

Chapitre 6

Copules et risques multiples

Arthur Charpentier

Les copules sont devenus en quelques années un outils important pour modéliser les risques multivariés (entre autres). Les copules permettent de « coupler » les lois marginales afin d’obtenir une loi multivariée, d’où le nom latin *copula* choisi par Abe Sklar en 1959 : « having worked out the basic properties of these functions, I wrote about them to Fréchet, in English. He asked me to write a note about them in French. While writing this, I decided I needed a name for those functions. Knowing the word copula as a grammatical term for a word or expression that links a subject and a predicate, I felt that this would make an appropriate name for a function that links a multidimensional distribution to its one-dimensional margins, and used it as such. » (raconté dans (89), et rapporté par (73)).

Ces fonctions sont simplement des fonctions de répartition en dimension quelconque dont les lois marginales sont uniformes sur $[0, 1]$. Comme le notait (39), les copules sont intéressantes « firstly, as a way of studying scale-free measures of dependence, and secondly, as a starting point for constructing families of bivariate distributions ». Ces fonctions ont depuis 60 ans (voir davantage si l’on attribue leur paternité à Wassily Hoeffding) quitté les espaces métriques de probabilité, et les statistiques pour être appliqués en assurance, en environnement, en finance, en fiabilité, etc.

6.1 Introduction

Avant de présenter les copules et plus généralement des outils permettant de modéliser la dépendance entre les risques, il convient de rappeler quelques notions sur l’indépendance. Et avant d’introduire les copules, nous reviendrons un peu sur la loi la plus classique pour modéliser des vecteurs aléatoires : le vecteur Gaussien.

6.1.1 De l'indépendance à la dépendance

Il y a 250 ans (dans (9)), Thomas Bayes définissait l'indépendance de la manière suivante : « events are independent when the happening of any one of the does neither increase nor abate the probability of the rest ». (11) était toutefois un peu plus précis lorsqu'il écrivait « two events are said independent when the probability of the happening of either of them is unaffected by our expectation of the occurrence of failure of the other ».

Pour une définition formelle, Laplace (dans (59)) affirmait que « si les évènements sont indépendants les uns des autres, la probabilité de l'existence de leur ensemble est le produit de leur probabilités particulières ». Autrement dit, deux évènements A et B sont indépendants si et seulement si :

$$\mathbb{P}(A \cap B) = \mathbb{P}(A \text{ et } B) = \mathbb{P}(A, B) = \mathbb{P}(A) \times \mathbb{P}(B).$$

En utilisant la définition de l'espérance conditionnelle, $\mathbb{P}(A \text{ et } B) = \mathbb{P}(A|B) \times \mathbb{P}(B)$, l'indépendance se caractérise également par la relation :

$$\mathbb{P}(A|B) = \mathbb{P}(A).$$

On retrouve alors l'idée énoncée par Boole, dans (11). Par opposition, on dira que deux évènements sont dépendants si

$$\mathbb{P}(A \text{ et } B) = \mathbb{P}(A, B) \neq \mathbb{P}(A) \times \mathbb{P}(B).$$

Par extension, on dira que deux variables aléatoires X et Y sont indépendantes si, pour tout évènement $\mathcal{E} = \mathcal{E}_X \times \mathcal{E}_Y$ de l'espace produit,

$$\mathbb{P}((X, Y) \in \mathcal{E}_X \times \mathcal{E}_Y) = \mathbb{P}(X \in \mathcal{E}_X, Y \in \mathcal{E}_Y) = \mathbb{P}(X \in \mathcal{E}_X) \times \mathbb{P}(Y \in \mathcal{E}_Y).$$

En particulier, pour tout $x, y \in \mathbb{R}$, en posant $\mathcal{E} = (-\infty, x] \times (-\infty, y]$, on doit avoir

$$\begin{aligned} \mathbb{P}((X, Y) \in (-\infty, x] \times (-\infty, y]) &= F(x, y) = \mathbb{P}(X \leq x) \times \mathbb{P}(Y \leq y) \\ &= F_X(x) \times F_Y(y). \end{aligned}$$

Dans le cas des variables discrètes, en considérant $\mathcal{E} = \{(x, y)\}$, l'indépendance se caractérise par

$$\mathbb{P}((X, Y) = (x, y)) = f(x, y) = \mathbb{P}(X = x) \times \mathbb{P}(Y = y) = f_X(x) \times f_Y(y).$$

Cette dernière relation se généralisant aux densités dans le cas de variables continues (en dérivant les fonctions de répartition).

6.1.2 Le vecteur Gaussien

Comme le notait (65), « in multivariate analysis, the only distribution leading to tractable inference is the multivariate normal ». Il reprenait ainsi une

affirmation d'Anscombe ((4)), « the only type of bivariate distribution with which most of us feel familiar (other than the joint distribution of a pair of independent random variables) is the bivariate normal distribution ». Aussi, avant d'étudier les lois multivariées générales, il convient de bien comprendre la construction du vecteur Gaussien.

En dimension d , $\mathbf{X} = (X_1, \dots, X_d)$ admet une distribution Gaussienne si et seulement si, pour tout $\mathbf{a} \in \mathbb{R}^d$, $\mathbf{a}'\mathbf{X}$ suit une loi normale (univariée). Un vecteur Gaussien est alors caractérisé par deux paramètres, $\boldsymbol{\mu} = [\mu_i]_{i=1, \dots, d}$ (correspondant à l'espérance, $\mathbb{E}(\mathbf{X})$) et $\boldsymbol{\Sigma} = [\Sigma_{i,j}]_{i,j=1, \dots, d}$ (correspondant à la variance $\text{Var}(\mathbf{X})$). Comme les lois marginales d'un vecteur Gaussien sont également Gaussiennes, $\mathcal{N}(\mu_i, \Sigma_{i,i})$, la structure de dépendance, qui, ajoutée aux comportements marginaux permet de décrire la loi jointe est alors caractérisée par une matrice de corrélation, $\mathbf{R} = [r_{i,j}]_{i,j=1, \dots, d}$, où $\Sigma_{i,j} = r_{i,j} \sqrt{\Sigma_{i,i} \Sigma_{j,j}}$, en notant que $r_{i,i} = 1$. Autrement dit, la structure de dépendance, en dimension $d \geq 2$, est caractérisée par un ensemble de dépendances deux à deux, $r_{i,j}$.

En allant un peu plus loin, l'introduction de la dépendance à l'aide de cette matrice de corrélation permet de mieux comprendre comment passer d'un vecteur indépendant à un vecteur corrélé. En utilisant la décomposition de Cholesky, on a que si $\mathbf{X}_0^\perp = (X_{0,1}^\perp, \dots, X_{0,d}^\perp)$ est un vecteur Gaussien centré réduit à composantes indépendantes, alors $\mathbf{X} = \boldsymbol{\mu} + \mathbf{A}\mathbf{X}_0^\perp$ est un vecteur Gaussien $\mathcal{N}(\boldsymbol{\mu}, \mathbf{A}\mathbf{A}')$, où \mathbf{A} est une matrice triangulaire inférieure, correspondant à la décomposition de Cholesky de la matrice $\boldsymbol{\Sigma}$ (c'est-à-dire $\mathbf{A}\mathbf{A}' = \boldsymbol{\Sigma}$). On retrouve, avec cette relation, une propriété énoncée par Rosenblatt, à savoir qu'une loi multivariée peut s'écrire à l'aide des lois conditionnelles itérées (on parle parfois de chain rule, ou de la transformation de Rosenblatt, introduite dans (81)),

$$\begin{aligned} f(x_1, x_2, \dots, x_d) &= f_1(x_1) \cdot f_{2|1}(x_2|x_1) \cdot f_{3|2,1}(x_3|x_2, x_1) \cdots \\ &\quad \cdots f_{d|d-1, \dots, 2,1}(x_d|x_{d-1}, \dots, x_2, x_1). \end{aligned}$$

Nous verrons réapparaître ces différentes idées régulièrement dans ce chapitre.

6.1.3 Plan du chapitre

Après cette rapide introduction, nous reviendrons dans la section 6.2 sur les concepts importants pour modéliser des vecteurs aléatoires, en particulier les classes de Fréchet (c'est à dire les classes de lois jointes dont les lois marginales sont données), mais aussi deux formes de symétrie : par rotation (on parlera de lois sphériques, et par extension elliptiques) et par permutation des marges (on parlera d'échangeabilité). La section 6.3 posera les bases sur les copules, en dimension 2 tout d'abord, puis en dimension plus grande ensuite. Nous présenterons des propriétés algébriques de l'espace des copules, tout en présentant quelques extensions importantes lors de l'agrégation des risques ou de l'étude de problèmes de vieillissement (avec les quasi-copules, et les semi-copules). La section 6.4 verra une présentation des principales familles de copules, à commencer par les grandes familles que sont les copules elliptiques, les copules Ar-

chimédiennes, et les copules extrêmes. Nous présenterons alors plus en détails quelques familles à un ou deux paramètres, en particulier les copules de Clayton, de Marshall-Olkin, de Gumbel, et la copule Gaussienne. Ensuite, dans la section 6.5, nous verrons comment quantifier la dépendance, en présentant des mesures de corrélation (d'association ou de concordance, si l'on souhaite utiliser un terme moins connoté que la corrélation). Les corrélations de Pearson, de Spearman, et le τ de Kendall seront ainsi présentés. Nous évoquerons également les mesures de dépendance construites par analyse canonique nonlinéaire (comme extension de la corrélation maximale). Dans la section 6.6 nous aborderons le problème de l'aggrégation des risques multiples, et de la recherche de bornes pour des risques dans une classe de Fréchet donnée. Enfin, nous conclurons, dans la section 6.7 par quelques points sur l'inférence et les tests d'ajustement de copules.

Remarque 6.1 Les copules seront présentées ici afin d'étudier les risques multiples, représentés sous la forme de vecteurs aléatoires $\mathbf{X} = (X_1, \dots, X_d)$. On imagine aisément qu'il serait possible de les utiliser pour étudier des séries temporelles, ou des processus, (X_t) , où l'on s'intéresserait à la structure de dépendance des lois fini-dimensionnelles $(X_{t_1}, X_{t_2}, \dots, X_{t_d})$, où $t_1 \leq t_2 \leq \dots \leq t_d$. La dépendance temporelle ne sera pas détaillée dans ce chapitre, mais peut se trouver par ailleurs (par exemple le chapitre 8 de (55), ou (26) qui établissait un parallèle entre les processus Markoviens et les copules).

6.2 Vecteurs aléatoires de \mathbb{R}^d

Un vecteur aléatoire $\mathbf{X} = (X_1, \dots, X_d)$, à valeurs dans \mathbb{R}^d est une collection de d variables aléatoires réelles. La loi de \mathbf{X} peut être caractérisée par

- sa fonction de répartition, pour $\mathbf{x} = (x_1, \dots, x_d) \in \mathbb{R}^d$:

$$F(\mathbf{x}) = \mathbb{P}(\mathbf{X} \leq \mathbf{x}) = \mathbb{P}(X_1 \leq x_1, \dots, X_d \leq x_d)$$

- sa densité (si elle existe), pour $\mathbf{x} = (x_1, \dots, x_d) \in \mathbb{R}^d$:

$$f(\mathbf{x}) = f(x_1, \dots, x_d) = \left. \frac{\partial^d F(\mathbf{z})}{\partial \mathbf{z}} \right|_{\mathbf{z}=\mathbf{x}},$$

ou sa fonction de probabilité, pour un vecteur à valeurs discrètes

$$f(\mathbf{x}) = \mathbb{P}(X_1 = x_1, \dots, X_d = x_d).$$

- sa fonction génératrice, pour $\mathbf{z} = (z_1, \dots, z_d) \in \mathbb{R}^d$:

$$\psi(\mathbf{z}) = \mathbb{E} \left(e^{\mathbf{z}' \mathbf{X}} \right) = \mathbb{E} \left(e^{z_1 X_1 + \dots + z_d X_d} \right)$$

dès lors que l'espérance existe.

Dans presque tout ce chapitre, nous travaillerons sur les fonctions de répartition pour caractériser la loi d'un vecteur aléatoire.

Remarque 6.1 Sous des hypothèses de dérivabilités, en dimension $d = 1$, les moments de X peuvent être reliés aux dérivées de la fonction génératrice en 0, c'est-à-dire :

$$\mathbb{E}(X^n) = \left. \frac{d^n \psi(z)}{dz^n} \right|_{z=0}.$$

On peut obtenir des notions similaires en dimension $d \geq 2$, en introduisant les co-moments :

$$\mathbb{E}(X_1^{n_1} X_2^{n_2} \dots X_d^{n_d}) = \left. \frac{\partial^{n_1} \psi(z_1, \dots, z_d)}{\partial z_1^{n_1} \partial z_2^{n_2} \dots \partial z_d^{n_d}} \right|_{\mathbf{z}=\mathbf{0}},$$

où $n = n_1 + \dots + n_d$. Les premiers moments sont donc

- un vecteur de taille d pour caractériser l'espérance, $[\mathbb{E}(X_i)]_{i=1, \dots, d}$,
- une matrice $d \times d$ pour caractériser la variance, $[\mathbb{E}(X_i X_j)]_{i,j=1, \dots, d}$,
- un tableau $d \times d \times d$ (parfois mis sous la forme $d \times d^2$) pour caractériser la skewness $[\mathbb{E}(X_i X_j X_k)]_{i,j,k=1, \dots, d}$, etc.

On note que travailler sur les moments d'ordre supérieur à 2 en dimension d grande va rapidement être compliqué.

