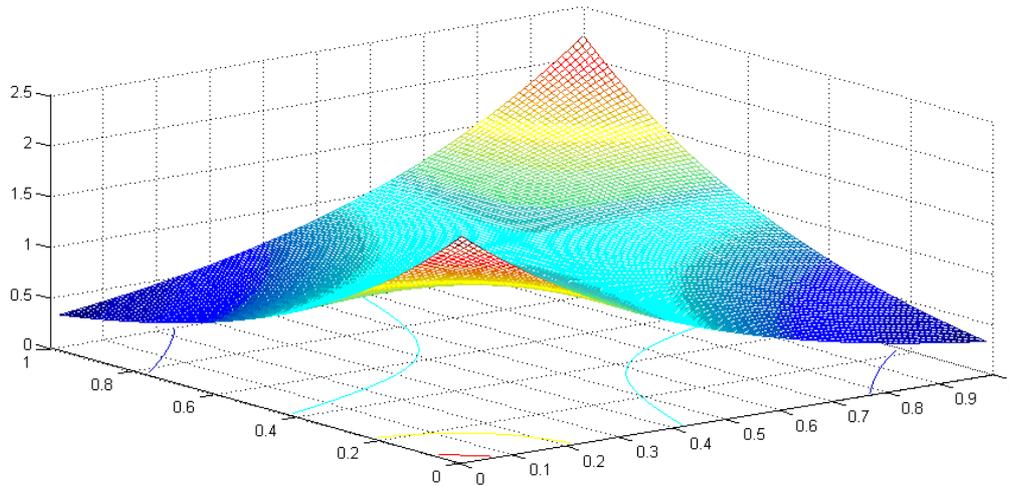
$$C(u_1, \dots, u_d) = -\frac{1}{\alpha} \ln\left(1 + \frac{1}{(e^{-\alpha} - 1)^{d-1}} \prod_{i=1}^d (e^{-\alpha u_i} - 1)\right)$$

Dans la section 6 de ce chapitre nous avons calculé la dérivée du générateur de la copule de Franck pour des variables dont la dimension est 2 et 5. Cela permet de calculer la densité de la copule à ces ordres. Pour le cas bi-varié la densité est :

$$c(u_1, u_2) = \frac{\alpha(1 - e^{-\alpha})e^{-\alpha(u_1+u_2)}}{\left((1 - e^{-\alpha}) - (1 - e^{-\alpha u_1})(1 - e^{-\alpha u_2})\right)^2}$$

La figure suivante correspond à la densité de la copule de Franck bi-variée pour $\alpha = 2$.

Figure 11 : Densité de la copule de Franck bi-variée pour $\alpha=2$



3.2.3. Copule de Gumbel

Contrairement à la copule gaussienne la copule de Gumbel, parfois appelée copule de Gumbel-Hougaard, permet de modéliser les dépendances extrêmes. En effet, la copule de Gumbel appréhende les dépendances positives et possède la caractéristique de pouvoir représenter des risques dont la structure de dépendance est plus accentuée sur la queue supérieure. Elle est à ce titre particulièrement adaptée en assurance et en finance pour étudier l'impact de la survenance d'événements de forte intensité sur la dépendance entre branches d'assurance ou actifs financiers.

Cette copule appartient à la famille des copules archimédiennes et son générateur s'écrit :

$$\varphi(u) = (-\ln(u))^\alpha$$

avec $\alpha > 1$ et $u \in]0,1]$.

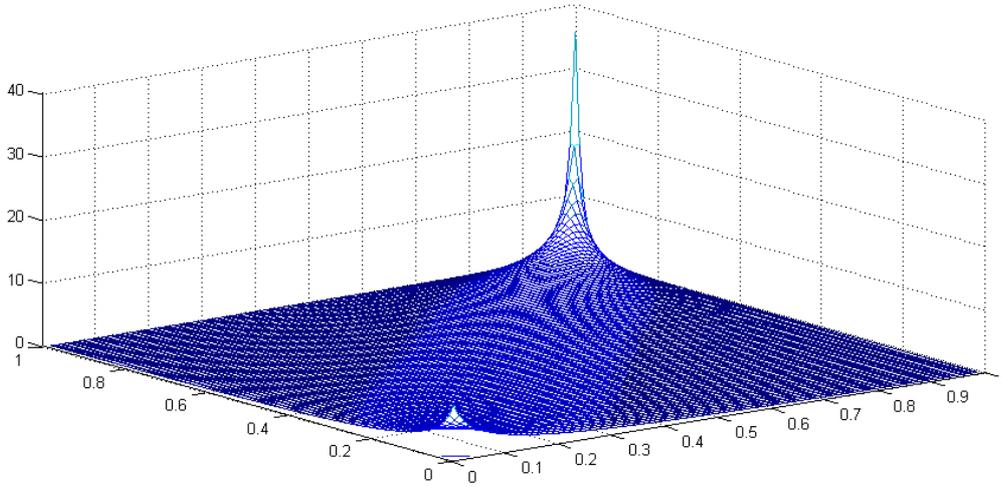
La copule de Gumbel s'écrit donc :

$$C(u_1, \dots, u_d) = \exp\left(-\left[\sum_{i=1}^d (-\ln(u_i))^\alpha\right]^{1/\alpha}\right)$$

Bien qu'existante pour tout entier positif d , l'expression explicite de la densité de cette copule est généralement complexe notamment pour des lois multi-variées. Dans la section 6 de ce chapitre nous présentons une méthode nous permettant de calculer numériquement cette densité.

La figure suivante correspond à la densité de la copule de Gumbel bi-variée pour $\alpha = 2$.

Figure 12 : Densité de la copule de Gumbel bi-variée pour $\alpha=2$



Maintenant que nous avons présenté les cinq copules paramétriques que nous avons retenues dans le cadre de nos travaux. Il est important de disposer de la copule empirique (construite directement à partir des séries de données) pour pouvoir la comparer à ces copules paramétriques. La section suivante présente deux méthodes de construction de la copule empirique à partir d'un échantillon de données.

4. Copules empiriques

Nous examinons dans cette section deux procédures de construction de la copule empirique d'un ensemble de variables à partir de leurs historiques. La première procédure repose sur la statistique de rang (copule de Deheuvels) et la seconde est fondée sur l'estimation empirique des densités de probabilité marginales et de la densité jointe.

4.1. Copule de Deheuvels

Deheuvels [1979] a introduit la notion de copule empirique. Celle-ci se base sur le rang des observations pour extraire ensuite la structure de dépendance.

Soit $(x_1^t, \dots, x_d^t)_{1 \leq t \leq T}$ un échantillon d'observations de taille T de (X_1, \dots, X_d) , et soit $(r_1^t, \dots, r_d^t)_{1 \leq t \leq T}$ la statistique de rang associée à cet échantillon multi-varié. La statistique de rang représente tout simplement le rang d'une observation, et donc sa place dans l'échantillon une fois ordonné. Cela signifie que pour tout $1 \leq i \leq d$, r_i^t est le rang de x_i^t dans $(x_i^t)_{1 \leq t \leq T}$. La copule empirique introduite par Deheuvels est définie sur l'ensemble $\left\{ \left(\frac{t_1}{T}, \dots, \frac{t_d}{T} \right), 1 \leq i \leq d, t_i = 0, 1, \dots, T \right\}$ par l'équation suivante :

$$\hat{C}_T \left(\frac{t_1}{T}, \dots, \frac{t_d}{T} \right) = \frac{1}{T} \sum_{t=1}^T \prod_{i=1}^d 1_{r_i^t \leq t_i}$$

On peut montrer que cette copule empirique satisfait plusieurs propriétés, la convergence asymptotique vers C en est un exemple. Cependant, la copule empirique de Deheuvels peut représenter plusieurs points de discontinuité ce qui limite son utilisation notamment lorsque sa densité est requise.

Il est toutefois possible d'estimer la fonction de densité de la copule empirique de Deheuvels à partir de la copule \hat{C}_T , avec la relation suivante (Nelsen [1998]) :

$$\hat{c}_T\left(\frac{t_1}{T}, \dots, \frac{t_d}{T}\right) = \sum_{j_1=1}^2 \dots \sum_{j_d=1}^2 (-1)^{j_1 + \dots + j_d} \hat{C}_T\left(\frac{t_1 - j_1 + 1}{T}, \dots, \frac{t_d - j_d + 1}{T}\right)$$

4.2. Estimation empirique par les densités de probabilité

Cette procédure consiste à estimer empiriquement les densités marginales f_1, \dots, f_d et la densité jointe f d'un vecteur de variables aléatoires $X = (X_1, \dots, X_d)$ ce qui va permettre ensuite d'estimer la densité de sa copule en utilisant la formule suivante (en supposant qu'aucun terme au dénominateur ne s'annule) :

$$c(u_1, \dots, u_d) = \frac{f(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d))}{f_1(F_1^{-1}(u_1)) \dots f_d(F_d^{-1}(u_d))}$$

Cette approche est détaillée au chapitre II-3.

5. Calibrage des copules paramétriques

Le calibrage des copules est une étape incontournable dans le cadre de l'étude de la structure de dépendance. Cependant, cette étape est généralement complexe notamment si l'on ne dispose pas directement de la densité de la copule étudiée (cf. section 6.1 concernant la copule de Gumbel). Ci-après, nous présentons trois méthodes de calibrage. Il est à noter que, lors du calibrage d'une copule, le choix d'une méthode ou de l'autre dépend fortement de la complexité et la nature de la copule (cf. chapitre II-3).

5.1. Estimation par le maximum de vraisemblance

Comme nous l'avons souligné précédemment, la densité jointe d'un vecteur aléatoire $X = (X_1, \dots, X_d)$ peut s'écrire :

$$f(x_1, \dots, x_d; \theta) = f_1(x_1; \theta) \dots f_d(x_d; \theta) c(F_1(x_1; \theta), \dots, F_d(x_d; \theta); \theta)$$

où θ est le vecteur des paramètres à estimer, c la densité de la copule associée à la loi de $X = (X_1, \dots, X_d)$ et pour $1 \leq i \leq d$, F_i est la fonction de répartition de la variable X_i dont la densité de probabilité est f_i .

Le calibrage de la copule paramétrique sur un échantillon $(x_1^t, \dots, x_d^t)_{1 \leq t \leq T}$ de taille T peut être obtenu directement par le maximum de la log-vraisemblance qui s'écrit :

$$L(\theta) = \sum_{t=1}^T \ln(f(x_1^t, \dots, x_d^t; \theta))$$

Ce qui est équivalent à :

$$L(\theta) = \sum_{t=1}^T \sum_{i=1}^d \ln(f_i(x_i^t; \theta)) + \sum_{t=1}^T \ln(c(F_1(x_1^t; \theta), \dots, F_d(x_d^t; \theta); \theta))$$

Bien que possible théoriquement, la recherche du vecteur θ qui maximise la log-vraisemblance, pose plusieurs problèmes. En effet, la recherche d'une expression analytique de θ se caractérise en général par des calculs assez lourds ce qui la rend difficile voire impossible. En outre, la recherche du maximum par des procédures numériques peut s'avérer limitée et complexe car le temps de calcul peut être assez considérable notamment quand la taille T de l'échantillon ou la dimension d du modèle multi-varié augmentent.

Joe et Xu [1996] proposent de calibrer séparément les densités marginales et ensuite calibrer la copule. La section suivante présente brièvement cette méthode.

5.2. Procédure d'estimation de Joe & Xu [1996]

Cette approche concerne les modèles multi-variés où l'on peut associer chaque élément du vecteur θ à une loi marginale unique. L'approche consiste à estimer les paramètres du modèle multi-varié en estimant les paramètres des densités marginales par le maximum de la log-vraisemblance. Dans la suite, on appellera cette méthode : IFM (pour *inference function for marging*).

La méthode IFM est adaptée au cas où le vecteur des paramètres à estimer, θ , peut être découpé en sous vecteurs où chacun est associé à l'une des fonctions de densités marginales ou à la densité de la copule. Cela est par exemple le cas dans le cadre d'une loi gaussienne où l'on peut associer chaque moyenne et chaque variance à une seule marginale et associer la matrice de corrélation à la densité de la copule gaussienne.

Supposons donc qu'on dispose d'un vecteur de variables aléatoires $X = (X_1, \dots, X_d)$ dont les densités marginales f_1, \dots, f_d sont paramétrés respectivement par les paramètres $\theta_1, \theta_2, \dots, \theta_d$ et dont la copule est paramétré par α de telle sorte que : $\theta = (\theta_1, \theta_2, \dots, \theta_d, \alpha)$.

On peut écrire donc :

$$f(x_1, \dots, x_d; \theta) = f_1(x_1; \theta_1) \dots f_d(x_d; \theta_d) c(F_1(x_1; \theta_1), \dots, F_d(x_d; \theta_d); \alpha)$$

Pour un ensemble $(x_1^t, \dots, x_d^t)_{1 \leq t \leq T}$ de T observations du vecteur $X = (X_1, \dots, X_d)$ on considère les d fonctions log-vraisemblance notée $(L_k)_{1 \leq k \leq d}$ des lois marginales et la fonction de log-vraisemblance de la loi jointe L définies par:

$$L_k(\theta_k) = \sum_{t=1}^T \ln(f_k(x_k^t; \theta_k))$$

$$L(\theta_1, \theta_2, \dots, \theta_d, \alpha) = \sum_{t=1}^T \ln(f(x_1^t, \dots, x_d^t; \theta_1, \theta_2, \dots, \theta_d, \alpha))$$

La méthode IFM consiste à trouver les paramètres $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_d$ qui maximisent respectivement les d fonctions log-vraisemblance L_1, L_2, \dots, L_d et ensuite chercher le paramètre $\hat{\alpha}$ qui maximise L . Plus précisément :

- Pour chaque $1 \leq k \leq d$ on cherche $\hat{\theta}_k$ qui maximise L_k . Donc après cette étape on aura une estimation $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_d$ des paramètres $\theta_1, \theta_2, \dots, \theta_d$:

$$\hat{\theta}_k = \arg \max(L_k(\theta_k))$$

- Ensuite on cherche le paramètre $\hat{\alpha}$ qui maximise $L(\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_d, \alpha)$. Un simple développement permet de s'apercevoir que cela revient à trouver $\hat{\alpha}$ de telle sorte que : $\sum_{t=1}^T \ln(c(F_1(x_1^t; \hat{\theta}_1), \dots, F_d(x_d^t; \hat{\theta}_d); \alpha))$ soit maximal (la partie concernant les marginales est constante).

En d'autres termes :

$$\hat{\alpha} = \arg \max \sum_{t=1}^T \ln(c(F_1(x_1^t; \hat{\theta}_1), \dots, F_d(x_d^t; \hat{\theta}_d); \alpha))$$

La méthode proposée par Joe et Xu est plus pratique que l'estimation directe par le maximum de la vraisemblance classique. En effet, elle permet de réduire considérablement la complexité des calculs et rend ainsi le calibrage de plusieurs modèles multi-variés faisable et convergent.

Par ailleurs, il est important de préciser que la comparaison directe de l'IFM et la méthode d'estimation par le maximum de vraisemblance classique est un point assez difficile à cause du temps de calcul nécessaire pour obtenir les estimations. Cependant, un nombre de tests comparatifs effectués dans Xu [1996] sur plusieurs modèles multi-variés affirment que l'IFM est une méthode très efficace et donne des résultats extrêmement proches de la méthode classique.

5.3. Méthode des moments : Calibrage par le tau de Kendall

Pour certaines copules, nous disposons des expressions analytiques liant le tau de Kendall à leurs paramètres de dépendance (exemple : la copule de Student). L'estimation du tau de Kendall permet donc, dans certains cas, de calibrer ces copules. Pour mieux appréhender cette procédure, se référer à la partie relative au calibrage de la copule de Student au chapitre II-3 ou à Demarta et McNeil [2004].

6. Compléments : calcul des densités de la copule de Gumbel et de la copule de Franck

6.1. Densité de la copule de Gumbel

Afin de calibrer la copule de Gumbel dans un modèle multi-varié, le calcul de sa densité s'avère incontournable. Nous proposons dans la suite une méthode que nous avons développée afin de pouvoir calculer cette densité.

La copule de Gumbel est définie par :

$$C(u_1, \dots, u_d) = \exp \left(- \left[\sum_{i=1}^d (-\ln(u_i))^\alpha \right]^{1/\alpha} \right)$$

La densité de la copule de Gumbel s'écrit :

$$c(u_1, \dots, u_d) = (\varphi^{-1})^d (\varphi(u_1) + \dots + \varphi(u_d)) \prod_{i=1}^d \varphi'(u_i)$$

où $\varphi(u) = (-\ln(u))^\alpha$ et $\varphi^{-1}(u) = \exp(-u^{1/\alpha}) = \exp(-u^\beta)$ et $\beta = \frac{1}{\alpha}$.

La détermination de la forme analytique de la densité de la copule de Gumbel pose un réel problème car le calcul de la dérivée d'ordre d de φ^{-1} est lourd et complexe. La majorité des travaux réalisés à ce sujet se contentent généralement d'étudier la densité bi-variée et ne présentent aucune démarche pour calculer la densité multi-variée pour des dimensions supérieures à 2.

Le but de cette section est de présenter la solution que nous proposons pour remédier à ce problème. La démarche que nous proposons sous forme d'algorithme, permet de calculer numériquement la dérivée d'ordre d de φ^{-1} et nous permet donc de calculer numériquement la densité de la copule.

La méthode consiste à calculer, par récurrence, pour tout réel $u \in \mathfrak{R}$ la dérivée d'ordre d en fonction des dérivées jusqu'à l'ordre $(d-1)$. Plus précisément, nous procédons pour chaque $1 \leq n \leq d$ au calcul des dérivées de φ^{-1} d'ordre $1, \dots, n-1$ pour en déduire ensuite la dérivée d'ordre n . Par récurrence on arrive à calculer la dérivée d'ordre d pour chaque $u \in \mathfrak{R}$.

Cette démarche se base sur la formule de Leibniz qui s'écrit pour deux fonctions dérivables en $u \in \mathfrak{R}$:

$$(fg)^{(n)}(u) = \sum_{k=0}^n \binom{n}{k} f^{(k)}(u) g^{(n-k)}(u)$$

Or on peut démontrer facilement que :

$$(\varphi^{-1})'(u) = -\beta \varphi^{-1}(u) u^{\beta-1}$$

Donc pour tout $1 \leq n$:

$$(\varphi^{-1})^{(n)}(u) = ((\varphi^{-1})')^{(n-1)}(u) = (-\beta \varphi^{-1} \cdot g)^{(n-1)}(u) = -\beta \sum_{k=0}^{n-1} \binom{n-1}{k} (\varphi^{-1})^{(k)}(u) g^{(n-k-1)}(u)$$

Posons $g(u) = u^{\beta-1}$ alors pour $1 \leq k$:

$$g^{(k)}(u) = u^{\beta-1-k} \prod_{i=1}^k (\beta - i)$$

Nous remarquons que la dérivée d'ordre d de φ^{-1} est entièrement définie par les dérivées d'ordres inférieurs de φ^{-1} et par les dérivées de g . Donc par récurrence nous arrivons à calculer la dérivée d'ordre d de φ^{-1} .

Ce résultat nous donne la possibilité de calculer numériquement la densité d-variée de la copule de Gumbel et nous permet d'étudier donc la structure de dépendance qu'elle implique.

6.2. Densité de la copule de Franck

La densité de la copule de Franck s'écrit :

$$c(u_1, \dots, u_d) = (\varphi^{-1})^d (\varphi(u_1) + \dots + \varphi(u_d)) \prod_{i=1}^d \varphi'(u_i)$$

où $\varphi(u) = -\ln\left(\frac{e^{-\alpha u} - 1}{e^{-\alpha} - 1}\right)$ et son inverse $\varphi^{-1}(u) = -\frac{\ln((e^{-\alpha} - 1)e^{-u} + 1)}{\alpha}$

La détermination de la densité de la copule de Franck pour un ordre d passe par le calcul de la dérivée première de φ et de la dérivée d'ordre d de φ^{-1} .

La dérivée de φ s'écrit :

$$\varphi'(u) = \frac{\alpha e^{-\alpha u}}{(e^{-\alpha u} - 1)}$$

Les dimensions des modèles multi-variés qui nous intéressent sont la dimension 2 et la dimension 5. Soit donc $\beta = e^{-\alpha} - 1$, les dérivées d'ordre 2 et 5 de φ^{-1} sont :

$$(\varphi^{-1})^2(u) = \frac{\beta^2 e^{-2u}}{\alpha(\beta e^{-u} + 1)^2} - \frac{\beta e^{-u}}{\alpha(\beta e^{-u} + 1)} = -\frac{\beta e^{-u}}{\alpha(\beta e^{-u} + 1)^2}$$

$$(\varphi^{-1})^5(u) = \frac{\beta e^{-u}}{\alpha(\beta e^{-u} + 1)} - \frac{15\beta^2 e^{-2u}}{\alpha(\beta e^{-u} + 1)^2} + \frac{50\beta^3 e^{-3u}}{\alpha(\beta e^{-u} + 1)^3} - \frac{60\beta^4 e^{-4u}}{\alpha(\beta e^{-u} + 1)^4} + \frac{24\beta^5 e^{-5u}}{\alpha(\beta e^{-u} + 1)^5}$$

La densité de la copule multi-variée d'ordres 2 et 5 est naturellement déduite du calcul des dérivées.

Fonction de répartition et densité de probabilité empirique

Dans le but de déterminer empiriquement la structure de dépendance optimale parmi les cinq copules paramétriques présentées au chapitre II-1 (la copule gaussienne, Student, Cook-Johnson, Gumbel et Franck), il est indispensable de disposer des densités marginales et de la densité jointe des indices financiers et macroéconomiques étudiés. Ces densités permettent, dans un premier lieu, de calibrer les cinq copules paramétriques et de décrire, ensuite, la copule empirique. La question qui s'ensuit est donc de savoir quelles sont les lois marginales et la loi jointe des indices financiers et macroéconomiques que nous allons retenir pour, d'un côté, calibrer l'ensemble des copules paramétriques et, d'un autre, déterminer la copule empirique ?

La démarche présentée au long de ce chapitre est purement empirique et vise donc à estimer les densités de probabilité marginales et la densité jointe en se basant uniquement sur l'historique des indices financiers et macroéconomiques étudiés. La première section présente une méthode d'estimation non biaisée des fonctions de répartition empiriques à partir d'un échantillon de données. Bien qu'ayant peu d'intérêt pour notre étude, celle-ci permet de mieux appréhender les points développés dans les sections qui suivent. La section 2 présente un ensemble de démarches permettant de construire des densités de probabilité empiriques. Cette section repose principalement sur le livre de Silverman [1986]. Enfin, il est à noter qu'à la fin de la section 2, l'estimation des densités de probabilité par la méthode des noyaux est présentée. Cette méthode est importante pour les chapitres suivants car elle est utilisée pour calibrer les cinq copules paramétriques et pour la détermination de la copule empirique. Les tests d'adéquations de Kolmogov-Smirnov sont réalisés au chapitre II-4 pour les estimations réalisées sur les données étudiées et présentées au chapitre I-3.

1. La fonction de répartition empirique classique

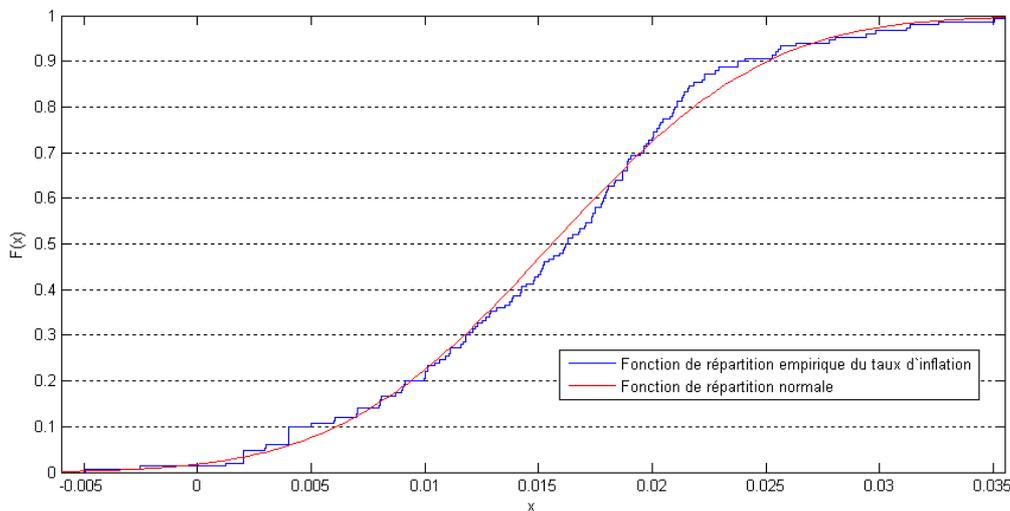
Dans cette section nous nous intéressons aux variables aléatoires unidimensionnelles. Le cas multivarié n'est autre qu'une généralisation simple du cas uni-varié. Soit donc Y une variable aléatoire univariée et $(y_t)_{1 \leq t \leq T}$ un ensemble de T observations de Y . La fonction de répartition de Y peut être approximée par une fonction de répartition attribuant la probabilité $1/T$ à chacun des éléments de l'ensemble des observations.

