

QUELLE STRUCTURE DE DEPENDANCE POUR UN GENERATEUR DE SCENARIOS ECONOMIQUES EN ASSURANCE ?

IMPACT SUR LE BESOIN EN CAPITAL

Kamal ARMEL¹

Frédéric PLANCHET²

Aymric KAMEGA

Université de Lyon, Université Lyon 1 – ISFA, Laboratoire SAF³

WINTER & Associés⁴ - TELECOM Bretagne⁵ - EURIA⁶

Résumé

Cet article propose une analyse théorique et empirique de l'impact du choix de la structure de dépendance intégrée à un générateur de scénarios économiques sur le niveau du capital de solvabilité dans le cadre du dispositif Solvabilité 2.

1. INTRODUCTION

PLANCHET et al. [2009] définissent un scénario économique comme suit : « un scénario économique correspond à une projection de grandeurs économiques et financières sur un horizon d'intérêt ». Les générateurs de scénarios économiques (GSE) 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 ou encore le taux de chômage (*cf.* AHLGRIM et al. [2005]). La projection de ces indices financiers et

¹ Kamal Armel est actuaire et ingénieur de l'Ecole Nationale Supérieure des Télécommunications de Bretagne.
Contact : kamal.armel@telecom-bretagne.eu.

² Frédéric Planchet et Aymric Kamega sont membres du laboratoire SAF, EA n°2429 et actuaires chez WINTER & Associés. Contact : fplanchet@winter-associes.fr

³ Institut de Science Financière et d'Assurances (ISFA) - 50 avenue Tony Garnier - 69366 Lyon Cedex 07 - France.

⁴ WINTER & Associés 55 avenue René Cassin - 69009 Lyon - France.

⁵ Ecole Nationale Supérieure des Télécommunications de Bretagne (TELECOM Bretagne) – Technopole Brest-Iroise – 29200 Brest.

⁶ EURo Institut d'Actuariat (EURIA) - 6 avenue le Gorgeu - CS93837- 29238 Brest

macroéconomiques permet d'avoir une estimation des valeurs futures des couples rendements/risques des actifs gérés (actions, immobilier, obligations...), de la valeur future du passif (provisions techniques, dettes financières...) ou du niveau du risque de ruine. C'est à ce titre que les GSE constituent un exercice incontournable pour les modèles ALM en assurance vie ou DFA en assurance non vie, pour l'évaluation d'*Embedded Value* (EV) ou encore pour la détermination d'un besoin de capital, notamment dans le cadre du dispositif Solvabilité 2.

Si l'identification des risques à intégrer à un générateur de scénarios économiques peut être considérée comme relativement aboutie (*cf.* PLANCHET et al. [2009] pour une présentation synthétique et FALEH et al. [2010] pour une discussion sur l'état de l'art en la matière.), la crise financière de 2008 a mis en évidence certaines faiblesses dans leur modélisation. Deux éléments sont ainsi mis en évidence et vont sans doute donner lieu à de nombreux développements dans les prochaines années :

- la structure de dépendance entre les actifs ;
- le risque de liquidité, intimement associé à la gestion efficace des couvertures financières.

On s'intéresse ici plus particulièrement à la structure de dépendance retenue et aux conséquences de ce choix sur les résultats obtenus pour le calcul du capital économique ou prudentiel en référence à un critère de contrôle de la probabilité de ruine.

La quasi-totalité des modèles actuels s'appuient sur des structures de dépendance dans lesquels la corrélation tient une place centrale. De nombreux travaux en finance, dont on pourra trouver une synthèse dans KHAROUBI-RAKOTOMALALA [2009], mettent toutefois en évidence le caractère dynamique de l'intensité de la dépendance et sa structure non linéaire. EMBRECHTS et al. [1999] listent les principaux défauts du coefficient de corrélation. On retrouve notamment les deux points suivants :

- Le coefficient de corrélation de deux variables aléatoires peut s'annuler sans que les deux variables soient indépendantes. L'exemple classique consiste à considérer une variable normale $X \approx N(0,1)$, et à calculer sa corrélation avec la variable $Y = X^2$. La corrélation est nulle bien que X et X^2 soient parfaitement dépendants.
- La corrélation linéaire n'est pas invariante par une transformation croissante non linéaire, ce qui signifie d'un point de vue économique et informationnel que 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 $\ln(Y)$. Les coefficients de corrélation sont différents alors que le contenu informationnel dans les deux couples est le même.

Par ailleurs, en pratique, l'intensité de la dépendance augmente dans les situations défavorables, ce qui limite l'efficacité des mesures de diversification calibrées avec des structures ne prenant pas en compte cet effet. L'introduction de structures de dépendance non linéaires intégrant de la dépendance de queue apparaît ainsi comme un élément incontournable de l'évolution des générateurs de scénarios économiques. Dans ce contexte, ARMEL [2010] propose une démarche systématique pour la sélection d'une copule adaptée dans le cadre du modèle d'Ahlgrim (cf. AHLGRIM et al. [2005]).

Le présent article est organisé de la manière suivante : dans une première étape on présente d'un point de vue formel le cadre de travail proposé (cf. annexe et la partie 2) et la seconde étape est consacrée à un exemple d'application à partir d'un portefeuille d'actifs. L'approche proposée consiste à partir de données historiques d'un ensemble d'indices financiers et macroéconomiques pour étudier leur structure de dépendance. Cette étude conduit à choisir, parmi un ensemble de copules, la copule adaptée à ces données. Cette copule optimale peut alors être utilisée pour projeter les indices en passant par un modèle décrivant leurs dynamiques. La figure suivante schématise cette approche :

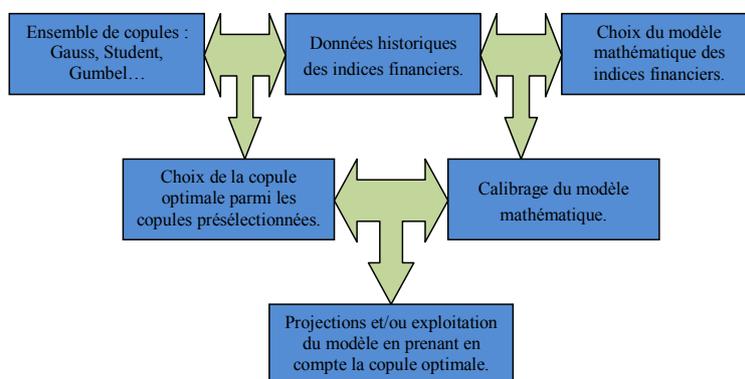


Figure 1 - Schéma de l'approche globale

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.