Remarque 6.2 Les notions d'ordre sur \mathbb{R}^d et les notions de croissance des fonctions $\mathbb{R}^d \rightarrow \mathbb{R}$ sont étroitement liées. En reprenant les notations de (75), un vecteur $\mathbf{a} = (a_1, \dots, a_d) \in \mathbb{R}^d$ est dominé par $\mathbf{b} = (b_1, \dots, b_d) \in \mathbb{R}^d$ pour l'ordre de la majorisation, noté $\mathbf{a} \preceq \mathbf{b}$ si $\sum_{i=1}^d a_i = \sum_{i=1}^d b_i$ et $\sum_{i=1}^k a_{i:d} \leq \sum_{i=1}^k b_{i:d}$, pour tout $k \in \{1, \dots, d\}$, où $a_{i:d}$ est la i plus grande valeur du vecteur \mathbf{a} . De manière plus classique, on peut considérer un ordre par orthant, au sens où $\mathbf{a} \leq \mathbf{b}$ si et seulement si $a_i \leq b_i$ pour tout $i = 1, \dots, d$. Une fonction $h : \mathbb{R}^d \rightarrow \mathbb{R}$ croissante (au sens usuel) si :

$$\mathbf{x} \leq \mathbf{y} \text{ implique } h(\mathbf{x}) \leq h(\mathbf{y}).$$

Mais si on change la notion d'ordre utilisée dans \mathbb{R}^d , on obtient une toute autre notion de croissance. Aussi :

$$\mathbf{x} \preceq \mathbf{y} \text{ implique } h(\mathbf{x}) \leq h(\mathbf{y}),$$

correspond à la notion de Schur convexité (qui est une notion plus forte que la notion de croissance). Pour les fonctions continues $\mathbb{R} \rightarrow \mathbb{R}$, dérivables, la croissance peut être caractérisée par $h'(x) \geq 0$. En dimension supérieure, on peut imaginer une notion liée à une relation de la forme $\partial^d h(\mathbf{x}) / \partial \mathbf{x} \geq 0$ en tout point \mathbf{x} . Cette notion sera appelée d croissance par la suite. Dans le cas non-dérivable, la notion de 2 croissance par exemple se traduira sous la forme :

$$h(b_1, b_2) + h(a_1, a_2) - h(b_1, a_2) - h(a_1, b_2) \geq 0,$$

pour tout $a_i \leq b_i$.

6.2.1 Classes de Fréchet

La modélisation multivariée se voit souvent comme une extension de la modélisation univariée. On suppose donc que l'on connaît les comportements marginaux, c'est la dire les lois univariées. L'outil central est alors la classe des lois multivariées dont les lois marginales sont données, souvent appelées classe de Fréchet.

Définition 6.1 Soient F_1, \dots, F_d d fonctions de répartition $\mathbb{R} \rightarrow [0, 1]$. On notera $\mathcal{F}(F_1, \dots, F_d)$ l'ensemble des fonctions de répartition $\mathbb{R}^d \rightarrow [0, 1]$ dont les lois marginales sont précisément F_1, \dots, F_d .

Comme l'ont montré Fréchet et Hoeffding, les classes de Fréchet sont bornées,

Proposition 6.1 Pour tout $F \in \mathcal{F}(F_1, \dots, F_d)$, et

$$F^-(\mathbf{x}) \leq F(\mathbf{x}) \leq F^+(\mathbf{x}) \text{ pour tout } \mathbf{x} \in \mathbb{R}^d,$$

où

$$F^+(\mathbf{x}) = \min\{F_i(x_i), i = 1, \dots, d\},$$

et

$$F^-(\mathbf{x}) = \max\{0, F_1(x_1) + \dots + F_d(x_d) - (d - 1)\}.$$

$F^+ \in \mathcal{F}(F_1, \dots, F_d)$, et généralement $F^- \notin \mathcal{F}(F_1, \dots, F_d)$, sauf dans le cas où $d = 2$.

Remarque 6.3 En fait, il est possible que $F^- \in \mathcal{F}(F_1, \dots, F_d)$ alors que $d > 2$. Une condition nécessaire et suffisante est d'avoir soit

$$\sum_{i=1}^d F_i(x_i) \leq 1 \text{ pour tout } 0 < F_i(x_i) < 1 \text{ où } i = 1, \dots, d,$$

soit

$$\sum_{i=1}^d F_i(x_i) \geq d - 1 \text{ pour tout } 0 < F_i(x_i) < 1 \text{ où } i = 1, \dots, d.$$

Notons de plus que si $F^- \notin \mathcal{F}(F_1, \dots, F_d)$, F^- n'en est pas moins une borne atteignable.

Cette écriture d'une loi multivariée (en dimension d) à partir de d lois marginales unidimensionnelles n'est pas la seule. Parmi les autres classes, on pourrait aussi s'intéresser aux classes définies par deux paires de lois, $\mathcal{F}(F_1, \dots, F_{d-1}, F_d)$ ou de manière un peu plus générale $\mathcal{F}(F_1, \dots, F_k, F_{k+1}, \dots, F_d)$. En dimension $d = 3$, cela revient à définir la classe $\mathcal{F}(F_{1,2}, F_3)$. Ces classes avaient été étudiées par (44).

Moyennant quelques hypothèses de compatibilité, il est aussi possible de considérer une intersection non-vide, $\mathcal{F}(F_1, \dots, F_k, F_{k+1}, \dots, F_d)$. En dimension $d = 3$,

on peut ainsi considérer $\mathcal{F}(F_{1,2}, F_{2,3})$. A condition que les lois marginales pour la seconde composante soient identiques. Cette classe est équivalente à

$$\mathcal{F}(F_{1|2}, F_{3|2}, F_2).$$

Sous cette forme, on peut noter que cette classe n'est pas vide car elle contient en particulier le vecteur obtenu en supposant X_1 et X_3 indépendants, conditionnellement à X_2 , c'est-à-dire

$$F(x_1, x_2, x_3) = \int_{-\infty}^{x_2} F_{1|2}(u) F_{3|2}(u) dF_2(u).$$

Enfin, parmi les sous-classes de cette classe de Fréchet, on peut aussi s'intéresser aux classes définies par les lois des paires, c'est à dire, en dimension 3,

$$\mathcal{F}(F_{12}, F_{23}, F_{13})$$

moyennant quelques hypothèses de compatibilité. Il n'est pas évident que cette classe soit non-vide, et si elle n'est pas vide, elle n'est pas forcément réduite à un singleton.

Exemple 6.1 Soient $\mathbf{X} = (X_1, X_2, X_3)$ un vecteur Gaussien centré réduit dont les corrélations croisées sont $r + \varepsilon$, où $r \in [-1, 1)$ et $\varepsilon \in [0, 1 - r]$, et $\mathbf{Y} = (Y_1, Y_2, Y_3)$ dont les lois marginales sont des lois $\mathcal{N}(0, 1)$, et telle que

$$\mathbb{P}(Y_1 \leq y_1, Y_2 \leq y_2, Y_3 \leq y_3) = \Phi(y_1)\Phi(y_2)\Phi(y_3)[1 + \theta(1 - \Phi(y_1))(1 - \Phi(y_2))(1 - \Phi(y_3))],$$

avec $\theta \in (-1, 1)$ quelconque. Cette loi est telle que les paires sont non-corrélées (quel que soit θ). On supposera les vecteurs \mathbf{X} et \mathbf{Y} indépendants. Alors

$$\mathbf{Z} = (Z_1, Z_2, Z_3) = \sqrt{\frac{r}{r + \varepsilon}}(X_1, X_2, X_3) + \sqrt{\frac{\varepsilon}{r + \varepsilon}}(Y_1, Y_2, Y_3)$$

est un vecteur dont les couples (Z_i, Z_j) suivent des lois Gaussiennes bivariées de corrélation r (comme cela est montré dans (55) ou (63)).

6.2.2 Symétrie, échangeabilité et indépendance conditionnelle

Un vecteur aléatoire est dit symétrique si la loi jointe est invariante par permutation des composantes,

Définition 6.2 Soit $\mathcal{S}(n)$ la classe des permutations de $\{1, \dots, n\}$. Si $\mathbf{X} = (X_1, \dots, X_n)$ est un vecteur tel que

$$(X_{\sigma(1)}, \dots, X_{\sigma(n)}) \stackrel{\mathcal{L}}{=} (X_1, \dots, X_n), \text{ pour tout } \sigma \in \mathcal{S}(n),$$

alors \mathbf{X} sera dit n -échangeable, ou symétrique.

De manière équivalente, \mathbf{X} sera dit symétrique si

$$H\mathbf{X} \stackrel{\mathcal{L}}{=} \mathbf{X} \text{ pour tout } H \in \mathcal{P}(n),$$

où $\mathcal{P}(n)$ désigne l'ensemble des matrices $n \times n$ de permutation.

La notion d'échangeabilité est aussi appelé interchangeabilité. (8) disait des risques échangeables qu'ils étaient « indistingables ».

Définition 6.3 Une suite finie de variables aléatoires $\{X_1, \dots, X_n\}$ est dite échangeable, ou n -échangeable si

$$(X_1, \dots, X_n) \stackrel{\mathcal{L}}{=} (X_{\sigma(1)}, \dots, X_{\sigma(n)})$$

pour tout permutation σ de $\{1, \dots, n\}$. Par extension, une suite infinie $\{X_1, X_2, \dots\}$ de variables aléatoire est dite échangeable si

$$(X_1, X_2, \dots) \stackrel{\mathcal{L}}{=} (X_{\sigma(1)}, X_{\sigma(2)}, \dots)$$

pour tout permutation finie σ de \mathbb{N}^* (c'est-à-dire $\text{Card}\{i, \sigma(i) \neq i\} < \infty$).

Définition 6.4 Une suite n -échangeable (X_i) sera dite m -extensible (pour $m > n$) si $(X_1, \dots, X_n) \stackrel{\mathcal{L}}{=} (Z_1, \dots, Z_n)$ où (Z_i) est une suite de variables m -échangeable.

Exemple 6.2 Soit X_1, \dots, X_n une séries de variables aléatoires telles que

$$\text{Var}(X_i) = \sigma^2 \text{ et } \text{cov}(X_i, X_j) = \rho\sigma^2,$$

pour tout $i = 1, \dots, n$, et $j \neq i$. Alors

$$\begin{aligned} 0 &\leq \text{Var}\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \text{Var}(X_i) + \sum_{i \neq j} \text{cov}(X_i, X_j) \\ &= n\sigma^2 + n(n-1)\rho\sigma^2, \end{aligned}$$

Aussi, forcément :

$$\rho \geq -\frac{1}{n-1}.$$

L'échangeabilité implique une forme de dépendance positive entre les composantes (en fait, l'échangeabilité est une notion relativement forte de dépendance positive).

Le théorème de de Finetti (démontré dans (28) dans le cas de variables de Bernoulli et étendu par (51)) affirme que les mesures échangeables peuvent être écrites comme des mélanges de mesures produits, ce qui correspond à une propriété d'indépendance conditionnelle : il existe \mathcal{S} telle que pour tout A_i ,

$$\mathbb{P}(X_1 \in A_1, \dots, X_n \in A_n | \mathcal{S}) = \mathbb{P}(X_1 \in A_1 | \mathcal{S}) \times \dots \times \mathbb{P}(X_n \in A_n | \mathcal{S}).$$

Théorème 6.1 *Soit X_1, X_2, \dots une suite échangeable, alors il existe \mathcal{S} telle que, sachant \mathcal{S} les X_i sont indépendants.*

La preuve de ce théorème est établie dans (19).

Corollaire 6.1 *Soit X_1, X_2, \dots une suite échangeable de variables suivant une loi de Bernoulli, et notons $S_n = X_1 + \dots + X_n$. Alors la distribution de S_n est un mélange de lois binomiales, c'est-à-dire il existe une loi H définie sur $[0, 1]$ telle que*

$$\mathbb{P}(S_n = k) = \int_0^1 \binom{n}{k} \omega^k (1 - \omega)^{n-k} dH(\omega).$$

Exemple 6.3 Considérons un portefeuille constitué de n polices, observées pendant une période de temps identique. Soit I_i la variable aléatoire qui indique si la police i a eu des sinistres, ou pas (c'est-à-dire I_i suit une loi de Bernoulli). La variable d'intérêt pour l'assureur est $I = I_1 + \dots + I_n$, le nombre de polices qui ont eu un sinistre. Supposer les polices indépendantes peut être irréaliste dans certains cas. (24) suggère d'introduire de la dépendance en supposant que chaque police est soumise à un choc commun, qui pourrait être une inondation, ou un tremblement de terre, par exemple. Aussi $I_i = \min\{J_i + J_0, 1\}$, où les variables J_i sont indépendantes. En suivant l'idée de (83), supposons que les variables I_1, \dots, I_n aient été extraites d'un ensemble infini de variables échangeables de variables indicatrices (I_i). Le théorème de de Finetti permet de garantir qu'il existe une variable Θ telle que les I_i soient indépendantes conditionnellement à Θ . Aussi

$$\mathbb{P}(I_1 + \dots + I_n = k) = \binom{n}{k} \int \theta^k (1 - \theta)^{n-k} dG(\theta)$$

où G est une fonction de répartition sur $[0, 1]$. (61) propose ainsi de transposer les modèles d'épidémie en assurance, où l'on suppose souvent les variables n -échangeables. Supposons que m catastrophes puissent affecter n polices d'un portefeuille d'un assureur. Chaque catastrophe j peut toucher D_j polices, indépendamment, dans le portefeuille, parmi les n . Un assuré est touché par la catastrophe j avec une probabilité p_j . On note I_i la variable indicatrice valant 1 si l'assuré i a subi une perte (au moins une fois). Notons que $(1 - I_1, \dots, 1 - I_n)$ sont des variables de Bernoulli n -échangeables. Le nombre de polices touchées par la catastrophe j est $N_j \sim \mathcal{B}(D_j, 1 - p_j)$, et donc

$$\pi_k = \mathbb{P}(I_1 = 0, \dots, I_n = 0) = \prod_{j=1}^m \mathbb{E} \left(\binom{n-k}{N_j} \binom{n}{N_j}^{-1} \right)$$

pour tout $k = 0, \dots, n$.

6.2.3 Lois sphériques et elliptiques

Si les symmétries constituent une notion naturelle d'invariance, on peut aussi définir une invariance par rotations (centrées sur l'origine pour des vecteurs centrés).

Définition 6.5 *On dira que \mathbf{X} a une distribution sphérique si*

$$H\mathbf{X} \stackrel{\mathcal{L}}{=} \mathbf{X} \text{ pour toute matrice } H \in \mathcal{O}(n),$$

où $\mathcal{O}(n)$ désigne l'ensemble des matrices $n \times n$ orthogonales (c'est-à-dire $H'H = \mathbb{I}$).