On définit donc la fonction de répartition empirique de Y par :

$$\hat{F}(y) = \frac{1}{T} \sum_{i=1}^T 1_{y_i \leq y}$$

Cette formule signifie que $\hat{F}(y)$ est égale au nombre d'éléments de $(y_t)_{1 \leq t \leq T}$ inférieurs à y divisé par la taille T . La figure suivante illustre la fonction de répartition empirique du taux d'inflation dont l'historique se situe entre janvier 1997 et juin 2009 (cf. chapitre I-3) et celle d'une loi normale dont la moyenne et la volatilité sont égales à celles de l'inflation. Nous observons en particulier que la fonction de répartition empirique de l'inflation est proche de celle de la loi normale. Cela est cohérent avec le test de normalité de Jarque-Bera (cf. chapitre I-3).

Figure 13 : Fonction de répartition empirique du taux d'inflation



Il est à noter toutefois que, bien que simple, l'estimation de la fonction de répartition par cette approche demeure limitée. En effet, cette estimation n'est significative que si la taille de l'ensemble des T observations est importante. En outre, \hat{F} présente plusieurs points de discontinuité ce qui rend son utilisation compliquée quand il s'agit d'estimer la densité de probabilité.

2. Densité de probabilité empirique

Nous appelons la fonction de densité de probabilité empirique, l'estimation de la densité de probabilité d'une variable aléatoire à partir d'un ensemble d'observations. Par exemple, si on suppose

que les observations suivent la densité paramétrique d'une loi normale, la densité est totalement déterminée par l'estimation de la moyenne et de l'écart-type. Ce genre d'estimation ne fait pas l'objet de cette section ; l'approche présentée ci-après est totalement empirique et ne se base que sur les observations. Bien que la majorité des applications concernent des modèles multi-variés, les développements suivants ne vont concerner que les densités uni-variées. Le cas multi-varié n'étant que la généralisation du cas uni-variés. Dans ce qui suit on considère que Y est une variable aléatoire univariée dont $(y_t)_{1 \leq t \leq T}$ est un historique de T observations. Cette section se base sur Silverman [1986].

2.1. Histogramme

L'un des estimateurs les plus simples de la densité est l'histogramme. Considérons un point d'origine y_0 , un diamètre h et définissons un ensemble d'intervalles $(I_m)_{m \in \mathbb{Z}}$ dont les éléments s'écrivent, pour tout entier relatif m :

$$I_m = [y_0 + mh, y_0 + (m+1)h[$$

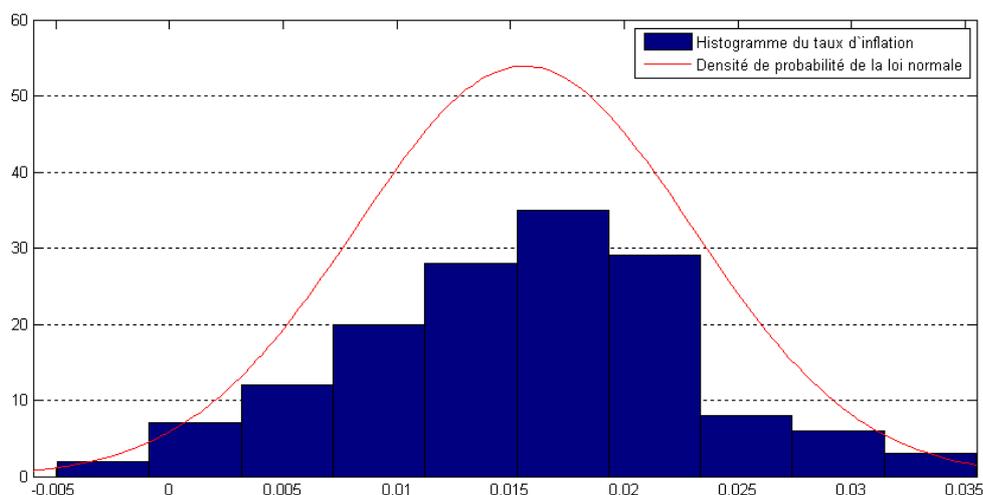
Alors l'histogramme consiste à estimer la densité de probabilité de Y par :

$$\hat{f}(y) = \frac{1}{Th} (\text{nombre d'éléments de } (y_t)_{1 \leq t \leq T} \text{ dans } I_m^y) .$$

où $I_m^y \in (I_m)_{m \in \mathbb{Z}}$ et $y \in I_m^y$.

Cela revient à dire que la valeur que prend la densité à y est égale au nombre d'éléments de $(y_t)_{1 \leq t \leq T}$ appartenant au même intervalle que y divisé par la taille de l'historique et par le diamètre de l'intervalle. Il est à noter que la valeur du diamètre h détermine considérablement la qualité du lissage de la densité estimée (cf. Silverman [1986]). Ci-après l'histogramme du taux d'inflation réalisé sur un historique datant de janvier 1997 à juin 2009. Une seconde courbe représentant la densité d'une loi normale dont la moyenne est la volatilité correspondent à celle du taux d'inflation est tracée.

Figure 14 : Histogramme du taux d'inflation



L'histogramme peut être généralisé en prenant des intervalles à diamètre variable et dans ce cas la densité s'écrit :

$$\hat{f}(y) = \frac{(\text{nombre d'éléments de } (y_t)_{1 \leq t \leq T} \text{ dans } I_m^y)}{T(\text{diamètre de } I_m^y)}$$

Bien que simple à appréhender et à mettre en pratique, cette estimation est limitée. En effet, la densité estimée peut représenter plusieurs points de discontinuité ce qui limite son utilisation notamment lorsque la dérivée est requise. Aussi, l'histogramme n'offre une bonne estimation de la densité de probabilité que lorsque l'on dispose d'une série de données assez longue.

2.2. L'estimateur naturel de la densité de probabilité

Supposons que la variable aléatoire Y admet une densité de probabilité f , alors celle-ci s'écrit pour tout $y \in \mathfrak{R}$:

$$f(y) = \lim_{h \rightarrow 0} \left(\frac{1}{2h} P(y-h < Y < y+h) \right)$$

Pour chaque réel h on peut estimer $P(y-h < Y < y+h)$ par la proportion des éléments de l'échantillon $(y_t)_{1 \leq t \leq T}$ dans l'intervalle $]y-h, y+h[$. Un estimateur naturel de la densité est donc :

$$\hat{f}(y) = \frac{1}{2Th} (\text{nombre d'éléments de } (y_t)_{1 \leq t \leq T} \text{ dans }]y-h, y+h[)$$

Cet estimateur peut s'écrire aussi sous la forme :

$$\hat{f}(y) = \frac{1}{T} \sum_{i=1}^T \frac{1}{h} w\left(\frac{y-y_i}{h}\right)$$

avec : pour tout $z \in \mathfrak{R}$, $w(z) = \frac{1}{2}$ si $-1 < z < 1$ et $w(z) = 0$ sinon.

Par ailleurs, cette méthode d'estimation est relativement comparable à celle de l'histogramme. En effet, considérons un histogramme construit à partir d'intervalles de largeur $2h$. Si on suppose qu'aucune des observations ne soit aux bornes de ces intervalles et si y est au centre de l'un de ces intervalles alors la valeur estimée $\hat{f}(y)$ est la même dans les deux méthodes. On peut dire donc que l'estimation naturelle de la densité peut être vue comme une manière de construire un histogramme où chaque point est le centre d'un intervalle de l'ensemble des intervalles $(I_m)_{m \in \mathbb{Z}}$ de l'histogramme.

Aussi, l'estimateur naturel présente des points de discontinuité et la qualité de l'estimation dépend du choix du paramètre h . La section suivante généralise l'estimation naturelle et propose un estimateur qui ne présente pas de discontinuité.

2.3. L'estimateur à noyau

L'estimateur naturel de la densité de probabilité peut être généralisé en remplaçant la fonction w par une fonction noyau notée K . Cette fonction vérifie :

$$\int_{\mathfrak{R}} K(y) dy = 1$$

En général la fonction K est une densité de probabilité symétrique (par exemple la densité d'une variable normale). Par analogie avec l'estimateur naturel, l'estimateur de Kernel de la densité de probabilité est défini par :

$$\hat{f}(y) = \frac{1}{Th} \sum_{i=1}^T K\left(\frac{y - y_i}{h}\right)$$

Le paramètre h représente la largeur de bande « *bandwidth* » de l'estimateur (pour mieux visualiser l'impact de ce paramètre sur la qualité de l'estimation se référer à Silverman [1986]) et la fonction K est appelé le noyau de l'estimateur.

Par ailleurs, le passage d'un estimateur de Kernel uni-varié à un estimateur multi-varié se fait généralement sans complication. Par exemple si nous désirons estimer les densités empiriques uni-variées et multi-variées de la variable $X = (X_1, \dots, X_d)$ à partir d'un estimateur de Kernel à noyau gaussien, alors sa densité empirique, estimée sur un historique $(x_1^t, \dots, x_d^t)_{1 \leq t \leq T}$ de taille T , est donnée pour tout $x = (x_1, \dots, x_d) \in \mathfrak{R}^d$ par la formule suivante :

$$f(x) = \frac{1}{Th^d} \sum_{t=1}^T K\left(\frac{1}{h}(x - x^t)\right)$$

avec :

- L'élément $x^t = (x_1^t, \dots, x_d^t)$ représente la valeur historique de $X = (X_1, \dots, X_d)$ à la date t .
- Le noyau est gaussien et s'écrit pour tout $z \in \mathfrak{R}^d$: $K(z) = (2\pi)^{-d/2} \exp(-zz'/2)$
- Le paramètre h est la largeur de bande caractérisant la densité. On peut l'estimer par le maximum de vraisemblance.

Les densités marginales s'écrivent pour tout $1 \leq i \leq d$ et pour tout $y \in \mathfrak{R}$:

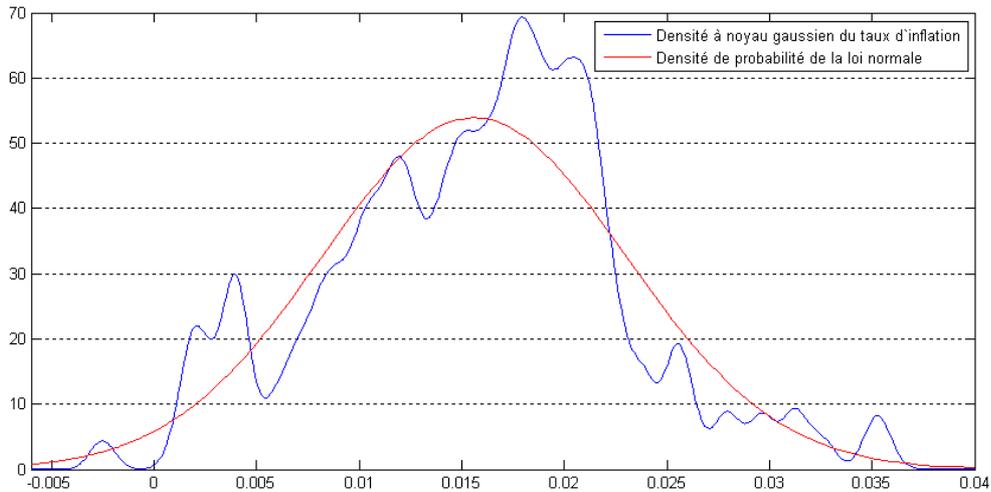
$$f_i(y) = \frac{1}{Th_i} \sum_{t=1}^T K\left(\frac{1}{h_i}(y - x_i^t)\right) = (2\pi)^{-1/2} \frac{1}{Th_i} \sum_{t=1}^T \exp\left(-\frac{1}{2h_i^2}(y - x_i^t)^2\right)$$

Les fonctions de répartition peuvent être ensuite estimées en intégrant les densités formulées ci-dessus. Les fonctions de répartition marginales s'écrivent donc :

$$F_i(y) = \frac{1}{Th_i \sqrt{2\pi}} \sum_{t=1}^T \int_{-\infty}^y \exp\left(-\frac{(y - x_i^t)^2}{2h_i^2}\right) dy$$

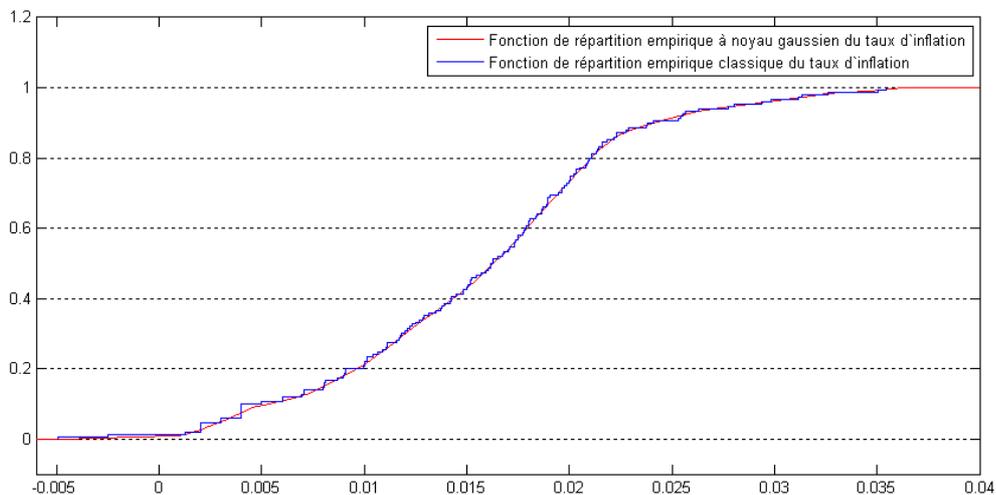
La figure ci-après illustre la densité de probabilité du taux d'inflation estimée à partir d'un noyau gaussien (la largeur de bande est estimée par le maximum de vraisemblance : $h = 6,1646E - 04$). Une seconde courbe représentant la densité d'une loi normale dont la moyenne est la volatilité correspondent à celle du taux d'inflation est tracée.

Figure 15 : Densité à noyau gaussien du taux d'inflation



La figure suivante illustre la courbe de la fonction de répartition calculée à partir des densités de probabilité à noyau gaussien et la courbe de la fonction de répartition empirique classique (présenté au début de ce chapitre) du taux d'inflation.

Figure 16 : Comparaison de la fonction de répartition empirique à noyau gaussien et la fonction de répartition empirique classique pour le taux d'inflation



Nous observons que les deux courbes sont quasiment superposées. Cela signifie que l'estimateur empirique de la densité de probabilité à noyau gaussien capte la dynamique et la distribution de l'indice tout en gardant des propriétés mathématiques importantes comme la continuité et la dérivabilité.

Maintenant que l'on sait comment estimer empiriquement les densités de probabilité marginales et la densité jointe d'un ensemble de variables aléatoires, nous procédons au chapitre suivant à présenter la procédure de sélection de la copule paramétrique optimale parmi les copules présélectionnées présentées au chapitre II-1.

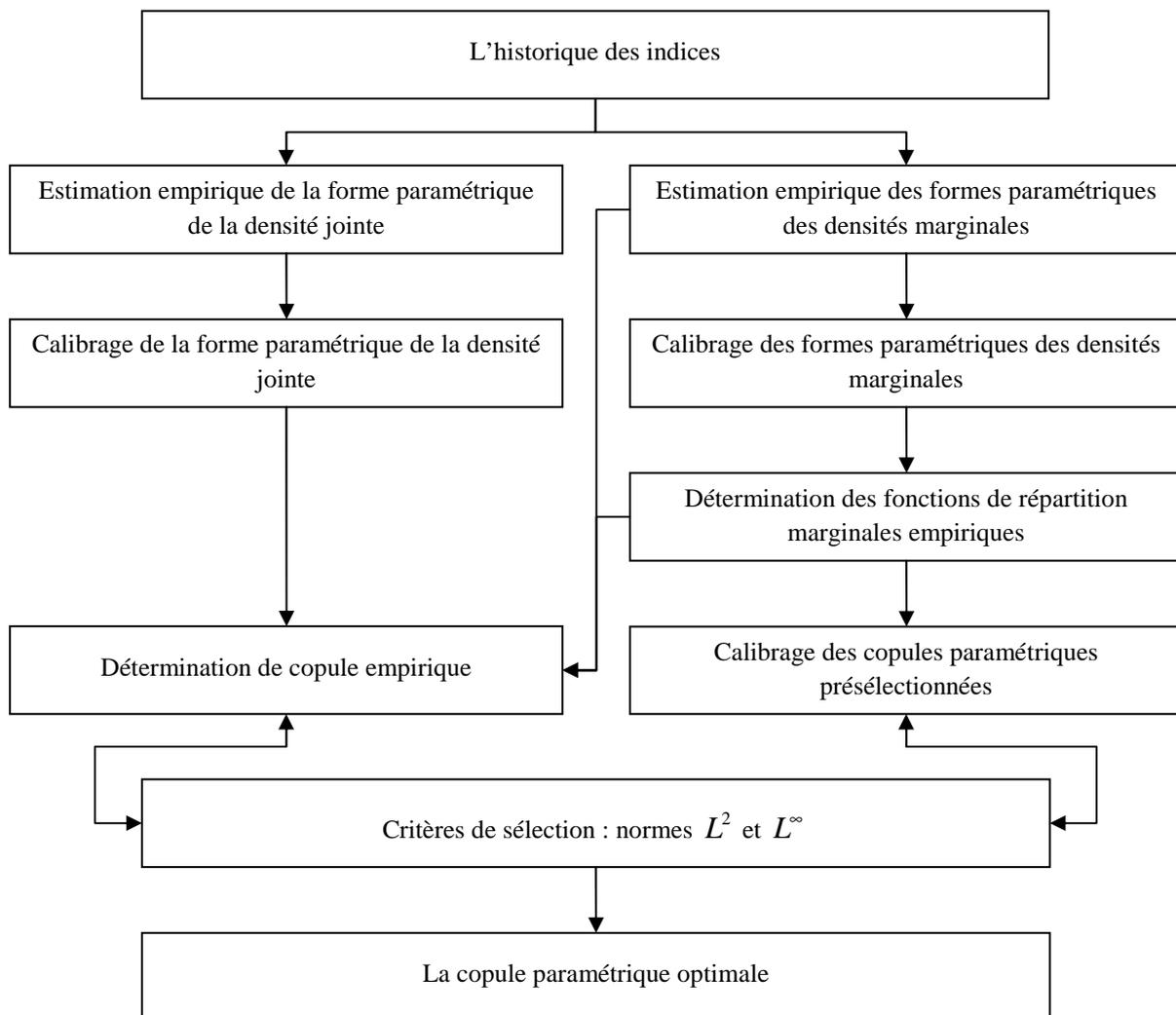
Procédure du choix de la copule optimale

Dans ce chapitre nous présentons la procédure de sélection de la copule optimale. Cette copule sera choisie parmi les cinq copules présentées au chapitre II-1. A titre de rappel, ces copules sont : la copule gaussienne, la copule de Student, la copule de Frank, la copule de Gumbel et la copule de Cook-Johnson.

Bien que la mise en œuvre ne concerne que le cas bi-varié, Kharoubi-Rakotomalala [2008] et Durrleman, Nikeghbali & Roncalli [2000] ont été d'un grand intérêt pour la conception de la structure générale de la procédure présentées dans ce chapitre. En effet, nous nous sommes inspirés du cas bi-varié pour l'étendre au cas multi-varié.

Le graphe suivant résume la démarche que nous proposons pour sélectionner la copule optimale :

Figure 17 : Sélection de la copule optimale par une démarche empirique



1. Estimation et calibrage des densités de probabilité

1.1. Estimation de la densité de probabilité empirique

Dans le cadre de la présente étude, les formes paramétriques des densités empiriques uni-variées et multi-variées sont estimées à partir d'un estimateur de Kernel à noyau gaussien. Cette estimation ressemble étroitement à l'estimation paramétrique des densités de probabilité réalisée par Kharoubi-Rakotomalala [2008] sur l'historique de quelques indices boursiers internationaux.

Par ailleurs, un test d'adéquation de Kolmogorov-Smirnov doit être réalisé pour s'assurer que l'historique est adéquat aux fonctions de répartition déduites des densités estimées. Nous rappelons que ce test est un test d'hypothèse utilisé pour déterminer si un échantillon suit bien une loi donnée connue par sa fonction de répartition continue, ou bien si deux échantillons suivent la même loi.

La forme paramétrique de la densité jointe empirique à noyau gaussien d'un vecteur de variables aléatoires $X = (X_1, \dots, X_d)$ estimée sur un historique $(x_1^t, \dots, x_d^t)_{1 \leq t \leq T}$ de taille T est donnée, pour tout $x = (x_1, \dots, x_d) \in \mathfrak{R}^d$, par la formule suivante :

$$\hat{f}(x) = \frac{1}{Th^d} \sum_{t=1}^T K\left(\frac{1}{h}(x - x^t)\right)$$

avec :

- L'élément $x^t = (x_1^t, \dots, x_d^t)$ représente la valeur historique de $X = (X_1, \dots, X_d)$ à la date t .

- Le noyau est gaussien et s'écrit pour tout $z \in \mathfrak{R}^d$: $K(z) = (2\pi)^{-d/2} \exp(-zz'/2)$

- Le paramètre h caractérise la densité. On peut l'estimer par le maximum de vraisemblance.

Les densités marginales s'écrivent pour tout $1 \leq i \leq d$ et pour tout $y \in \mathfrak{R}$:

$$\hat{f}_i(y) = \frac{1}{Th_i} \sum_{t=1}^T K\left(\frac{1}{h_i}(y - x_i^t)\right) = (2\pi)^{-1/2} \frac{1}{Th_i} \sum_{t=1}^T \exp\left(\frac{-1}{2h_i^2}(y - x_i^t)^2\right)$$

Le paramètre h_i est la largeur de bande caractérisant la densité de la variable X_i . On peut l'estimer par le maximum de vraisemblance.

Contrairement aux densités de probabilité, seules les fonctions de répartition uni-variées sont utilisées par la suite. Celles-ci s'obtiennent en intégrant les densités marginales formulées ci-dessus. Les fonctions de répartition marginales s'écrivent donc :

$$\hat{F}_i(y) = \frac{1}{Th_i \sqrt{2\pi}} \sum_{t=1}^T \int_{-\infty}^y \exp\left(-\frac{(y - x_i^t)^2}{2h_i^2}\right) dy$$

1.2. Calibrage des densités marginales et de la densité jointe

Dans la section précédente nous avons présenté les formes paramétriques des densités de probabilité et des fonctions de répartition. Cependant, pour que ces densités soient calibrées, les largeurs de bande $(h_i)_{1 \leq i \leq d}$ et h caractérisant les lois marginales et la densité jointe doivent être estimées. Pour ce faire, nous avons opté pour une estimation numérique basée sur la méthode du maximum de vraisemblance.

Par ailleurs, il est important de ne pas utiliser le même historique pour estimer la forme paramétrique de la densité de probabilité et pour la calibrer. En effet, dans ce cas, on peut prouver facilement que lorsque le paramètre à estimer ($(h_i)_{1 \leq i \leq d}$ ou h) tend vers 0 la log-vraisemblance de la densité concernée tend positivement vers l'infini et est donc divergente. Pour remédier à ce problème, nous proposons de considérer un historique de taille $(T+1)$. Les T premières valeurs serviront à estimer la forme paramétrique de la densité et les T dernières serviront à la calibrer. Soit donc $(x_1^t, \dots, x_d^t)_{1 \leq t \leq T+1}$ un tel historique. Nous appelons désormais « historique de calibrage » l'ensemble $(x_1^t, \dots, x_d^t)_{2 \leq t \leq T+1}$ et on écrit :

$$\hat{h} = \arg \max \left(\sum_{j=2}^{T+1} \ln(\hat{f}(x_1^j, \dots, x_d^j)) \right) = \arg \max \left(\sum_{j=2}^{T+1} \ln \left(\frac{1}{Th^d} \sum_{t=1}^T K \left(\frac{1}{h} (x^j - x^t) \right) \right) \right)$$

et

$$\hat{h}_i = \arg \max \left(\sum_{j=2}^{T+1} \ln(\hat{f}_i(x_i^j)) \right) = \arg \max \left(\sum_{j=2}^{T+1} \ln \left(\frac{1}{Th_i} \sum_{t=1}^T K \left(\frac{1}{h} (x_i^j - x_i^t) \right) \right) \right)$$

L'élément $x^t = (x_1^t, \dots, x_d^t)$ représente la valeur historique de $X = (X_1, \dots, X_d)$ à la date t .