La description de la structure de dépendance du modèle est effectuée dans cet article de manière naturelle *via* l'utilisation de copules. Les notations et les notions nécessaires à la compréhension de ce papier sont présentées en annexe.

2. CHOIX DE LA COPULE OPTIMALE

Dans cette section nous présentons la procédure de sélection de la copule optimale. Cette copule sera choisie parmi les cinq copules présentées dans la suite. Bien que la mise en œuvre ne concerne que le cas bi-varié, KHAROUBI-RAKOTOMALALA [2008] et DURRLEMAN et al. [2000] ont été d'un grand intérêt pour la conception de la structure globale de la procédure présentée ici. Nous nous sommes inspirés du cas bi-varié pour l'étendre au cas multi-varié. La démarche que nous proposons pour sélectionner la copule optimale en partant des séries historiques des indices financiers et macroéconomiques étudiés est détaillée dans la suite et se schématise par :

- Estimation empirique et calibrage des formes paramétriques des densités marginales et de la densité jointe des indices étudiés. Cette étape permet également d'avoir l'estimation empirique des fonctions de répartition.
- Calibrage des copules paramétriques présélectionnées.
- Estimation de la copule empirique.
- Calcul des distances L^2 entre la copule empirique et les copules paramétriques. La copule optimale est celle qui minimise cette distance.

2.1 Estimation et calibrage des densités de probabilité empiriques

2.1.1 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 à noyau gaussien. 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 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 pour tout $z \in \mathfrak{R}^d$ $K(z) = (2\pi)^{-d/2} \exp(-zz'/2)$ et le paramètre h qui est la largeur de bande et caractérise la densité. On peut l'estimer par le maximum de vraisemblance. Les densités marginales s'écrivent de même pour tout $1 \leq i \leq d$ et pour tout $y \in \mathfrak{R}$:

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

Le paramètre h_i est la largeur de bande caractérisant la densité de la variable X_i .

2.1.2 Calibrage des densités marginales et de la densité jointe

Pour calibrer les densités de probabilité empiriques, les largeurs de bande $(h_i)_{1 \leq i \leq d}$ et h les caractérisant doivent être estimées. Cette estimation est 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 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.

2.1.3 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 :

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

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 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 : $\hat{F}_i(y) = \alpha$.

2.2 Estimation de la copule empirique

La densité de la copule empirique est estimée 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 les formes analytiques de la densité jointe \hat{f} , des densités marginales $(\hat{f}_i)_{1 \leq i \leq d}$ et des fonctions de répartition $(\hat{F}_i)_{1 \leq i \leq d}$ on peut estimer numériquement les inverses $(\hat{F}_i^{-1})_{1 \leq i \leq d}$ et ensuite estimer les valeurs numériques de la densité de la copule empirique.

Notre objectif est de pouvoir comparer la copule empirique aux cinq copules paramétriques pour en choisir la copule optimale. Pour ce faire, les cinq copules paramétriques doivent être calibrées. Cela fait l'objet de la section suivante.

2.3 Calibrage des copules paramétriques

Afin de réaliser ce calibrage, nous utilisons la méthode IFM. Cela signifie que nous procédons au calibrage des cinq copules paramétriques après avoir calibré les densités de probabilité empiriques 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 pour utiliser ensuite la méthode IFM pour l'estimation du nombre de degrés de liberté (*cf.* DEMARTA et MCNEIL [2004]).

2.3.1 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 $\beta = (\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_d))$ et Φ^{-1} est l'inverse de la distribution normale centrée réduite uni-variée. On considère $(x_1^t, \dots, x_d^t)_{1 \leq t \leq T}$ un échantillon 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 on note $(u_1^t, \dots, u_d^t)_{1 \leq t \leq T}$ l'échantillon uniformisé. On note

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

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

La log-vraisemblance admet un maximum. Celui-ci est atteint quand la matrice Σ est égale à (cf. 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 $(\Phi^{-1}(F_1(X_1)), \dots, \Phi^{-1}(F_d(X_d)))$. 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.

2.3.2 Calibrage de la copule de Student

Rappelons que 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 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 ($\Sigma = 1$). À titre de rappel sa densité s'écrit pour tout $x = (x_1, \dots, x_d) \in \mathfrak{R}^d$:

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

où Γ est la fonction gamma.

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 des 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. 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 degré de liberté ν . Cette méthode est utilisée dans les travaux de MASHAL et ZEEVI [2002] et donne des estimations très proches de celles trouvées par la méthode du maximum de vraisemblance.

2.3.2.1 Estimation de la matrice de corrélation

On peut montrer (cf. LINDSKOG et al. [2003]) qu'il existe une relation simple liant le tau de Kendall à la matrice Σ :

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

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, l'estimateur de Σ s'écrit :

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

Cependant, nous n'avons *a priori* aucune garantie que la matrice $\hat{\Sigma}$ soit définie positive. Si ce cas se présente, cette matrice peut être approchée par une matrice définie positive, ce qui revient à résoudre un problème d'optimisation.

2.3.2.2 Estimation du nombre de degrés de liberté de la copule de Student

La démarche la plus simple 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 :

$$\hat{\nu} = \arg \max \left(\sum_{i=1}^T \ln \left(\frac{f_{\nu, \hat{\Sigma}}(t_{\nu}^{-1}(u'_1), \dots, t_{\nu}^{-1}(u'_d))}{\prod_{i=1}^d f_{\nu}(t_{\nu}^{-1}(u'_i))} \right) \right)$$

2.3.3 Calibrage de la copule de Cook-Johnson

La copule de Cook-Johnson, connue aussi sous les noms de copule de Clayton ou copule de Kimeldorf-Sampson, est la copule archimédienne dont le générateur est défini, pour $\alpha > 0$ et pour $u \in]0,1]$, par $\phi(u) = \alpha^{-1}(u^{-\alpha} - 1)$. Elle 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 s'écrit :

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

Le calcul de la densité de la copule de Cook-Johnson nous permet de la calibrer par l'application de la méthode IFM. Notons que le calibrage de cette copule ne nécessite que l'estimation du paramètre α .

2.3.4 Calibrage de la copule de Franck

La copule de Franck ne présente pas de dépendance de queue. Le générateur de cette copule archimédienne est $\phi(u) = -\ln \left(\frac{e^{-\alpha u} - 1}{e^{-\alpha} - 1} \right)$ où $\alpha \neq 0$ et $u \in]0,1]$. La copule de

Franck d -variée s'écrit donc :

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

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} . Les dimensions des

modèles multi-variés qui nous intéressent sont la dimension 2 et la dimension 5. Les densités de la copule de Franck pour ces dimensions sont présentées en annexe.