Par exemple la distribution $\mathcal{N}(\mathbf{0}, \mathbb{I})$ est sphérique. Plusieurs caractérisations des distributions sphériques sont possibles, comme le montre. (37) :

1. $H\mathbf{X} \stackrel{\mathcal{L}}{=} \mathbf{X}$ pour tout $H \in \mathcal{O}(n)$,
2. la fonction caractéristique de \mathbf{X} , $\mathbb{E}(e^{i\mathbf{t}'\mathbf{X}})$ est de la forme $\psi : \mathbf{t} \mapsto \phi(\mathbf{t}'\mathbf{t})$, où ϕ est une fonction $\mathbb{R}_+ \rightarrow \mathbb{R}_+$,
3. \mathbf{X} admet la représentation $\mathbf{X} = R \cdot \mathbf{U}$ où R est une variable aléatoire positive, indépendante de \mathbf{U} , vecteur uniformément distribué sur la sphère unité de \mathbb{R}^n ,
4. pour tout $\mathbf{a} \in \mathbb{R}^n$, $\mathbf{a}'\mathbf{X} \stackrel{\mathcal{L}}{=} \|\mathbf{a}\|_2 X_i$ pour tout $i = 1, \dots, n$, où $\|\cdot\|_2$ désigne la norme euclidienne.

La loi sphérique est alors complètement caractérisée par ϕ , ou par la loi de R . La construction d'une loi sphérique à partir de la relation $\mathbf{X} = R \cdot \mathbf{U}$ se visualise sur la Figure 6.1. Notons que l'idée de la construction est d'écrire $\mathbf{X} = \|\mathbf{X}\|_2 \cdot \mathbf{U}$ où $\mathbf{U} = \mathbf{X}/\|\mathbf{X}\|_2$ (qui est effectivement un vecteur de norme 1).

En reprenant la construction du vecteur Gaussien $\mathcal{N}(\mu, \Sigma)$ à partir du vecteur Gaussien centré, à composantes indépendantes, $\mathcal{N}(\mathbf{0}, \mathbb{I})$, on va pouvoir construire des distributions elliptiques à partir de ces distributions sphériques.

Définition 6.6 *On dira que \mathbf{X} a une distribution elliptique si $\mathbf{X} \stackrel{\mathcal{L}}{=} \mu + \mathbf{A}\mathbf{Y}$ où \mathbf{Y} a une distribution sphérique, où $\mathbf{A}\mathbf{A}' = \Sigma$ est une matrice de variance-covariance.*

Par exemple, la distribution $\mathcal{N}(\mu, \Sigma)$ est elliptique. On dira que \mathbf{X} aura une distribution elliptique de paramètres μ , $\Sigma = \mathbf{A}\mathbf{A}'$ et ϕ (caractérisant la distribution sphérique sous-jacente).

Il existe là aussi plusieurs caractérisations de ces distributions sphériques, en particulier, \mathbf{X} admet une représentation de la forme $\mathbf{X} \stackrel{\mathcal{L}}{=} \mu + R \cdot \mathbf{A} \cdot \mathbf{U}$ où R est une variable positive, indépendante de \mathbf{U} uniformément distribuée sur la sphère unité de \mathbb{R}^n , et \mathbf{A} vérifie $\mathbf{A}\mathbf{A}' = \Sigma$.

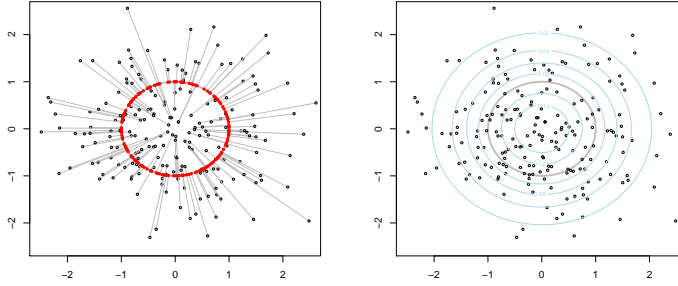


Figure 6.1 – Construction d'une loi sphérique par le relation $\mathbf{X} = R \cdot \mathbf{U}$, avec une distribution uniforme sur le cercle unité, \mathbf{U} et la composante radiale R , indépendante de \mathbf{U} .

Exemple 6.4 La distribution t de Student, de paramètres $\nu > 0$, $\boldsymbol{\mu}$ et $\boldsymbol{\Sigma}$, est obtenue comme loi du vecteur $\boldsymbol{\mu} + A\sqrt{\nu}\mathbf{Z}/S$, où $\mathbf{Z} \sim \mathcal{N}(0, \mathbb{I})$ et $S \sim \chi^2(\nu)$ sont indépendants. La densité de la loi t de Student, de paramètres ν , $\boldsymbol{\mu}$ et $\boldsymbol{\Sigma}$, est alors

$$\mathbf{x} \mapsto \frac{\Gamma((n+\nu)/2)}{(\pi\nu)^{n/2} \Gamma(\nu/2)} |\boldsymbol{\Sigma}|^{-1/2} \left(1 + \frac{1}{\nu} (\mathbf{x} - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right)^{-(n+\nu)/2}$$

(pour reprendre les notations de (37)). Cette densité est représentée sur la Figure 6.2. Notons que si $\nu = 1$, on retrouve la loi de Cauchy multivariée. On notera que si $\nu > 1$, $\boldsymbol{\mu}$ correspond à l'espérance de \mathbf{X} , et que si $\nu > 2$, $\frac{\nu}{\nu-2}\boldsymbol{\Sigma}$ correspond à la matrice de variance-covariance de \mathbf{X} .

Les lois elliptiques ont beaucoup utilisées compte tenu de leur flexibilité apparente. Elle vérifie également des propriétés intéressantes de stabilité par addition : si \mathbf{X} est elliptique de générateur ϕ , alors $\mathbf{a} + \mathbf{b}\mathbf{X}$ également. Ces lois ont également stable par sous indice, car si $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2)$ est elliptique de générateur ϕ , alors il en sera de même pour \mathbf{X}_1 et \mathbf{X}_2 . La loi conditionnelle $\mathbf{X}_1 | \mathbf{X}_2 = \mathbf{x}_2$ sera aussi elliptique, pour tout \mathbf{x}_2 (mais le générateur sera alors différent, comme le montre (37)). Les applications des lois elliptiques en finance sont développées dans (?).

6.2.4 La distribution de Dirichlet

Parmi les autres distributions construites à partir d'une loi symétrique, on peut aussi penser à la loi de Dirichlet. Notons $\mathcal{G}(\alpha)$ la loi Gamma, de densité

$$y \mapsto \frac{y^{\alpha-1} \exp[-y]}{\Gamma(\alpha)} \text{ pour } y \geq 0.$$

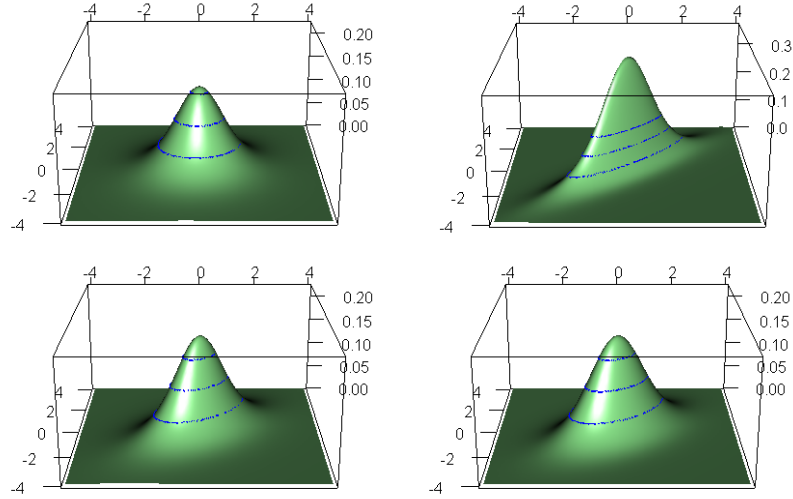


Figure 6.2 – Exemples de densités de loi t de Student, avec en haut, de gauche à droite, $(r = 0.1, \nu = 4)$ et $(r = 0.9, \nu = 4)$, et en bas $(r = 0.5, \nu = 4)$ et $(r = 0.5, \nu = 10)$.

Soient Y_1, \dots, Y_n, Y_{n+1} des variables indépendantes de loi Gamma, de paramètres α_i (distincts dans le cas général, identiques pour construire une loi symétrique). Le vecteur

$$\mathbf{X} = \left(\frac{Y_1}{Y_1 + \dots + Y_n + Y_{n+1}}, \dots, \frac{Y_n}{Y_1 + \dots + Y_n + Y_{n+1}} \right)$$

est une loi de Dirichlet, de paramètre $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_n, \alpha_{n+1})$. Cette loi est définie dans le simplexe de \mathbb{R}^{n+1} :

$$\mathcal{S} = \{\mathbf{u} \in [0, \infty)^n, \sum_{i=1}^n u_i \leq 1\}.$$

Cette loi est aussi appelée loi Beta multivariée, car les composantes marginales suivent des lois Beta (comme (30) dans la section XI.4).

6.2.5 Modélisation des portefeuilles homogènes de crédit

Dans les modèles standard, comme KMV (détaillé dans (23)) ou CreditMetrics (détaillé dans (48)), la modélisation des défauts joints dans un portefeuille d'émetteurs obligataires est relativement simple, souvent basée sur une hypothèse de normalité. Ces approches ont beaucoup été critiquées compte tenu de leur difficulté à modéliser les gros portefeuilles, où des phénomènes de contagion peuvent être observés (dans une région géographique, ou un secteur d'activité).

Considérons m émetteurs, que l'on suit sur un horizon temporel donné T (classiquement une année). Soit Y_i l'indicateur du défaut de l'émetteur i à horizon T , c'est-à-dire $Y_i = 1$ si l'entreprise i fait défaut avant la date T , et $Y_i = 0$ sinon. Notons également que des modèles multinomiaux (voire multinomiaux ordonnés, dans le cas des ratings) ont aussi été considérés.

Deux formes de modèles ont généralement été proposés dans les années 90,

1. des modèles à variables latentes : les défauts surviennent quand la variable latente X_i (interprétable comme la valeur de l'entreprise i à horizon T) passe en dessous d'un seuil D_i (signifiant que sa dette est devenu beaucoup trop grance). Le modèle de Merton ((68)), CreditMetrics ou KMV sont basés sur cette idée.
2. des modèles de mélange : les probabilités de défaut sont fonction d'un facteur Θ , ou plus précisément, Y_i sachant Θ_i suit une loi de Bernoulli $\mathcal{B}(\Theta_i)$, où p_i est une variable aléatoire à valeurs dans $[0, 1]$. Les variables $\Theta_1, \dots, \Theta_n$ sont généralement supposées indépendantes, comme par exemple CreditRisk+ (détaillé dans (94)).

Dans les modèles latents, on suppose qu'il existe $\mathbf{Y}^* = (Y_1^*, \dots, Y_d^*)$, à marges continues, de lois respectives F_1, \dots, F_d , et considérons des seuils D_1, \dots, D_d (en dessous desquels les compagnies sont en faillite, et font défaut). Posons $Y_i = \mathbf{1}\{Y_i^* \leq D_i\}$, de telle sorte que la probabilité de défaut pour l'émetteur i s'écrit

$$p_i = \mathbb{P}(Y_i = 1) = \mathbb{P}(Y_i^* \leq D_i) = F_i(D_i), \text{ pour } i = 1, \dots, d.$$

(68) suppose que X_i est la valeur de l'entreprise i à la date T , et D_i est la valeur de dette acceptable. On suppose également que $\mathbf{Y}^* \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Le modèle KMV est très proche de ce modèle de Merton, les D_i étant choisis de telle sorte que la probabilité de défaut p_i soit égale à la moyenne des probabilités de défauts des entreprises situées à la même distance du défaut que l'entreprise i . Dans CreditMetrics, $\mathbf{Y}^* \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma})$ et les seuils D_i étant choisis de telle sorte que les probabilités de défaut p_i soit égales à la moyenne des probabilités de défauts des entreprises ayant le même rating que l'entreprise i . Finalement, (62) suggère de supposer que Y_i^* soit interprété comme la durée de vie résiduelle de l'entreprise i . Dans ce cas, Y_i^* suit une loi exponentielle de paramètre λ_i , tel que $\mathbb{P}(Y_i^* \leq T) = p_i$, où p_i est déterminé comme dans la méthodologie développée par CreditMetrics.

On notera que les portefeuilles homogènes considérés par KMV et CreditMetrics sont à reliés de la notion de modèle échangeable : le vecteur latent \mathbf{Y}^* est échangeable, et que les probabilités de défaut sont supposées égales $\mathbb{P}(Y_i^* \leq D_i) = p$, pour tout $i = 1, \dots, d$, et donc \mathbf{Y} est échangeable.

Considérons les deux cas suivants,

$$\begin{cases} \mathbf{Y}^* \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}), \text{ distribution Gaussienne} \\ \mathbf{Y}^* \sim t_\nu(\mathbf{0}, \boldsymbol{\Sigma}), \text{ distribution Student } t \end{cases}$$

où $\boldsymbol{\Sigma}$ est une matrice de corrélation, avec une unique corrélation croisée $\rho > 0$, telle que \mathbf{Y}^* soit échangeable. Les seuls sont fixés, et donc, \mathbf{Y} est échangeable.

Ces modèles de mélange ont été intensivement utilisés car ils permettent d'obtenir des modèles facilement interprétables.

Supposons que le vecteur des indicatrices de défaut \mathbf{Y} suit une mélange échangeable de Bernoulli, c'est-à-dire il existe Θ à valeurs dans $[0, 1]$ telle que, conditionnellement à Θ les variables Y_1, \dots, Y_d soient i.i.d., Bernoulli, de paramètre Θ , c'est-à-dire

$$p = \mathbb{P}(\text{défaut}) = \mathbb{P}(Y_i = 1) = \mathbb{E}(Y_i) = \mathbb{E}(\Theta)$$

$$p_k = \mathbb{P}(k \text{ défauts}) = \mathbb{P}(Y_{i_1} = 1, \dots, Y_{i_k} = 1) = \mathbb{E}(\Theta^k) = \int_0^1 q^k dH(q),$$

où H est la distribution de la variable de mélange Θ . Afin de construire un modèle, plusieurs lois peuvent être considérées

$$\begin{cases} \text{loi Beta : } dH(q) = \beta(\alpha, \beta)^{-1} q^{\alpha-1} (1-q)^{\beta-1} dq \\ \text{Probit-normal : } \Phi^{-1}(\Theta) \sim \mathcal{N}(\mu, \sigma^2) \\ \text{Logit-normal : } \log(\Theta/(1-\Theta)) \sim \mathcal{N}(\mu, \sigma^2) \end{cases}$$

Le modèle Probit avait été retenu par CreditMetrics et KMV, alors que CreditRisk+ était basé sur un modèle Logit.