2. Fonctions de répartition uni-variées et leurs inverses généralisés

On peut déduire les fonctions de répartition marginales en intégrant la densité de probabilité empirique après avoir estimé sa forme et l'avoir calibré. On écrit donc :

$$\hat{F}_i(y) = \frac{1}{Th_i \sqrt{2\pi}} \sum_{t=1}^T \int_{-\infty}^y \exp \left(-\frac{(y - x_i^t)^2}{2h_i^2} \right) dy$$

Remarquons que cela revient à une somme de T fonctions de répartition normales dont les moyennes respectives sont x_i^1, \dots, x_i^T et dont les volatilités sont égales à h_i . Cette remarque facilite considérablement l'implémentation de la fonction de répartition empirique.

En outre, l'estimation numérique de la copule empirique présentée dans la suite, nécessite le calcul de l'ensemble des fonctions $(\hat{F}_i^{-1})_{1 \leq i \leq d}$ définies par :

$$\hat{F}_i^{-1}(\alpha) = \inf \{ y \in \mathfrak{R}, \hat{F}_i(y) \geq \alpha \}$$

Remarquons aussi que la fonction de répartition est continue et strictement monotone, elle est donc bijective. Ainsi, l'inverse généralisé et l'inverse coïncident ce qui signifie que $\hat{F}_i^{-1}(\alpha)$ est la solution de l'équation : $\hat{F}_i(y) = \alpha$. Cette remarque simplifie considérablement le calcul numérique des $(\hat{F}_i^{-1})_{1 \leq i \leq d}$ qui revient à résoudre numériquement l'équation : $\hat{F}_i(y) = \alpha$.

3. La copule empirique

La densité de la copule est estimée empiriquement en s'appuyant sur la formule suivante (il est facile de vérifier qu'aucun terme au dénominateur ne s'annule) :

$$\hat{c}(u_1, \dots, u_d) = \frac{\hat{f}(\hat{F}_1^{-1}(u_1), \dots, \hat{F}_d^{-1}(u_d))}{\hat{f}_1(\hat{F}_1^{-1}(u_1)) \dots \hat{f}_d(\hat{F}_d^{-1}(u_d))}$$

En effet, après avoir déterminé empiriquement (comme nous l'avons présenté ci-dessus) la forme analytique de la densité jointe \hat{f} , des densités marginales $(\hat{f}_i)_{1 \leq i \leq d}$ et les formes analytiques des $(\hat{F}_i)_{1 \leq i \leq d}$ on peut estimer numériquement les inverses généralisées $(\hat{F}_i^{-1})_{1 \leq i \leq d}$ et ensuite estimer les valeurs numériques de la densité de la copule empirique.

4. Calibrage des copules paramétriques

Notre objectif est de pouvoir comparer la copule empirique aux cinq copules paramétriques pour en choisir la copule optimale. Pour ce faire, les modèles multi-variés associés aux cinq copules doivent être calibrés permettant ainsi de calibrer la copule paramétrique. Afin de réaliser ce calibrage, nous utilisons la méthode IFM présentée dans le chapitre II-1. Cela signifie que nous procédons au calibrage des cinq copules paramétriques après avoir calibré les densités présentées ci-dessus.

Pour la copule de Student la démarche est légèrement différente car la méthode IFM n'est pas directement utilisée pour la calibrer. En effet, nous procédons d'abord à l'estimation de la matrice de corrélation par la méthode des moments et nous utilisons ensuite la méthode IFM pour estimer le nombre de degrés de liberté (Demarta et McNeil [2004]).

4.1. Calibrage de la copule gaussienne

Rappelons que la densité de la copule gaussienne s'écrit :

$$c(u_1, \dots, u_d) = \frac{1}{\det(\Sigma)^{1/2}} \exp\left(-\frac{1}{2} \beta(\Sigma^{-1} - I_d) \beta'\right)$$

où I_d est la matrice unité de $M_d(\mathfrak{R})$, $\beta = (\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_d))$ et Φ^{-1} est l'inverse généralisé de la distribution normale centrée réduite uni-variée.

Dans un cadre plus général, soit $(x_1^t, \dots, x_d^t)_{1 \leq t \leq T}$ un historique d'une variable aléatoire vectorielle $X = (X_1, \dots, X_d)$ dont les fonctions de répartitions marginales calibrées sont F_1, \dots, F_d et soit $(u_1^t, \dots, u_d^t)_{1 \leq t \leq T}$ la transformation de ces observations en un ensemble de réalisations de la variable uniforme $U = (F_1(X_1), \dots, F_d(X_d))$. Notre objectif est de calibrer la copule gaussienne sur l'historique $(x_1^t, \dots, x_d^t)_{1 \leq t \leq T}$.

Notons $\beta_t = (\Phi^{-1}(u_1^t), \dots, \Phi^{-1}(u_d^t))$, $1 \leq t \leq T$, alors la log-vraisemblance de c s'écrit :

$$L(\Sigma) = \frac{-T}{2} \ln(\det(\Sigma)) - \frac{1}{2} \sum_{t=1}^T \beta_t (\Sigma^{-1} - I_d) \beta_t'$$

La log-vraisemblance admet un maximum. Celui-ci est atteint quand la matrice Σ prend la forme suivante (Cherubini [2004]) :

$$\hat{\Sigma} = \frac{1}{2} \sum_{t=1}^T \beta_t \beta_t'$$

Nous remarquons que cette écriture a exactement la même forme que celle utilisée pour le calcul empirique des corrélations linéaires. Donc, calibrer la copule gaussienne revient à estimer la matrice de variance-covariance (ou de corrélation) de la variable aléatoire vectorielle $(\Phi^{-1}(F_1(X_1)), \dots, \Phi^{-1}(F_d(X_d)))$ dont l'historique $(\Phi^{-1}(F_1(x_1^t)), \dots, \Phi^{-1}(F_d(x_d^t)))_{1 \leq t \leq T}$ est construit à partir de l'historique de $X = (X_1, \dots, X_d)$. Il est important de préciser que cette matrice est généralement différente de la matrice des corrélations linéaires de $X = (X_1, \dots, X_d)$. Les deux matrices sont égales si par exemple les lois marginales des variables X_1, \dots, X_d sont normales centrées et réduites.

4.2. Calibrage de la copule de Student

La fonction de densité de la copule de Student d -variée, s'écrit pour tout $(u_1, \dots, u_d) \in [0,1]^d$:

$$c(u_1, \dots, u_d) = \frac{f_{\nu, \Sigma}(t_\nu^{-1}(u_1), \dots, t_\nu^{-1}(u_d))}{\prod_{i=1}^d f_\nu(t_\nu^{-1}(u_i))}$$

Le calibrage de cette copule consiste à estimer les paramètres : ν et Σ . Contrairement à la copule gaussienne, nous ne disposons pas de formule explicite donnant les paramètres maximisant la log-vraisemblance de la densité de cette copule. En outre, la recherche numérique de ses paramètres n'est pas particulièrement simple, notamment quand la dimension du modèle est importante (estimation de ν et des $d(d-1)/2$ paramètres de la matrice Σ).

Ainsi, afin de surmonter cette difficulté, nous proposons de calibrer la copule de Student en passant par la méthode des moments en utilisant le tau de Kendall. Cette procédure consiste, dans un premier temps, à estimer la matrice Σ à partir de l'estimation du tau de Kendall et d'utiliser, dans un second temps, la matrice estimée pour trouver une estimation du degrés de liberté ν . Cette méthode est utilisée dans les travaux de Mashal et Zeevi [2002] et donne des estimations très proches de ceux trouvées par la méthode du maximum de vraisemblance.

Dans un cadre plus général, soit $(x_1^t, \dots, x_d^t)_{1 \leq t \leq T}$ un historique d'une variable aléatoire vectorielle $X = (X_1, \dots, X_d)$ dont les fonctions de répartition marginales calibrées sont F_1, \dots, F_d et soit $(u_1^t, \dots, u_d^t)_{1 \leq t \leq T}$ la transformation de ces observations en un ensemble de réalisations de la variable uniforme $U = (F_1(X_1), \dots, F_d(X_d))$.

4.2.1. Estimation de la matrice de corrélation

La relation liant le tau de Kendall à la matrice Σ est présentée ci-après. Une démonstration plus générale de ce résultat est donnée dans Lindskog et al. [2003].

$$\tau(X_m, X_n) = \frac{2}{\pi} \arcsin(\Sigma_{m,n})$$

La méthode d'estimation de la matrice de corrélation Σ consiste à estimer empiriquement le tau de Kendall pour chaque couple des d variables aléatoires X_1, \dots, X_d par l'estimateur non biaisé suivant :

$$\hat{\tau}(X_m, X_n) = \frac{2}{T(T-1)} \sum_{j=2}^T \sum_{i=1}^{j-1} \text{sign}\{(x_m^j - x_m^i)(x_n^j - x_n^i)\}$$

Ensuite, nous définissons l'estimateur de Σ par :

$$\hat{\Sigma}_{m,n} = \sin\left(\frac{\pi}{2} \hat{\tau}(X_m, X_n)\right)$$

Cependant, nous n'avons *a priori* aucune garantie que la matrice $\hat{\Sigma}$ soit définie positive (ce problème ne se pose pas dans le cadre de notre étude). Dans ce cas, cette matrice peut être approchée par une matrice définie positive, ce qui revient à résoudre un problème d'optimisation.

4.2.2. Estimation du nombre de degrés de liberté de la copule

La démarche la plus facile pour estimer le nombre de degrés de liberté ν est de considérer le maximum de la vraisemblance de la copule de Student dont la matrice de corrélation Σ est égale à la valeur estimée par la méthode des moments $\hat{\Sigma}$. Cela se traduit par :

$$\hat{\nu} = \arg \max \left(\sum_{i=1}^T \ln \left(\frac{f_{\nu, \hat{\Sigma}}(t_{\nu}^{-1}(u_1^i), \dots, t_{\nu}^{-1}(u_d^i))}{\prod_{i=1}^d f_{\nu}(t_{\nu}^{-1}(u_i^i))} \right) \right)$$

4.3. Calibrage des copules : Cook-Johnson, Franck et Gumbel

Nous appliquons directement la méthode IFM pour calibrer le modèle multi-varié dont la structure de dépendance est l'une de ces trois copules archimédiennes. Notons que quelque soit la dimension de la variable aléatoire vectorielle $X = (X_1, \dots, X_d)$, le calibrage de chacune de ces trois copules nécessite l'estimation d'un seul paramètre (cf. chapitre II-1). Le calibrage est fait numériquement en partant des densités de chaque copule.

5. Critères de sélection de la copule optimale

Bien que ne présentant aucune démarche de calcul ou de mise en œuvre opérationnelle, nous nous sommes inspirés du livre Kharoubi-Rakotomalala [2008] pour définir les critères de sélection.

Ainsi, dans le but de trouver la copule optimale parmi nos cinq copules, nous allons étudier la distance, au sens de la norme L^2 , entre la densité empirique et les densités des cinq copules paramétriques.

Il est rappelé que la copule empirique présentée dans la section 3 de ce chapitre a une forme paramétrique continue sur $[0,1]$ et est intégrable.

Si c est la densité d'une copule paramétrique parmi les copules présélectionnées et si \hat{c} est la densité de la copule empirique alors la norme L^2 s'écrit :

$$N_2 = \left(\int_{[0,1]^d} (\hat{c}(u_1, \dots, u_d) - c(u_1, \dots, u_d))^2 du_1 \dots du_d \right)^{1/2}$$

La densité de la copule paramétrique optimale doit être *a priori* proche de la densité empirique et donc elle doit minimiser la norme L^2 . Ce critère est le critère central retenu pour la sélection de la

copule optimale. En outre, la norme infinie L^∞ , est calculée pour se faire une idée sur la distance maximale entre les copule paramétrique et la copule empirique. Cette norme s'écrit :

$$N_\infty = \sup_{[0,1]^d} \left(\left| \hat{c}(u_1, \dots, u_d) - c(u_1, \dots, u_d) \right| \right)$$

Il est toutefois très complexe de calculer, même numériquement, les deux normes présentées ci-dessus notamment lorsque la dimension du modèle est importante. La solution que nous apportons à ce problème consiste à approximer numériquement ces deux normes par des séries.

Pour ce faire, nous nous appuyons sur le théorème de la convergence dominée en théorie de l'intégration (cf. Jacod [2001-2002]). Soit donc f une fonction bornée (la mesure est notée μ) définie sur un ensemble dont la mesure est finie I . Soit δ_n le plus grand diamètre d'une partition finie $E = \{E_i\}_{1 \leq i \leq n}$ de I . Soit e_i un point de E_i , pour $1 \leq i \leq n$. Alors en utilisant le théorème de la convergence dominée on peut écrire :

$$\int_I f d\mu = \lim_{\delta_n \rightarrow 0} \sum_{i=1}^n f(e_i) \mu(E_i)$$

Nous utilisons ce résultat pour approximer les normes définies ci-dessus. Soit n un entier, $I =]0,1]^d$ et soit $I_i = \left] \frac{i-1}{n}, \frac{i}{n} \right]$ pour $1 \leq i \leq n$. Définissons $E_{i_1, \dots, i_d} = \prod_{j=1}^d \left] \frac{i_j-1}{n}, \frac{i_j}{n} \right]$ pour $1 \leq i_1, \dots, i_d \leq n$. On remarque que ces ensembles forment une partition de I . Alors, on peut approximer la norme L^2 de la différence entre la copule empirique et paramétrique par la formule suivante (l'approximation est d'autant plus fine que n est grand et on peut remarquer aussi que $\frac{2i-1}{2n} \in I_i = \left] \frac{i-1}{n}, \frac{i}{n} \right]$):

$$\tilde{N}_2 = \left(\frac{1}{n^d} \sum_{1 \leq i_1, \dots, i_d \leq n} \left(\hat{c}\left(\frac{2i_1-1}{2n}, \dots, \frac{2i_d-1}{2n}\right) - c\left(\frac{2i_1-1}{2n}, \dots, \frac{2i_d-1}{2n}\right) \right)^2 \right)^{1/2}$$

Par ailleurs, la norme L^∞ de la différence entre la copule empirique et paramétrique peut être approximée par :

$$\tilde{N}_\infty = \sup_{1 \leq i_1, \dots, i_d \leq n} \left(\left| \hat{c}\left(\frac{2i_1-1}{2n}, \dots, \frac{2i_d-1}{2n}\right) - c\left(\frac{2i_1-1}{2n}, \dots, \frac{2i_d-1}{2n}\right) \right| \right)$$

La mise en œuvre de cette procédure a nécessité l'implémentation de plusieurs centaines de lignes de codes sur Matlab et le calcul des normes par approximation numérique prend un temps considérable. Le chapitre suivant donne plus de détail sur ces aspects et présente les résultats de cette procédure appliquée aux données présentées au chapitre I-3.

Chapitre II-4

Le choix de la copule optimale

Nous essayons à travers ce chapitre d'appliquer la procédure de détermination de la structure de dépendance détaillée au chapitre II-3 sur les données présentées au chapitre I-3. Nous examinerons alors la structure de dépendance du taux d'inflation, du taux d'intérêt nominal long, du taux d'intérêt nominal court, du rendement d'un investissement en actions et du rendement d'un investissement en immobilier.

La première section est consacrée au calibrage des densités de probabilité empiriques, la seconde section présentera les résultats du calibrage des copules paramétriques et la dernière section est consacrée à la présentation de la copule optimale.

Comme nous l'avons précisé au chapitre I-3, les séries historiques utilisées datent du mois de janvier 1997 au mois de juin 2009. Mis à part la série de l'immobilier dont le taux de rendement annuel est à fréquence trimestrielle, toutes les autres séries représentent des taux annuels à fréquences mensuelles.

1. Densité de probabilité et fonction de répartition empiriques

La construction des formes paramétriques des densités de probabilité marginales des indices financiers et macroéconomiques étudiés et leur densité de probabilité jointe, s'appuie sur l'estimateur de Kernel à noyau gaussien (cf. chapitre II-3). Rappelons que cela consiste à placer un noyau gaussien sur chaque point des échantillons retenus :

$$\hat{f}_i(y) = \frac{1}{Th_i} \sum_{t=1}^T K\left(\frac{y - x_i^t}{h_i}\right) = (2\pi)^{-1/2} \frac{1}{Th_i} \sum_{t=1}^T \exp\left(-\frac{1}{2h_i^2}(y - x_i^t)^2\right)$$

Ensuite les cinq largeurs de bande, h_i , caractérisant chaque densité, f_i , sont estimées par le maximum de vraisemblance. Pour chaque indice les valeurs historiques, dont la taille est $(T + 1)$, sont scindées en deux parties : les T premières valeurs (de 1 à T) servent à estimer la forme paramétrique de la densité et les T dernières (de 2 à $T+1$) servent à la calibrer (cf. chapitre II-3). Le tableau suivant présente les valeurs des largeurs de bandes des densités empiriques des cinq indices ainsi que la largeur de bande de la densité jointe :

Tableau 3 : Calibrage des densités empiriques

Densité de probabilité	Largeur de bande « h »
Taux d'inflation	6,1646E-04
Taux long nominal	1,6379E-05
Taux court nominal	8,8596E-05
Taux de rendement des actions	2,8152E-04
Taux de rendement de l'immobilier	8,0078E-03
Loi jointe des cinq indices	7,7148E-03

Nous observons que les largeurs de bande relatives au taux d'intérêt nominal long, au taux d'intérêt nominal court et au rendement des actions sont significativement inférieures aux autres largeurs de bande. Cette remarque permet d'expliquer l'observation des pics autour des valeurs historiques de ces trois indices comme le montre les figures suivantes.

Figure 18 : Densités de probabilité empiriques à noyau gaussien du taux d'inflation et des taux d'intérêt nominaux

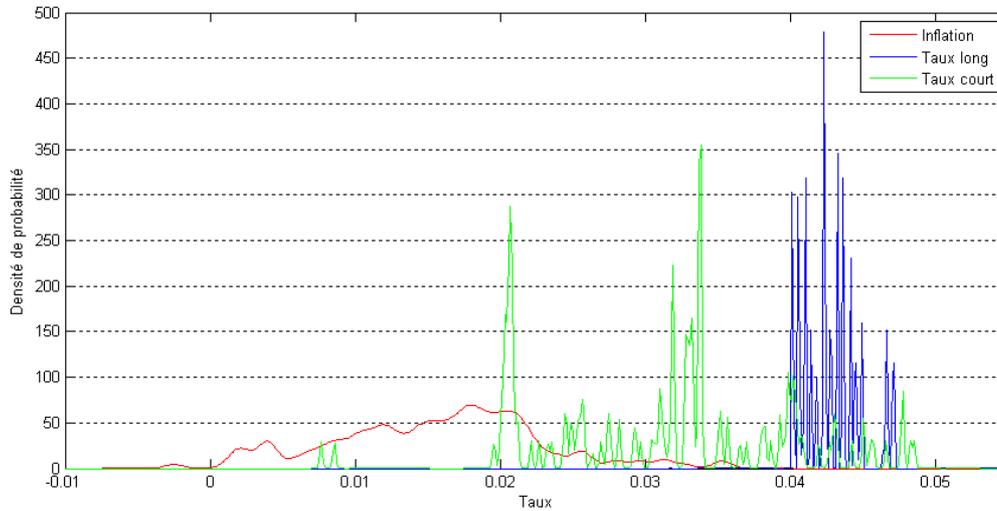
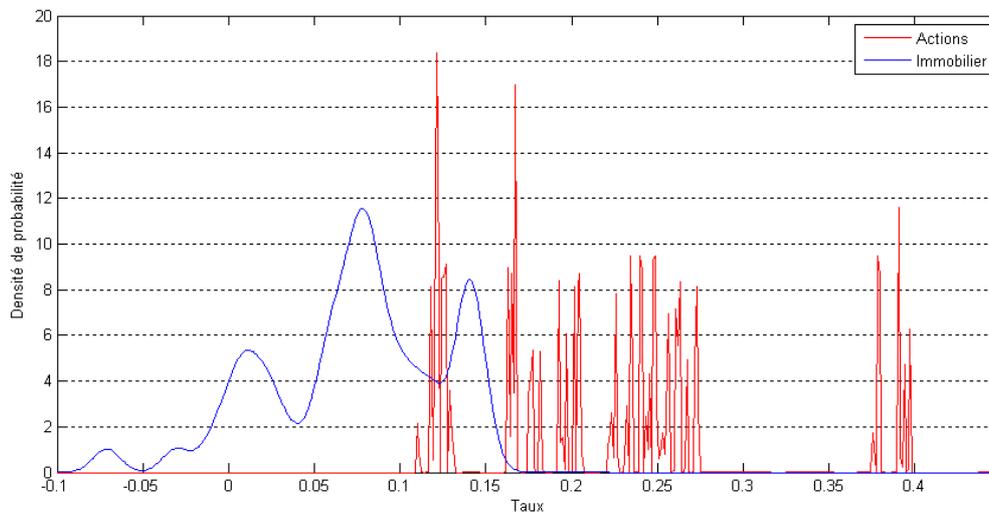


Figure 19 : Densités de probabilité empiriques à noyau gaussien du rendement des actions et du rendement de l'immobilier



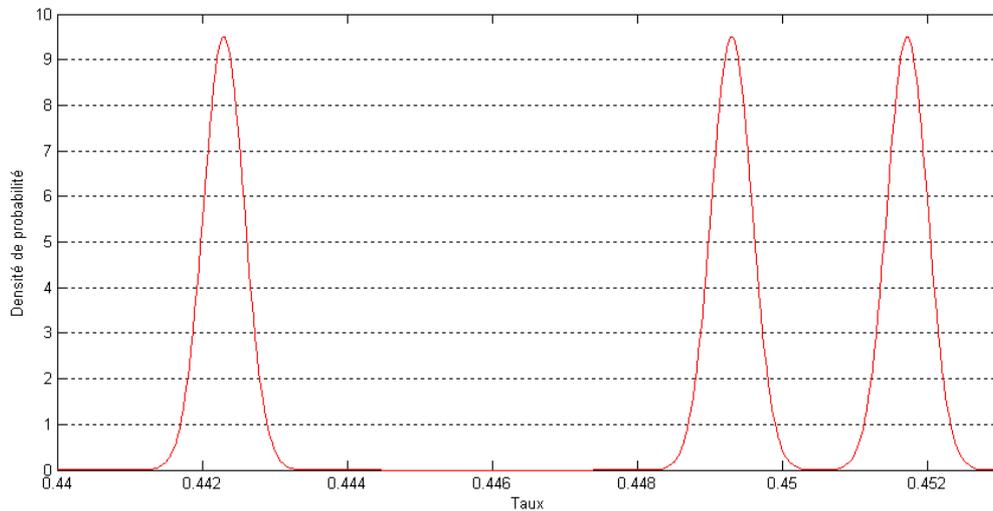
En effet, quand la largeur de bande est significativement faible, le noyau gaussien placé autour d'une valeur historique s'annule « rapidement » dès qu'on s'écarte du voisinage de cette valeur historique. Cela se justifie par la valeur de la limite du noyau gaussien, placé autour de la valeur historique x_i^t , quand la largeur de bande tend vers zéro :

$$\lim_{h \rightarrow 0} (2\pi)^{-1/2} \frac{1}{h} \exp\left(\frac{-1}{2h^2} (x - x_i^t)^2\right) = \begin{cases} 0, & \text{si } x \neq x_i^t \\ +\infty, & \text{si } x = x_i^t \end{cases}$$

La figure suivante illustre la densité de probabilité du rendement des actions sur un intervalle réduit (un zoom). On observe clairement qu'au voisinage d'une valeur historique, la densité de probabilité

est pratiquement égale au noyau gaussien pondéré (par $\frac{1}{T}$) placé autour de cette valeur et quasiment nulle ailleurs.

Figure 20 : Densité de probabilité empirique à noyau gaussien du rendement des actions sur un intervalle réduit



Par ailleurs, les figures ci-dessus montrent que les valeurs non nulles de la densité de probabilité de chaque indice se situent entre sa valeur historique minimale et sa valeur historique maximale ce qui est statistiquement cohérent avec le fait que la probabilité d'observer une réalisation qui ne se situe pas entre ces deux valeurs soit nulle. Aussi, on peut observer que, pour chaque indice, la fonction de répartition est constante sur deux intervalles : le premier concerne les valeurs inférieures à la valeur historique minimale où la fonction de répartition est nulle et le second concerne les valeurs supérieures à la valeur historique maximale où elle vaut 1.