Le calcul de la densité de la copule de Franck multi-variée d'ordre 5 nous permet de la calibrer par l'application de la méthode IFM. Notons que le calibrage de cette copule ne nécessite que l'estimation du paramètre α .

2.3.5 Calibrage de la copules de Gumbel

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. Le générateur de cette copule archimédienne s'écrit $\phi(u) = (-\ln(u))^\alpha$ 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. Nous proposons dans la suite une méthode afin de calculer cette densité, ce qui permettra ensuite de calibrer la copule. On part de l'expression générale :

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

où $\phi(u) = (-\ln(u))^\alpha$ et $\phi^{-1}(u) = \exp(-u^{1/\alpha}) = \exp(-u^\beta)$, $\beta = \frac{1}{\alpha}$. La détermination de la forme analytique de cette densité est délicate 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. 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 calcule ainsi 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 $u \in \mathfrak{R}$:

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

que l'on utilise avec $(\phi^{-1})'(u) = -\beta\phi^{-1}(u)u^{\beta-1}$. On en déduit en notant $g(u) = u^{\beta-1}$ que, pour tout $1 \leq n$:

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

On a par ailleurs $g^{(k)}(u) = u^{\beta-1-k} \prod_{i=1}^k (\beta-i)$. On note 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 il est simple de 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.

Le calcul de la densité de la copule de Gumbel nous permet de la calibrer (estimation du paramètre α) par l'application de la méthode IFM.

2.4 Critères de sélection de la copule optimale

Pour choisir la copule optimale parmi les cinq copules présélectionnées, la distance, au sens de la norme L^2 , entre la densité empirique et les densités de ces cinq copules est étudiée. Il est rappelé que la densité de la copule empirique présentée ci-dessus a une forme paramétrique continue sur $[0,1]$ et est donc intégrable. Si c est la densité d'une copule paramétrique calibrée 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 . En outre, la norme infinie, L^∞ , est calculée pour se faire une idée sur la distance maximale entre une copule paramétrique et la copule empirique. Cette norme s'écrit :

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

Il est toutefois très complexe de calculer directement les deux normes présentées ci-dessus notamment lorsque la dimension du modèle est importante. La solution alternative

est d'approximer numériquement ces deux normes par des séries :

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

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

Maintenant que la procédure de sélection de la copule optimale est présentée, la partie suivante (*cf.* section 3.1) illustre son application pour la détermination de la copule optimale d'un ensemble d'indices financiers et macroéconomiques sur un historique datant de janvier 1997 à juin 2009. Cette copule optimale servira ensuite à simuler des trajectoires du modèle multi-varié dont les dynamiques uni-variées sont spécifiées par le modèle d'Ahlgrim. Pour plus de détails sur la méthode de simulation des copules, le lecteur pourra se référer au document annexe à cet article.

3. APPLICATION ET ETUDE DE L'IMPACT DU CHANGEMENT DE STRUCTURE DE DEPENDANCE

Cette partie consiste en une réflexion sur la modélisation d'indices financiers et macroéconomiques en intégrant le choix de la copule optimale pour prendre en compte la structure de dépendance. Le modèle qui est à la base de cette réflexion est le modèle d'Ahlgrim (*cf.* AHLGRIM et al. [2005]). Le lecteur pourra trouver une présentation détaillée de ce modèle dans l'article original de ses concepteurs ainsi que dans PLANCHET et al. [2009]. Nous utilisons ici une variante dans laquelle l'excès de rendement des actions est gaussien. Outre sa flexibilité, ce modèle présente l'avantage de fournir des formules fermées pour le prix des zéro-coupon (*cf.* HIBBERT [2001]).

Après avoir choisi la copule optimale (*cf.* section 3.1), on projette ici les indices étudiés en substituant la copule optimale à la copule gaussienne utilisée par défaut dans le modèle retenu (*cf.* sections 3.2 et 3.3), ce qui permet ensuite d'évaluer les performances d'un fonds composé d'obligations, d'actions et d'immobilier et de calculer le besoin en capital associé à la perte maximale au niveau seuil de 99,5 % (*cf.* section 3.4).

3.1 Choix de la copule optimale : application

Nous examinons dans cette section 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 entre 1997 et 2009. Le premier paragraphe est consacré à la présentation des données, le second au calibrage des densités de probabilité empiriques et le troisième présente les résultats du calibrage des copules paramétriques. Le dernier paragraphe est consacré à la présentation de la copule optimale.

3.1.1 *Présentation des données utilisées*

Les séries de données retenues datent de janvier 1997 à juin 2009. Mis à part la série de l'immobilier dont le taux de rendement annuel est à fréquence trimestrielle, toutes les séries représentent des taux annuels à fréquences mensuelles :

- L'historique du taux annuel d'inflation est construit à partir de l'indice des prix à la consommation (IPC) publié sur le site de l'INSEE¹.
- 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 France².
- L'historique du 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 France³.
- L'historique des rendements des actions correspond à l'historique de l'indice SBF 250 avec dividendes réinvestis (*total return*). Cet historique est en téléchargement libre sur le site d'Euronext⁴.
- 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'INSEE⁵.

Les historiques des taux d'intérêt réels (à long et à court terme) sont construits à partir des valeurs historiques des taux d'intérêt nominaux auxquels on retranche les valeurs historiques du taux d'inflation. Le tableau suivant présente les statistiques descriptives de base de l'ensemble des séries historiques⁶ :

¹ www.indices.insee.fr/bsweb/servlet/bsweb.

² www.banque-france.fr/fr/statistiques/base/statmon/html/tmf_mens_france_fr_tauxmarfinan.htm.

³ www.banque-france.fr/fr/statistiques/base/statmon/html/tmf_mens_zeuro_fr_tauxmarmonet.htm.

⁴ www.euronext.com/editorial/wide/editorial-7350-FR.html.

⁵ www.indices.insee.fr/bsweb/servlet/bsweb?action=BS_RECHGUIDEE&BS_IDARBO=0500000000000000.

⁶ Test de Jarque-Bera : l'hypothèse nulle est celle de la normalité de l'indice. Ce test vaut 1 si on peut refuser l'hypothèse nulle avec un niveau de significativité de 5 % et vaut 0 sinon.

Statistiques descriptives	Taux d'inflation	Taux long nominal	Taux court nominal	Actions	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

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 (actions et taux long) ne répondent pas au test de normalité de Jarque-Bera. Cela implique notamment que la distribution multi-variée n'est pas normale. Ainsi, se contenter de décrire la structure de dépendance par la mesure des coefficients de corrélation s'avère insuffisant.

3.1.2 Calibrage des densités de probabilité empiriques

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 :

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

Afin de tester l'adéquation des fonctions de répartition estimées aux séries de données, cinq tests de Kolmogov-Smirnov¹, 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.