Remarque 6.4 Plus généralement, (90) s'est penché sur la modélisation des durées de vie résiduelles échangeables, et (64) sur la modélisation de variables indicatrice (défaut ou décès).

6.3 Les copules

La notion de copule a été introduite par Sklar en 1959 (qui avait alors utilisé le mot latin *copula* comme rappelé dans l'introduction), motivé par les travaux de Fréchet dans les années 50 sur les tableaux de contingence. Les copules sont aussi été appelées « fonction de dépendance » par Deheuvels en 1979 (dans (29)), ou « représentation uniforme » par Kimeldorf et Sampson en 1975 (dans (56)).

Remarque 6.2 On notera qu'elles apparaissaient déjà dans les travaux de Hoeffding, mais qui avait considéré comme loi de référence des lois uniformes sur $[-1/2, +1/2]$. Ormis cette petite différence, toute la théorie des copules se trouve introduite dans (52).

6.3.1 Les copules en dimension 2

Définition 6.7 Une copule $C : [0, 1]^2 \rightarrow [0, 1]$ est une fonction de répartition dont les lois marginales sont uniformes sur $[0, 1]$.

De manière équivalente, on a la caractérisation suivante :

Théorème 6.2 Une copule $C : [0, 1]^2 \rightarrow [0, 1]$ est une fonction qui vérifie les trois conditions suivantes :

- $C(u_1, 0) = C(0, u_2) = 0$ pour tout $u_1, u_2 \in [0, 1]$,
- $C(u_1, 1) = u_1$ et $C(1, u_2) = u_2$ pour tout $u_1, u_2 \in [0, 1]$,
- C est une fonction 2-croissante, c'est-à-dire pour tout $0 \leq u_i \leq v_i \leq 1$,

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

Les deux premières conditions se traduisent graphiquement par les conditions de bords de la Figure 6.3,

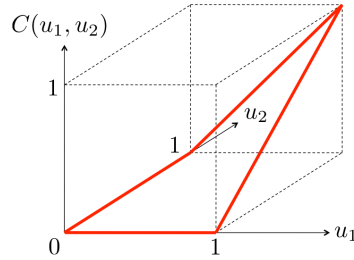


Figure 6.3 – Conditions de bord d'une copule en dimension $d = 2$ (avec pour la partie inférieure $C(u_1, 0) = C(0, u_2) = 0$, et pour la partie supérieure $C(u_1, 1) = u_1$ et $C(1, u_2) = u_2$).

Si $\mathbf{U} = (U_1, U_2)$ admet pour fonction de répartition C , alors $\mathbf{1} - \mathbf{U} = (1 - U_1, 1 - U_2)$ a pour fonction de répartition C^* , que l'on appellera copule de survie, ou copule duale, associée à C :

Définition 6.8 La copule de survie C^* associée à la copule C est la copule définie par

$$C^*(u_1, u_2) = u_1 + u_2 - 1 + C(1 - u_1, 1 - u_2).$$

Exemple 6.5 Soit U une variable uniformément répartie sur $[0, 1]$. Alors $\mathbf{U} = (U, U)$ admet pour fonction de répartition C^+ (en utilisant les notations de la Proposition 6.1), où :

$$C^+(u_1, u_2) = \min\{u_1, u_2\},$$

alors que $\mathbf{U} = (U, 1 - U)$ admet pour fonction de répartition C^- où :

$$C^-(u_1, u_2) = \max\{0, u_1 + u_2 - 1\}.$$

Ces fonctions sont des copules puisqu'il s'agit de fonction de répartition de couples dont les lois marginales sont uniforme sur $[0, 1]$. Il s'agit des bornes de Fréchet-Hoffding de l'ensemble des copules en dimension 2, au sens de la Proposition 6.1 : pour toute copule C ,

$$C^-(u_1, u_2) \leq C(u_1, u_2) \leq C^+(u_1, u_2) \text{ pour tout } 0 \leq u_1, u_2 \leq 1.$$

Exemple 6.6 Si \mathbf{X} est un vecteur à composantes marginales continues, de loi $F \in \mathcal{F}(F_1, F_2)$, on peut construire une copule induite par F en posant $U_i = F_i(X_i)$, $i = 1, 2$. Ces deux lois étant uniformes sur $[0, 1]$, comme rappelé dans le chapitre 3, la fonction de répartition du couple $\mathbf{U} = (U_1, U_2)$ est une copule, qui peut s'écrire :

$$C(u_1, u_2) = F(F_1^{-1}(u_1), F_2^{-1}(u_2)), \text{ pour tout } u_1, u_2 \in [0, 1].$$

De manière générale, en notant h^{-1} l'inverse généralisé d'une fonction $h : \mathbb{R} \rightarrow \mathbb{R}$ monotone, définie par $h^{-1}(t) = \inf\{x, h(x) \geq t, t \in \mathbb{R}\}$, on peut construire une copule en posant $C(u_1, u_2) = F(F_1^{-1}(u_1), F_2^{-1}(u_2))$.

Exemple 6.7 Soit $\{X_1, \dots, X_n\}$ un échantillon i.i.d. de loi F , et notons $X_{i:n}$ la i ème plus grande valeur. Rappelons que les fonctions de répartition de $X_{1:n} = \min\{X_1, \dots, X_n\}$ et $X_{n:n} = \max\{X_1, \dots, X_n\}$ sont respectivement données par

$$F_{1:n}(x) = 1 - [1 - F(x)]^n \text{ et } F_{n:n}(x) = [F(x)]^n \text{ pour tout } x,$$

et de plus, la fonction de répartition du couple est donnée par

$$F_{1,n}(x, y) = \begin{cases} F(x)^n - (F(y) - F(x))^n & , \text{ si } x < y \\ F(y)^n & , \text{ si } x \geq y \end{cases}.$$

On en déduit alors que le copule du couple $(X_{1:n}, X_{n:n})$, noté C_n est de la forme suivante

$$C_{1,n}(u, v) = \begin{cases} v - \left(v^{1/n} + (1-u)^{1/n} - 1\right)^n & , \text{ si } 1 - (1-u)^{1/n} < v^{1/n} \\ v & , \text{ si } 1 - (1-u)^{1/n} \geq v^{1/n} \end{cases},$$

appelé copule min-max par (84). On peut noter que si $n \rightarrow \infty$ alors $C_n \rightarrow C^\perp$ (par rapport à la norme uniforme).

Les copules sont des fonctions continues. Plus précisément, elles vérifient une condition de Lipschitz : pour tout $0 \leq u_i, v_i \leq 1$,

$$|C(u_1, u_2) - C(v_1, v_2)| \leq |u_1 - v_1| + |u_2 - v_2|.$$

6.3.2 Les copules en dimension $d > 2$

La propriété de croissance de la fonction de répartition est liée à la propriété de d -croissance, qui s'interprète comme le fait que $\mathbb{P}(\mathbf{x} \leq \mathbf{X} \leq \mathbf{y}) \geq 0$ pour $\mathbf{X} = (X_1, \dots, X_d)$ de loi F , pour tout $\mathbf{x} \leq \mathbf{y}$ (au sens où $x_i \leq y_i$).

Définition 6.9 Une fonction $h : \mathbb{R}^d \rightarrow \mathbb{R}$ est dite d -croissante si pour tout hyper-rectangle $[\mathbf{a}, \mathbf{b}]$ de \mathbb{R}^d , $V_h([\mathbf{a}, \mathbf{b}]) \geq 0$, où :

$$V_h([\mathbf{a}, \mathbf{b}]) = \Delta_{\mathbf{a}}^{\mathbf{b}} h(\mathbf{t}) = \Delta_{a_d}^{b_d} \Delta_{a_{d-1}}^{b_{d-1}} \dots \Delta_{a_2}^{b_2} \Delta_{a_1}^{b_1} h(\mathbf{t}), \quad (6.1)$$

pour tout \mathbf{t} , où :

$$\Delta_{a_i}^{b_i} h(\mathbf{t}) = h(t_1, \dots, t_{i-1}, b_i, t_{i+1}, \dots, t_n) - h(t_1, \dots, t_{i-1}, a_i, t_{i+1}, \dots, t_n). \quad (6.2)$$

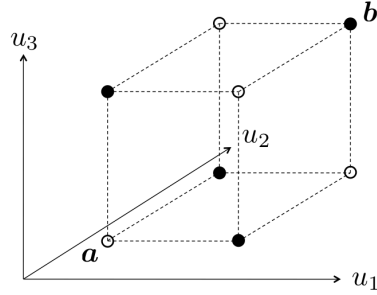


Figure 6.4 – La notion de 3-croissance : on somme la valeur aux différents sommets de l'hyperrectangle $[a, b]$, avec un signe positif (ici •) ou négatif (ici ◦).

Définition 6.10 Une copule en dimension d est une fonction de répartition sur $[0, 1]^d$ dont les lois marginales sont uniformes sur $[0, 1]$.

De manière équivalente, les copules sont des fonctions $C : [0, 1]^d \rightarrow [0, 1]$ telles que pour tout $0 \leq u_i \leq 1$ pour tout $i = 1, \dots, d$ les conditions suivantes sont vérifiées,

$$C(1, \dots, 1, u_i, 1, \dots, 1) = u_i, \quad (6.3)$$

$$C(u_1, \dots, u_{i-1}, 0, u_{i+1}, \dots, u_d) = 0, \quad (6.4)$$

$$C \text{ est } d\text{-croissante.} \quad (6.5)$$

En fait les équations 6.3 et 6.4 impliquent que les marges sont uniformément réparties sur $[0, 1]$. De plus, C vérifie une propriété de Lipschitz, garantissant la continuité de la fonction C (comme en dimension 2). Le résultat le plus important sur les copules est le théorème de Sklar (démontré dans (88) ou (73)),

Théorème 6.3 1. Si C est une copule, et F_1, \dots, F_d des fonctions de répartition univariées, alors pour tout $(x_1, \dots, x_d) \in \mathbb{R}^d$,

$$F(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d)) \quad (6.6)$$

est une fonction de répartition de $\mathcal{F}(F_1, \dots, F_d)$.

2. Réciproquement, si $F \in \mathcal{F}(F_1, \dots, F_d)$, il existe une copule C satisfaisant l'équation 6.6. Cette copule n'est pas forcément unique, mais si les lois marginales F_1, \dots, F_d sont continues, elle le sera, avec, pour tout $(u_1, \dots, u_d) \in [0, 1]^d$,

$$C(u_1, \dots, u_d) = F(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d)), \quad (6.7)$$

où les fonctions quantiles $F_1^{-1}, \dots, F_d^{-1}$ sont les inverses généralisés (continues à gauche) des fonctions F_i .

Ce théorème a permis de motiver l'utilisation des copules en tant que « fonctions de dépendance », permettant de capturer des propriétés de dépendance invariants par changement d'échelle,

Proposition 6.2 *Soit (X_1, \dots, X_d) un vecteur aléatoire de copule C . Soient $\phi_1, \dots, \phi_d, \phi_i : \mathbb{R} \rightarrow \mathbb{R}$ des fonctions continues strictement croissantes, alors C est la copule du vecteur $(\phi_1(X_1), \dots, \phi_d(X_d))$.*

Une démonstration de ce résultat se trouve dans (55), ou (35) dans un cadre plus général : en effet, notons que l'hypothèse de continuité des ϕ_i n'est pas nécessaire si les marges sont absolument continues.

Définition 6.11 *Soit C une copule, alors la fonction*

$$C^*(u_1, \dots, u_d) = \sum_{k=0}^d \left((-1)^k \sum_{i_1, \dots, i_k} C(1, \dots, 1, 1 - u_{i_1}, 1, \dots, 1, 1 - u_{i_k}, 1, \dots, 1) \right), \quad (6.8)$$

pour tout $(u_1, \dots, u_d) \in [0, 1] \times \dots \times [0, 1]$, est une copule, appelée copule de survie, ou copule duale, de C .

Si $\mathbf{U} = (U_1, \dots, U_d)$ a pour fonction de répartition C , alors C^* est la fonction de répartition du vecteur $\mathbf{1} - \mathbf{U} = (1 - U_1, \dots, 1 - U_d)$. Et si

$$\mathbb{P}(X_1 \leq x_1, \dots, X_d \leq x_d) = C(\mathbb{P}(X_1 \leq x_1), \dots, \mathbb{P}(X_d \leq x_d)),$$

pour tout $(x_1, \dots, x_d) \in \mathbb{R}$, alors

$$\mathbb{P}(X_1 > x_1, \dots, X_d > x_d) = C^*(\mathbb{P}(X_1 > x_1), \dots, \mathbb{P}(X_d > x_d)).$$

Exemple 6.8 Soit U une variable uniformément répartie sur $[0, 1]$. Alors $\mathbf{U} = (U, U, \dots, U)$ admet pour fonction de répartition C^+ (en utilisant les notations de la Proposition 6.1 et de l'exemple 6.5), où :

$$C^+(u_1, \dots, u_d) = \min\{u_1, \dots, u_d\},$$

On peut aussi définir

$$C^-(u_1, \dots, u_d) = \max\{0, u_1 + \dots + u_d - (d - 1) - 1\}$$

qui correspond à la borne inférieure de Fréchet-Hoffding de l'ensemble des copules au sens de la Proposition 6.1 : pour toute copule C ,

$$C^-(u_1, \dots, u_d) \leq C(u_1, \dots, u_d) \leq C^+(u_1, \dots, u_d) \text{ pour tout } 0 \leq u_1, \dots, u_d \leq 1.$$

6.3.3 Propriétés de l'ensemble des copules

Proposition 6.3 *La classe des copules est convexe, c'est-à-dire que si $\{C_\theta, \theta \in \Theta\}$ est une famille de copules, et que H est une fonction de répartition sur Θ , alors*

$$C(x_1, \dots, x_n) = \int_{\Theta} C_\theta(x_1, \dots, x_n) dH(\theta)$$

est une copule.

