# ANNEXE A L'ARTICLE : QUELLE STRUCTURE DE DEPENDANCE POUR UN GENERATEUR DE SCENARIOS ECONOMIQUES EN ASSURANCE ?
## IMPACT SUR LE BESOIN EN CAPITAL

Version 1.1 du 03/10/2010

**Kamal Armel**[α] **Frédéric Planchet**[*] **Aymric Kamega**

**Université de Lyon, université Lyon 1**

**ISFA - Laboratoire SAF** [β]

**WINTER & Associés** [γ]

**TELECOM Bretagne**[+]

**EURIA**[++]

RESUME

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.

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

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

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

[γ] WINTER & Associés 55 avenue René Cassin - 69009 Lyon - France.

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

[++] EURo Institut d'Actuariat (EURIA) - 6 avenue le Gorgeu - CS93837- 29238 BREST

**SOMMAIRE**

## 1. Structure de dependance 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 $C:[0,1]^d \to [0,1]$ 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 $U=(U_1,\ldots,U_d)$ à valeurs dans $\Re^d$, où les variables aléatoires $U_1,...,U_d$ sont de lois uniformes sur $[0,1]$ : $C(u)=P(U \leq u)$ pour tout $u \in [0,1]^d$. Soit alors $X=(X_1,\ldots,X_d)$ une variable aléatoire à valeurs dans $\Re^d$ 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 $F_1,...,F_d$ et la distribution jointe $F$ :

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

pour tout $x=(x_1,...,x_d) \in \Re^d$. 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 :

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

Si on dispose d'un échantillon d'observations de taille $T$ de $(X_1, X_2)$, $(x_1^t, x_2^t)_{1 \le t \le T}$, on peut construire l'estimateur empirique suivant du tau de Kendall :

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

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

## 1.3. PRESENTATION DES COPULES PARAMETRIQUES 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 $M_d(\Re)$ l'ensemble des matrices carrées réelles de taille $d^2$. Une loi continue est dite elliptique de paramètre de position $\mu = (\mu_1, \ldots, \mu_d) \in \Re^d$ et de matrice de forme symétrique définie positive $\Sigma \in M_d(\Re)$ si sa densité $f$ peut s'écrire pour tout $x = (x_1, \ldots, x_d) \in \Re^d$ :

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

où l'on désigne par $z'$ la transposée de $z$ et $g$ est une fonction à valeurs positives vérifiant $\int_{\Re^d} g(xx')\,dx = 1$. 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 $M_d(\Re)$. 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 $\Sigma$ la « matrice de corrélation ».

La loi d'un vecteur gaussien est un exemple classique de loi elliptique, associé au choix $g(y) = (2\pi)^{-d/2} \mathbf{exp}(-y/2)$. 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 $(u_1,\ldots,u_d) \in [0,1]^d$ :

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

La fonction de distribution $\Phi^{-1}$ est l'inverse de la distribution normale centrée réduite uni-variée. La fonction $\phi_\Sigma(x) = \dfrac{\mathbf{exp}\left(-\frac{1}{2}x\Sigma^{-1}x'\right)}{(2\pi)^{d/2}\mathbf{det}(\Sigma)^{1/2}}$ est la fonction de répartition de la loi normale centrée réduite et $\Sigma$ 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 :

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

où $I_d$ est la matrice unité de $M_d(\Re)$ et $\beta = \left(\Phi^{-1}(u_1),\ldots,\Phi^{-1}(u_d)\right)$.

#### 1.3.1.2 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 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 $(u_1,\dots,u_d)\in[0,1]^d$ :

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

La fonction de distribution $t_\nu^{-1}$ est l'inverse de la distribution de Student centrée réduite uni-variée à $\nu$ degrés de liberté. La fonction $f_{\nu,\Sigma}$ est la densité de probabilité de la loi de Student centrée réduite, $\Sigma$ sa matrice de corrélation et $f_\nu$ est la densité uni-variée de la loi de Student centrée réduite ($\Sigma=1$). À titre de rappel sa densité s'écrit pour tout $x=(x_1,\dots,x_d)\in\Re^d$ :