¹ Test de Kolmogov-Smirnov : 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.

3.1.3 Calibrage des copules

3.1.3.1 La copule gaussienne

La détermination des fonctions de répartition à partir de l'estimation de leurs densités par l'estimateur à noyau gaussien permet de calculer numériquement la matrice de corrélation et ainsi de calibrer la copule gaussienne.

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 %

3.1.3.2 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.

Student : corrélation	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 $\nu = 14,55$.

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

Le calibrage des copules de 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 :

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, on peut calculer et puis comparer les distances entre les densités de ces copules paramétriques et la densité de la copule empirique.

3.1.4 Choix de la copule optimale

Les normes sont estimées sur 9 765 625 vecteurs de dimension 5, ce qui correspond à un pas de discrétisation $n = 25$. 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 :

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

On observe en premier lieu que la copule de Student n'est pas adaptée aux données. Parmi les possibilités restantes, la norme infinie n'est pas discriminante. En outre, à partir de la norme L^2 , il apparaît que les distances 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). La copule optimale est donc une copule archimédienne. Au final, le calcul des normes L^2 place en tête la copule de Cook-Johnson.

On en conclut que la copule de Cook-Johnson représente une forme paramétrique appropriée pour modéliser la structure de dépendance des indices financiers et macroéconomiques étudiés sur la période 1997-2009.

Dans ce qui suit, on propose de projeter les indices financiers et macroéconomiques étudiés en combinant cette structure de dépendance avec le modèle d'Ahlgren et al. [2005].

3.2 Étude de l'impact : calibrage du modèle

Le modèle est calibré sur les séries historiques présentées *supra*. Le tableau suivant résume les résultats du calibrage :

Calibrage du modèle d'Ahlgrim	Inflation	Taux long	Taux court	Immobilier
Espérance à long terme : μ	1,27 %	3,04 %	NA	5,80 %
Écart type de l'erreur : σ	1,00 %	1,04 %	1,00 %	2,69 %
Vitesse retour à la moyenne : k	0,61	0,46	0,10	0,09
R ² ajusté	85,64 %	91,90 %	-0,45 %	91,63 %
P-value Fisher	< 2.2e-16	< 2.2e-16	0,561	< 2.2e-16
Test Jarque-Bera sur les résidus	0	0	1	0

Nous remarquons que pour l'inflation, les taux long et l'immobilier, l'ensemble des résultats et des tests valide les hypothèses posées *a priori* pour la validation des modèles et du calibrage. Nous observons en particulier que les trois résidus répondent au test de normalité de Jarque-Bera. Cependant, en ce qui concerne le taux réel à court terme, nous observons que les tests ne valident pas les hypothèses du modèle et de son calibrage. Aussi, les résidus associés au taux d'intérêt réel à court terme ne répondent pas au test de normalité de Jarque-Bera.

L'excès de rendement est calibré sur l'indice SBF 250 dividende réinvesti. Le tableau suivant présente les résultats obtenus :

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

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

Corrélations historiques	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 cette section, la section suivante illustre les projections des indices financiers et macroéconomiques en prenant en compte deux structures de dépendance, la copule gaussienne et la copule de Cook-Johnson.

3.3 Étude de l'impact : construction de projections

Les projections sont réalisées sur une période de 10 ans à fréquence trimestrielle. 100 000 trajectoires sont simulées à chaque trimestre. 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 :

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	5 %

Mis à part l'immobilier, ces valeurs initiales correspondent aux valeurs observées au mois de juin 2009. La valeur initiale de l'immobilier est ajustée à 5 % au lieu de -9,75 % observée au mois de juin 2009 afin de se mettre dans un univers de projection plus réaliste. Plusieurs tests réalisés montrent que l'impact du changement des valeurs initiales sur les indicateurs numériques retenus pour comparer la copule gaussienne et la copule de Cook-Johnson est faible (*cf.* section 3.2.3 pour la présentation de ces indicateurs).

3.3.1 Projection avec la copule gaussienne

Dire que la structure de dépendance des résidus et de l'excès de rendement est décrite par la corrélation 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 ($\varepsilon_{q,t}$, $\varepsilon_{l,t}$, $\varepsilon_{r,t}$, $\varepsilon_{lm,t}$, x_t) en prenant en compte leur matrice de corrélation historique (décomposition de Cholesky). Ensuite les projections des indices financiers et macroéconomiques s'en déduisent par les formules discrètes des modèles associés. À titre d'illustration, le graphique ci-dessous présente la projection du taux d'inflation lorsque l'on utilise la copule gaussienne.

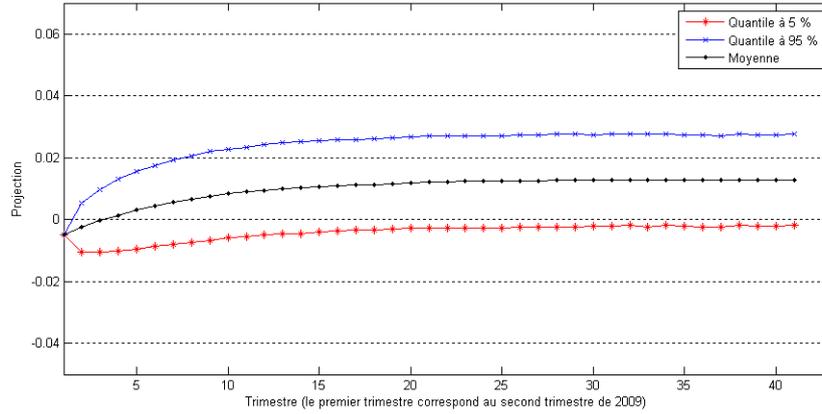


Figure 2 - Projection du taux d'inflation par l'approche basée sur la copule gaussienne

3.3.2 Projection avec la copule de Cook-Johnson

Afin de combiner le modèle d'Ahlgrim et la copule de Cook-Johnson, on simule d'abord des trajectoires de variables uniformes (U_1, \dots, U_5) dont la fonction de répartition est la copule de Cook-Johnson. Ensuite, on utilise l'inverse des fonctions de répartition des indices financiers et macroéconomiques données par le modèle d'Ahlgrim pour avoir les trajectoires simulées de ces indices. Les indices modélisés dans AHLGRIM et al. [2005] 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 qui 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. Afin d'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 :

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

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 $u_1^{t+\delta}$. Le graphique ci-dessous présente la projection du taux d'inflation lorsque l'on utilise la copule de Cook-Johnson.