Une preuve de ce résultat peut être trouvée dans (73).

Exemple 6.9 Les fonctions $C^\perp(u_1, \dots, u_d) = \prod_{i=1}^d u_i$ et $C^+(u_1, \dots, u_d) = \min\{u_1, \dots, u_d\}$, sont des copules (respectivement la copule indépendante et la borne supérieure de Fréchet-Hoeffding), et pour tout $\alpha \in [0, 1]$,

$$C(u_1, \dots, u_d) = \alpha C^\perp(u_1, \dots, u_d) + [1 - \alpha] C^+(u_1, \dots, u_d),$$

est une copule.

6.3.4 Copules, quasi-copules et semi-copules

Quelques définitions

Les copules ont été présentées en introduisant une notion de d -croissance (étendant la notion de croissance dans le cas univarié) légitimée car elle se traduit pas le fait qu'elle est équivalente à avoir une masse positive sur tout hyperrectangle, c'est-à-dire

$$\mathbb{P}(U \in [a, b]) = \mathbb{P}(a \leq U \leq b) \geq 0,$$

où $[a, b] = [a_1, b_1] \times \dots \times [a_d, b_d]$.

D'autres notions de "croissance" ont été définies dans la littérature. Une condition plus faible est de supposer une relation de ce genre mais uniquement pour les hyperrectangles touchant un bord, c'est-à-dire il existe $i \in \{1, \dots, d\}$ tel que $a_i = 0$ ou $b_i = 1$. Comme l'a montré (3), cette condition est équivalente à demander une croissance par composante, et que la condition de Lipschitz soit vérifiée.

Définition 6.12 *Une fonction $Q : [0, 1]^d \rightarrow [0, 1]$ est une quasi-copule si pour tout $0 \leq u_i \leq 1$, $i = 1, \dots, d$:*

$$Q(1, \dots, 1, u_i, 1, \dots, 1) = u_i, \quad (6.9)$$

$$Q(u_1, \dots, u_{i-1}, 0, u_{i+1}, \dots, u_d) = 0, \quad (6.10)$$

et $s \mapsto Q(u_1, \dots, u_{i-1}, s, u_{i+1}, \dots, u_d)$ est une fonction croissante, pour tout i , et :

$$|Q(u_1, \dots, u_d) - Q(v_1, \dots, v_d)| \leq |u_1 - v_1| + \dots + |u_d - v_d|.$$

Exemple 6.10 Les copules sont des quasi-copules. De plus, C^- , introduite dans l'exemple 6.8, qui n'était pas une copule est en revanche une quasi-copule.

Les quasi-copules sont intéressantes car elles apparaissent naturelles dès que l'on cherche des bornes à des ensembles de copules.

Proposition 6.4 Soit \mathcal{C} un ensemble non vide de copules (éventuellement fermé), et notons \mathcal{C}^- et \mathcal{C}^+ les bornes de \mathcal{C} , au sens où

$$\mathcal{C}^-(\mathbf{u}) = \inf\{C(\mathbf{u}), C \in \mathcal{C}\} \text{ et } \mathcal{C}^+(\mathbf{u}) = \sup\{C(\mathbf{u}), C \in \mathcal{C}\}.$$

Alors \mathcal{C}^- et \mathcal{C}^+ sont les meilleures bornes possibles pour \mathcal{C} , mais en général, \mathcal{C}^- et \mathcal{C}^+ ne sont pas des copules. En revanche, \mathcal{C}^- et \mathcal{C}^+ sont des quasi-copules.

Remarque 6.5 Cette façon d'écrire les quasi-copules (comme des bornes inférieures d'ensembles de copules) rappelle la construction des capacités comme des bornes inférieures d'ensemble de probabilités, évoquée dans les Chapitres 2 et 3.

Définition 6.13 Une fonction $S : [0, 1]^d \rightarrow [0, 1]$ est une semi-copule si pour tout $0 \leq u_i \leq 1$, $i = 1, \dots, d$,

$$S(1, \dots, 1, u_i, 1, \dots, 1) = u_i, \quad (6.11)$$

$$S(u_1, \dots, u_{i-1}, 0, u_{i+1}, \dots, u_d) = 0, \quad (6.12)$$

et $s \mapsto S(u_1, \dots, u_{i-1}, s, u_{i+1}, \dots, u_d)$ sur $[0, 1]$ est croissante.

Exemple 6.11 Soit \mathcal{H} l'ensemble de fonctions de distortion (introduites dans le Chapitre 3) c'est à dire l'ensemble des fonctions continues, strictement croissantes $[0, 1] \rightarrow [0, 1]$ telles que $h(0) = 0$ et $h(1) = 1$. On notera que $h \in \mathcal{H}$ si et seulement si $h^{-1} \in \mathcal{H}$, autrement dit (\mathcal{H}, \circ) est un groupe, où \circ est l'opérateur de composition. L'élément neutre est la fonction identité sur $[0, 1]$.

Pour tout $h \in \mathcal{H}$ et $C \in \mathcal{C}$ (l'ensemble des copules), posons

$$\Psi_h(C)(u_1, u_2) = h^{-1}(C(h(u_1), h(u_2))), 0 \leq u_1, u_2 \leq 1.$$

On parlera de h -copule. L'ensembles des \mathcal{H} -copules sont les fonctions $\Psi_h(C)$ pour une fonction de distorsion h et une copule C . Notons de plus que pour $h, h' \in \mathcal{H}$,

$$\Psi_{h \circ h'}(C)(u_1, u_2) = (\Psi_h \circ \Psi_{h'})(C)(u_1, u_2), 0 \leq u_1, u_2 \leq 1.$$

Les \mathcal{H} -copules échangeables sont des semi-copules. De plus, $\Psi_h(C^\perp)$ est une quasi-copule si et seulement si h est une fonction de distorsion telle que $-\log h$ soit convexe.

Pour décrire les phénomènes de vieillissement, il est naturel de s'intéresser aux semi-copules suivantes

$$\mathcal{A} = \{\Psi_h(C^\perp), h \in \mathcal{H}\}.$$

où les éléments de \mathcal{A} s'écrivent

$$C(u, v) = h^{-1}(h(u)h(v)) = \phi^{-1}(\phi(u) + \phi(v)),$$

où $h \in \mathcal{H}$ est le générateur multiplicatif, et $\phi = -\log h$ le générateur additif. Les éléments de \mathcal{A} sont appelées des t -normes Archimédiennes dans (85), et correspondent à des semi-copules Archimédiennes. Les copules Archimédiennes seront présentées en détails dans la section 6.4.3.

Remarque 6.6 Comme l'a montré (7), la famille \mathcal{S} est stable par distortion, au sens où $\Psi_h(C) \in \mathcal{S}$, pour tout $C \in \mathcal{S}$ et $h \in \mathcal{H}$. Mais cette propriété n'est pas vérifiée pour l'ensemble des copules \mathcal{C} .

Définition 6.14 Une capacité sur un espace (Ω, \mathcal{A}) est une fonction d'ensemble $\nu : \mathcal{A} \rightarrow [0, 1]$ telle que $\nu(\emptyset) = 0$, $\nu(\Omega) = 1$, et si $A \subset B$, alors $\nu(A) \leq \nu(B)$. De plus, la capacité ν est convexe si et seulement si pour tout A, B , $\nu(A) + \nu(B) \leq \nu(A \cup B) + \nu(A \cap B)$.

Comme noté dans le Chapitre 2, les capacités sont des notions plus faibles que les probabilités : une probabilité \mathbb{P} sur (Ω, \mathcal{A}) est une fonction d'ensemble $\mathbb{P} : \mathcal{A} \rightarrow [0, 1]$ telle que $\mathbb{P}(\emptyset) = 0$, $\mathbb{P}(\Omega) = 1$, et $\mathbb{P}(A) = \sum \mathbb{P}(A_n)$ pour toute famille dénombrable disjointe A_n telle que $A = \cup A_n$.

Remarque 6.7 Soit f une fonction de distorsion, c'est-à-dire $f : [0, 1] \rightarrow [0, 1]$, strictement convexe, croissante et dérivable, telle que $f(0) = 0$ et $f(1) = 1$, alors, pour toute mesure de probabilité \mathbb{P} , $f \circ \mathbb{P}$ est une capacité convexe.

Exemple 6.12 Comme l'a noté (7), comprendre le vieillissement joint (bivariate aging) est à l'intersection des problèmes de dépendance, de copule, et de vieillissement univarié. En particulier, trois classes de fonctions arrivent naturellement pour modéliser les comportements joints, pour des variables identiquement distribuées de loi F_1

1. la copule de survie

$$C^*(u_1, u_2) = \overline{F}\left(\overline{F}_1^{-1}(u_1), \overline{F}_1^{-1}(u_2)\right), 0 \leq u_1, u_2 \leq 1.$$

2. la fonction de vieillissement multivarié (introduite par (6)),

$$B(u_1, u_2) = \exp\left(-\overline{F}_1^{-1}\left(\overline{F}(-\log u_1, -\log u_2)\right)\right), 0 \leq u_1, u_2 \leq 1.$$

B est une semi-copule, que l'on peut relier à $h(x_1, x_2) = \overline{F}_1^{-1}(\overline{F}(x_1, x_2))$ introduite par (5) sur des problèmes similaires.

3. La semi-copule Archimédienne de générateur (additif) \overline{F}_1^{-1} ,

$$A(u_1, u_2) = \overline{F}_1 \left(\overline{F}_1^{-1}(u_1) + \overline{F}_1^{-1}(u_2) \right) = \Gamma^{-1}(\Gamma(u_1)\Gamma(u_2)), 0 \leq u_1, u_2 \leq 1,$$

où $\Gamma(x) = \exp(-\overline{F}_1^{-1}(x))$ est une fonction de distorsion $[0, 1] \rightarrow [0, 1]$.

Or comme $\Gamma \in \mathcal{H}$, alors $A = \Psi_\Gamma(C^\perp)$.

(7) a proposé plusieurs interprétations de ces fonctions sur des problèmes de vieillissement joint.

6.3.5 Simulations aléatoires et copules

Les copules sont particulièrement intéressantes pour simuler des vecteurs aléatoires, car les copules sont des lois de vecteurs dont les marges sur uniformément réparties sur $[0, 1]$. Or la simulation de lois repose sur les générateurs de nombres aléatoires (les fonctions **Random**) qui génèrent précisément des variables uniformément réparties sur $[0, 1]$. Ou plus précisément, on impose deux conditions aux générateurs de nombres aléatoires,

- qu'ils génèrent des nombres uniformément distribués sur $[0, 1]$,
- que des appels consécutifs de fonctions **Random** génèrent des variables indépendantes,

ou qui génère des échantillons qui sont, statistiquement, uniformes et indépendants.

Autrement dit, pour générer $\mathbf{U}^\perp = (U_1^\perp, \dots, U_d^\perp)$ avec des composantes indépendantes, il suffit de faire d appels de fonction **Random**. Or nous cherchons à simuler ici un vecteur $\mathbf{U} = (U_1, \dots, U_d)$ dont la loi soit une copule C , supposée dérivable. L'idée naturelle est d'utiliser la décomposition à l'aide des lois conditionnelles itérées

$$\begin{aligned} \mathbb{P}(U_1 \leq u_1, \dots, U_d \leq u_d) &= \mathbb{P}(U_d \leq u_d | U_1 \leq u_1, \dots, U_{d-1} \leq u_{d-1}) \\ &\quad \times \mathbb{P}(U_{d-1} \leq u_{d-1} | U_1 \leq u_1, \dots, U_{d-2} \leq u_{d-2}) \\ &\quad \times \dots \\ &\quad \times \mathbb{P}(U_3 \leq u_3 | U_1 \leq u_1, U_2 \leq u_2) \\ &\quad \times \mathbb{P}(U_2 \leq u_2 | U_1 \leq u_1) \times \mathbb{P}(U_1 \leq u_1). \end{aligned}$$

En commençant par la fin, $\mathbb{P}(U_1 \leq u_1) = u_1$ puisque U_1 suit une loi uniforme. Ensuite, notons que

$$\begin{aligned}
& \mathbb{P}(U_2 \leq u_2 | U_1 = u_1) \\
&= \mathbb{P}(U_2 \leq u_2, U_3 \leq 1, \dots, U_d \leq 1 | U_1 = u_1) \\
&= \lim_{h \rightarrow 0} \mathbb{P}(U_2 \leq u_2, U_3 \leq 1, \dots, U_d \leq 1 | U_1 \in [u_1, u_1 + h]) \\
&= \lim_{h \rightarrow 0} \frac{\mathbb{P}(u_1 \leq U_1 \leq u_1 + h, U_2 \leq u_2, U_3 \leq 1, \dots, U_d \leq 1)}{\mathbb{P}(U_1 \in [u_1, u_1 + h])} \\
&= \lim_{h \rightarrow 0} \frac{C(u_1 + h, u_2, 1, \dots, 1) - C(u_1, u_2, 1, \dots, 1)}{h} \\
&= \left. \frac{\partial C(x_1, u_2, 1, \dots, 1)}{\partial x_1} \right|_{x_1 = u_1},
\end{aligned}$$

que l'on pourra noter de manière plus synthétique

$$\mathbb{P}(U_2 \leq u_2 | U_1 = u_1) = \partial_1 C(u_1, u_2, 1, \dots, 1).$$

Et de manière plus générale, nous aurions

$$\begin{aligned}
& \mathbb{P}(U_k \leq u_k | U_1 = u_1, \dots, U_{k-1} = u_{k-1}) \\
&= \left. \frac{\partial^{k-1} C(x_1, \dots, x_{k-1}, u_k, 1, \dots, 1)}{\partial x_1 \dots \partial x_{k-1}} \right|_{x_1 = u_1, \dots, x_{k-1} = u_{k-1}},
\end{aligned}$$

que l'on notera aussi

$$\partial_{1, \dots, k-1} C(u_1, \dots, u_{k-1}, u_k, 1, \dots, 1).$$

Nous avons alors un algorithme simple pour générer $\mathbf{U} = (U_1, \dots, U_n)$ suivant C à partir d'un vecteur $\mathbf{U}^\perp = (U_1^\perp, \dots, U_d^\perp)$ (généré à partir d'appels successifs de fonction **Random**),

- pour $i = 1, \dots, d$, posons $U_i^\perp \leftarrow \mathbf{Random}$.
- générons U_1 uniformément sur $[0, 1]$,

$$u_1 \leftarrow U_1^\perp$$

- générons U_2 suivant la loi conditionnelle $\partial_1 C(u_1, \cdot, 1, \dots, 1)$,

$$u_2 \leftarrow [\partial_1 C(u_1, \cdot, 1, \dots, 1)]^{-1}(U_2^\perp),$$

- générons U_k suivant la loi conditionnelle $\partial_{1, \dots, k-1} C(u_1, \dots, u_{k-1}, \cdot, 1, \dots, 1)$,

$$u_k \leftarrow [\partial_{1, \dots, k-1} C(u_1, \dots, u_{k-1}, \cdot, 1, \dots, 1)]^{-1}(U_k^\perp),$$

... etc.