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

où $\Gamma$ 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 $\phi$ 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 $\phi$ est $d$ fois dérivable sur $]0,1[$ et si pour tout $1\le i\le d$, $\phi^{(i)}>0$ pour $i$ pair et $\phi^{(i)}<0$ sinon. Cette fonction définit alors la copule archimédienne de générateur $\varphi$. De la définition précédente on déduit que la densité d'une copule archimédienne *d*-variée s'écrit :

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

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 $\phi(u) = \alpha^{-1}\left(u^{-\alpha} - 1\right)$. Elle s'écrit donc :

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

Cette copule est différentiable et sa densité s'écrit :

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

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 $\phi(u) = -\mathbf{ln}\left(\frac{e^{-\alpha u} - 1}{e^{-\alpha} - 1}\right)$ où $\alpha \neq 0$ et $u \in ]0,1]$. La copule de Franck *d*-variée s'écrit donc :

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

Pour le cas bi-varié la densité est :

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

Par ailleurs, la densité *d*-variée de la copule de Franck s'écrit à l'aide de l'expression générale présentée *supra* avec $\phi^{-1}(u) = -\frac{\mathbf{ln}\left(\left(e^{-\alpha} - 1\right)e^{-u} + 1\right)}{\alpha}$. La détermination de la densité de la copule de Franck pour un ordre $d$ passe par le calcul de la dérivée première de $\phi$, $\phi'(u)=\frac{\alpha e^{-\alpha u}}{\left(e^{-\alpha u}-1\right)}$, et de la dérivée d'ordre $d$ de $\phi^{-1}$. 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 $\phi^{-1}$ sont égales à :

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

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

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 $\phi(u)=\left(-\mathbf{ln}(u)\right)^{\alpha}$ avec $\alpha>1$ et $u\in\left]0,1\right]$. La copule de Gumbel s'écrit donc :

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

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 :

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

où $\phi(u)=\left(-\mathbf{ln}(u)\right)^{\alpha}$ et $\phi^{-1}(u)=\mathbf{exp}\left(-u^{1/\alpha}\right)=\mathbf{exp}\left(-u^{\beta}\right)$, $\beta=\frac{1}{\alpha}$. La détermination de la forme analytique de cette densité est délicate car le calcul de la dérivée d'ordre $d$ de $\phi^{-1}$ 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 $\phi^{-1}$ et nous permet donc de calculer numériquement la densité de la copule. La méthode consiste à calculer, par récurrence, pour tout réel $u \in \Re$ 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 $\phi^{-1}$ d'ordre $1,...,n-1$ pour en déduire ensuite la dérivée d'ordre *n*. Par récurrence on calcule ainsi la dérivée d'ordre $d$ pour chaque $u \in \Re$. Cette démarche se base sur la formule de Leibniz qui s'écrit pour deux fonctions dérivables en $u \in \Re$ :

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

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 $1 \leq n$ :

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

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 $\phi^{-1}$ est entièrement définie par les dérivées d'ordres inférieurs de $\phi^{-1}$ et par les dérivées de *g*. Donc par récurrence il est simple de calculer la dérivée d'ordre *d* de $\phi^{-1}$. 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 PARAMETRIQUES

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 $X = (X_1, \ldots, X_d)$ peut s'écrire :

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

où $\theta$ est le vecteur des paramètres à estimer, $c$ la densité de la copule associée à la loi de $X = (X_1,...,X_d)$ 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

$\left(x_1^t,\ldots,x_d^t\right)_{1\le t\le T}$ de taille $T$ peut être obtenu directement par le maximum de la log-vraisemblance qui s'écrit $\ln L(\theta)=\sum_{t=1}^{T}\ln\left(f\left(x_1^t,\ldots,x_d^t;\theta\right)\right)$, ce qui est équivalent à :

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

Bien que possible théoriquement, la recherche du vecteur $\theta$ 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, $\theta$, 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 $X=\left(X_1,\ldots,X_d\right)$ dont les densités marginales $f_1,\ldots,f_d$ sont paramétrés respectivement par les paramètres $\theta_1,\theta_2,\ldots,\theta_d$ et dont la copule est paramétré par $\alpha$ de telle sorte que $\theta=\left(\theta_1,\theta_2,\ldots,\theta_d,\alpha\right)$. Les $d$ log-vraisemblance $\ln L_k\left(\theta_k\right)=\sum_{t=1}^{T}\ln f_k\left(x_k^t;\theta_k\right)$ des lois marginales sont maximisées afin d'obtenir des estimateurs $\hat{\theta}_1,\hat{\theta}_2,\ldots,\hat{\theta}_d$ que l'on utilise ensuite pour maximiser en $\alpha$ $\ln L(\alpha)=\sum_{t=1}^{T}\ln c\left(F_1\left(x_1^t;\hat{\theta}_1\right),\ldots,F_d\left(x_d^t;\hat{\theta}_d\right);\alpha\right)$.