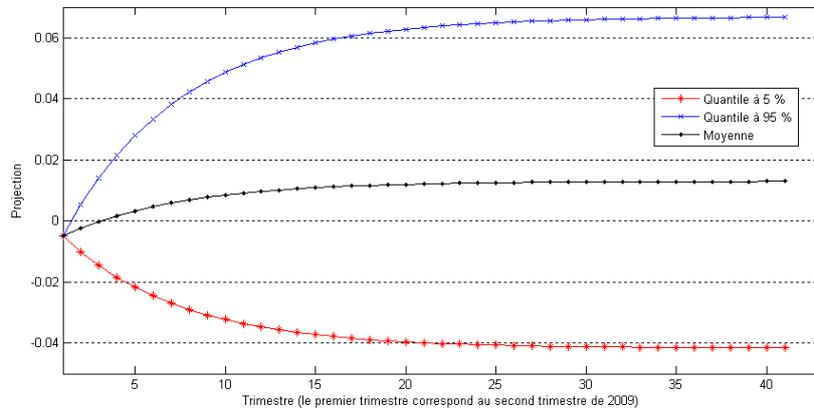


Fig. 2 - Projection du taux d'inflation par l'approche basée sur la copule de Cook-Johnson

3.4 ÉTUDE DE L'IMPACT : COMPARAISON DES COPULES GAUSSIENNE ET DE COOK-JOHNSON

Globalement, 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 ne varient pas par changement de la structure de dépendance. Ce constat était attendu car 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.

Afin de mesurer l'impact des différences entre ces deux structures de dépendance, on peut considérer deux types de comparaison. D'une part, dans le cadre de la gestion d'actifs, la comparaison des deux quantiles à 0,5 % de la valeur d'un portefeuille générique géré en visant une allocation cible est un indicateur qui peut se révéler utile (ces quantiles correspondent à la VaR à 0,5 %). D'autre part, dans un contexte solvabilité 2, la comparaison des deux quantiles à 99,5 % de l'opposé de la perte du portefeuille permet de vérifier si la structure de dépendance gaussienne surestime ou sous-estime le capital à

constituer. En effet, le montant du capital à constituer est égal à l'opposé de la perte maximale au seuil de 99,5 %.

Ces deux types d'analyse sont respectivement étudiés dans les paragraphes 3.4.2 et 3.4.3. En outre, dans le paragraphe 3.4.4, une comparaison entre les besoins de capitaux obtenus dans chacun des deux cas et celui obtenu avec la formule standard des spécifications techniques du QIS5 dans le cadre du projet Solvabilité 2 est proposée. Au préalable dans le paragraphe 3.4.1 on présente le portefeuille générique retenu et quelques outils utiles à la comparaison.

3.4.1 Composition du portefeuille et présentation des outils de comparaison

Nous proposons dans la suite d'étudier un cas simple où le portefeuille est composé de trois actifs : obligations zéro-coupons de maturité 5 ans, actions et immobilier. Nous supposons que le fonds se compose de :

- 80 % d'obligations d'Etat EEA de maturité 5 ans ;
- 10 % d'actions diversifiées sur le marché français ;
- 10 % d'immobilier.

Nous considérons, sans perte de généralité, que la valeur du portefeuille à la date initiale est égale à 1. Nous supposons également qu'à chaque trimestre le gestionnaire fait en sorte de maintenir stable la structure du portefeuille, dans la logique du respect d'une allocation stratégique. Le rendement du portefeuille correspond au rapport diminué de 1 du prix de marché des actifs sur leurs prix d'achat au trimestre précédent.

Les deux paragraphes suivants présentent les deux indicateurs numériques retenus pour comparer les deux structures de dépendance sur un horizon de 3 ans. Ces deux indicateurs consistent en une comparaison des deux *VaR* à 0,5 % du fonds et des deux montants du capital à constituer au seuil de 99,5 % issus des deux approches : l'approche par la structure de dépendance gaussienne et l'approche par la copule de Cook-Johnson. Les deux indicateurs numériques se basent sur le calcul de la *VaR* à 0,5 % du portefeuille (*cf.* ci-après). Pour les deux approches et à chaque trimestre t , la *VaR* du portefeuille est alors calculée en deux étapes :

- À chaque trimestre 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* à 0,5 % 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 de 1 à l'instant 0.

3.4.2 Comparaison de la VaR du portefeuille

Le premier indicateur est le rapport des VaR à 0,5 %. Ce rapport correspond au trimestre t au rapport des deux VaR du portefeuille calculées sur la période $[0,t]$ et s'écrit :

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

Les moyennes du portefeuille, les courbes des VaR à 0,5 % et le rapport sont présentés dans la figure ci-après (cf. Fig. 4) sur un horizon de 3 ans.

Il apparaît que la VaR à 0,5 % de la valeur du portefeuille évaluée par l'approche basée sur la copule gaussienne est supérieure à celle évaluée avec l'approche basée sur la copule de Cook-Johnson. En effet, le rapport des VaR sur un an, $R(1 \text{ an})$, est de 1,086. Cela signifie que la plus faible valeur du portefeuille à un an, au quantile 0,5 %, est 8,60 % supérieure avec la corrélation qu'avec la copule de Cook-Johnson.

Plusieurs scénarios de tests sur les valeurs initiales des indices modélisés et sur leurs moyennes ont été réalisés. Ces scénarios donnent tous des résultats proches de ceux du scénario central (cf. section 3.1).

3.4.3 Comparaison du besoin de capital

Le second indicateur numérique retenu est le rapport du montant du capital à constituer. Ce montant est égal au quantile à 99,5 % de l'opposée de la perte et est donc égal à 1 (valeur initiale du portefeuille) diminué de la VaR à 0,5 % de la valeur du portefeuille. Cet indicateur se base donc également sur le calcul de la VaR à 0,5 %. Le second indicateur s'écrit ainsi :

$$R'(t) = \frac{1 - VaR_{[0,t]}^{Gauss}}{1 - VaR_{[0,t]}^{Cook-Johnson}}$$

Les moyennes du portefeuille, les courbes des VaR à 0,5 % et le rapport sont présentés dans la figure ci-après sur un horizon de 3 ans (pour mémoire, le rapport des VaR correspondant au premier indicateur y est également présenté).