Figure 21 : Fonctions de répartition empiriques à noyau gaussien du taux d'inflation et des taux d'intérêt

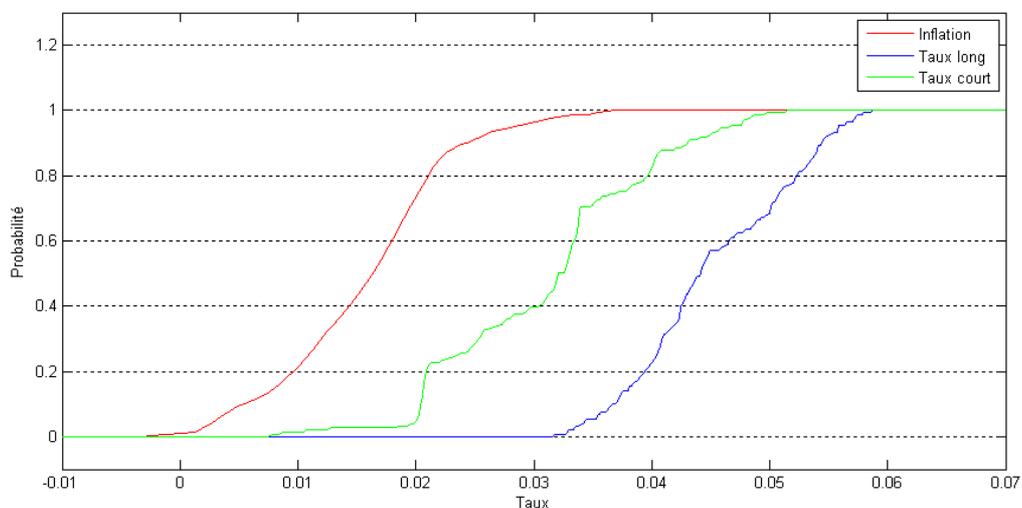
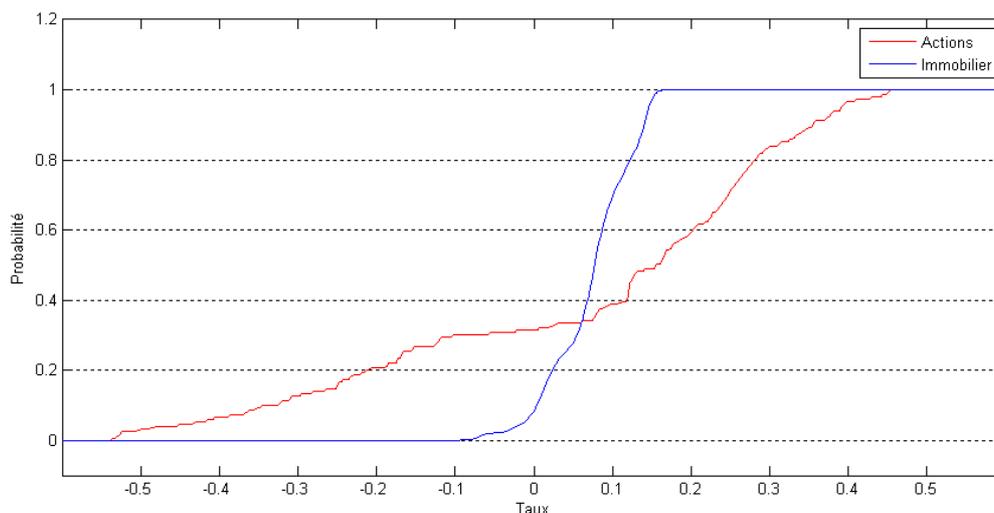


Figure 22 : Fonctions de répartition empiriques à noyau gaussien du rendement des actions et de l'immobilier



Afin de tester l'adéquation des fonctions de répartition estimées aux séries de données, cinq tests de Kolmogov-Smirnov comparant chaque série de données historiques à sa fonction de répartition estimée empiriquement sont réalisés. Tous ces tests acceptent l'hypothèse de l'adéquation des séries et des fonctions de répartition avec un niveau de significativité de 5 % et donc on peut dire que les fonctions de répartition dont les densités sont estimées à partir de noyaux gaussiens sont adaptées aux données étudiées.

Comme nous l'avons expliqué au chapitre II.3, l'estimation de la forme paramétrique et le calibrage des densités de probabilité marginales et de la densité jointe permet de calculer les valeurs numériques de la densité de la copule empirique :

$$\hat{c}(u_1, \dots, u_d) = \frac{\hat{f}(\hat{F}_1^{-1}(u_1), \dots, \hat{F}_d^{-1}(u_d))}{\hat{f}_1(\hat{F}_1^{-1}(u_1)) \dots \hat{f}_d(\hat{F}_d^{-1}(u_d))}$$

Le choix de la structure de dépendance optimale, parmi les cinq copules paramétriques retenues, nécessite le calibrage de celles-ci. La section suivante présente les résultats relatifs à ce calibrage.

2. Calibrage des copules

2.1. La copule gaussienne

Rappelons que le calibrage de la copule gaussienne revient à estimer la matrice de corrélation de la variable aléatoire vectorielle $(\Phi^{-1}(F_1(X_1)), \dots, \Phi^{-1}(F_d(X_d)))$ dont l'historique $(\Phi^{-1}(F_1(x_1^t)), \dots, \Phi^{-1}(F_d(x_d^t)))_{1 \leq t \leq T}$ est construit à partir de l'historique de $X = (X_1, \dots, X_d)$ (où Φ^{-1} est l'inverse généralisé de la distribution normale centrée réduite uni-variée). Dans notre cas, la dimension du modèle multi-varié est 5 et la détermination des fonctions de répartition à partir de l'estimation de leurs densités par l'estimateur de Kernel à noyau gaussien permet de calculer numériquement la matrice de corrélation et ainsi le calibrage de la copule gaussienne. Remarquons que cette matrice est différente de la matrice de corrélation des indices étudiés présentés au chapitre I-3 ce qui s'explique par la non normalité des fonctions de répartition marginales estimées.

Tableau 4 : Matrice de corrélation de la copule gaussienne

	Taux d'inflation	Taux long nominal	Taux court nominal	Actions	Immobilier
Taux d'inflation	100,00 %	-3,09 %	35,25 %	-19,85 %	38,20 %
Taux long nominal	-3,09 %	100,00 %	45,05 %	25,52 %	-28,29 %
Taux court nominal	35,25 %	45,05 %	100,00 %	7,73 %	-12,61 %
Actions	-19,85 %	25,52 %	7,73 %	100,00 %	23,89 %
Immobilier	38,20 %	-28,29 %	-12,61 %	23,89 %	100,00 %

2.2. La copule de Student

Le calibrage de la densité de la copule de Student nécessite l'estimation de la matrice de corrélation et le nombre de degrés de liberté. Dans un premier temps, la matrice de corrélation est estimée à partir de l'estimation du tau de Kendall. Cette matrice sert dans un second temps à l'estimation du nombre de degrés de liberté par le maximum de vraisemblance.

Tableau 5 : Matrice de corrélation de la copule de Student

	Taux d'inflation	Taux long nominal	Taux court nominal	Actions	Immobilier
Taux d'inflation	100,00 %	-5,97 %	8,05 %	-31,70 %	55,52 %
Taux long nominal	-5,97 %	100,00 %	54,74 %	21,74 %	-43,27 %
Taux court nominal	8,05 %	54,74 %	100,00 %	-1,87 %	-37,17 %
Actions	-31,70 %	21,74 %	-1,87 %	100,00 %	19,75 %
Immobilier	55,52 %	-43,27 %	-37,17 %	19,75 %	100,00 %

Le nombre de degrés de liberté estimé est : $\nu = 14,55$.

2.3. Le calibrage des copules archimédiennes : Cook-Johnson, Gumbel et Franck

Le calibrage des copules : Cook-Johnson, Gumbel et Franck ne nécessite que l'estimation d'un seul paramètre (noté α). Ainsi aucun problème d'optimisation n'apparaît et la recherche du maximum de vraisemblance converge rapidement. Le tableau suivant présente les résultats obtenus :

Tableau 6 : Calibrage des copules Cook-Johnson, Gumbel et Franck

La copule	α
Cook Johnson	9,0918E-02
Gumbel	1,0026E+00
Franck	4,8828E-05

Maintenant que les cinq densités des copules paramétriques sont calibrées sur l'historique des cinq indices, nous appliquons la procédure de sélection de la copule optimale en calculant et puis en comparant les distances entre les densités de ces copules paramétriques et la densité de la copule empirique.

3. Le choix de la copule optimale

Afin de trouver la copule optimale parmi nos cinq copules paramétriques, les normes suivantes sont étudiées pour chacune d'elle :

$$N_2 = \left(\int_{[0,1]^d} (\hat{c}(u_1, \dots, u_d) - c(u_1, \dots, u_d))^2 du_1 \dots du_d \right)^{1/2}$$

$$N_\infty = \sup_{[0,1]^d} \left(|\hat{c}(u_1, \dots, u_d) - c(u_1, \dots, u_d)| \right)$$

Dans notre cas la dimension est égale à 5. Il est donc très complexe de procéder directement au calcul des intégrales d'ordre cinq pour retrouver les distances. Nous avons donc considéré l'alternative numérique consistant à approximer ces normes par :

$$\tilde{N}_2 = \left(\frac{1}{n^d} \sum_{1 \leq i_1, \dots, i_d \leq n} \left(\hat{c}\left(\frac{2i_1-1}{2n}, \dots, \frac{2i_d-1}{2n}\right) - c\left(\frac{2i_1-1}{2n}, \dots, \frac{2i_d-1}{2n}\right) \right)^2 \right)^{1/2}$$

$$\tilde{N}_\infty = \sup_{1 \leq i_1, \dots, i_d \leq n} \left(\left| \hat{c}\left(\frac{2i_1-1}{2n}, \dots, \frac{2i_d-1}{2n}\right) - c\left(\frac{2i_1-1}{2n}, \dots, \frac{2i_d-1}{2n}\right) \right| \right)$$

Cette approximation est d'autant plus fine que n est grand. Cependant, ces normes sont calculées sur n^d vecteurs ce qui rend les calculs parfois lents et complexes. Des tests consistant à comparer l'estimation numérique et le calcul direct de l'intégrale de quelques fonctions à plusieurs variables classiques (produits et sommes d'un ensemble de variables, puissance de la somme, inverse...) montrent que l'estimation numérique est une très bonne approximation dès que $n \geq 20$. Ainsi, en ce qui concerne notre étude, nous avons estimé ces normes sur 9 765 625 vecteurs de dimension 5, ce qui correspond à $n = 25$. Le tableau suivant donne la valeur de la norme L^2 et la norme L^∞ pour les cinq copules paramétriques étudiées. Les valeurs minimales sont soulignées.

Tableau 7 : Distances entre les copules paramétriques et la copule empirique

La copule	La norme L^2	La norme L^∞
Gaussienne	1,408	<u>13,476</u>
Student	1,916	58,224
Cook Johnson	<u>0,974</u>	13,786
Gumbel	0,989	13,865
Franck	0,988	13,857

Nous observons que le calcul des normes L^2 place en tête la copule de Cook-Johnson. Aussi, les distances L^2 entre la copule empirique et les trois copules archimédiennes sont significativement inférieures à celles entre la copule empirique et les deux copules elliptiques (gaussienne et Student). Cela prouve encore une fois que la structure de dépendance gaussienne est loin d'être représentative de la structure de dépendance des indices financiers et macroéconomiques étudiés.

Nous concluons ainsi que la copule de Cook-Johnson représente une forme paramétrique appropriée pour modéliser la structure de dépendance des indices financiers et macroéconomiques étudiés. On rappelle que cette copule présente une dépendance asymptotique à gauche. Cela implique que la dépendance des fortes valeurs positives des indices étudiés au delà d'un seuil donné tend

asymptotiquement à s'annuler quand le niveau du seuil augmente. Ceci n'est cependant pas le cas pour les fortes valeurs négatives dont la dépendance tend à augmenter avec le seuil considéré.

Maintenant que nous avons déterminé la copule appropriée à la structure de dépendance du taux d'inflation, des taux d'intérêt nominaux à long et à court terme, du rendement des actions et du rendement de l'immobilier, il serait intéressant de pouvoir combiner cette structure avec un modèle décrivant la dynamique de ces indices. La partie III tente d'illustrer une telle combinaison et propose de projeter les indices financiers et macroéconomiques étudiés en utilisant le modèle d'Ahlgrim et al. [2005].

Partie III

Modélisation et simulation

Cette partie consiste en une réflexion sur la modélisation des indices financiers et macroéconomiques étudiés en intégrant les conclusions formulées à l'issue de la deuxième partie. Le modèle qui est à la base de cette réflexion est le modèle d'Ahlgrim et al. [2005]. L'un des objectifs de ce volet est de projeter les indices en question ce qui peut servir ensuite à évaluer les performances d'un fonds composé d'obligations, d'actions et de l'immobilier.

Le premier chapitre de cette partie présente le modèle d'Ahlgrim et al. [2005] dont le calibrage est présenté au chapitre III.2. Le troisième chapitre présente une démarche simple de projection du modèle en se basant sur une structure de dépendance linéaire (la corrélation). Le dernier chapitre présente une procédure de projection du modèle en prenant en compte les conclusions formulées à l'issue du chapitre II.4 et donc, en prenant en compte la structure de dépendance optimale qui est la copule de Cook-Johnson.

Chapitre III-1

Présentation du modèle d'Ahlgrim

Les organismes assureurs s'intéressent de plus en plus aux modèles financiers. En effet, ces modèles sont une partie intégrante de toute analyse financière dynamique (*DFA* en anglais : *dynamic financial analysis*) et sont fréquemment utilisés pour l'évaluation de la solvabilité, l'allocation du capital ou le *pricing* des produits proposés par les assureurs. Ainsi, afin d'accentuer la sensibilisation des actuaires à la gestion des risques financiers, la *Casualty Actuarial Society (CAS)* et la *Society of Actuaries (SOA)* se sont associées pour solliciter des recherches sur le développement de modèles intégrés. Le modèle d'Ahlgrim et al. [2005] s'inscrit dans ce cadre.

Avant de présenter les dynamiques des indices financiers et macroéconomiques du modèle, Ahlgrim et al. [2005] reviennent sur les modèles intégrés de Wilkie [1995] et d'Hibbert et al. [2001] et sur les principaux modèles composites relatifs aux taux d'intérêt et aux actions. Ensuite, Ahlgrim et al. [2005] s'intéressent aux taux d'inflation, aux taux d'intérêt, aux rendements des actions, aux taux de dividende, aux rendements de l'immobilier et aux taux de chômage. Cependant, le taux chômage ne s'inscrit pas dans le cadre de notre étude et n'est donc pas présenté dans la suite.

Le modèle d'Ahlgrim et al. [2005] accorde une importance particulière à l'interdépendance entre les différentes séries économiques et financières étudiées mais cette interdépendance reste limitée car elle se base sur la corrélation.

Nous présentons au long de ce chapitre le modèle d'Ahlgrim. Notre objectif est de se mettre en capacité de déterminer, en partant du modèle, les lois du taux d'inflation, du taux d'intérêt nominal à long terme, du taux d'intérêt nominal à court terme, du rendement d'un investissement en actions et du rendement d'un investissement en immobilier.

1. Le taux d'inflation

Le taux d'inflation à l'instant t , noté q_t , est supposé suivre un processus d'Ornstein-Uhlenbeck (modèle de type Vasicek à un facteur) :

$$dq_t = k_q(\mu_q - q_t)dt + \sigma_q dB_{q,t}$$

- k_q est la vitesse de retour à la moyenne.

- μ_q est le taux moyen à long terme.

- σ_q est l'écart type de l'erreur.

- $B_{q,t}$ est un mouvement brownien.

Afin de calibrer et projeter la dynamique du taux d'inflation, Ahlgrim et al. [2005] propose une résolution approximative de cette équation reposant sur la discrétisation d'Euler. Cette démarche n'est pas retenue et nous proposons donc la solution exacte de cette équation différentielle stochastique qui s'écrit entre deux instants t et t_0 (voir annexe A pour la démonstration) :

$$q_t = q_{t_0} \exp[-k_q(t - t_0)] + \mu_q(1 - \exp[-k_q(t - t_0)]) + \sigma_q \sqrt{\frac{1 - \exp[-2k_q(t - t_0)]}{2k_q}} \varepsilon_{q,t}^{t_0}$$

- $\varepsilon_{q,t}^{t_0}$ est une variable aléatoire de loi normale centrée réduite.

La réécriture de cette formule à l'instant $t_0 = 0$ montre que la loi du taux d'inflation à chaque instant t est une loi normale :

$$q_t \approx N\left(q_0 \exp[-k_q t] + \mu_q(1 - \exp[-k_q t]), \sigma_q^2 \frac{1 - \exp[-2k_q t]}{2k_q}\right).$$

Aussi, on s'aperçoit que le taux d'inflation peut s'écrire sous la forme récurrente suivante :

$$q_{t+\delta} = q_t \exp[-k_q \delta] + \mu_q(1 - \exp[-k_q \delta]) + \sigma_q \sqrt{\frac{1 - \exp[-2k_q \delta]}{2k_q}} \varepsilon_{q,t} = \alpha q_t + \beta + \gamma \varepsilon_{q,t}$$

où δ représente le pas de la discrétisation. Par exemple, pour une discrétisation mensuelle : $\delta = 1/12$.

Ainsi pour calibrer la loi du taux d'inflation, une simple régression linéaire peut être effectuée pour estimer α , β et γ et en déduire ensuite les paramètres du modèle en s'appuyant sur les relations suivantes :

$$k_q = -\ln(\alpha)/\delta, \mu_q = \frac{\beta}{1-\alpha} \text{ et } \sigma_q = \gamma \sqrt{\frac{2k_q}{1 - \exp[-2k_q \delta]}}.$$

2. Les taux d'intérêt

Le modèle Ahlgrim et al. [2005] s'intéresse plus particulièrement au taux d'intérêt réels à long et à court terme. Nous présentons dans cette partie les modèles des taux réels tels qu'ils sont présentés dans le modèle d'Ahlgrim et al. [2005] pour en déduire ensuite les lois des taux nominaux.

Ahlgrim et al. [2005] proposent de reprendre un cas particulier du modèle Hull-White [1994] pour décrire la dynamique du taux d'intérêt réel long noté l et le taux d'intérêt réel court, noté r . Les équations différentielles stochastiques que vérifient ces deux variables sont (modèle de type Vasicek à deux facteurs) :

$$dl_t = k_l(\mu_l - l_t)dt + \sigma_l dB_{l,t}$$

$$dr_t = k_r(l_t - r_t)dt + \sigma_r dB_{r,t}$$

Comme pour le taux d'inflation, le taux réel à long terme s'écrit entre deux instants t et t_0 :

$$l_t = l_{t_0} \exp[-k_l(t - t_0)] + \mu_l(1 - \exp[-k_l(t - t_0)]) + \sigma_l \sqrt{\frac{1 - \exp[-2k_l(t - t_0)]}{2k_l}} \varepsilon_{l,t}^{t_0}$$

Pour le taux réel à court terme, la solution, entre deux instants t et t_0 suffisamment proches pour considérer que le taux réel long est constant sur la période, s'écrit (voir annexe A pour la démonstration)¹² :

$$r_t = r_{t_0} \exp[-k_r(t - t_0)] + l_{t_0}(1 - \exp[-k_r(t - t_0)]) + \sigma_r \sqrt{\frac{1 - \exp[-2k_r(t - t_0)]}{2k_r}} \varepsilon_{r,t}^{t_0}$$

La réécriture de ces deux formules montre que les lois du taux réel long et du taux réel court sont, à chaque instant t , des lois normales. Aussi, la forme récurrente des taux réels s'écrit (voir annexe A pour les démonstrations) :

$$l_{t+\delta} = l_t \exp[-k_l\delta] + \mu_l(1 - \exp[-k_l\delta]) + \sigma_l \sqrt{\frac{1 - \exp[-2k_l\delta]}{2k_l}} \varepsilon_{l,t}$$

$$r_{t+\delta} = r_t \exp[-k_r\delta] + l_t(1 - \exp[-k_r\delta]) + \sigma_r \sqrt{\frac{1 - \exp[-2k_r\delta]}{2k_r}} \varepsilon_{r,t}$$

où δ représente le pas de la discrétisation. Par exemple, pour une discrétisation mensuelle : $\delta = 1/12$.

Le calibrage du modèle des taux d'intérêt réels pose certaines difficultés. D'abord, les séries historiques de ces taux ne sont pas directement observables sur les marchés financiers. Ahlgrim et al. [2005] propose donc de construire ces séries de données en retranchant les taux d'inflation aux taux d'intérêt nominaux. Aussi, une autre difficulté qui concerne la méthode de calibrage se pose. En effet, au regard de la dépendance entre les taux longs et les taux courts, l'utilisation de la méthode des moindres carrés ordinaires (MCO) pour l'estimation des paramètres n'est pas appropriée. Ahlgrim et al. [2005] propose de retenir une procédure d'estimation s'appuyant sur une application en deux étapes de la méthode des moindres carrés ordinaires. Lors de la première étape, la méthode des MCO est appliquée pour estimer les paramètres du taux d'intérêt réel long. Les valeurs estimées de ce dernier serviront comme variable explicative lors de la seconde étape pour estimer les paramètres du taux d'intérêt réel court par la méthode des MCO.

¹² Ici aussi on ne retient pas la résolution proposée par Ahlgrim et al. [2005] reposant sur la méthode d'Euler.

$$\text{Etape 1 : } l_{t+1} = \alpha_1 l_t + \beta_1 + \gamma_1 \varepsilon_{l,t}$$

$$\text{Etape 2 : } r_{t+1} = \alpha_2 \hat{l}_t + \beta_2 r_t + \gamma_2 \varepsilon_{l,t}$$

Les paramètres du modèle du taux d'intérêt réel long sont estimés en s'appuyant sur les formules suivantes :

$$k_l = -\ln(\alpha_1) / \delta, \mu_l = \frac{\beta_1}{1 - \alpha_1} \text{ et } \sigma_l = \gamma_1 \sqrt{\frac{2k_l}{1 - \exp[-2k_l \delta]}}.$$

À la différence des modèles précédents, le taux d'intérêt réel court présente deux variables endogènes car $\alpha_2 + \beta_2 = 1$. Dans ce contexte on estime les paramètres du modèle par la méthode MCO entre la série historique $(r_{t+1} - r_t)_t$ et la série $(\hat{l}_t - r_t)_t$:

$$\Delta r_{t+1} = (r_{t+1} - r_t) = \alpha_2 (\hat{l}_t - r_t) + \gamma_2 \varepsilon_{l,t}.$$

On peut ainsi estimer les paramètres du modèle par les formules suivantes :

$$k_r = -\ln(1 - \alpha_2) / \delta \text{ et } \sigma_r = \gamma_2 \sqrt{\frac{2k_r}{1 - \exp[-2k_r \delta]}}.$$

Contrairement au taux long réel, le paramètre k_r représente la vitesse de convergence vers les taux d'intérêt long.

Pour déterminer les lois des taux d'intérêt nominaux, nous nous basons sur les lois du taux d'inflation et des taux d'intérêt réels :

$${}_{no\ min\ al} l_t = {}_{réel} l_t + q_t$$

$${}_{no\ min\ al} r_t = {}_{réel} r_t + q_t$$

Ainsi les dynamiques proposées par Ahlgrim et al. [2005] pour les taux réels et le taux d'inflation font que les lois marginales des taux d'intérêt nominaux à l'instant t soient normales. Il est à noter toutefois que le calibrage des densités de probabilité des taux nominaux nécessite l'estimation de la corrélation entre les résidus des taux réels et de l'inflation.