Exemple 6.13 On retrouve ici l'idée de l'utilisation de la décomposition de Cholesky : on génère un vecteur gaussien centré réduit à composantes indépendantes, \mathbf{X}_0^\perp , puis on pose $\mathbf{X} = \boldsymbol{\mu} + \mathbf{A}\mathbf{X}_0^\perp$, où \mathbf{A} est une matrice triangulaire

triangulaire inférieur telle que $\mathbf{A}\mathbf{A}' = \mathbf{\Sigma}$, qui permet de générer un vecteur $\mathcal{N}(\boldsymbol{\mu}, \mathbf{\Sigma})$.

Aussi, en dimension 2, pour générer un vecteur de corrélation $r \in (-1, 1)$, on pose

- pour la première composante $x_1 \leftarrow X_1^\perp$,
- pour la seconde $x_2 \leftarrow rx_1 + \sqrt{1-r^2}X_2^\perp$.

Un exemple de simulation de copule Gaussienne peut se visualiser sur la Figure 6.5 à gauche..

Exemple 6.14 Considérons la copule de Clayton en dimension $d = 2$ (cette copule sera présentée plus en détails dans la section 6.4.6),

$$C(u_1, u_2) = (u_1^{-\alpha} + u_2^{-\alpha} - 1)^{-1/\alpha}, \text{ où } \alpha \geq 0.$$

Alors (U_1, U_2) a pour distribution C si et seulement si U_1 suit une loi uniforme sur $[0, 1]$ et que $U_2|U_1 = u_1$ a pour loi conditionnelle

$$\mathbb{P}(U_2 \leq u_2 | U_1 = u_1) = \partial_2 C(u_1, u_2) = (1 + u_1^\alpha [u_2^{-\alpha} - 1])^{-1-1/\alpha}.$$

L'algorithme pour générer une copule de Clayton est alors de la forme suivante,

- générons U_1 uniformément sur $[0, 1]$,

$$u_1 \leftarrow U_1^\perp,$$

- générons U_2 suivant la loi conditionnelle $\partial_2 C(u_1, u_2)$,

$$u_2 \leftarrow [\partial_1 C(u_1, \cdot)]^{-1}(U_2^\perp) = \left([(U_2^\perp)^{-\alpha/(1+\alpha)} - 1]u_1^{-\alpha} + 1 \right)^{-1/\alpha}.$$

Un exemple de simulations de copule de Clayton peut se visualiser sur la Figure 6.5 à droite.

Exemple 6.15 En fait, plus généralement, il est possible de trouver un algorithme simple pour générer une copule Archimédienne (qui seront décrite dans la section 6.4.3) en dimension 2. Pour cela,

- générons U_1 uniformément sur $[0, 1]$,

$$u_1 \leftarrow U_1^\perp,$$

- générons U_2 suivant la loi conditionnelle $\partial_2 C(u_1, \cdot)$,

$$u_2 \leftarrow \phi^{-1}(\phi(u) - \phi(u_1)),$$

$$\text{où } u = \phi'^{-1}(\phi'(u_1)/U_2^\perp).$$

(47) propose d'utiliser le fait que $C(U_1, U_2)$ et $\phi(U_1)/[\phi(U_1) + \phi(U_2)]$ sont deux variables indépendantes, la première de loi K , où $K(t) = t - \phi(t)/\phi'(t)$, et la seconde est uniforme sur $[0, 1]$. Aussi,

- générons U de loi K , $u \leftarrow K^{-1}(U_1^\perp)$,
- posons $u_1 \leftarrow \phi^{-1}(U_2^\perp \phi(u))$ et $u_2 \leftarrow \phi^{-1}([1 - U_2^\perp]\phi(u))$.

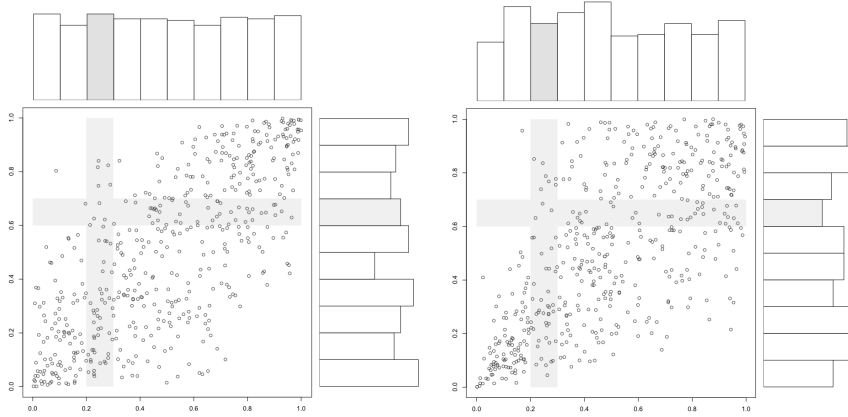


Figure 6.5 – Simulation de vecteurs (U_1, U_2) suivant une copule Gaussienne à gauche, et une copule de Clayton à droite. En haut et sur le côté sont représentés les histogrammes des marges.

6.4 Les familles usuelles de copules

Avant de présenter les copules paramétriques les plus usuelles, on appellera C^\perp la copule indépendante,

Définition 6.15 On appellera copule indépendante C^\perp la copule définie par

$$C^\perp(u_1, \dots, u_d) = u_1 \cdots u_d = \prod_{i=1}^d u_i.$$

Remarque 6.8 Soit $\mathbf{X} \in \mathcal{F}(F_1, \dots, F_d)$. On notera $\mathbf{X}^\perp \in \mathcal{F}(F_1, \dots, F_d)$ un vecteur dont la copule est C^\perp . On dira qu'il s'agit d'une version indépendante de \mathbf{X} .

6.4.1 Les bornes de Fréchet-Hoeffding et la comonotonie

La famille des copules est bornée : pour toute copule C ,

$$C^-(u_1, \dots, u_d) \leq C(u_1, \dots, u_d) \leq C^+(u_1, \dots, u_d),$$

pour tout $(u_1, \dots, u_d) \in [0, 1] \times \dots \times [0, 1]$, où

$$\begin{cases} C^-(u_1, \dots, u_d) = \max\{0, u_1 + \dots + u_d - (d-1)\} \\ C^+(u_1, \dots, u_d) = \min\{u_1, \dots, u_d\}. \end{cases}$$

Les bornes sont appelées « bornes de Fréchet-Hoeffding ». On notera que si C^+ est une copule, en revanche C^- est une copule uniquement dans le cas $d = 2$.

Définition 6.16 La copule comonotone C^+ , ou copule de dépendance (positive) maximale, est définie par $C^+(u_1, \dots, u_d) = \min\{u_1, \dots, u_d\}$.

Définition 6.17 La borne inférieure de Fréchet C^- la fonction définie par $C^-(u_1, \dots, u_d) = \max\{0, u_1 + \dots + u_d - (d - 1)\}$. En dimension $d = 2$, C^- sera appelée copule anticomonotone.

Les graphiques associées à ces copules sont représentées sur la Figure 6.6.

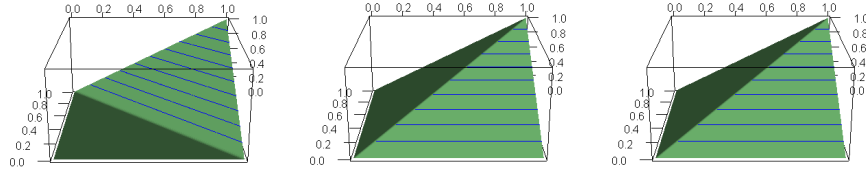


Figure 6.6 – Copule anticomonotone C^- , copule indépendante C^\perp , et copule comonotone C^+ , en dimension $d = 2$.

- Proposition 6.5**
1. Si $d = 2$, C^- est la fonction de répartition du couple $(U, 1 - U)$ où U est uniformément distribuée sur $[0, 1]$.
 2. (X_1, X_2) a pour copule C^- si et seulement s'il existe ϕ strictement croissante et ψ strictement décroissante telles que $(X_1, X_2) \stackrel{\mathcal{L}}{=} (\phi(Z), \psi(Z))$ où Z est une variable aléatoire.
 3. C^+ est la fonction de répartition du vecteur (U, \dots, U) où U est uniformément distribuée sur $[0, 1]$.
 4. (X_1, \dots, X_n) a pour copule C^+ si et seulement s'il existe des fonctions ϕ_i strictement croissantes telles que $(X_1, \dots, X_n) \stackrel{\mathcal{L}}{=} (\phi_1(Z), \dots, \phi_n(Z))$ où Z est une variable aléatoire.

Une preuve de ce résultat est donnée dans (73). La borne supérieure de Fréchet-Hoeffding est correspond au cas de variables « comonotones », mentionnées dans le Chapitre 3.

Exemple 6.16 Ces bornes sur les copules peuvent en fait fournir des bornes sur certaines quantités. En particulier, si $d = 2$, et que $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}$ est 2-croissante, alors pour tout $(X_1, X_2) \in \mathcal{F}(F_1, F_2)$

$$\mathbb{E}(\phi(F_1^{-1}(U), F_2^{-1}(1 - U))) \leq \mathbb{E}(\phi(X_1, X_2)) \leq \mathbb{E}(\phi(F_1^{-1}(U), F_2^{-1}(U))),$$

où U est uniformément répartie sur $[0, 1]$ (dans un cadre plus général, ce résultat a été établi par (91), et sera développé dans le Section 6.6).

Exemple 6.17 Comme évoqué auparavant, il convient de faire attention entre la dépendance en dimension d et la dépendance par paires. En particulier, certains résultats, qui pourrait sembler intuitifs (car ils sont vrais dans le cas Gaussien) sont en fait faux,

- si X_1 et X_2 sont comonotones, ainsi que X_2 et X_3 , X_1 et X_3 peuvent ne pas être comonotone. Par exemple, si

$$\begin{aligned} (X_1, X_2, X_3) &= (1, 1, 1) \text{ avec probabilité } 1/4, \\ &= (1, 2, 3) \text{ avec probabilité } 1/4, \\ &= (3, 2, 1) \text{ avec probabilité } 1/4, \\ &= (3, 3, 3) \text{ avec probabilité } 1/4, \end{aligned}$$

alors X_1 et X_3 sont indépendants.

- si X_1 et X_2 sont comonotones, et que X_2 et X_3 sont indépendants, X_1 et X_3 peuvent ne pas être indépendants. Aussi, par exemple, avec

$$\begin{aligned} (X_1, X_2, X_3) &= (1, 1, 3) \text{ avec probabilité } 1/4, \\ &= (2, 1, 1) \text{ avec probabilité } 1/4, \\ &= (2, 3, 3) \text{ avec probabilité } 1/4, \\ &= (3, 3, 1) \text{ avec probabilité } 1/4, \end{aligned}$$

X_1 et X_2 sont comonotones, X_2 et X_3 sont indépendants, et X_1 et X_3 sont indépendants.

Remarque 6.9 Soit $\mathbf{X} \in \mathcal{F}(F_1, \dots, F_d)$. On notera $\mathbf{X}^+ \in \mathcal{F}(F_1, \dots, F_d)$ un vecteur dont la copule est C^+ . On dira qu'il s'agit d'une version comonotone de \mathbf{X} . De manière similaire, en dimension $d = 2$, on notera $\mathbf{X}^- \in \mathcal{F}(F_1, F_2)$ une version anticomonotone de \mathbf{X} .

6.4.2 Les copules elliptiques

Nous avons vu dans la Section 6.2.3 qu'un vecteur \mathbf{X} avait une distribution elliptique, de moyenne $\boldsymbol{\mu} \in \mathbb{R}^d$, de matrice de variance $\boldsymbol{\Sigma}$ et de générateur $g : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ si la fonction caractéristique du \mathbf{X} était

$$\mathbb{E}(e^{it'\mathbf{X}}) = e^{it'\boldsymbol{\mu}} g(t'\boldsymbol{\Sigma}t).$$

On peut noter que $\mathbf{X} = \boldsymbol{\mu} + R \cdot \mathbf{A} \cdot \mathbf{U}$, où $\mathbf{A}\mathbf{A}'$ est la décomposition de Cholesky de la matrice $\boldsymbol{\Sigma}$, où \mathbf{U} est une variable aléatoire uniformément distribuée sur la sphère unité de \mathbb{R}^d , et où R est une variable positive, indépendante de \mathbf{U} , (dont la loi peut être reliée à g , comme le rappelle (37)). On notera $\mathcal{E}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, g)$ la classe des lois elliptiques.

Deux cas particuliers importants sont la copule Gaussienne, et la copule de Student à $\nu > 2$ degrés de liberté (définie ici en dimension 2, mais qui se généralise en dimension quelconque puisque les lois elliptiques existent en dimension $d \geq 1$).