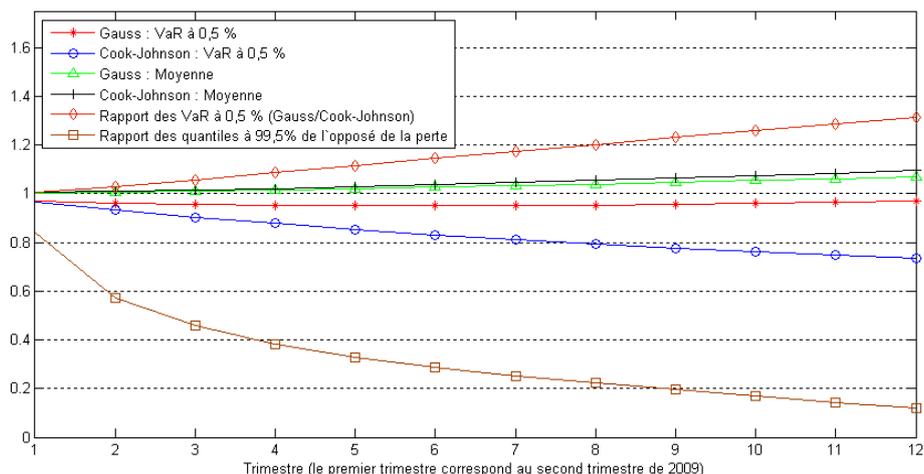


Fig. 3 - Indicateurs numériques de comparaison sur un horizon de 3 ans

On observe que le montant du capital à constituer est plus faible dans l'approche basée sur la copule gaussienne. En pratique, le rapport des montants du capital à constituer sur un an, $R^1(1 \text{ an})$, est de 38,30 %. Ce qui signifie que le montant du capital à constituer avec la copule gaussienne représente moins de 40 % du montant du capital à constitué avec la copule de Cook-Johnson (en pratique, le besoin en capital est égal à 4,7 % avec la copule gaussienne, contre 12,2 % avec la copule de Cook-Johnson).

Ici également, plusieurs scénarios de tests sur les valeurs initiales des indices modélisés et sur leurs moyennes ont été réalisés. Ces scénarios donnent tous des résultats proches de ceux du scénario central (cf. section 3.1).

Il apparaît donc que la corrélation conduit à sous-estimer très sensiblement la valeur des fonds que l'assureur doit mobiliser pour respecter les contraintes réglementaires et donne donc une image surévaluée des marges de solvabilité de l'assureur.

3.4.4 Comparaison du besoin de capital avec le SCR marché de la formule standard du QIS5 (Solvabilité 2)

Dans le cadre de la présente étude, il est intéressant de positionner le besoin en capital obtenu avec la copule de Gauss ou avec la copule de Cook-Johnson par rapport au SCR du risque marché issu de la formule standard du QIS 5 (cf. CEIOPS [2010]).

Compte tenu de la structure du portefeuille retenue (cf. paragraphe 3.4.1), pour le calcul du SCR au titre des risques de marché on considère les risques sur les taux d'intérêt,

les actions et l'immobilier. Les résultats du SCR marché sont présentés dans le tableau suivant.

Risque	Charge de capital - formule standard En % de la VM globale
Taux d'intérêt	5,6 %
Action	3,0 %
Immobilier	2,5 %
SCR Marché (scénario Up)	7,6 %

Le SCR issu de la formule standard du QIS 5 est donc égal à 7,6 % de la valeur de marché du portefeuille. Pour mémoire, le besoin en capital estimé à partir des projections est égal à 4,7 % lorsque l'on utilise la copule gaussienne et 12,2 % avec la copule de Cook-Johnson. On note ainsi que dans le cadre de l'exemple retenue (pour mémoire le portefeuille est composé à 80 % d'obligations de maturité 5 ans, 10 % d'actions et 10 % d'immobilier) :

- le besoin en capital obtenu à partir de la copule gaussienne minore de 39 % celui obtenu à partir de la formule standard du QIS 5 pour le risque de marché ;
- le besoin en capital obtenu à partir de la copule de Cook-Johnson majore de 60 % celui obtenu à partir de la formule standard du QIS 5 pour le risque de marché.

4. CONCLUSION

Construire un générateur de scénarios économiques (GSE) est devenu indispensable dans le cadre de la mise en œuvre du dispositif Solvabilité 2 et, plus largement, dans le cadre du pilotage technique d'une activité d'assurance. De nombreux travaux ont été consacrés au choix de modèle pour les risques marginaux (actions, taux, immobilier, etc.) mais la manière de combiner les risques sous-jacents pour construire la dynamique globale reste souvent basée sur de simples corrélations, dans un cadre gaussien.

Nous avons montré dans ce travail l'insuffisance de cette approche et proposé une démarche générale pour substituer à une structure de dépendance gaussienne une structure non linéaire en choisissant la copule optimale parmi une liste déterminée *a priori*.

Si le processus de sélection et de calibrage de la copule est sensiblement plus complexe que la simple estimation de coefficients de corrélation, la génération de trajectoires prenant en compte la nouvelle structure de dépendance ne conduit à modifier le modèle existant qu'à la marge.

L'ensemble de la démarche est illustré en généralisant le modèle classique proposé dans AHLGRIM et al. [2005] qui fournit un cadre dans lequel s'inscrivent de fait de nombreux générateurs utilisés en pratique. Dès lors, il nous semble que les modèles mis en place se doivent d'opter pour une structure de dépendance non linéaire sous peine de sous-estimer le montant du capital de solvabilité et d'induire ainsi une vision biaisée de la solvabilité réelle de l'organisme.

En parallèle de l'impact sur le calcul d'un quantile de la distribution de la valeur du portefeuille se pose la question des conséquences sur l'estimation des provisions *best estimate* en présence d'un dispositif de participation aux bénéfices : des travaux sont en cours sur ce sujet.

5. BIBLIOGRAPHIE

AHLGRIM K. C., D'ARCY S. P., GORVETT R. W. [2005] Modeling Financial Scenarios: A Framework for the Actuarial Profession. *Proceedings of the Casualty Actuarial Society* 92. (<http://www.casact.org/pubs/proceed/proceed05/05187.pdf>).

ARMEL K. [2010] *Structure de dépendance des générateurs de scénarios économiques - Modélisation et Simulation*, Mémoire d'actuariat, EURIA.

BERA A., JARQUE C. [1980] *Efficient Tests for Normality, Heteroscedasticity and Serial Independence of Regression Residuals*, Economic letter.

CEIOPS [2010] *QIS 5 Technical Specifications*, European Commission

CHERUBINI U. [2004] *Copula methods in finance*, Wiley finance.

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.

DURRELMAN V., NIKEGHBALI A., RONCALLI T. [2000] *Which copula is the right one*, Working paper, Crédit Lyonnais.

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.

FALEH A., PLANCHET F., RULLIERE D. [2010] Les générateurs de scénarios économiques : de la conception à la mesure de la qualité, *Assurances et Gestion des Risques*, Vol. 78.

FANG K., KOTZ S. & NG W. [1990] *Symmetric multivariate and related distributions*, Chapman & Hall.