3. Le taux de rendement d'un investissement en actions

3.1. Le rendement hors dividendes

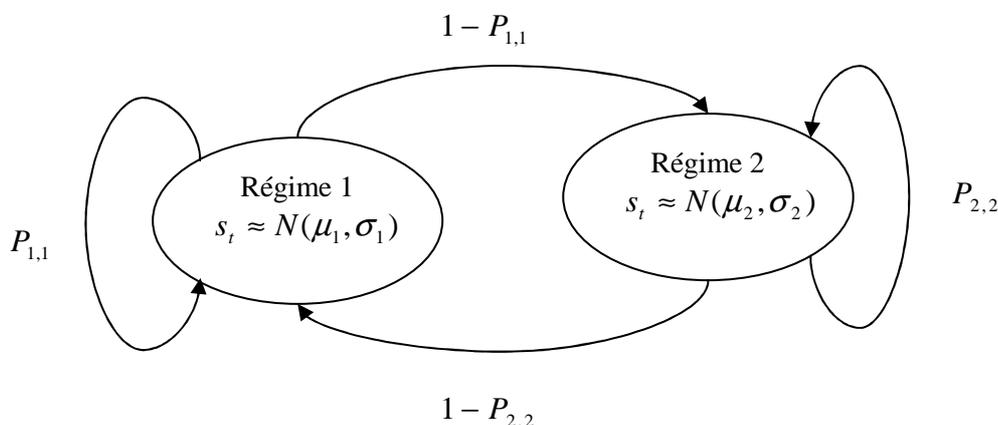
Le rendement hors dividendes des actions, notés s_t , est égal au taux d'intérêt nominal (taux sans risque), soit $r_t + q_t$, augmenté d'un excès de rendement attribuable à l'appréciation du capital et noté x_t :

$$s_t = r_t + q_t + x_t.$$

Ahlgrim et al. [2005] proposent une approche de type Hardy [2001] : le modèle à changement de régime est appliqué à l'excès de rendement des actions x_t . Le modèle à changement de régime proposé par Hardy [2001] modélise le rendement des actions (dividendes compris) en prenant en compte deux périodes (régimes) : une période où le marché connaît une forte volatilité ce qui correspond à une période de crise et une période où le marché connaît une volatilité normale ce qui

correspond à une période normale. Hardy [2001] reprend le modèle de Black-Scholes pour modéliser les prix des actions lors de chaque période. Le rendement suit donc, lors de chaque régime, une loi normale. Le modèle à changement de régime peut se schématiser par :

Figure 23 : Chaîne de Markov du modèle de type Hardy à deux régimes



3.2. Le taux de dividende

Ahlgrim et al. [2005] supposent que le taux de dividende, noté y_t , est modélisé par :

$$d \ln(y_t) = k_y (\mu_y - \ln(y_t)) dt + \sigma_y dB_{y,t}$$

- k_y est la vitesse de retour à la moyenne.

- μ_y est le logarithme du taux de dividende moyen à long terme.

Toutefois, une des principales difficultés de cette modélisation est la disponibilité des données. En effet, les séries de données sur les taux de dividende avec un historique suffisamment long ne sont pas disponibles publiquement. Aussi, pour estimer les paramètres de ce sous-modèle, Ahlgrim et al. [2005] utilisent des données privées.

3.3. Rendement des actions avec dividendes réinvestis : approche alternative

Au regard des difficultés susceptibles d'être rencontrées dans le cadre des spécifications du modèle d'Ahlgrim et al. [2005] (modélisation à partir de l'excès de rendement sur les actions et des taux de dividendes), une approche alternative est proposée pour modéliser le rendement des actions.

Cette alternative repose sur une modélisation des rendements totaux des actions avec dividendes réinvestis. En termes de choix du modèle, l'approche consistant à considérer que l'excès de rendement suit un seul régime a été privilégiée aux dépens du modèle à deux régimes de type Hardy [2001] retenu par Ahlgrim et al [2005]. Dans ce cadre l'excès de rendement suit une loi normale de paramètres μ_x , σ_x et le rendement des actions suit donc une loi normale. Cette alternative est aussi retenue dans Planchet, Théron et Kamega [2009].

Afin d'estimer les paramètres μ_x , σ_x , on définit les valeurs historique de l'excès de rendement par : $x_t = \ln\left(\frac{ind_t}{ind_{t-1}}\right) - r_t - q_t$, où ind_t représente la valeur d'un investissement en action avec dividendes réinvestis.

4. Le taux de rendement de l'immobilier

Le taux de rendement de l'immobilier à l'instant t , noté Im_t , est supposé suivre un processus d'Ornstein-Uhlenbeck :

$$d Im_t = k_{im} (\mu_{im} - q_{im}) dt + \sigma_{im} dB_{im,t}$$

La forme récurrente de la dynamique du rendement de l'immobilier s'écrit :

$$Im_{t+\delta} = Im_t \exp[-k_{im} \delta] + \mu_{im} (1 - \exp[-k_{im} \delta]) + \sigma_{im} \sqrt{\frac{1 - \exp[-2k_{im} \delta]}{2k_{im}}} \varepsilon_{im,t}$$

En termes de méthodes de calibrage, l'approche est similaire à celle présentée ci-dessus pour le taux d'inflation.

5. Le prix d'une obligation

Ahlgrim et al [2005] reprennent la relation de Fisher pour calculer le prix d'une obligation zéro-coupon à l'instant t qui paie 1 en terme nominal à la date T . Ils considèrent ainsi que si des obligations sont tarifées à partir des taux d'inflation et des taux d'intérêt réels attendus jusqu'à la maturité, alors on a la relation suivante :

$$P_{nom}(t, T) = P_{réel}(t, T) P_{inf}(t, T)$$

où :

- $P_{nom}(t, T)$ est le prix à la date t d'une obligation zéro-coupon payant 1 en terme nominal à la date T .

- $P_{inf}(t, T)$ est le prix à la date t d'une obligation calculée à partir des taux d'inflations attendus payant 1 à la date T par la formule suivante (Hull [2007] page 698) :

$$P_{inf}(t, T) = A_q(t, T) \exp(-B_q(t, T)q(t))$$

où :

$$\begin{cases} B_q(t, T) = \frac{1 - \exp\left(-k_q(T-t)\right)}{k_q} \\ A_q(t, T) = \exp\left(\frac{(B_q(t, T) - T + t)(k_q^2 \mu_q - \sigma_q^2 / 2)}{k_q^2} - \frac{\sigma_q^2 B_q(t, T)^2}{4k_q}\right) \end{cases}$$

$-P_{réel}(t, T)$ est le prix d'une obligation zéro-coupon à l'instant t qui paie 1 en termes réels à la date T . Ce prix est calculé par la formule suivante (Hibbert [2001]) :

$$P_{réel}(t, T) = \exp(A_{réel}(t, T) - B_r(t, T)r(t) - B_l(t, T)l(t))$$

où :

$$\begin{cases} B_r(t, T) = \frac{1 - \exp(-k_r(T-t))}{k_r} \\ B_l(t, T) = \frac{k_r}{k_r - k_l} \left(\frac{1 - \exp(-k_l(T-t))}{k_l} - \frac{1 - \exp(-k_r(T-t))}{k_r} \right) \end{cases}$$

et

$$\begin{aligned} A_{réel}(t, T) = & (B_r(t, T) - T + t) \left(\mu_r - \frac{\sigma_r^2}{2k_r} \right) + B_l(t, T) \mu_l - \frac{\sigma_r^2 B_r(t, T)^2}{4k_r} \\ & + \frac{\sigma_l^2}{2} \left(\frac{T-t}{k_l^2} - 2 \frac{B_l(t, T) + B_r(t, T)}{k_l^2} + \frac{1 - \exp(-2k_r(T-t))}{2k_r(k_r - k_l)^2} - \frac{2k_r(1 - \exp(-(k_r + k_l)(T-t)))}{k_l(k_r - k_l)^2(k_r + k_l)} + \frac{k_r^2(1 - \exp(-2k_l(T-t)))}{2k_l^3(k_r - k_l)^2} \right) \end{aligned}$$

On note dans la formule ci-dessus, qui pour mémoire permet de calculer à l'instant t le prix d'une obligation zéro-coupon qui paie 1 en termes nominal à la date T , qu'il n'apparaît pas de prime de risque dans l'évaluation du prix des obligations. En effet, il est supposé qu'il n'y a pas de risque de liquidité et que les obligations sont sans risque de défaut et ont toutes la même maturité, et à ce titre la prime de risque pour les obligations est nulle.

Par ailleurs, la formule ci-dessus (retenue par Ahlgrim et al. [2005]) s'appuie implicitement sur une hypothèse d'indépendance entre les taux d'intérêt réels et l'inflation. En pratique, dans le traitement des taux nominaux une attention particulière doit toutefois être portée à l'analyse des corrélations entre les taux d'intérêt réels et le taux d'inflation. Lorsque l'hypothèse d'indépendance n'est pas vérifiée un terme s'ajoute à la formule précédente. Ce point est détaillé dans Hibbert [2001].

Afin de pouvoir projeter les indices macroéconomiques et financiers étudiés dans le cadre du modèle d'Ahlgrim il est nécessaire d'estimer les paramètres du modèle. Le chapitre suivant présente l'estimation des paramètres du modèle d'Ahlgrim sur les séries de données présentées au chapitre I-3.

Chapitre III-2

Calibrage du modèle d'Ahlgrim

Ce chapitre présente les résultats du calibrage du modèle d'Ahlgrim et al. [2005]. Rappelons que cela revient à estimer les paramètres de deux types de modèles. D'une part les modèles de type Vasicek qui s'apparentent, sous leurs formes récurrentes, à des modèles autorégressifs d'ordre 1 (modèles sur l'inflation, les taux d'intérêt réel et les rendements de l'immobilier), et d'autre part le modèle retenu pour modéliser la dynamique des rendements des actions. Il est toutefois à noter qu'en ce qui concerne ce modèle, l'approche alternative présentée au chapitre III-1 est retenue.

Le calibrage des modèles de type Vasicek à un facteur est réalisé par une estimation par les moindres carrés simples. Pour les modèles de type Vasicek à deux facteurs, l'estimation est réalisée par la méthode des doubles moindres carrés (DMC). Ensuite, la qualité d'ajustement est évaluée à partir du coefficient de détermination ajusté (R^2 ajusté, représente la part de la variance de la variable endogène expliquée par la variance du modèle corrigée par le degré de liberté de l'estimation) et la significativité globale est mesurée à partir d'un test de Fisher (test dont l'hypothèse nulle H_0 correspond à la nullité de tous les coefficients du modèle).

Aussi, une attention particulière est accordée à l'analyse des résidus et de l'excès de rendement supposés normaux par le modèle. À cet effet, le test de Jarque-Bera, fondé sur la notion d'asymétrie (Skewness) et d'aplatissement (Kurtosis), est réalisé. Pour ce test, l'hypothèse nulle H_0 est celle de la normalité des résidus. Ce test vaut 1 si on peut refuser H_0 avec un niveau de significativité de 5 % et 0 sinon

Dans la suite de ce chapitre, les notations sont identiques à celles présentées au chapitre III-1. Les séries de données utilisées pour le calibrage, sont présentées au chapitre I-3. Pour mémoire, nous avons retenu des séries de données datant de janvier 1997 à juin 2009, couvrant donc une période d'environ 13 ans et mis à part la série historique du rendement annuel de l'immobilier, qui est à fréquence trimestrielle, toutes les autres séries concernent des taux annuels à fréquences mensuelles.

1. Le taux d'inflation

À titre de rappel, la forme récurrente de la dynamique de l'inflation est :

$$q_{t+\delta} = q_t \exp[-k_q \delta] + \mu_q (1 - \exp[-k_q \delta]) + \sigma_q \sqrt{\frac{1 - \exp[-2k_q \delta]}{2k_q}} \varepsilon_{q,t} = \alpha q_t + \beta + \gamma \varepsilon_{q,t}$$

où δ représente le pas de la série historique. La fréquence des données historiques de l'inflation est mensuelle et donc : $\delta = 1/12$.

L'estimation des paramètres est réalisée à partir de la méthode des moindres carrés. Le tableau suivant présente les valeurs estimées :

Tableau 8 : Calibrage du taux d'inflation

Espérance de l'inflation : μ_q	1,27 %
Écart type de l'erreur : σ_q	1,00 %
Vitesse retour à la moyenne : k_q	0,61

Bien que les résultats du calibrage soient cohérents avec les statistiques descriptives présentées au chapitre I-3, les paramètres présentés ne sont pas directement estimés mais sont issus d'une transformation non linéaire de ceux-ci. Cette approche est de nature à introduire un biais dans les paramètres et invite à apprécier la précision des estimateurs avec prudence.

Concernant les tests sur le modèle et les coefficients, nous obtenons les résultats suivants :

Tableau 9 : Tests sur le modèle de l'inflation

R^2 ajusté	85,64 %	Le pouvoir explicatif du modèle est de 85,64 %
P-value Fisher	< 2.2e-16	Le modèle global est significatif
Test Jarque-Bera	0	Hypothèse de normalité des résidus est vérifiée

Nous remarquons que l'ensemble des résultats et des tests valident les hypothèses posées *a priori* pour la validation du modèle et du calibrage.

2. Les taux d'intérêt

À titre de rappel, les formes récurrentes des dynamiques des taux d'intérêt réels longs et courts sont :

$$l_{t+\delta} = l_t \exp[-k_l \delta] + \mu_l (1 - \exp[-k_l \delta]) + \sigma_l \sqrt{\frac{1 - \exp[-2k_l \delta]}{2k_l}} \varepsilon_{l,t}$$

$$r_{t+\delta} = r_t \exp[-k_r \delta] + l_r (1 - \exp[-k_r \delta]) + \sigma_r \sqrt{\frac{1 - \exp[-2k_r \delta]}{2k_r}} \varepsilon_{r,t}$$

où δ représente le pas des séries historiques. La fréquence des séries historiques des taux d'intérêt est mensuelle et donc : $\delta = 1/12$.

L'estimation des paramètres est réalisée à partir de la méthode des doubles moindres carrés. Le tableau suivant présente les valeurs estimées des paramètres.

Tableau 10 : Calibrage du modèle des taux réels

Écart type de l'erreur du taux réel court : σ_r	1,00 %
Vitesse retour à la moyenne du taux réel court: k_r	0,10
Espérance de du taux réel long terme : μ_l	3,04 %
Écart type de l'erreur du taux réel long : σ_l	1,04 %
Vitesse retour à la moyenne du taux réel long: k_l	0,46

Comme pour l'inflation, on notera que les paramètres présentés ne sont pas les paramètres estimés mais sont issus d'une transformation non linéaire de ceux-ci ce qui invite à apprécier la précision des estimateurs avec prudence.

Le tableau suivant reprend les principaux résultats sur les tests d'adéquation réalisés pour la première étape des MCO (modèle sur les taux d'intérêt réels longs).

Tableau 11 : Tests sur le modèle du taux réel long

R^2 ajusté	91,90 %	Le pouvoir explicatif du modèle est de 91,90 %
P-value Fisher	< 2.2e-16	Le modèle global est significatif
Test Jarque-Bera	0	Hypothèse de normalité des résidus est vérifiée

Nous remarquons dans cette première étape que l'ensemble des résultats et des tests valident les hypothèses posées *a priori* pour la validation du modèle et du calibrage.

Le tableau suivant reprend les principaux résultats sur les tests d'adéquation réalisés pour la deuxième étape des MCO.

Tableau 12 : Tests sur le modèle du taux réel court

R^2 ajusté	-0,45 %	Le pouvoir explicatif du modèle est très faible
P-value Fisher	0,561	Le modèle global n'est pas significatif
Test Jarque-Bera	1	Hypothèse de normalité des résidus non vérifiée

Nous observons que ces tests ne valident pas l'ensemble des hypothèses du modèle et du calibrage. Il conviendra donc de considérer avec prudence l'ensemble de ces résultats et en particulier les intervalles de confiance lors des simulations et des projections.

3. Le taux de rendement d'un investissement en actions

Le rendement d'un investissement en actions s'écrit comme la somme de l'inflation, du taux d'intérêt court réel et de l'excès de rendement. Pour calibrer le modèle des actions il convient donc de calibrer l'excès de rendement.

L'approche consistant à considérer que l'excès de rendement suit une loi normale à un seul régime est retenue. L'excès de rendement est calibré sur l'indice SBF 250 dividende réinvestis. Le tableau suivant résume les résultats observés.

Tableau 13 : Calibrage de l'excès de rendement

Moyenne	3,60 %
Volatilité	26,38 %
Test Jarque-Bera	1

4. Le taux de rendement de l'immobilier

À titre de rappel, la forme récurrente de la dynamique du rendement de l'immobilier est :

$$Im_{t+\delta} = Im_t \exp[-k_{Im} \delta] + \mu_{Im} (1 - \exp[-k_{Im} \delta]) + \sigma_{Im} \sqrt{\frac{1 - \exp[-2k_{Im} \delta]}{2k_{Im}}} \varepsilon_{Im,t}$$

où δ représente le pas de la série historique. La fréquence de l'historique retenu pour l'immobilier est trimestrielle et donc : $\delta = 1/4$.

L'estimation des paramètres peut être réalisée à partir de la méthode des moindres carrés ordinaires. Cependant, contrairement au taux d'inflation et aux taux d'intérêt réels, le calibrage du rendement de l'immobilier sur l'historique datant du premier trimestre 1997 au second trimestre 2009 pose un problème au niveau de la vitesse de retour à la moyenne. En effet, cette valeur est négative et si elle est retenue pour les projections et les simulations on se retrouvera avec un rendement divergent à long terme. L'origine de ce problème réside dans la forte baisse des rendements pendant les deux derniers trimestres de 2009. En effet, le taux de rendement passe de -3,04 % au dernier trimestre de 2008 à -7,11 % et puis à -9,75 % lors des deux premiers trimestres de 2009. Ces dernières valeurs sont les deux valeurs minimales sur tout l'historique considéré et s'éloignent significativement de la tendance globale des autres valeurs historiques. Ainsi, ces deux valeurs sont retirées afin d'estimer les paramètres du modèle. Le tableau suivant présente les paramètres estimés :

Tableau 14 : Calibrage du modèle sur le rendement de l'immobilier

Espérance du taux de rendement de l'immobilier : μ_{Im}	5,80 %
Écart type de l'erreur : σ_{Im}	2,69 %
Vitesse retour à la moyenne : k_{Im}	0,09

Bien que les résultats du calibrage soient cohérents avec les statistiques descriptives présentées au chapitre I-3, les paramètres présentés ne sont pas directement estimés mais sont issus d'une transformation non linéaire de ceux-ci. Cette approche est de nature à introduire un biais dans les paramètres et invite à apprécier la précision des estimateurs avec prudence.

Concernant les tests sur le modèle et les coefficients, nous obtenons les résultats suivants :

Tableau 15 : Tests sur le modèle de l'immobilier

R^2 ajusté	91,63 %	Le pouvoir explicatif du modèle est de 91,63 %
P-value Fisher	< 2.2e-16	Le modèle global est significatif
Test Jarque-Bera	0	Hypothèse de normalité des résidus est vérifiée

Nous remarquons que l'ensemble des résultats et des tests sont cohérent avec les hypothèses posées *a priori* pour la validation du modèle et du calibrage.

5. Matrice de corrélation des résidus

La matrice suivante est la matrice de corrélations entre les résidus $\varepsilon_{q,t}$, $\varepsilon_{l,t}$, $\varepsilon_{r,t}$ et $\varepsilon_{lm,t}$ ainsi que l'excès de rendement x_t . Cette matrice servira pour la projection du modèle d'Ahlgrim dans le cas où l'on suppose que la structure de dépendance est décrite par la corrélation.

Tableau 16 : Matrice de corrélation des résidus et de l'excès de rendement

	Inflation	Taux long réel	Taux court réel	Excès de rendement	Immobilier
Inflation	100,00 %	-83,73 %	-79,87 %	18,87 %	25,01 %
Taux long réel	-83,73 %	100,00 %	69,22 %	-11,39 %	-22,39 %
Taux court réel	-79,87 %	69,22 %	100,00 %	12,19 %	-21,30 %
Excès de rendement	18,87 %	-11,39 %	12,19 %	100,00 %	21,73 %
Immobilier	25,01 %	-22,39 %	-21,30 %	21,73 %	100,00 %

Sur la base du calibrage présenté dans ce chapitre, les chapitres III-3 et III-4 illustrent les projections des indices financiers et macroéconomiques étudiés. Le chapitre suivant présente les projections dans le cas où la structure de dépendance des résidus et de l'excès de rendement est décrite par la corrélation. Cela est équivalent à dire que la structure de dépendance des indices financiers et macroéconomiques étudiés est décrite aussi par la corrélation puisque les relations qui lient ces indices aux résidus et à l'excès de rendement sont linéaires. Le chapitre III-4 est consacré aux projections avec la prise en compte de la structure de dépendance optimale.

Chapitre III-3

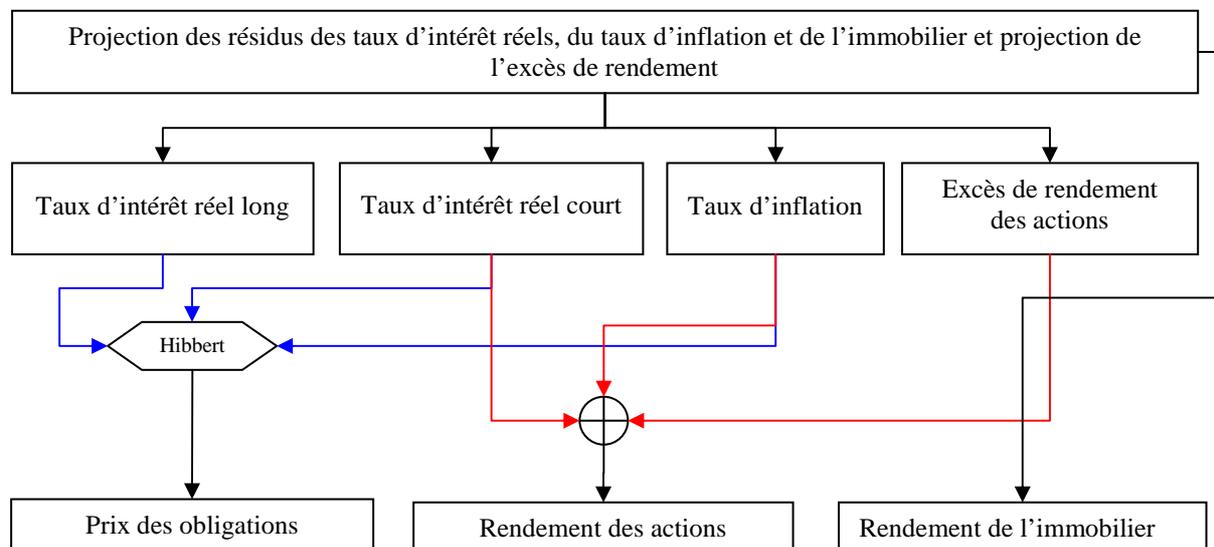
Projection du modèle d'Ahlgrim avec une structure de dépendance gaussienne

Cette section est relative à la projection des actifs sur la base du calibrage technique (aucun ajustement économique des paramètres estimés n'est réalisé) issu du modèle d'Ahlgrim. Ces projections portent sur trois classes d'actifs : le rendement des actions, les prix des zéro-coupon (en réel et en nominal) et le rendement de l'immobilier.

Ce chapitre présente les projections dans le cas où la structure de dépendance des résidus et de l'excès de rendement est décrite par la corrélation. Cela est équivalent à dire que la structure de dépendance des indices financiers et macroéconomiques étudiés est décrite aussi par la corrélation (et est donc gaussienne) puisque les relations qui lient ces indices aux résidus et à l'excès de rendement sont linéaires.

La démarche globale consiste à projeter dans un premier temps les résidus du taux d'inflation, des taux d'intérêt réels et de l'immobilier ainsi que l'excès de rendement. Ensuite les projections des indices financiers et macroéconomiques s'en déduisent par les formules présentées au chapitre III-1. La démarche peut se schématiser par la figure suivante :

Figure 24 : Schéma de la démarche de projection à partir des résidus et de l'excès de rendement



1. Paramètre d'initialisation des projections

Toutes les projections présentées sont une moyenne des valeurs obtenues à partir de $N = 100\,000$ simulations et sont réalisées sur une période de 10 ans à fréquence trimestrielle. Par ailleurs, pour l'ensemble des projections, on suppose que la date initiale, $t = 0$, correspond à la fin du mois de juin 2009. Les valeurs initiales retenues sont donc :

Tableau 17 : Valeurs initiales

Indices	Valeurs initiales
Taux d'inflation	-0,49 %
Taux d'intérêt réel à long terme	4,45 %
Taux d'intérêt réel à court terme	1,18 %
Taux de rendement de l'immobilier	-9,75 %

En outre, pour le prix des zéro-coupon, on considère que l'on souhaite évaluer à chaque instant le prix de deux obligations zéro-coupon : une qui paie 1 en réel à sa maturité restante égale à 5 ans, et une autre qui paie 1 en nominal à sa maturité restante égale à 5 ans.

2. Projection des résidus et de l'excès de rendement

Lors des projections, il convient de tenir compte des coefficients de corrélation observés entre les résidus des modèles calibrés ($\varepsilon_{q,t}$, $\varepsilon_{l,t}$, $\varepsilon_{r,t}$, $\varepsilon_{\text{Im},t}$) et l'excès de rendement des actions x_t . Cette matrice est présentée au chapitre III-2.