Exemple 6.18 Soit $r \in (-1, +1)$, alors la copule Gaussienne de paramètre r (en dimension 2) est

$$C(u_1, u_2) = \frac{1}{2\pi\sqrt{1-r^2}} \int_{-\infty}^{\Phi^{-1}(u_1)} \int_{-\infty}^{\Phi^{-1}(u_2)} \exp\left(\frac{x^2 - 2rxy + y^2}{2(1-r^2)}\right) dx dy$$

où Φ est la fonction de répartition de la loi normale centrée réduite

$$\Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) dz.$$

Exemple 6.19 Soit $r \in (-1, +1)$, et $\nu \geq 2$, alors la copule de Student de paramètres r et ν est

$$\int_{-\infty}^{T_\nu^{-1}(u_1)} \int_{-\infty}^{T_\nu^{-1}(u_2)} \frac{1}{\pi\nu\sqrt{1-r^2}} \frac{\Gamma(\frac{\nu}{2} + 1)}{\Gamma(\frac{\nu}{2})} \left(1 + \frac{x^2 - 2rxy + y^2}{\nu(1-r^2)}\right)^{-\frac{\nu}{2}+1} dx dy.$$

où T_ν est la fonction de répartition de la loi de Student (univariée) à ν degrés de liberté, c'est-à-dire

$$T_\nu(x) = \int_{-\infty}^x \frac{\Gamma(\frac{\nu+1}{2})}{\sqrt{\nu\pi} \Gamma(\frac{\nu}{2})} \left(1 + \frac{z^2}{\nu}\right)^{-\frac{\nu+1}{2}} dz$$

Comme pour l'exemple précédant, ces copules ont été présentées ici en dimension $d = 2$ pour des questions de simplicité, mais elle se généralise aisément en dimension supérieure, avec comme paramètre une matrice de corrélation \mathbf{R} , et un degré de liberté ν .

6.4.3 Les copules Archimédiennes

Les copules Archimédiennes strictes en dimension 2

Définition 6.18 Soit ϕ une fonction décroissante convexe sur $(0, 1] \rightarrow [0, \infty]$ telle que $\phi(1) = 0$ et $\phi(0) = \infty$. On appellera copule Archimédienne stricte de générateur ϕ la copule définie par

$$C(u_1, u_2) = \phi^{-1}(\phi(u_1) + \phi(u_2)), u_1, u_2 \in [0, 1].$$

Exemple 6.20 Soit $\phi(t) = t^{-\alpha} - 1$ pour $\alpha \geq 0$; la copule associée est la copule de Clayton (introduite dans (20), mais utilisée par de nombreux auteurs sous d'autres noms, et décrite plus en détails dans la section 6.4.6).

Notons que le générateur n'est pas unique : en multipliant le générateur par une constante (positive) on obtient la même copule. De plus, les copules Archimédiennes sont des copules symétriques, au sens où $C(u_1, u_2) = C(u_2, u_1)$.

Exemple 6.21 Il est aussi possible de chercher des copules Archimédiennes vérifiant une propriété de symétrie radiale, c'est-à-dire $C(u_1, u_2) = C^*(u_1, u_2)$. Ceci n'est possible que si $\phi(t) = \log \frac{e^{-\alpha t} - 1}{e^{-\alpha} - 1}$. Il s'agit de la copule dite de Frank (introduite dans (40)).

Remarque 6.10 Dans certains articles (comme (69) ou (34)) une représentation multiplicative des copules Archimédiennes est préférée

$$C(u_1, u_2) = h^{-1}[h(u_1) \cdot h(u_2)].$$

On retrouve la forme précédente en posant $h(t) = \exp[\phi(t)]$, ou réciproquement $\phi(t) = h(\log(t))$.

Remarque 6.11 Une autre caractérisation des copules Archimédienne peut se faire à l'aide de la fonction de Kendall :

$$K(t) = \mathbb{P}(C(U_1, U_2) \leq t) = t - \lambda(t) \text{ où } \lambda(t) = \frac{\phi(t)}{\phi'(t)}$$

et où (U_1, U_2) a pour loi C . De façon réciproque,

$$\phi(t) = \exp \left(\int_{t_0}^t \frac{ds}{\lambda(s)} \right),$$

où $t_0 \in (0, 1)$ est une constante arbitraire (nous avons noté que les générateurs étaient définis à une constante multiplicative près). Cette fonction est particulièrement importante en inférence statistique, comme nous le verrons dans la section 6.7.

Il est également possible de définir une classe un peu plus large en autorisant $\phi(0) \leq \infty$.

Définition 6.19 Soit ϕ une fonction décroissante convexe sur $(0, 1] \rightarrow [0, \infty]$ telle que $\phi(1) = 0$. Définissons l'inverse de ϕ par

$$\phi^{-1}(t) = \begin{cases} \phi^{-1}(t), & \text{pour } 0 \leq x \leq \phi(0) \\ 0, & \text{pour } \phi(0) < t < \infty. \end{cases}$$

On appellera copule Archimédienne de générateur ϕ la copule définie par

$$C(u_1, u_2) = \phi^{-1}(\phi(u_1) + \phi(u_2)), \text{ pour } u_1, u_2 \in [0, 1].$$

Les copules Archimédiennes non strictes présente un ensemble nul

$$\{(u_1, u_2), C(u_1, u_2) = 0\}$$

non vide, pour lequel

$$\mathbb{P}((U_1, U_2) \in \{(u_1, u_2), C(u_1, u_2) = 0\}) = 0.$$

Cet ensemble est majoré par une courbe nulle, $\{(u_1, u_2), \phi(u_1) + \phi(u_2) = 0\}$, de masse

$$\mathbb{P}((U_1, U_2) \in \{(u_1, u_2), \phi(u_1) + \phi(u_2) = 0\}) = -\frac{\phi(0)}{\phi'(0^+)},$$

qui est non nulle dès lors que $-\phi'(0^+)$ est finie.

Exemple 6.22 Soit $\phi(t) = t^\alpha - 1$, où $\alpha \in [-1, \infty)$, avec le cas limite $\phi(t) = -\log(t)$ si $\alpha = 0$; la copule associée est la copule de Clayton. Le générateur est strict si $\alpha > 0$ (comme noté dans la section précédente).

Le tableau 6.1 présente quelques familles archimédiennes classiques, avec la notation utilisée dans (73).

Les copules Archimédiennes en dimension $d > 2$

Les copules Archimédiennes sont très utilisées en analyse fonctionnelle (comme en témoignent (85) ou (2)) car elles sont associatives, au sens où, si C est une copule Archimédienne, alors

$$C(C(u_1, u_2), u_3) = C(u_1, C(u_2, u_3)), \text{ pour tout } 0 \leq u_1, u_2, u_3 \leq 1.$$

Cette propriété devrait nous permettre de construire aisément des copules en dimension quelconque (moyennant quelques hypothèses supplémentaires).

Définition 6.20 Soit ϕ une fonction décroissante convexe, $(0, 1] \rightarrow [0, \infty]$ telle que $\phi(1) = 0$. Définissons le quasi-inverse si $\phi(0) < \infty$ par

$$\phi^{-1}(t) = \begin{cases} \phi^{-1}(t) & , \text{ pour } 0 \leq t \leq \phi(0) \\ 0 & , \text{ pour } \phi(0) < t < \infty. \end{cases}$$

Si $d > 2$, supposons que ϕ^{-1} est d -complètement monotone (où, pour rappels, ψ est d -complètement monotone si elle est continue, et que ses dérivées sont monotones, de signe alterné, c'est-à-dire pour tout $k = 0, 1, \dots, d$, $(-1)^k d^k \psi(t)/dt^k \geq 0$). Une copule Archimédienne est alors définie par

$$C(u_1, \dots, u_n) = \phi^{-1}(\phi(u_1) + \dots + \phi(u_n)), \text{ pour tout } u_1, \dots, u_n \in [0, 1].$$

Table 6.1 – Générateurs des principales copules Archimédiennes.

| | $\psi(t)$ | ensemble θ | nom |
|------|---|---------------------------------|-----------------|
| (1) | $\frac{1}{\theta}(t^{-\theta} - 1)$ | $[-1, 0) \cup (0, \infty)$ | Clayton |
| (2) | $(1 - t)^\theta$ | $[1, \infty)$ | |
| (3) | $\log \frac{1 - \theta(1 - t)}{t}$ | $[-1, 1)$ | Ali-Mikhail-Haq |
| (4) | $(-\log t)^\theta$ | $[1, \infty)$ | Gumbel |
| (5) | $-\log \frac{e^{-\theta t} - 1}{e^{-\theta} - 1}$ | $(-\infty, 0) \cup (0, \infty)$ | Frank |
| (6) | $-\log\{1 - (1 - t)^\theta\}$ | $[1, \infty)$ | Joe |
| (7) | $-\log\{\theta t + (1 - \theta)\}$ | $(0, 1]$ | |
| (8) | $\frac{1 - t}{1 + (\theta - 1)t}$ | $[1, \infty)$ | |
| (9) | $\log(1 - \theta \log t)$ | $(0, 1]$ | Gumbel-Barnett |
| (10) | $\log(2t^{-\theta} - 1)$ | $(0, 1]$ | |
| (11) | $\log(2 - t^\theta)$ | $(0, 1/2]$ | |
| (12) | $(\frac{1}{t} - 1)^\theta$ | $[1, \infty)$ | |
| (13) | $(1 - \log t)^\theta - 1$ | $(0, \infty)$ | |
| (14) | $(t^{-1/\theta} - 1)^\theta$ | $[1, \infty)$ | |
| (15) | $(1 - t^{1/\theta})^\theta$ | $[1, \infty)$ | Genest - Ghoudi |
| (16) | $(\frac{\theta}{t} + 1)(1 - t)$ | $[0, \infty)$ | |

Ces copules sont obtenues par itérations successives, en posant

$$C_2(u_1, u_2) = \phi^{-1}(\phi(u_1) + \phi(u_2))$$

et ensuite, pour $n \geq 2$,

$$C_{n+1}(u_1, \dots, u_{n+1}) = C_2(C_n(u_1, \dots, u_n), u_{n+1}).$$

Exemple 6.23 Soit ψ la transformée de Laplace d'une variable positive Θ , alors d'après le théorème de Bernstein, ψ est complètement montone, et $\psi(0) = 1$. Alors $\phi = \psi^{-1}$ est complètement montone, et permet d'engendrer une copule Archimédienne de dimension d pour tout $d \geq 2$. Par exemple si Θ suit une loi Gamma d'espérance a et de variance a , alors $\psi(t) = (1+t)^{-\alpha}$, et $\phi(t) = t^{-1/\alpha} - 1$. On retrouve la copule de Clayton (cette construction par variable latente dans les modèles de vieillissement pour définir la copule de Clayton sera détaillée dans la section 6.4.6). Cette notation est celle retenue dans (55) pour définir les copules Archimédiennes,

$$C(u_1, \dots, u_d) = \psi(\psi^{-1}(u_1) + \dots + \psi^{-1}(u_d)).$$

Exemple 6.24 Soit $\mathbf{X} = (X_1, \dots, X_d)$ un vecteur de durées de vies résiduelles, dont la loi de survie jointe est supposée Schur-constante (c'est-à-dire à la fois Schur-concave et Schur-convexe). Alors il existe $S : \mathbb{R}_+ \rightarrow [0, 1]$ telle que

$$\mathbb{P}(X_1 > x_1, \dots, X_d > x_d) = S(x_1 + \dots + x_d).$$

Les lois marginales X_i sont alors Schur-constantes (c'est-à-dire, puisque l'on est en dimension 1, que les lois marginales suivent des lois exponentielles), et la copule de survie de \mathbf{X} est une copule Archimédienne de générateur S^{-1} . Notons également que les durées de vies conditionnelles sont identiques, au sens où

$$\mathbb{P}(X_i - x_i > t | \mathbf{X} > \mathbf{x}) = \mathbb{P}(X_j - x_j > t | \mathbf{X} > \mathbf{x}),$$

pour tout $t > 0$ et $\mathbf{x} \in \mathbb{R}_+^d$. En particulier, si S est la fonction de survie d'une loi de Pareto, on obtient la copule de Clayton, et si S est la fonction de survie d'une distribution de Weibull, on obtient la copule de Gumbel, comme l'a montré (72).

Exemple 6.25 Il existe une représentation des copules Archimédiennes qui évoquera les modèles sphériques, obtenue par (67) à partir des travaux de (?) et (?). En effet, considérons une variable aléatoire positive R , indépendante de \mathbf{U} , uniformément distribué sur le simplexe de \mathbb{R}^d . Alors la copule de survie de $\mathbf{X} = R \cdot \mathbf{U}$ est Archimédienne, et son générateur est l'inverse de la d transformée de Williamson, défini par

$$\phi^{-1}(t) = \int_x^\infty \left(1 - \frac{x}{t}\right)^{d-1} dF_R(t),$$

où F_R est la fonction de répartition de R . On pourra noter que

$$R \stackrel{\mathcal{L}}{=} \phi(U_1) + \dots + \phi(U_d).$$

Remarque 6.3 En dimension d , la copule indépendante vérifie

$$C^\perp(u_1, \dots, u_d) = u_1 \cdots u_d = \prod_{i=1}^d u_i = \exp \left(\sum_{i=1}^d \log(u_i) \right)$$

soit

$$\log C^\perp(u_1, \dots, u_d) = \sum_{i=1}^d \log(u_i)$$

L'idée des copules Archimédienne était de considérer une fonction ϕ^{-1} d -complètement monotone sur $[0, \infty]$, et de faire une distorsion, en remplaçant u_i par $h(x_i) = \exp[-\phi(x_i)]$, où $x_i \in [0, 1]$ (comme introduit par (57) ou (14)). Alors

$$\log C^\perp(\exp \phi(x_1), \dots, \exp \phi(x_d)) = \sum_{i=1}^d \phi(x_i)$$

soit, $h(C(x_1, \dots, x_d)) = C^\perp(h(x_1), \dots, h(x_d))$, ou encore

$$C(x_1, \dots, x_d) = h^{-1}[C^\perp(h(x_1), \dots, h(x_d))], \text{ où } x_i \in [0, 1]$$

Les copules Archimédiennes sont alors une distorsion de la copule indépendante. Il pourrait paraître légitime de tenter des transformation sur une copule autre que C^\perp .