GENEST C., MACKAY R. J. [1986] *The joy of copulas: Bivariate distributions with uniform marginals*, *The American Statistician*, 40, 280-283.

HIBBERT J., MOWBRAY P., TURNBULL C. [2001] *A Stochastic Asset Model & Calibration for Long-Term Financial Planning Purposes*, Barrie & Hibbert Limited.

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. [2009] *Les fonctions copule en finance*, Paris : Presses 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.

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.E., KAMEGA A. [2009] *Scénarios économiques en assurance - Modélisation et simulation*, Paris : Economica.

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.

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.

**ANNEXE A L'ARTICLE : QUELLE STRUCTURE DE DÉPENDANCE POUR UN
GÉNÉRATEUR DE SCÉNARIOS ÉCONOMIQUES EN ASSURANCE ? IMPACT
SUR LE BESOIN EN CAPITAL**

Résumé

Cette annexe est rattachée à l'article « Quelle structure de dépendance pour un générateur de scénarios économiques en assurance ? Impact sur le besoin en capital » et reprend les notations et les notions nécessaires à la compréhension de cet article.

1. STRUCTURE DE DÉPENDANCE NON LINEAIRE

La description de la structure de dépendance du modèle est effectuée ici de manière naturelle *via* l'utilisation de copules. Une attention particulière est accordée aux cinq copules qui nous intéressent dans le cadre de cette étude, à savoir : la copule gaussienne, la copule de Student, la copule de Frank, la copule de Gumbel et la copule de Cook-Johnson. Le lecteur intéressé par des présentations complètes de la théorie des copules pourra se référer aux livres de NELSON [1999] ou de JOE [1997].

1.1 Notations

On rappelle qu'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 alors $X = (X_1, \dots, X_d)$ 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 . Le théorème de Sklar (*cf.* SKLAR [1959]) précise le lien que définit la copule C entre les fonctions de répartition marginales F_1, \dots, F_d et la distribution jointe F :

$$F(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d))$$

pour tout $x = (x_1, \dots, x_d) \in \mathfrak{R}^d$. On peut trouver la démonstration de ce théorème dans NELSON [1999] ou SKLAR [1959].

D'autre part, il est rappelé que 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 le tau de Kendall.

1.2 Tau de Kendall

Le tau de Kendall est une mesure de concordance bien connue en statistiques. Elle donne 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. Elle offre par ailleurs l'avantage de s'exprimer simplement en fonction de la copule associée au couple de variables aléatoires. Pour plus de détail le lecteur pourra se référer à EMBRECHTS et al. [2002], DEMARTA et MCNEIL [2004] et LINDSKOG et al. [2003]. Soient (X_1, X_2) et (X_1', X_2') deux couples de vecteurs aléatoires de même loi, le tau de Kendall est défini par :

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

qui est donc 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 :

$$\begin{aligned} \tau(X_1, X_2) &= 4E(C(U_1, U_2)) - 1 \\ &= 4 \iint_{[0,1]^2} C(u_1, u_2) dC(u_1, u_2) - 1 = 1 - 4 \iint_{[0,1]^2} \partial_1 C(u_1, u_2) \partial_2 C(u_1, u_2) du_1 du_2 \end{aligned}$$

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 l'estimateur empirique suivant du tau de Kendall :

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

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

1.3 Présentation des copules paramétriques retenues

Dans le cadre de notre article, dont ce document est l'annexe, la recherche de la copule optimale est réalisée sur un ensemble de cinq copules sélectionnées *a priori* et présentées dans cette section. 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.

1.3.1 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 en général 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 la « matrice de corrélation ».

La loi d'un vecteur gaussien est un exemple classique de loi elliptique, associé au choix $g(y) = (2\pi)^{-d/2} \exp(-y/2)$. Ces lois vérifient, comme les vecteurs gaussiens, des propriétés algébriques intéressantes. La propriété la plus importante est que les lois elliptiques forment une classe stable par transformation affine. Pour une présentation détaillée de ces lois, le lecteur pourra se référer à FANG et al. [1990]. Les copules elliptiques sont définies à partir des familles des lois elliptiques. Une copule est dite elliptique si elle est la copule d'une loi elliptique. On en considère ici deux cas particuliers, la copule gaussienne et la copule de Student.

1.3.1.1 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. 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 de la distribution normale centrée réduite uni-

variée. La fonction $\phi_{\Sigma}(x) = \frac{\exp\left(-\frac{1}{2}x\Sigma^{-1}x'\right)}{(2\pi)^{d/2} \det(\Sigma)^{1/2}}$ 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). En dérivant la formule définissant la copule gaussienne on peut facilement extraire la densité de la copule gaussienne d -variée qui s'écrit :

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

où I_d est la matrice unité de $M_d(\mathfrak{R})$ et $\beta = (\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_d))$.

1.3.1.2 La copule de Student

La copule de Student (*t copula*) est la copule sous-jacente à une distribution multivariée de Student. Cette structure de dépendance capte les dépendances extrêmes positives et négatives. Elle est construite de la même manière que la copule gaussienne mais 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 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 ($\Sigma = 1$). À titre de rappel sa densité s'écrit pour tout $x = (x_1, \dots, x_d) \in \mathfrak{R}^d$:

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

où Γ est la fonction gamma.

1.3.2 Copules archimédiennes

Les copules archimédiennes, définies par GENEST et MACKAY [1986], permettent de décrire des structures de dépendance très diverses. 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 $C(u_1, \dots, u_d) = \phi^{-1}(\phi(u_1) + \dots + \phi(u_d))$ est une copule si ϕ est d fois dérivable sur $]0,1[$ et si pour tout $1 \leq i \leq d$, $\phi^{(i)} > 0$ pour i pair et $\phi^{(i)} < 0$ sinon. Cette fonction définit alors la copule archimédienne de générateur ϕ . De la définition précédente on déduit que la densité d'une copule archimédienne d -variée s'écrit :

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

Les copules archimédiennes présentent un double intérêt. D'une part, elles permettent de construire une grande variété de familles de copules et permettent donc de 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 [1999]. Les trois familles de copules archimédiennes qui nous intéressent dans le cadre de notre étude sont présentées ci-après.

1.3.2.1 La copule de Cook-Johnson

La copule de Cook-Johnson, connue aussi sous les noms de copule de Clayton ou copule de Kimeldorf-Sampson, est la copule archimédienne dont le générateur est défini, pour $\alpha > 0$ et pour $u \in]0,1]$, par $\phi(u) = \alpha^{-1}(u^{-\alpha} - 1)$. Elle 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 s'écrit :

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

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.