Les résidus et l'excès de rendement des actions suivent des lois normales. Si on suppose que la structure de dépendance de ces lois est gaussienne, la simulation de ces cinq variables revient à simuler cinq variables gaussiennes dont la matrice de corrélation est égale à la matrice de corrélation historique. Numériquement cela revient à simuler cinq variables normales indépendantes, et ensuite, en déduire les trajectoires des résidus et l'excès de rendement en passant par la factorisation de Cholesky.

3. Projection de l'inflation et des taux nominaux

La simulation des trajectoires des résidus permet de trouver les projections du taux d'inflation et des taux d'intérêt réels en utilisant les formes discrètes présentées au chapitre III.1.

Sur la base des valeurs initiales présentées au début de ce chapitre et sur la base du calibrage statistique du chapitre III-2, on obtient les projections suivantes pour le taux d'inflation et les taux d'intérêt réels :

Figure 25 : Projection du taux d'inflation

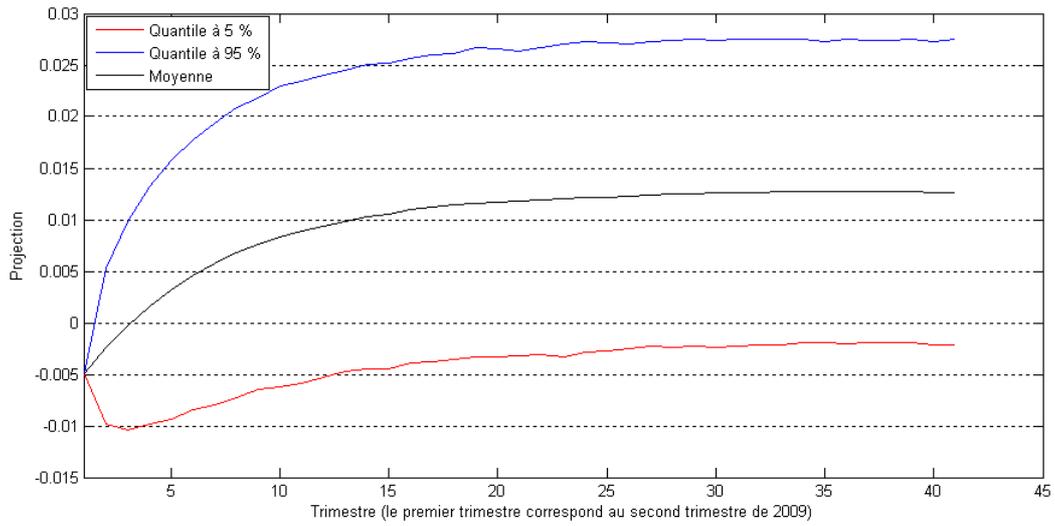


Figure 26 : Projection du taux d'intérêt réel à long terme

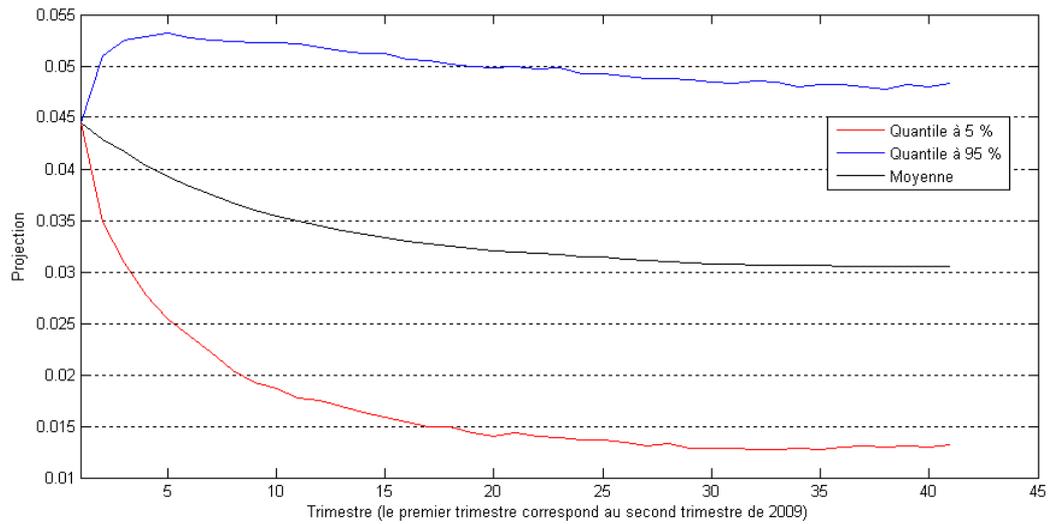
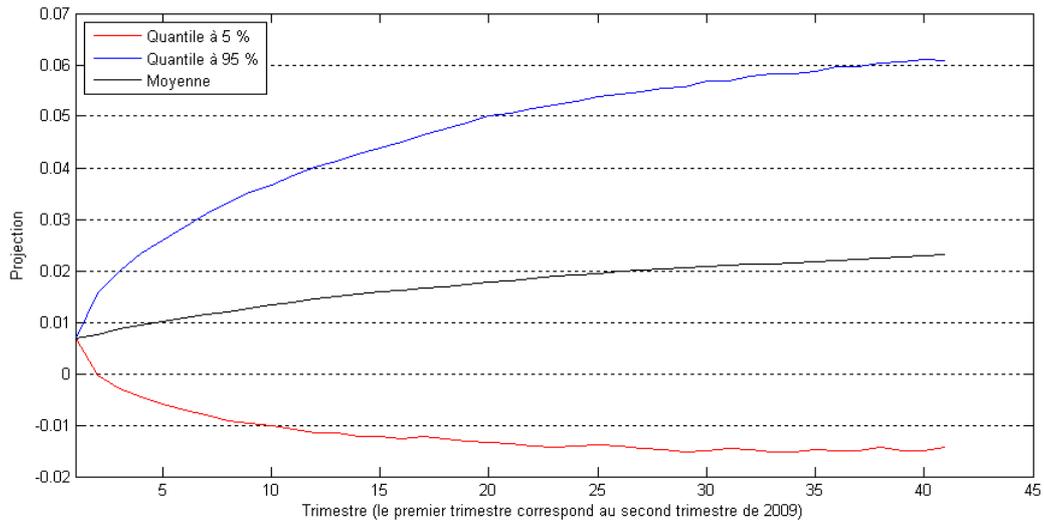


Figure 27 : Projection du taux d'intérêt réel à court terme



Nous observons que les courbes moyennes des trois indices partent de leurs valeurs initiales pour rejoindre le niveau des moyennes à long terme avec une vitesse égale à la vitesse de retour à la moyenne. La projection des taux d'intérêt nominaux à court terme est une combinaison de la projection des taux d'inflation et des taux d'intérêt réels. Les figures suivantes illustrent les projections des taux d'intérêt nominaux à long et à court terme.

Figure 28 : Projection du taux d'intérêt nominal à long terme

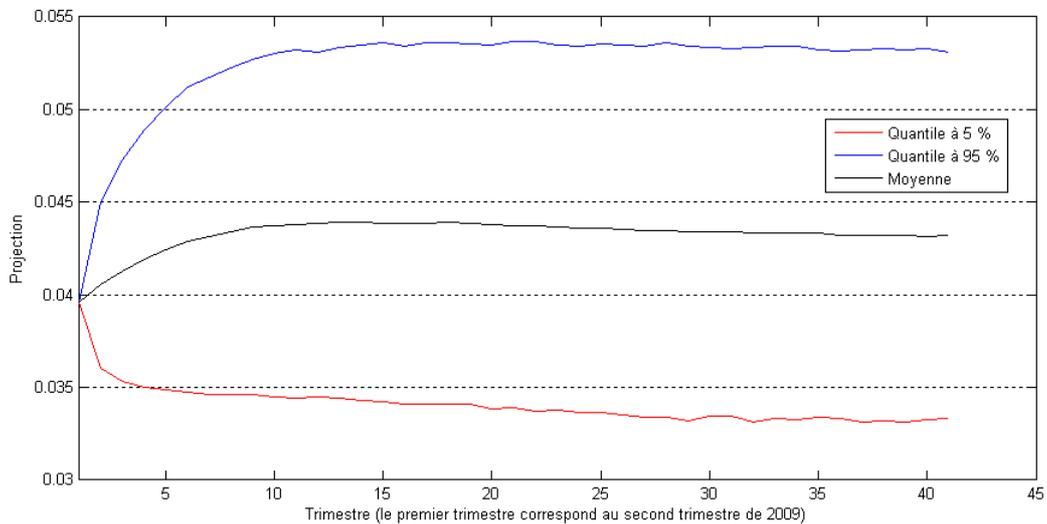
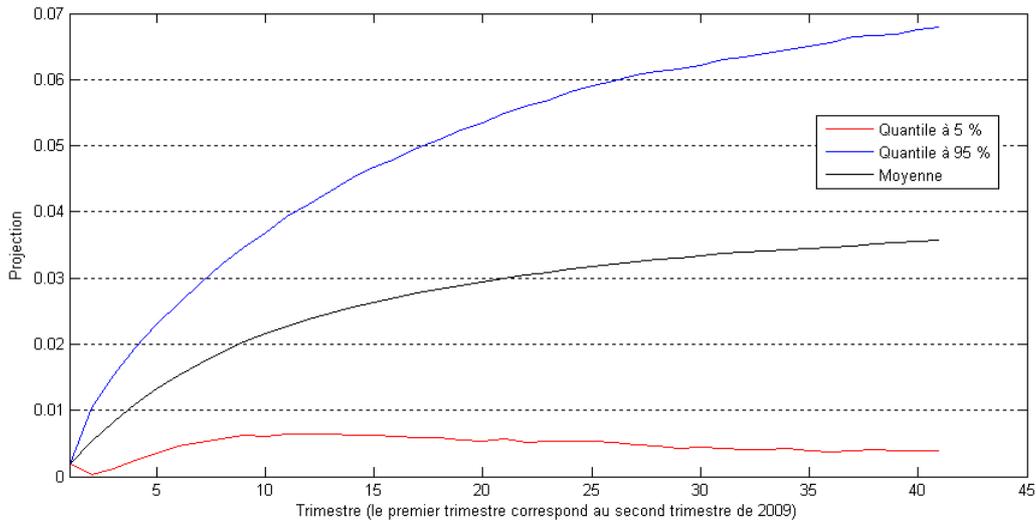


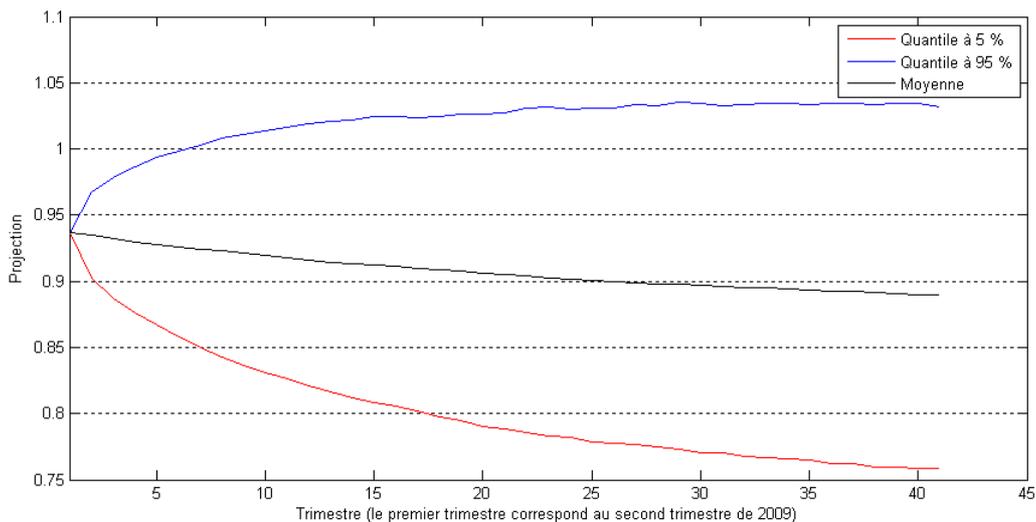
Figure 29 : Projection du taux d'intérêt nominal à court terme



4. Projection du prix d'une obligation zéro-coupon

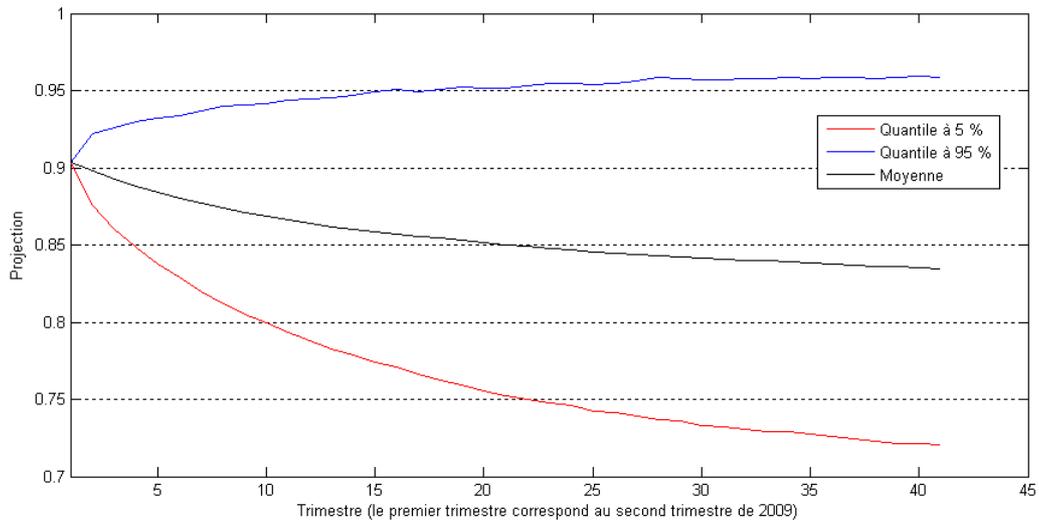
Dans le cadre du modèle d'Ahlgrim, le prix d'une obligation zéro-coupon en réel se déduit des paramètres des modèles de taux à deux facteurs (vitesses de retour à la moyenne, moyennes de long terme, écart types) à partir des formules explicites présentées au chapitre III-1. La projection illustrée dans la figure suivante correspond, à chaque date, au prix d'une obligation réelle payant 1 à sa maturité restante égale à 5 ans.

Figure 30 : Projection du prix d'une obligation zéro-coupon réelle à maturité 5 ans



Pour évaluer le prix d'une obligation zéro-coupon en nominal, il convient de tenir compte de l'inflation. En pratique, pour établir le prix d'un zéro-coupon en nominal, nous reprenons la relation suivante, présentée dans Ahlgrim et al. [2005] et au chapitre III.1 : $P_{nom}(t, T) = P_{réel}(t, T)P_{inf}(t, T)$. La projection illustrée dans la figure suivante correspond, à chaque date, au prix d'une obligation nominale payant 1 à sa maturité restante égale à 5 ans.

Figure 31 : Projection du prix d'une obligation zéro-coupon nominale à maturité 5 ans



5. Projection du taux de rendement d'un investissement en actions

Pour le modèle d'Ahlgrim et al, le taux de rendement des actions s'écrit sous la forme :

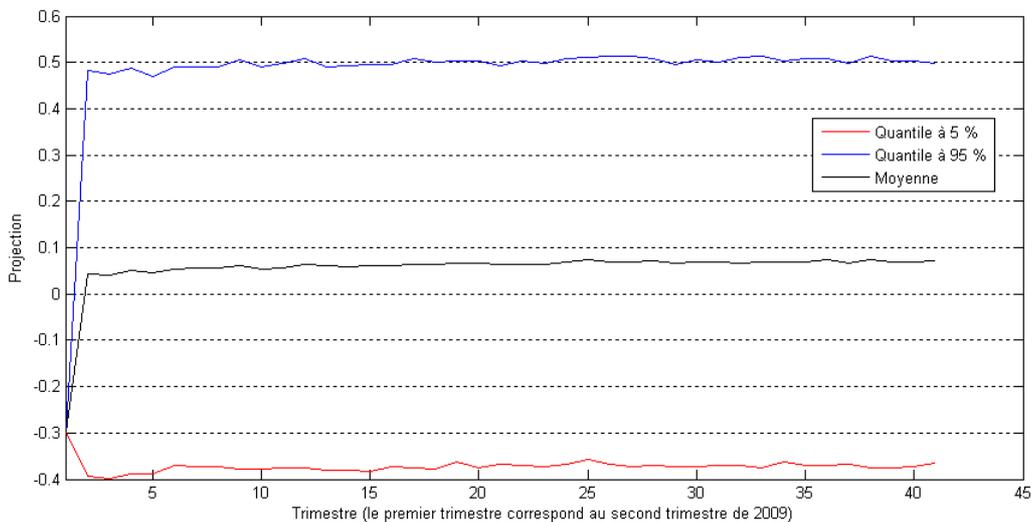
$$s_t = r_t + q_t + x_t$$

où :

- $r_t + q_t$: le taux d'intérêt nominal.
- x_t : excès de rendement attribuable à l'appréciation du capital.

Dans le cadre de l'approche alternative présentée au chapitre III-1, l'excès de rendement suit une loi normale. On obtient donc pour le rendement espéré des actions :

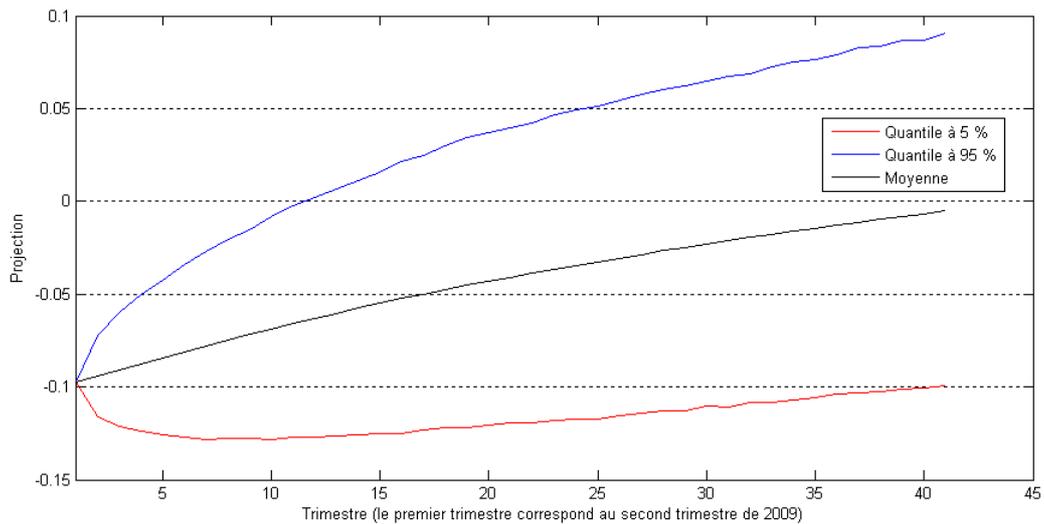
Figure 32 : Projection du rendement des actions



6. Projection du taux de rendement de l'immobilier

Sur la base de la valeur initiale présentée au début de ce chapitre et sur la base du calibrage statistique du chapitre III-2, on obtient la projection suivante pour le taux de rendement de l'immobilier :

Figure 33 : Projection du taux de rendement de l'immobilier



Nous observons que le rendement de l'immobilier rejoint lentement sa moyenne à long terme. Cela s'explique par la valeur relativement faible de sa vitesse de retour à la moyenne.

Projections du modèle d’Ahlgrim en prenant en compte la structure de dépendance optimale : la copule de Cook-Johnson

Le chapitre précédent présente une méthode de projection du modèle d’Ahlgrim en supposant que la structure de dépendance des indices macroéconomiques et financiers étudiés correspond à la copule gaussienne. Cela revient à considérer que la dépendance du taux d’inflation, des taux d’intérêt nominaux, du rendement des actions et du rendement de l’immobilier est totalement décrite par leur matrice de corrélation.

Cependant, les résultats présentés au chapitre II-4 montrent que cette structure de dépendance n’est pas adaptée aux indices macroéconomiques et financiers étudiés et que la copule de Cook-Johnson décrit mieux cette structure.

L’objet de ce chapitre est d’expliquer, dans un premier temps, une démarche qui permet de simuler des variables aléatoires dont la structure de dépendance est une copule paramétrique prédéfinie. Nous expliquons dans un second temps la procédure qui nous permet de projeter le modèle d’Ahlgrim en prenant en compte la copule de Cook-Johnson. Les dernières sections présentent et commentent les résultats de ces projections.

1. Simulation des copules paramétriques : Généralité

Dans le cas où les fonctions de répartition marginales F_1, \dots, F_d et une copule $C(u_1, \dots, u_d)$ sont spécifiées, on peut construire un unique modèle multi-varié dont la distribution jointe est $F(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d))$. Nous essayons au long de cette section d'expliquer comment on peut simuler les réalisations d'un vecteur aléatoire (X_1, \dots, X_d) dont les fonctions de répartition marginales sont respectivement F_1, \dots, F_d et dont la structure de dépendance est la copule C . Globalement, la démarche que nous proposons consiste à simuler dans un premier temps les trajectoires d'un vecteur aléatoire de variables uniformes (U_1, \dots, U_d) dont la fonction de répartition est la copule C . Ensuite, les trajectoires simulées de (X_1, \dots, X_d) sont obtenues en s'appuyant sur la propriété selon laquelle le vecteur aléatoire $(F_1^{-1}(U_1), \dots, F_d^{-1}(U_d))$ admet l'unique fonction de répartition jointe F qui vérifie : $F(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d))$ (et donc $(X_1, \dots, X_d) \approx (F_1^{-1}(U_1), \dots, F_d^{-1}(U_d))$).

Disposant des lois marginales, on remarque ainsi que la difficulté se résume à la simulation d'un vecteur de variables uniformes dont la fonction de répartition est C . Dans ce qui suit, on suppose que l'on dispose de la forme paramétrique de la copule C à l'image de la copule de Cook-Johnson et les autres copules paramétriques présentées au chapitre II-1. La méthode que nous proposons repose sur des simulations récursives utilisant les distributions uni-variées conditionnelles (cf. Embrechts et al. [2002]).

On considère le cas général $d \geq 2$ et on introduit la notation suivante pour tout $2 \leq i \leq d-1$:

$$C_i(u_1, \dots, u_i) = C(u_1, \dots, u_i, 1, \dots, 1)$$

On écrit aussi $C_1(u_1) = u_1$ et $C_d(u_1, \dots, u_d) = C(u_1, \dots, u_d)$.

Embrechts et al. [2002] proposent de considérer les distributions conditionnelles pour simuler les trajectoires d'un vecteur de variables aléatoires uniformes (U_1, \dots, U_d) . En effet, si l'on suppose que la fonction de répartition jointe de (U_1, \dots, U_d) est C alors (cf. Embrechts et al. [2002]) :

$$C_i(u_i / u_1, \dots, u_{i-1}) = P(U_i \leq u_i / U_1 = u_1, \dots, U_{i-1} = u_{i-1}) = \frac{\partial^{i-1} C_i(u_1, \dots, u_i)}{\partial u_1 \dots \partial u_{i-1}} \bigg/ \frac{\partial^{i-1} C_{i-1}(u_1, \dots, u_{i-1})}{\partial u_1 \dots \partial u_{i-1}}$$

Dans le cas où l'on dispose de la forme paramétrique de la copule, le calcul du numérateur et du dénominateur nous permet de disposer de $C_i(u_i / u_1, \dots, u_{i-1})$ pour tout $2 \leq i \leq d-1$. Dans ce cas, nous utilisons l'algorithme suivant pour simuler une trajectoire de (U_1, \dots, U_d) dont la fonction de répartition est la copule C :

- Simuler une valeur u_1 dont la loi est uniforme sur $[0,1]$ notée dans la suite $U(0,1)$.
- Simuler une valeur u_2 dont la loi est $C_2(u_2 / u_1)$.
- Continuer la même procédure.
- Simuler une valeur u_d de loi $C_d(u_d / u_1, \dots, u_{d-1})$.

En général pour simuler une valeur de loi $C_i(u_i/u_1, \dots, u_{i-1})$ on simule une valeur u de loi uniforme $U(0,1)$ et on calcule ensuite $C_i^{-1}(u/u_1, \dots, u_{i-1})$.

Nous avons retenu cette démarche pour la simulation de la copule de Cook-Johnson qui représente la structure de dépendance la plus adaptée aux indices financiers et macroéconomiques étudiés. Nous avons utilisé ensuite les fonctions de répartition de ces indices pour simuler leurs trajectoires. Avant de présenter les résultats des simulations, la section suivante tente de présenter brièvement la procédure de simulation du modèle d'Ahlgrim en prenant en compte la copule de Cook-Johnson.