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 $\left(x_1^t,\ldots,x_d^t\right)_{1\le t\le T}$ un échantillon d'observations de taille $T$ de $X=\left(X_1,\ldots,X_d\right)$, et soit $\left(r_1^t,\ldots,r_d^t\right)_{1\le t\le T}$ la statistique de rang associée à cet échantillon multi-varié. Pour tout $1\le i\le d$, $r_i^t$ est le rang de $x_i^t$ dans $\left(x_i^t\right)_{1\le t\le T}$. La copule empirique introduite par Deheuvels est définie sur l'ensemble $\left\{\left(\frac{t_1}{T},\ldots,\frac{t_d}{T}\right), 1\le i\le d, t_i=0,1,\ldots,T\right\}$ par l'équation suivante :

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

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 $f_1,\ldots,f_d$ et la densité jointe $f$ d'un vecteur de variables aléatoires $X=\left(X_1,\ldots,X_d\right)$ 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) :

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

## 1.6. SIMULATION DES MODELES MULTI-VARIES

Dans le cas où les fonctions de répartition marginales $F_1,...,F_d$ et une copule $C(u_1,\ldots,u_d)$ sont spécifiées, on peut construire un unique modèle multi-varié dont la distribution jointe est $F(x_1,\ldots,x_d)=C(F_1(x_1),\ldots,F_d(x_d))$. On souhaite simuler les réalisations d'un vecteur aléatoire $(X_1,\ldots,X_d)$ dont les fonctions de répartition marginales sont respectivement $F_1,...,F_d$ 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 $(U_1,\ldots,U_d)$ dont la fonction de répartition est la copule *C*. Ensuite, les trajectoires simulée de $(X_1,\ldots,X_d)$ sont obtenues en s'appuyant sur la propriété selon laquelle le vecteur aléatoire $(F_1^{-1}(U_1),\ldots,F_d^{-1}(U_d))$ admet l'unique fonction de répartition jointe $F$ qui vérifie : $F(x_1,\ldots,x_d)=C(F_1(x_1),\ldots,F_d(x_d))$ et donc que $(X_1,\ldots,X_d)\approx(F_1^{-1}(U_1),\ldots,F_d^{-1}(U_d))$.

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 $2\leq i\leq d-1$ :

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

On écrit aussi $C_1(u_1)=u_1$ et $C_d(u_1,\ldots,u_d)=C(u_1,\ldots,u_d)$. EMBRECHTS et al. [2002] proposent de considérer les distributions conditionnelles pour simuler les trajectoires d'un vecteur de variables aléatoires uniformes $(U_1,\ldots,U_d)$. En effet, si l'on suppose que la fonction de répartition jointe de $(U_1,\ldots,U_d)$ est *C* alors (*cf*. EMBRECHTS et al. [2002]) :

$$\begin{aligned} C_i(u_i/u_1,\ldots,u_{i-1}) &= P(U_i\leq u_i/U_1=u_1,\ldots,U_{i-1}=u_{i-1}) \\ &= \frac{\partial^{i-1}C_i(u_1,\ldots,u_i)}{\partial u_1\ldots\partial u_{i-1}} \Big/ \frac{\partial^{i-1}C_{i-1}(u_1,\ldots,u_{i-1})}{\partial u_1\ldots\partial u_{i-1}} \end{aligned}$$

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 $C_i(u_i/u_1,\ldots,u_{i-1})$ pour tout $2\leq i\leq d-1$. Dans ce cas, nous utilisons l'algorithme suivant pour simuler une trajectoire de $(U_1,\ldots,U_d)$ 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 $C_2\left(u_2 / u_1\right)$ ;

- continuer la même procédure ;

- simuler une valeur $u_d$ de loi $C_d\left(u_d / u_1,\ldots,u_{d-1}\right)$.

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

## 2. 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 marginal*s, 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.