1.3.2.2 La copule de Franck

La copule de Franck ne présente pas de dépendance de queue. Le générateur de cette copule archimédienne est $\phi(u) = -\ln\left(\frac{e^{-\alpha u} - 1}{e^{-\alpha} - 1}\right)$ où $\alpha \neq 0$ et $u \in]0,1]$. La copule de

Franck d -variée s'écrit donc :

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

Pour le cas bi-varié la densité est :

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

Par ailleurs, la densité d -variée de la copule de Franck s'écrit à l'aide de l'expression générale présentée *supra* avec $\phi^{-1}(u) = -\frac{\ln((e^{-\alpha} - 1)e^{-u} + 1)}{\alpha}$. La détermination de la densité de la copule de Franck pour un ordre d passe par le calcul de la dérivée première de ϕ , $\phi'(u) = \frac{\alpha e^{-\alpha u}}{(e^{-\alpha u} - 1)}$, et de la dérivée d'ordre d de ϕ^{-1} . Les dimensions des modèles multi-variés qui nous intéressent sont la dimension 2 et la dimension 5. Soit $\beta = e^{-\alpha} - 1$, les dérivées d'ordre 2 et 5 de ϕ^{-1} sont égales à :

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

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

1.3.2.3 La 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 appartient à la famille des copules archimédiennes et son générateur s'écrit $\phi(u) = (-\ln(u))^\alpha$ 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. Nous proposons dans la suite une méthode afin de calculer cette densité. On part de l'expression générale :

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

où $\phi(u) = (-\ln(u))^\alpha$ et $\phi^{-1}(u) = \exp(-u^{1/\alpha}) = \exp(-u^\beta)$, $\beta = \frac{1}{\alpha}$. La détermination de la forme analytique de cette densité est délicate 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. 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 calcule ainsi 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 $u \in \mathfrak{R}$:

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

que l'on utilise avec $(\phi^{-1})'(u) = -\beta \phi^{-1}(u) u^{\beta-1}$. On en déduit en notant $g(u) = u^{\beta-1}$

que, pour tout $1 \leq n$:

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

On a par ailleurs $g^{(k)}(u) = u^{\beta-1-k} \prod_{i=1}^k (\beta-i)$. On note 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 il est simple de 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.

Maintenant que les cinq copules paramétriques retenues dans le cadre de notre article sont présentées, il est important de pouvoir les calibrer. Cela fait l'objet de la section suivante.

1.4 Calibrage des copules paramétriques

Trois méthodes de calibrage sont présentées. Il est à noter que le choix d'une méthode ou de l'autre dépend fortement de la complexité et de la nature de la copule.

1.4.1 Estimation par le maximum de vraisemblance

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 $(x'_1, \dots, x'_d)_{1 \leq t \leq T}$ de taille T peut être obtenu directement par le maximum de la

log-vraisemblance qui s'écrit $\ln L(\theta) = \sum_{t=1}^T \ln(f(x'_1, \dots, x'_d; \theta))$, ce qui est équivalent à :

$$\ln L(\theta) = \sum_{t=1}^T \sum_{i=1}^d \ln f_i(x'_i; \theta) + \sum_{t=1}^T \ln c(F_1(x'_1; \theta), \dots, F_d(x'_d; \theta); \theta)$$

Bien que possible théoriquement, la recherche du vecteur θ qui maximise la log-vraisemblance, pose plusieurs problèmes notamment en termes de complexité et du temps de calcul. JOE et XU [1996] proposent de calibrer séparément les densités marginales pour ensuite calibrer la copule. La section suivante présente brièvement cette méthode.

1.4.2 Procédure d'estimation IFM

L'approche consiste à estimer les paramètres du modèle multi-varié en estimant d'abord les paramètres des densités marginales par le maximum de la log-vraisemblance. Elle a été proposée dans JOE et XU [1996]. Dans la suite, on appellera cette méthode : IFM (pour *inference function for marging*). Elle 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.

Plus précisément, on suppose que l'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 $\theta = (\theta_1, \theta_2, \dots, \theta_d, \alpha)$. Les d log-vraisemblance $\ln L_k(\theta_k) = \sum_{t=1}^T \ln f_k(x'_t; \theta_k)$ des lois marginales sont maximisées afin d'obtenir des estimateurs $\hat{\theta}_1, \hat{\theta}_2, \dots, \hat{\theta}_d$ que l'on utilise

ensuite pour maximiser en α $\ln L(\alpha) = \sum_{t=1}^T \ln c\left(F_1\left(x_1^t; \hat{\theta}_1\right), \dots, F_d\left(x_d^t; \hat{\theta}_d\right); \alpha\right)$.

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 de 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, de nombreux tests comparatifs effectués dans XU [1996] sur plusieurs modèles multi-variés montrent que l'IFM est une méthode très efficace et donne des résultats proches de la méthode classique.

1.4.3 Méthode des moments : calibrage par le tau de Kendall

Pour certaines copules, nous disposons des expressions analytiques liant le tau de Kendall au paramètre de la copule (exemple : la copule de Student). L'estimation du tau de Kendall permet donc, dans ces situations, de calibrer les copules. Pour mieux appréhender cette procédure, se référer à DEMARTA-MCNEIL [2004].

1.5 Construction de 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.

1.5.1 La 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 $X = (X_1, \dots, X_d)$, et soit $(r_1^t, \dots, r_d^t)_{1 \leq t \leq T}$ la statistique de rang associée à cet échantillon multi-varié. 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 :

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

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.

1.5.2 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) :

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

1.6 Simulation des modèles multi-variés

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 $F(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d))$. On souhaite 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 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 que $(X_1, \dots, X_d) \approx (F_1^{-1}(U_1), \dots, F_d^{-1}(U_d))$.

Disposant des lois marginales, on remarque 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. La méthode 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 $C_i^{-1}(u / u_1, \dots, u_{i-1})$.

BIBLIOGRAPHIE

AHLGRIM K. C., D'ARCY S. P., GORVETT R. W. [2005] Modeling Financial Scenarios: A Framework for the Actuarial Profession. *Proceedings of the Casualty Actuarial Society* 92. (<http://www.casact.org/pubs/proceed/proceed05/05187.pdf>).

ARMEL K. [2010] *Structure de dépendance des générateurs de scénarios économiques - Modélisation et Simulation*, Mémoire d'actuariat, EURIA.

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.

FANG K., KOTZ S. & NG W. [1990] *Symmetric multivariate and related distributions*, Chapman & Hall.

GENEST C., MACKAY R. J. [1986] *The joy of copulas: Bivariate distributions with uniform marginals*, The American Statistician, 40, 280-283.

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.

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.

NELSEN R. [1999] *An introduction to copulas*, Springer Lecture notes in statistics.

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.

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.