2. Projection du modèle d'Ahlgrim et la copule de Cook-Johnson

Pour simuler les trajectoires du taux d'inflation, des taux d'intérêt nominaux, des rendements des actions et du rendement de l'immobilier en combinant le modèle d'Ahlgrim et la copule de Cook-Johnson calibrée sur les données historiques de ces indices (cf. chapitre II-4), nous avons procédé en deux étapes.

La première étape consiste à simuler $N = 100\,000$ trajectoires de variables uniformes (U_1, \dots, U_5) dont la fonction de répartition est la copule de Cook-Johnson.

La seconde étape consiste à inverser ces valeurs numériquement en utilisant les fonctions de répartition des indices financiers et macroéconomiques étudiés en s'appuyant sur le modèle d'Ahlgrim.

À titre de rappel, les indices modélisés dans ce modèle suivent tous une loi normale. Il est rappelé aussi que dans le cadre de ce modèle il faut combiner les lois des taux d'intérêt réels et la loi de l'inflation pour disposer des lois des taux nominaux. Il s'agit d'opérations additives simples qui nous permettent à la fin de disposer des écritures récurrentes des taux nominaux et aussi du rendement des actions en introduisant les corrélations historiques entre les résidus et l'excès de rendement.

Pour illustrer notre démarche de simulation, nous expliquons ci-après comment simuler une trajectoire de l'inflation en partant de sa formule récurrente.

Considérons le modèle d'Ahlgrim pour l'inflation et supposons que l'on dispose à l'instant t de la valeur simulée de l'inflation ainsi que la valeur $u_1^{t+\delta}$ d'une variable uniforme simulée à partir de la copule de Cook-Johnson :

$$q_{t+\delta} = q_t \exp[-k_q \delta] + \mu_q (1 - \exp[-k_q \delta]) + \sigma_q \sqrt{\frac{1 - \exp[-2k_q \delta]}{2k_q}} \varepsilon_{q,t} = \alpha q_t + \beta + \gamma \varepsilon_{q,t}$$

où $\delta = \frac{1}{4}$ représente le pas de la discrétisation.

Alors la valeur simulée de l'inflation à l'instant $t + \delta$ est l'inverse de la fonction de répartition de la loi normale $N(\alpha q_t + \beta, \gamma^2)$ de $u_1^{t+\delta}$.

La section suivante présente les projections des indices financiers et macroéconomiques étudiés.

3. Projection des indices financiers et macroéconomiques

Comme pour le chapitre précédent, toutes les projections présentées sont une moyenne des valeurs obtenues à partir de $N = 100\,000$ simulations et sont réalisées sur une période de 10 ans avec une

fréquence trimestrielle. Par ailleurs, pour l'ensemble des projections, on suppose que la date initiale, $t = 0$, correspond à la fin du mois de juin 2009. Les valeurs initiales retenues sont donc :

Tableau 18 : Valeurs initiales

Indices	Valeurs initiales
Taux d'inflation	-0,49 %
Taux d'intérêt réel à long terme	4,45 %
Taux d'intérêt réel à court terme	1,18 %
Taux de rendement de l'immobilier	-9,75 %

En outre, pour le prix des zéro-coupon, on considère que l'on souhaite évaluer à chaque instant le prix de deux obligations zéro-coupon : une qui paie 1 en réel à sa maturité restante égale à 5 ans, et une autre qui paie 1 en nominal à sa maturité restante égale à 5 ans.

3.1. Projection de l'inflation et des taux nominaux

Sur la base des valeurs initiales présentées au début de ce chapitre et sur la base du calibrage statistique du chapitre III-2, on obtient les projections suivantes pour le taux d'inflation et les taux d'intérêt réels et nominaux. Contrairement au chapitre précédent, où les projections des taux nominaux se déduisent des projections de l'inflation et des taux réels, les taux nominaux sont directement simulés en utilisant la copule de Cook-Johnson. Les projections des taux réels s'en déduisent en retranchant les valeurs projetées de l'inflation aux valeurs projetées des taux nominaux. Pour mémoire, la copule de Cook-Johnson correspond à la structure de dépendance de l'inflation, des taux nominaux, du rendement des actions et de l'immobilier et donc il faut retenir les taux nominaux lors des projections.

Figure 34 : Projection du taux d'inflation

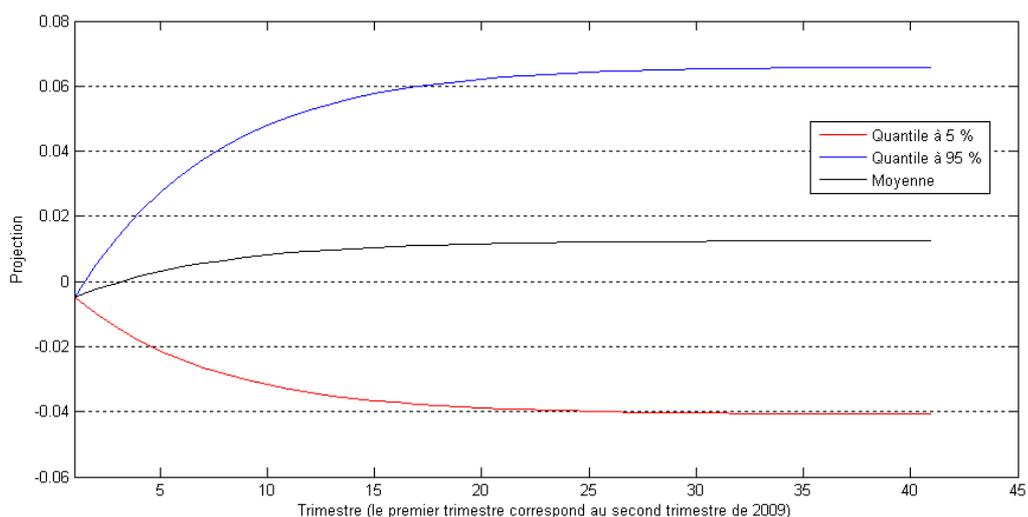


Figure 35 : Projection du taux d'intérêt réel à long terme

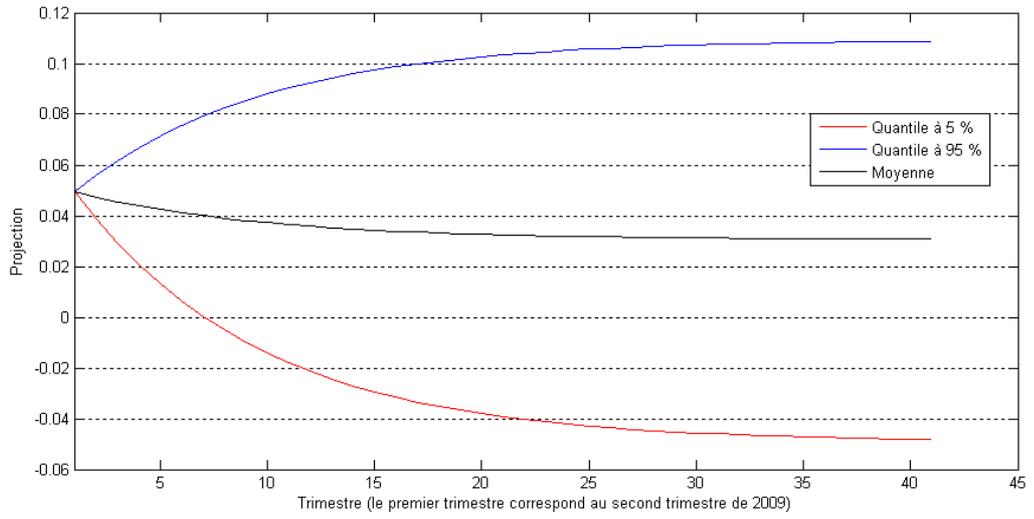


Figure 36 : Projection du taux d'intérêt réel à court terme

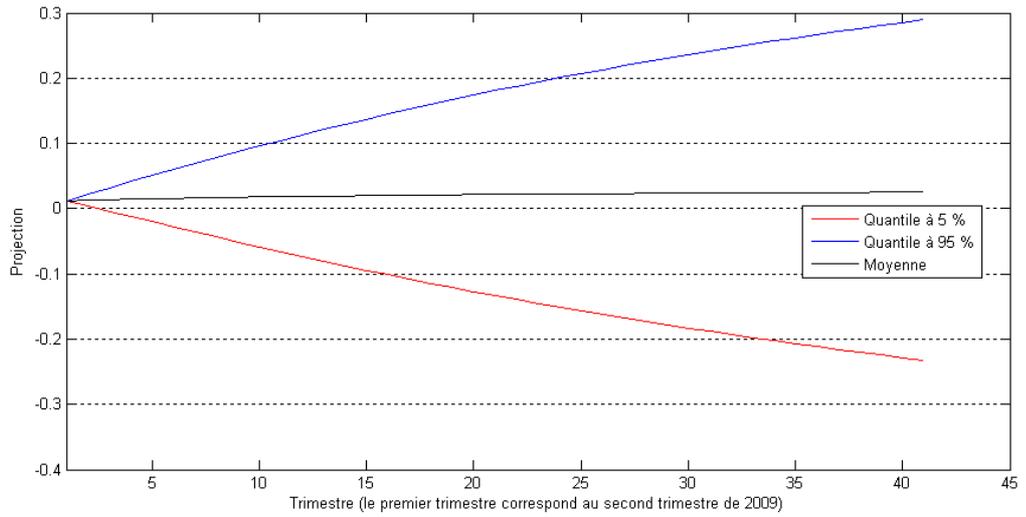


Figure 37 : Projection du taux d'intérêt nominal à long terme

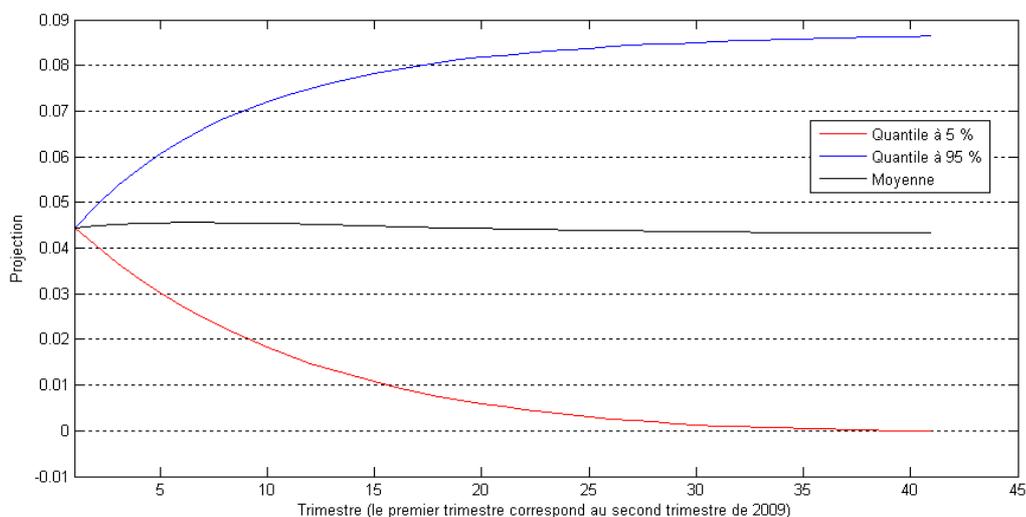
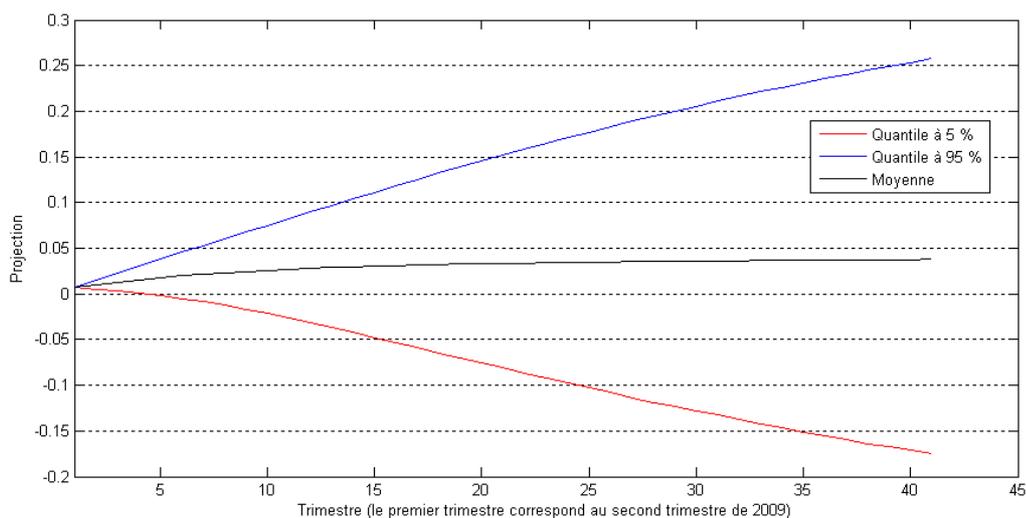


Figure 38 : Projection du taux d'intérêt nominal à court terme



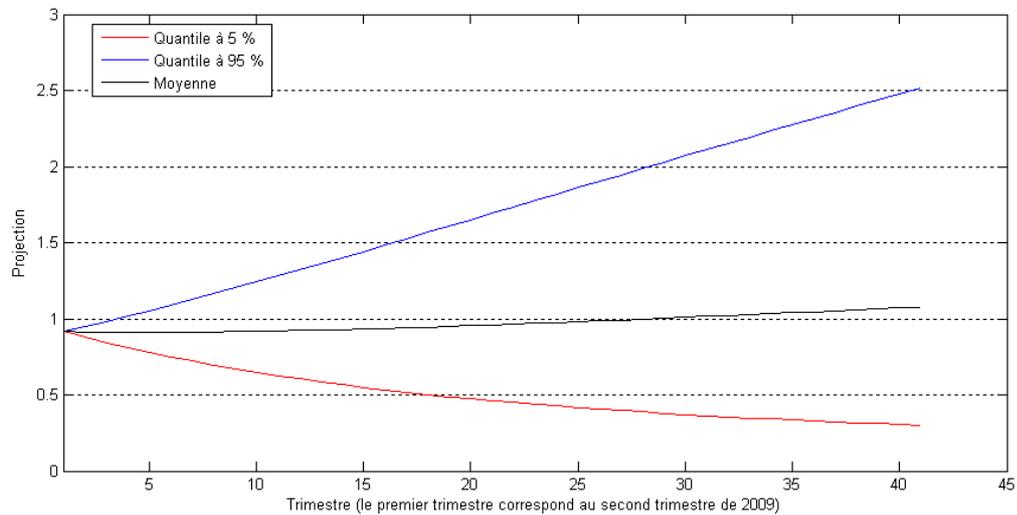
On observe que les courbes des moyennes des taux d'intérêt réels et de l'inflation partent de leurs valeurs initiales pour rejoindre le niveau des moyennes à long terme avec une vitesse égale à la vitesse de retour à la moyenne. On observe aussi que, dans le cadre de la simulation du modèle d'Ahlgrim en introduisant la copule de Cook-Johnson, les valeurs moyennes projetées du taux d'inflation et des taux d'intérêt (nominaux et réels) sont identiques à celles observées avec une structure de dépendance gaussienne (cela peut ne pas être visible sur certains graphes à cause du changement d'échelle). Par contre les quantiles d'ordre 95 % et 5 % sont significativement différents.

3.2. Projection du prix d'une obligation zéro-coupon en réel et en nominal, du rendement des actions et du rendement de l'immobilier

Dans le cadre du modèle d'Ahlgrim, le prix d'une obligation zéro-coupon en réel se déduit des paramètres des modèles des taux à deux facteurs (vitesses de retour à la moyenne, moyennes de long

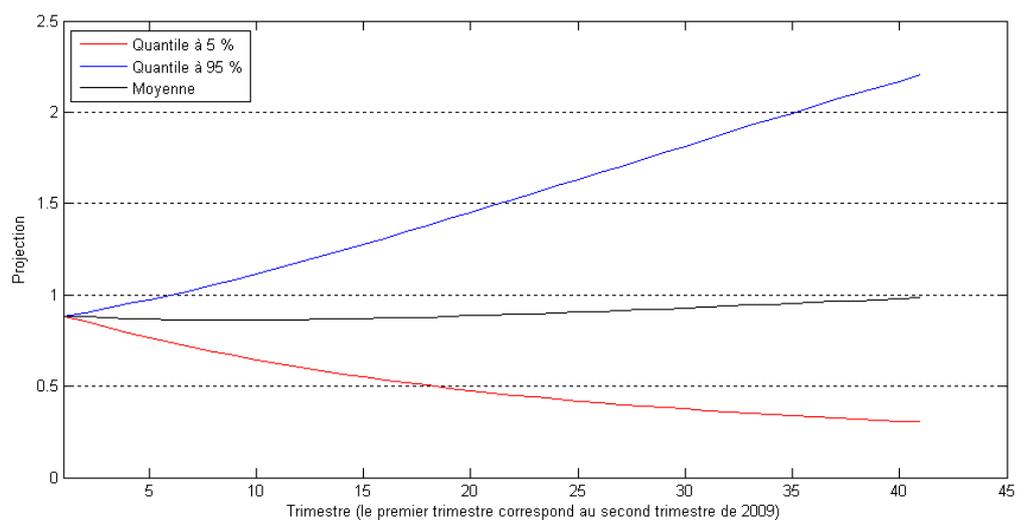
terme, écart types) à partir des formules explicites et des hypothèses présentées au chapitre III-1. La projection illustrée dans la figure suivante correspond, à chaque date, au prix d'une obligation réelle payant 1 à sa maturité restante égale à 5 ans.

Figure 39 : Projection du prix d'une obligation zéro-coupon réelle à maturité 5 ans



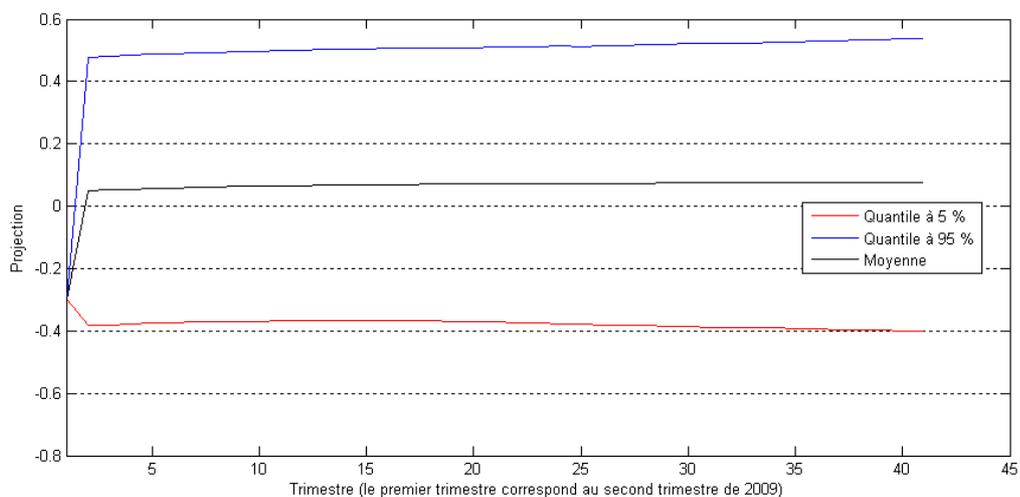
Pour évaluer le prix d'une obligation zéro-coupon en nominal, il convient de tenir compte de l'inflation. En pratique, pour établir le prix d'un zéro-coupon en nominal, nous reprenons la relation suivante, présentée dans Ahlgrim et al. [2005] et au chapitre III.1 : $P_{nom}(t,T) = P_{réel}(t,T)P_{inf}(t,T)$. La projection illustrée dans la figure suivante correspond, à chaque date, au prix d'une obligation nominale payant 1 à sa maturité restante égale à 5 ans.

Figure 40 : Projection du prix d'une obligation zéro-coupon nominale à maturité 5 ans



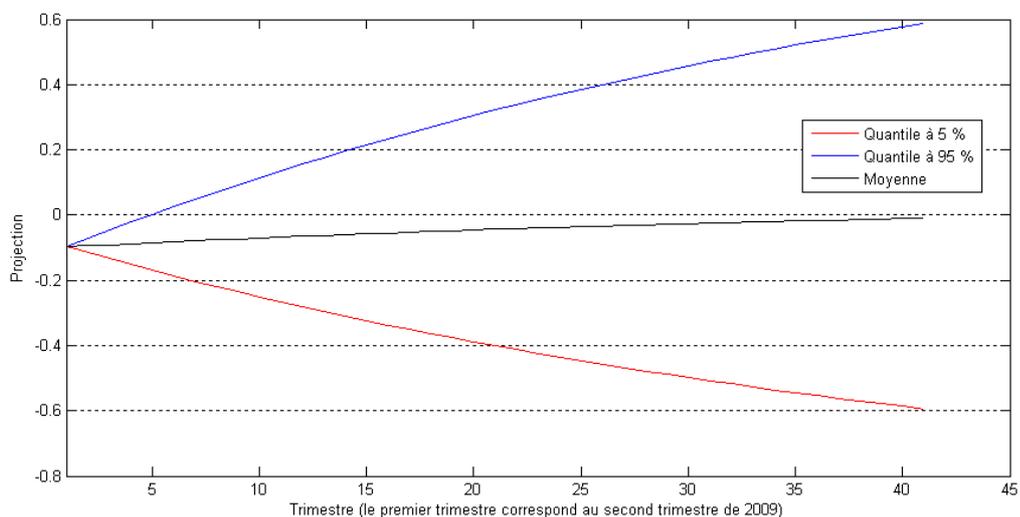
En ce qui concerne le rendement des actions et dans le cadre de l'approche alternative présentée au chapitre III-1, on obtient la projection suivante des rendements moyens espérés :

Figure 41 : Projection du rendement des actions



Sur la base de la valeur initiale présentée au début de ce chapitre et sur la base du calibrage statistique du chapitre III-2, on obtient la projection suivante pour le taux de rendement de l'immobilier :

Figure 42 : Projection du taux de rendement de l'immobilier



Nous observons que le rendement de l'immobilier rejoint lentement sa moyenne à long terme. Cela s'explique par la valeur relativement faible de sa vitesse de retour à la moyenne.

4. Comparaison de la copule gaussienne et la copule de Cook-Johnson dans le cadre du modèle d'Ahlgrim

Globalement, nous observons que, dans le cadre de la simulation du modèle d'Ahlgrim en introduisant la copule de Cook-Johnson, les valeurs moyennes projetées du taux d'inflation, des taux d'intérêt (nominaux et réels), du rendement des actions et du rendement de l'immobilier sont identiques à celles observées avec une structure de dépendance gaussienne (cela peut ne pas être visible sur certains graphes à cause du changement d'échelle). Ce constat est rassurant et tout à fait logique. En effet, la valeur moyenne de chaque indice correspond à l'espérance mathématique de sa loi marginale et est donc indépendante de la structure de dépendance.

Cependant, les quantiles d'ordre 95 % et 5 % sont significativement différents. On observe en particulier que l'intervalle formé par ces deux quantiles est plus large dans le cas où la structure de dépendance est décrite par la copule de Cook-Johnson. Les valeurs extrêmes sont donc mieux représentées par cette structure.

D'autre part, dans le cadre de la gestion d'actifs et afin de comparer les deux structures de dépendance, l'étude du rapport des deux VaR à 99,5 %, correspondant aux deux structures de dépendance combinées avec le modèle d'Ahlgrim, d'un portefeuille générique géré en visant une allocation cible, est un indicateur qui peut se révéler pertinent. En effet, les assureurs devront, à partir de 2012, estimer le niveau de fonds propre pour ne pas être en ruine sur un horizon d'un an avec une probabilité de 99,5 %. La comparaison des VaR à 99,5 % permettra de vérifier si la structure de dépendance gaussienne sous-estime la VaR du portefeuille (dans ce cas, les fonds à mobiliser sont surestimés), ou bien si elle la surestime (dans ce cas les fonds à mobiliser sont sous-estimés et l'organisme peut se retrouver contraint de faire un apport supplémentaire en termes de fonds propres). Avant de se lancer dans cette comparaison, nous proposons une brève présentation de ce que nous désignons par la VaR.

4.1. La Value at Risk

La VaR (de l'anglais *Value at Risk*) est une notion utilisée généralement pour mesurer le risque de marché d'un portefeuille d'instruments financiers. Elle correspond au montant de pertes qui ne devrait être dépassé qu'avec une probabilité donnée sur un horizon temporel donné.

Dans sa forme la plus générale, la VaR peut être déduite de la distribution de probabilité de la valeur future de l'actif à un niveau de confiance défini. Ce que nous cherchons dans le cadre de notre problématique est la valeur du portefeuille telle que la probabilité d'y être supérieur correspond au niveau de confiance choisi.

Etant donnée un niveau de confiance $\alpha \in [0,1]$, nous désignons dans la suite¹³ par la VaR de la valeur X d'un portefeuille au niveau de confiance α le montant suivant (la pire valeur à un niveau de confiance α) :

$$VaR_\alpha = \sup\{p \in \mathfrak{R}, P(X > p) \geq \alpha\}$$

Cette définition est équivalente à celle communément utilisée dans les milieux financiers et qui consiste à calculer le montant de perte maximale à un niveau de confiance prédéfini (car la perte maximale à un horizon donné est la différence entre la pire valeur du portefeuille X à cet horizon et la valeur initiale). Dans l'hypothèse que la distribution du portefeuille soit continue, la VaR_α correspond au quantile d'ordre $1 - \alpha$.

¹³ Cf. Kharoubi-Rakotomalala [2008]

Par ailleurs, si le portefeuille suit une loi normale, la VaR peut être directement calculée à partir de l'écart type de cette loi. Cependant, en dehors du cadre gaussien, il est rarement possible de trouver une formule analytique fermée pour la VaR.

Dans le cadre de nos travaux, nous nous posons dans un cadre non gaussien et nous sommes amenés donc à recourir à la simulation de Monte-Carlo de la valeur du portefeuille pour calculer sa VaR à 99,5 %.

4.2. Composition du portefeuille et horizon de comparaison

Afin de se faire une idée sur l'impact de la structure de dépendance gaussienne sur la valeur d'un portefeuille géré par l'assureur, nous proposons dans la suite d'étudier un cas simple où ce portefeuille est composé de trois actifs : obligations zéro-coupons de maturité 5 ans, actions et immobilier. Nous supposons que le fonds est réparti équitablement entre les actifs moins risqués (obligations) et les actifs plus risqués (actions et immobilier). Le fonds se compose finalement de :

- 50 % d'obligations de maturité 5 ans.
- 40 % d'actions diversifiées sur le marché français.
- 10 % d'immobilier. Pour calculer les performances de l'immobilier et afin de se mettre dans un univers plus réaliste, nous avons ajusté la valeur initiale à 5 % au lieu de -9,75 %.

Aussi, nous considérons, sans perte de généralité, que la valeur du portefeuille à la date initiale est de 1. Nous supposons également qu'à chaque trimestre le gestionnaire liquide tous les actifs à sa disposition (obligations de maturité 4,75 ans, actions et immobilier) et en achète des nouveaux avec les mêmes proportions définies ci-dessus (50 % d'obligations de maturité 5 ans, 40 % d'actions et 10 % d'immobilier). Le rendement du portefeuille correspond au rapport diminué de 1 du prix de vente des actifs sur leurs prix d'achat au semestre précédant la date de la vente.

Par ailleurs, nous proposons de comparer les VaR des deux approches (approche par la structure de dépendance gaussienne et approche par Cook-Johnson) sur un horizon de projection de 3 ans.

La comparaison sur un horizon à long terme (10 ans) peut être biaisée. En effet, l'incertitude autour de la valeur moyenne du portefeuille augmente avec l'horizon de projection. L'une des raisons de ce constat est la dépendance des performances des actifs du chemin simulé (*path dependant*)¹⁴. Cela se traduit par une VaR à 99,5 % plus importante qui a tendance à prendre des valeurs extrêmes quand on s'éloigne de la date initiale. En effet, pour les deux approches, la VaR à 99,5 % à un horizon à long terme (sur une période de 10 ans par exemple) correspond aux quantiles d'ordres 0,5 % de la valeur du fonds et correspond donc à un scénario défavorable sur l'ensemble de la période de projection (car la valeur du portefeuille est *path dependant*). Or, pour que ce scénario se réalise, il faudrait que les marchés financiers subissent une crise qui dure tout au long de cette période, ce qui est rarement observé.

Donc, il est plus pertinent de comparer les deux approches sur un horizon homogène avec la durée moyenne des périodes de crise (de l'ordre de 3 ans lors de la crise financière de 2000) et avec l'horizon d'estimation d'un an de la VaR à 99,5 % des réserves d'un organisme assureur. L'horizon de 3 ans semble répondre à ces critères.

¹⁴ Le prix des obligations, le rendement des actions et le rendement de l'immobilier sont évalués à partir de relation récurrentes entre les variables du modèle d'Ahlgrim et donc la valeur du portefeuille est *path dependant*.

4.3. Comparaison des deux VaR du portefeuille évalué par les deux approches

La figure suivante illustre les valeurs moyennes du portefeuille et les VaR à 99,5 % pour les deux approches : projection du modèle d'Ahlgrim avec la corrélation et la projection du modèle d'Ahlgrim avec la copule de Cook-Johnson. Le rapport des deux VaR est aussi illustré sur la figure.

Pour les deux approches et à chaque trimestre t , la VaR du portefeuille est calculée en deux étapes :

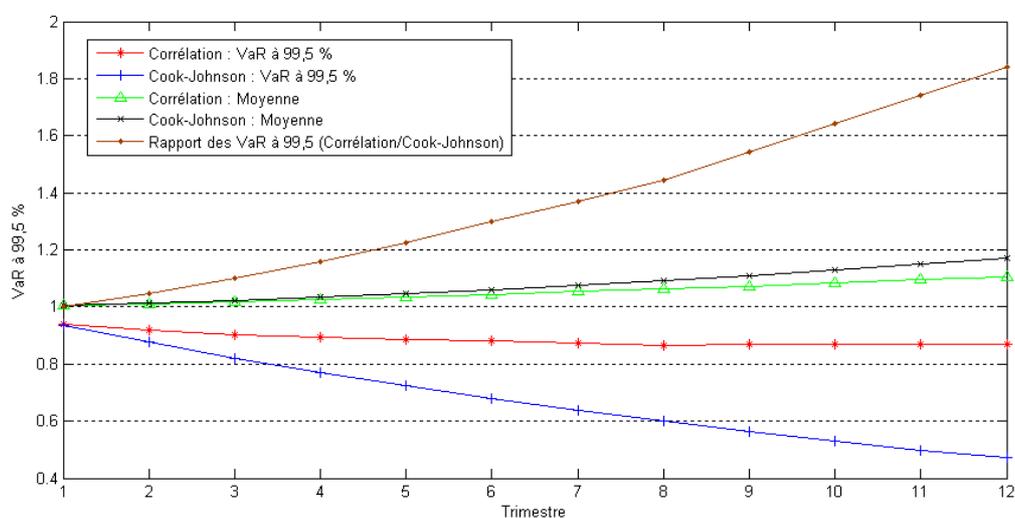
- A chaque trimestre (parmi les 12 correspondant à un horizon de 3 ans) et pour chaque simulation, la valeur du portefeuille est évaluée.
- On calcule ensuite le quantile à 0,5 % de la valeur du portefeuille à chaque trimestre. Ce quantile est la VaR du portefeuille sur la période $[0, t]$ et correspond à la pire valeur du portefeuille au niveau 99,5 % au trimestre t pour un investissement initial de 1 à l'instant 0.

Le rapport des VaR au trimestre t correspond au rapport des deux VaR du portefeuille calculées sur la période $[0, t]$ et évaluées par les deux approches :

$$R(t) = \frac{VaR_{[0,t]}^{Gauss}}{VaR_{[0,t]}^{Cook-Johnson}}$$

Les courbes des VaR à 99,5 % et la courbe du rapport des VaR sont présentées dans la figure ci-après sur un horizon de 3 ans.

Figure 43 : Comparaison des VaR à 99,5 % sur un horizon de 3 ans



Nous observons que les espérances moyennes de la valeur du portefeuille évaluées par les deux approches sont peu différentes. Cependant, nous observons que la VaR à 99,5 % du portefeuille évalué par l'approche basée sur la corrélation est supérieure à celle évaluée avec l'approche basée sur la copule de Cook-Johnson.

Nous observons en particulier que le rapport des VaR sur un an, $R(1 \text{ an})$, est de 1,158. Cela signifie que le besoin en capital à un an est 15,8 % supérieur avec la copule de Cook-Johnson qu'avec la corrélation. Sur trois ans, le rapport des VaR est de 1,841 et donc le besoin en capital est 84,1 % supérieur avec la copule de Cook-Johnson qu'avec la copule gaussienne.

Nous concluons donc que la corrélation surestime la VaR à 99,5 % du portefeuille. Cette surestimation conduit à sous-estimer la valeur des fonds que l'assureur doit mobiliser pour respecter les contraintes réglementaires et donne une image surévaluée des marges de solvabilité de l'assureur.

Conclusion générale

Globalement, nos travaux s'inscrivent dans un contexte général qui propose des outils théoriques et des démarches pour les mettre en œuvre afin d'étudier dans une première phase la structure de dépendance d'un ensemble d'indices financiers et macroéconomiques et, dans une seconde phase, de les projeter en s'appuyant sur un modèle mathématique adapté. Nos travaux ont nécessité une profonde réflexion et une vive recherche en modélisation mathématique ainsi qu'un investissement significatif en terme de mise en œuvre et donc d'implémentation sous Matlab.

Les apports des travaux présentés dans ce mémoire sont multiples. D'un point de vue théorique, nous proposons une approche consistante et générale permettant de sélectionner la structure de dépendance optimale d'un ensemble d'indices financiers et macroéconomiques. Cette approche nous permet d'une part de rejeter l'hypothèse de normalité de la structure de dépendance des variables financières et macroéconomiques étudiées, et nous permet d'autre part de la caractériser. Il apparaît que la dépendance entre ces variables est asymétrique avec une corrélation asymptotique sur les valeurs négatives qui n'est pas présente sur les valeurs positives (copule de Cook-Johnson).

Lors de l'étude de la structure de dépendance, nous avons traité le problème dans toute sa généralité sans faire d'hypothèses sur le nombre ou les lois marginales des variables étudiées. En particulier, nous avons proposé des méthodes de calibrage et de simulation dans un univers multi-varié alors que la majorité des travaux réalisés à ce sujet se contentent de travailler dans un univers bi-varié. Cela est, par exemple, le cas de la copule de Gumbel où nous avons pu développer un algorithme récurent pour la calibrer (cf. chapitre II-1).

Il est à noter toutefois que nous avons supposé implicitement que la structure de dépendance et ses paramètres de calibrage sont constants dans le temps. Kharoubi-Rakotomalala [2008] montre que la structure de dépendance des indices boursiers internationaux, pris deux à deux (donc dans un univers bi-varié), reste constante mais les paramètres du calibrage changent en fonction du temps (par exemple, si la structure de dépendance de deux indices boursiers est la copule de Cook-Johnson, alors Kharoubi-Rakotomalala [2008] montre que ces indices gardent cette structure au cours du temps et seul le paramètre de la copule α change).

Bien que nous ne proposons pas d'étude sur le changement, dans le temps, de la structure de dépendance des variables financières et macroéconomiques que nous avons étudiées, ce sujet est d'un grand intérêt et permet une modélisation plus sophistiquée.

Par ailleurs, la représentation de la dépendance par la copule de Cook-Johnson (la copule optimale) nous a permis, dans une seconde phase, de proposer une modélisation plus réaliste des indices étudiés en utilisant le modèle d'Ahlgrim.

Cette seconde phase s'inscrit dans un cadre qui intéresse en particulier les organismes assureurs. En effet, que ce soit dans le cadre prudentiel, pour la détermination des provisions et du capital de solvabilité, pour sa communication vers les tiers (*embedded value* et état comptable) ou pour ces besoins de pilotage techniques (allocations stratégiques, tests de rentabilité...), l'organisme assureur doit disposer d'un cadre rigoureux et cohérent prenant en compte l'ensemble des actifs de son bilan et les risques associés.

Dans ce cadre, la troisième partie a proposé un ensemble de méthodes pour générer des scénarios économiques. À cet effet, nous nous sommes appuyés sur le modèle d'Ahlgrim et nous avons proposé deux approches différentes. La première se base sur une structure de dépendance gaussienne (donc la matrice de corrélation) et la seconde se base sur la copule optimale (Cook-Johnson).

Dans le cadre de la gestion d'actifs et afin de comparer ces deux approches, l'étude du rapport des deux VaR à 99,5 %, (correspondant à ces deux approches) d'un portefeuille générique géré en visant une allocation cible, est un indicateur qui peut se révéler pertinent. Bien qu'elle soit d'un grand intérêt,

cette comparaison n'a été examinée que pour un cas simple, mais révélateur, présenté au chapitre III-4. Elle peut d'ailleurs faire l'objet d'une étude plus complète qui prend en compte : le choix de l'allocation stratégique, la réglementation, la stratégie d'investissement..., et qui indique l'impact de la structure gaussienne sur les performances des fonds multi-supports d'un organisme assureur.

Modèle de Vasicek

1. Modèle de Vasicek à un facteur

Soit $(x_t)_t$ une variable aléatoire continue vérifiant :

$$dx_t = k(\mu - x_t)dt + \sigma \cdot dB_t.$$

alors :

$$d(\exp(kt)x_t) = \exp(kt)dx_t + k \exp(kt)x_t dt = k\mu \exp(kt)dt + \exp(kt) \cdot \sigma \cdot dB_t$$

donc :

$$\exp(kt)x_t = x_{t_0} \exp(kt_0) + k\mu \int_{t_0}^t \exp(ks)ds + \sigma \cdot \int_{t_0}^t \exp(ks) \cdot dB_s$$

en simplifiant on obtient :

$$x_t = x_{t_0} \exp(-k(t-t_0)) + \mu(1 - \exp(-k(t-t_0))) + \sigma \cdot \int_{t_0}^t \exp(-k(t-s)) \cdot dB_s$$

Les propriétés de l'intégrale d'une fonction déterministe par rapport à un mouvement brownien permettent d'écrire :

$$x_t = x_{t_0} \exp[-k(t-t_0)] + \mu(1 - \exp[-k(t-t_0)]) + \sigma \sqrt{\frac{1 - \exp[-2k(t-t_0)]}{2k}} \mathcal{E}_t^{t_0}$$

- $\mathcal{E}_t^{t_0}$ est une variable aléatoire de loi normale centrée réduite.

2. Modèle de Vasicek à deux facteurs

Soit le système d'équations suivant :

$$dx_t = k_x(\mu_x - x_t)dt + \sigma_x \cdot dB_{x,t}$$

$$dy_t = k_y(x_t - y_t)dt + \sigma_y \cdot dB_{y,t}$$

Notre objectif est de déterminer la loi de y_t à l'instant t en fonction des variables x_{t_0} et y_{t_0} à l'instant t_0 .

2.1. Premier cas : t et t_0 très proches

En reprenant la même démarche que nous avons présentée pour le modèle à un facteur on peut écrire y_t comme suit :

$$\exp(k_y t) y_t = y_{t_0} \exp(k_y t_0) + k_y \int_{t_0}^t x_s \exp(k_y s) ds + \sigma_y \cdot \int_{t_0}^t \exp(k_y s) \cdot dB_{y,s}$$

Si on suppose que la variable x_t est constante sur la période $[t_0, t[$ on peut écrire :

$$y_t = y_{t_0} \exp(-k_y (t - t_0)) + x_{t_0} (1 - \exp(-k_y (t - t_0))) + \sigma_y \cdot \int_{t_0}^t \exp(-k_y (t - s)) \cdot dB_{y,s}$$

et donc :

$$y_t = y_{t_0} \exp(-k_y (t - t_0)) + x_{t_0} (1 - \exp(-k_y (t - t_0))) + \sigma_y \sqrt{\frac{1 - \exp[-2k_y (t - t_0)]}{2k_y}} \mathcal{E}_{y,t}^{t_0}$$

La formule ci-dessus est celle que nous avons proposée au chapitre III-1 dans le cadre de la résolution de l'équation différentielle stochastique retenue par le modèle d'Ahlgrim pour les taux d'intérêt réels à court terme.

2.2. Deuxième cas : cas général

Posons : $z_t = y_t + \frac{k_y}{k_x - k_y} (x_t - \mu_x)$. On cherche à déterminer la loi de z_t pour en déduire la loi

de y_t après la résolution du modèle à un facteur pour x_t .

Par construction on a : $dz_t = dy_t + \frac{k_y}{k_x - k_y} d(x_t - \mu_x)$ et donc :

$$dz_t = k_y (x_t - y_t) dt + \sigma_y dB_{y,t} + \frac{k_y}{k_x - k_y} (k_x (\mu_x - x_t) dt + \sigma_x \cdot dB_{x,t})$$

ainsi :

$$dz_t = k_y (x_t - y_t) dt + \frac{k_y}{k_x - k_y} k_x (\mu_x - x_t) dt + \sigma_y dB_{y,t} + \frac{k_y}{k_x - k_y} \sigma_x \cdot dB_{x,t}$$

donc :

$$dz_t = k_y [\mu_x - y_t + \frac{k_y}{k_x - k_y} (\mu_x - x_t)] dt + \sigma_y \cdot dB_{y,t} + \frac{k_y}{k_x - k_y} \sigma_x \cdot dB_{x,t}$$

et ainsi :

$$dz_t = k_y [\mu_x - z_t] dt + \sigma_y \cdot dB_{y,t} + \frac{k_y}{k_x - k_y} \sigma_x \cdot dB_{x,t}$$

en multipliant les deux cotés par $\exp(k_y t)$ et en développant on peut écrire :

$$z_t = z_{t_0} \exp(-k_y (t - t_0)) + \mu_x (1 - \exp(-k_y (t - t_0))) + \sigma_y \cdot \int_{t_0}^t \exp(-k_y (t - s)) dB_{y,s} + \frac{k_y \cdot \sigma_x}{k_x - k_y} \cdot \int_{t_0}^t \exp(-k_y (t - s)) dB_{x,s}$$

or :

$$x_t = x_{t_0} \exp(-k_x (t - t_0)) + \mu_x (1 - \exp(-k_x (t - t_0))) + \sigma_x \cdot \int_{t_0}^t \exp(-k_x (t - s)) \cdot dB_{x,s}$$

Ainsi, en remplaçant x_t par sa valeur dans l'expression de z_t on peut déduire la loi de y_t . Les propriétés de l'intégrale d'une fonction déterministe par rapport à un mouvement brownien permettent ensuite de calculer les intégrales figurant dans cette formule. Etant donc une déduction immédiate des développements précédent, nous laissons aux lecteurs intéressés le soin de retrouver l'expression finale de la loi y_t . Par ailleurs, il est à noter que le calibrage complet de y_t nécessite la détermination du coefficient de corrélation entre $B_{x,s}$ et $B_{y,s}$.

Bibliographie

- Ahlgrim K. C., d'Acy S. P., Gorvett R. W. [2008], "Modeling financial scenarios : A Frame work for the Actuarial Profession", Proceeding of the Casuality Actuarial Society.
- Banque de France : T4M. Consulté en ligne le 31-03-2010 : www.banque-france.fr/fr/statistiques/base/statmon/html/tmf_mens_zeuro_fr_tauxmarmonet.htm.
- Banque de France : TME. Consulté en ligne le 31-03-2010 : www.banque-france.fr/fr/statistiques/base/statmon/html/tmf_mens_france_fr_tauxmarfinan.htm.
- Bera A., Jarque C. [1980], "Efficient Tests for Normality, Heteroscedasticity and Serial Independence of Regression Residuals", Economic letter.
- Bouyé E. [2000], "Copulas for Finance : A Reading Guide and Some Applications", Financial Econometrics, Research Centre, City University Business School, Londres.
- Breymann W., Dias A., & Embrechts P. [2003], Dependence structures for multivariate high-frequency data in finance. *Quant. Finance* 3, 1–14.
- Cadoux D., Loizeau J-M. [2006], "Copules et dépendances : application pratique à la détermination du besoin en fonds propres d'un assureur non vie". Institut des actuaires.
- Commission de Contrôle des Assurances : Rapport d'Activité 2000/2001.
- Deheuvels P. [1979], "La fonction de dépendance empirique et ses propriétés. Un test non paramétrique d'indépendance", Académie Royale de Belgique - Bulletin de la Classe des Sciences, 5ème série.
- Demarta S, McNeil A-J. [2004], "The t Copula and Related Copulas", Department of Mathematics Federal Institute of Technology, ETH Zentrum, CH-8092 Zurich.
- Embrechts P., McNeil A., Straumann D. [1999], "Correlation : Pitfalls and Alternatives". Departement Mathematik, ETH Zentrum, CH-8092 Zurich.
- Embrechts P., McNeil A., Straumann D. [2002], "CORRELATION AND DEPENDENCE IN RISK MANAGEMENT : PROPERTIES AND PITFALLS", Departement Mathematik, ETH Zentrum, CH-8092 Zurich.
- ENPC : Ecole Nationale des Ponts et Chaussées [2008], "Fonction de répartition et copules". Consulté en ligne le 31-03-2010 : cermics.enpc.fr/~delmas/Enseig/copules-intro.pdf
- Euronext Paris indices : SBF 250 total return. Complete historical key figures, consulté le 31-03-2010 : www.euronext.com/editorial/wide/editorial-7350-FR.html.
- Fang K., Kotz S. & Ng W. [1990], "Symmetric multivariate and related distributions", Chapman & Hall.
- Friggit J. [2007] "Long Terme [1800-2005] Investment in Gold, Bonds, Stocks and Housing in France - with Insights into the USA and the UK : a Few Regularities", CGPC Working paper.
- Genest C., Mackay R. J. [1986], "The joy of copulas: Bivariate distributions with uniform marginals", *The American Statistician*, 40, 280-283.

- Genest C., Rivest L.P. [1993], "Stastical Inference Proecedures for Bivariate Archimedian Copulas". Journal of American Statistical Association, Volume 88, N°423.
- Hardy M. R. [2001] « A Regime-Switching Model of Long-Term Stock Returns », North American Actuarial Journal 5 (2).
- Hibbert J., Mowbray P., Turnbull C. [2001] « A Stochastic Asset Model & Calibration for Long-Term Financial Planning Purposes », Barrie & Hibbert Limited.
- Hull J. [2007], "Options, futures, et autres actifs dérivés", 6e édition. Pearson Education France.
- Hull J., White A. [1994], "Numerical Procedures for Implementing Term StructureModels II: Two-FactorModels", Journal of Derivatives.
- INSEE : Indice trimestriel des prix des logements anciens. Consulté en ligne le 31-03-2010 : www.indices.insee.fr/bsweb/servlet/bsweb?action=BS_RECHGUIDEE&BS_IDARBO=05000000000000
- INSEE : indice des prix à la consommation. Consulté en ligne le 31-03-2010 : www.indices.insee.fr/bsweb/servlet/bsweb.
- Jacod J. [2001-2002], "Théorie de l'intégration". Cours consulté en ligne le 31-03-2010 : www.proba.jussieu.fr/cours/Integr01.pdf.
- Joe H. [1997], "Multivariate models and dependence concepts", Chapman and Hall, London.
- Joe H., Xu J.J. [1996], "The estimation method of inference functions for margins for multivariate models", Department of Statistics, University of British Columbia, Technical Report.
- Kharoubi-Rakotomalala C. [2008], "Les fonction copules en finance", Paris : Publication de la Sorbonne.
- Lindskog F., McNeil A., Schmock U. [2003], "Kendall's tau for elliptical distributions", Research supported by Credit Suisse, Swiss Re and UBS through RiskLab, Switzerland.
- Longin F., Solnik B. [1995], "Is the correlation in international equity returns constant : 1960-1990?", Journal of International Money and Finance, Vol. 14, No, 1, pp. 3-26.
- Mashal R., Zeevi A. [2002], "Beyond correlation: extreme co-movements between financial assets". Unpublished, Columbia University.
- Nelsen R. [1999], "An introduction to copulas", Springer Lecture notes in statistics.
- Planchet F., Théron P., Kamega A. [2009], "Scénarios économiques en assurance : Modélisation et simulation". Paris : ECONOMICA
- Savu C., Trede M. [2006], "Hierarchical Archimedean Copulas", Institute of Econometrics, University of Münster.
- Schmidt T. [2006], "Coping with Copulas", Department of Mathematics, University of Leipzig.
- Shaw W. T., Lee K.T.A. [2007], "Copula Methods vs Canonical Multivariate Distributions : the multivariate Student T distribution with general degrees of freedom", Working Paper.
- Sklar A. [1959], Fonctions de répartition à n dimensions et leurs marges, Publications de l'Institut de Statistique de l'Université de Paris, 8, 229-231.
- Wilkie A. D. [1995] « More on a Stochastic Asset Model for Actuarial Use », British Actuarial Journal

1.

Xu J. J. [1996], "Statistical Modelling and Inference for Multivariate and Longitudinal Discrete Response Data", Ph.D. thesis. Department of Statistics, University of British Columbia.