Remarque 6.12 Un des résultats les plus anciens sur les copules Archimédienne est probablement le théorème de Ling : en dimension 2, les copules Archimédiennes sont les seules copules telles que $C(u, u) < u$ pour tout $u \in (0, 1)$, et l'équation fonctionnelle d'associativité (présentée en détails dans (1))

$$C(C(u_1, u_2), u_3) = C(u_1, C(u_2, u_3)) \text{ pour tout } u_1, u_2, u_3 \in [0, 1].$$

Cette construction itérative a été détaillée dans (85), pour construire des copules Archimédienne en dimension quelconque, au sens où

$$C_2(u_1, u_2) = \phi^{-1}(\phi(u_1) + \phi(u_2))$$

et

$$C_k(u_1, \dots, u_{k-1}, u_k) = C_2(C_{k-1}(u_1, \dots, u_{k-1}), u_k),$$

pour tout $k \geq 3$.

Proposition 6.6 Soit (C_n) une suite de copules absolument continues de copules Archimédiennes, de générateurs (ϕ_n) . La limite de C_n lorsque $n \rightarrow \infty$ est une copule Archimédienne si et seulement si une des conditions suivantes est satisfaite,

– il existe ϕ telle que $s, t \in [0, 1]$,

$$\lim_{n \rightarrow \infty} \frac{\phi_n(s)}{\phi'_n(t)} = \frac{\phi(s)}{\phi'(t)}. \quad (6.13)$$

– il existe λ continue telle que $\lim_{n \rightarrow \infty} \lambda_n(t) = \lambda(t)$.

– il existe K continue telle que $\lim_{n \rightarrow \infty} K_n(t) = K(t)$.

– il existe une suite de constante positive (c_n) telles que $\lim_{n \rightarrow \infty} c_n \phi_n(t) = \phi(t)$, pour tout $t \in [0, 1]$.

Ce résultat a été montré partiellement par (43) et généralisé par (17). Ces résultats de convergence sont intéressants pour faire de l'inférence sur les familles de copules Archimédiennes.

Exemple 6.26 Les copules archimédiennes sont intéressantes pour modéliser des risques échangeables. Rappelons que pour une série échangeable X_1, \dots, X_n de variables de Bernoulli (modélisant le défaut d'un émetteur), les variables sont conditionnellement indépendantes, et le vecteur (X_1, \dots, X_n) peut alors être modélisé naturellement par une copule Archimédienne (comme le fait (71)).

6.4.4 Les copules Archimédiennes généralisées et hiérarchiques

Considérons le cas où $C(u_1, \dots, u_d)$ peut s'écrire

$$\phi_1^{-1}[\phi_1[\phi_2^{-1}(\phi_2[\dots\phi_{d-1}^{-1}[\phi_{d-1}(u_1) + \phi_{d-1}(u_2)] + \dots + \phi_2(u_{d-1}))] + \phi_1(u_d)],$$

où les ϕ_i sont des générateurs de copules. C est une copule si $\phi_i \circ \phi_{i-1}^{-1}$ est l'inverse d'une transformée de Laplace. Cette copule est parfois appelée « fully nested Archimedean » (FNA). On commence par coupler deux composantes (ici U_1 et U_2), puis on couple cette paire avec U_3 , etc.

En dimension $d = 5$, cela donne

$$\begin{aligned} & \phi_1^{-1}[\phi_1(\phi_2^{-1}[\phi_2(\phi_3^{-1}[\phi_3(\phi_4^{-1}[\phi_4(u_1) + \phi_4(u_2)]) \\ & + \phi_3(u_3)]) + \phi_2(u_4)]) + \phi_1(u_5)]. \end{aligned}$$

On peut aussi envisager une construction plus hiérarchique, avec les copules dite « partially nested Archimedean » (PNA), en considérant un couplage complet pour (U_1, U_2, U_3) , un autre pour (U_4, U_5) , et en couplant les deux vecteurs,

$$\begin{aligned} & \phi_4^{-1}[\phi_4(\phi_1^{-1}[\phi_1(\phi_2^{-1}[\phi_2(u_1) + \phi_2(u_2)]) + \phi_1(u_3)]) \\ & + \phi_4(\phi_3^{-1}[\phi_3(u_4) + \phi_3(u_5)])] \end{aligned}$$

La condition pour avoir effectivement une copule est que $\phi_2 \circ \phi_1^{-1}$ soit l'inverse d'une transformée de Laplace, mais aussi $\phi_4 \circ \phi_1^{-1}$ ainsi que $\phi_4 \circ \phi_3^{-1}$. Ces deux constructions sont visualisées sur la Figure 6.7.

Enfin, on peut envisager une construction réellement hiérarchique, ne visant pas uniquement à faire du couplage (deux à deux),

$$\begin{aligned} & \phi_3^{-1}[\phi_3(\phi_1^{-1}[\phi_1(u_1) + \phi_1(u_2) + \phi_1(u_3)]) \\ & + \phi_3(\phi_2^{-1}[\phi_2(u_4) + \phi_2(u_5)])]. \end{aligned}$$

Dans ce cas, on a effectivement construit une copule si $\phi_3 \circ \phi_1^{-1}$ ainsi que $\phi_3 \circ \phi_2^{-1}$ sont des inverses de transformées de Laplace. On pourrait aussi envisager

$$\begin{aligned} & \phi_3^{-1}[\phi_3(\phi_1^{-1}[\phi_1(u_1) + \phi_1(u_2)] + \phi_3(u_3)) \\ & + \phi_3(\phi_2^{-1}[\phi_2(u_4) + \phi_2(u_5)])]. \end{aligned}$$

Exemple 6.27 Si les ϕ_i sont des générateurs de copules de Gumbel de paramètre θ_i , une condition suffisante pour que C soit une copule est que les θ_i soient croissants, et supérieurs à 1. De même, si les ϕ_i sont des générateurs de copules de Clayton de paramètre θ_i , une condition suffisante pour que C soit une copule est que les θ_i soient croissants, et supérieurs à 0.

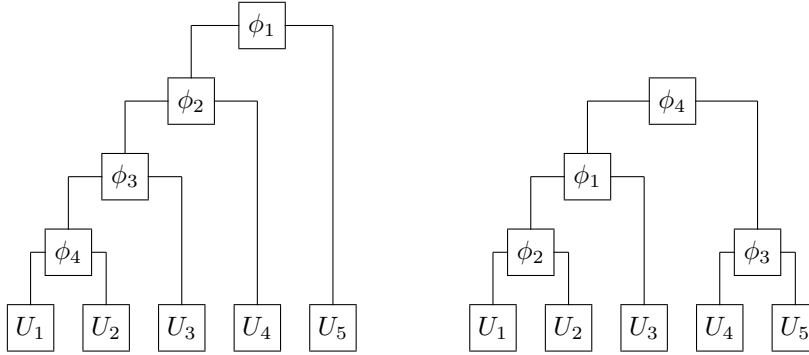


Figure 6.7 – Copules Archimédiennes « fully nested » (à gauche) et « partially nested » (à droite)

6.4.5 Les copules extrêmes

Soient \mathbf{X}^k , $k = 1, \dots, n$ un échantillon i.i.d. de vecteurs aléatoires de loi jointe F (supposée continue), de copule C et de lois marginales F_1, \dots, F_d . On note \mathbf{M}^n le vecteur des maximums, composante par composantes,

$$\mathbf{M}^n = (\max\{X_1^1, \dots, X_1^n\}, \dots, \max\{X_d^1, \dots, X_d^n\})$$

Soit C_n la copule associée à \mathbf{M}^n , alors

$$C_n(u_1, \dots, u_d) = C\left(u_1^{\frac{1}{n}}, \dots, u_d^{\frac{1}{n}}\right)^n$$

Les copules extrêmes sont les copules qui sont obtenues comme limites de telles copules.

Définition 6.21 Une copule C est une copule extrême s'il existe une copule Γ telle que

$$\Gamma\left(u_1^{\frac{1}{n}}, \dots, u_d^{\frac{1}{n}}\right)^n \rightarrow C(u_1, \dots, u_d) \text{ lorsque } n \rightarrow \infty$$

pour tout $(u_1, \dots, u_d) \in [0, 1]^d$. On dira de plus que Γ est dans le max-domaine d'attraction de C .

L'ensemble des copules extrêmes coïncide avec l'ensemble des copules max-stables, c'est-à-dire qui vérifient

$$C\left(u_1^{\frac{1}{n}}, \dots, u_d^{\frac{1}{n}}\right)^n = C(u_1, \dots, u_d)$$

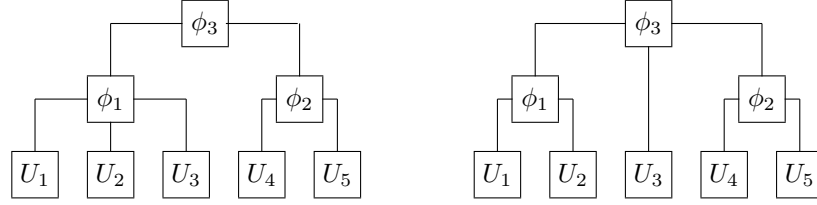


Figure 6.8 – Copules Archimédiennes hiérarchiques en dimension 5, avec deux constructions différentes.

pour tout entier $n \geq 1$, et pour tout $(u_1, \dots, u_d) \in [0, 1]^d$.

Il est possible d'énoncer le théorème de Pickands-Balkema-de Haan sous la forme suivante, afin de caractériser l'ensemble des copules extrêmes,

Théorème 6.4 *Une copule $C : [0, 1]^d \rightarrow [0, 1]$ est une copule extrême si et seulement si il existe une mesure finie H sur le simplexe de \mathbb{R}^d , \mathcal{S}_d , appelée mesure spectrale, telle que*

$$C(u_1, \dots, u_d) = \exp[-\ell(-\log u_1, \dots, -\log u_d)]$$

où la fonction de dépendance de queue ℓ est donnée par

$$\ell(x_1, \dots, x_d) = \int_{\mathcal{S}_d} \max\{\omega_1 x_1, \dots, \omega_d x_d\} dH(\omega_1, \dots, \omega_d)$$

pour tout $(x_1, \dots, x_d) \in \mathbb{R}_+^d$. On suppose que $\int_{\mathcal{S}_d} \omega_i dH(\omega_1, \dots, \omega_d) = 1$ pour $i = 1, \dots, n$.

Notons que

$$\ell(x_1, \dots, x_d) = \lim_{t \rightarrow 0} [1 - C(1 - tx_1, \dots, 1 - tx_d)]$$

Cette fonction de dépendance de queue ℓ est convexe, homogène de degré 1 (c'est-à-dire $\ell(tx_1, \dots, tx_d) = t\ell(x_1, \dots, x_d)$ pour tout $t > 0$) et vérifie

$$\max\{x_1, \dots, x_d\} \leq \ell(x_1, \dots, x_d) \leq x_1 + \dots + x_d.$$

Comme tenu de cette propriété d'homogénéité, on peut aussi écrire

$$\ell(x_1, \dots, x_d) = [x_1 + \dots + x_d] A\left(\frac{x_1}{x_1 + \dots + x_d}, \dots, \frac{x_d}{x_1 + \dots + x_d}\right)$$

où A est une fonction $\mathcal{S}_d \rightarrow [d^{-1}, 1]$ appelée fonction de dépendance de Pickands (définie dans (77) et (78)).

La copule extrême s'écrit alors

$$C(u_1, \dots, u_d) = \exp \left[\left(\sum_{i=1}^d \log u_i \right) A \left(\frac{u_1}{\sum_{i=1}^d \log u_i}, \dots, \frac{u_d}{\sum_{i=1}^d \log u_i} \right) \right]$$

Cette fonction A est convexe, et vérifie

$$\max\{\omega_1, \dots, \omega_d\} \leq A(\omega_1, \dots, \omega_d) \leq 1$$

pour tout $(\omega_1, \dots, \omega_d) \in \mathcal{S}_d$. Cette dernière propriété caractérise l'ensemble des fonctions de dépendance en dimension $d = 2$ (mais pas au delà). Dans ce cas, on peut alors écrire

$$C(u_1, u_2) = [u_1 u_2]^{A(\log u_2 / \log u_1 u_2)}$$

où, par abus de notation, $A : [0, 1] \rightarrow [1/2, 1]$ est convexe et vérifie

$$\max\{\omega, 1 - \omega\} \leq A(\omega) \leq 1.$$

Exemple 6.28 Si C est une copule Archimédienne de générateur ϕ , tel qu'il existe une limite

$$\lim_{s \rightarrow 0} \frac{s\phi'(1-s)}{\phi(1-s)} = -\theta \in (-\infty, -1]$$

alors C est dans le domaine d'attraction de la copule dont la fonction de dépendance est

$$\ell(x_1, \dots, x_d) = [x_1^\theta + \dots + x_d^\theta]^{\frac{1}{\theta}}.$$

La copule limite est alors

$$C(u_1, \dots, u_d) = \exp \left[- \left([-\log u_1]^\theta + \dots + [-\log u_d]^\theta \right)^{\frac{1}{\theta}} \right]$$

qui correspond à la copule de Gumbel.

Exemple 6.29 En dimension 2, la copule de Student t de corrélation r et à ν degrés de liberté s'écrit

$$\int_{-\infty}^{T_\nu^{-1}(u_1)} \int_{-\infty}^{T_\nu^{-1}(u_2)} \frac{1}{\pi\nu\sqrt{1-r^2}} \frac{\Gamma(\frac{\nu}{2}+1)}{\Gamma(\frac{\nu}{2})} \left(1 + \frac{x^2 - 2rxy + y^2}{\nu(1-r^2)} \right)^{-\frac{\nu}{2}+1} dx dy.$$

Cette copule est dans le domaine d'attraction de la copule extrême dont la fonction de dépendance de Pickands est définie par

$$A(\omega) = \omega T_{\nu+1}(z_\omega) + (1-\omega) T_{\nu+1}(z_{1-\omega})$$

où

$$z_\omega = \sqrt{\frac{1+\nu}{1-r^2}} \left(\left[\frac{\omega}{1-\omega} \right]^{\frac{1}{\nu}} - r \right)$$

pour tout $\omega \in [0, 1]$.

Dans la section 6.5.4, nous présenterons des mesures de dépendance permettant de décrire la dépendance dans les queues d'une distribution jointe. Nous n'irons toutefois pas plus loin dans cette section, les modèles d'extrêmes multivariés seront détaillés davantage dans le chapitre 7.

6.4.6 La copule de Clayton

Dans cette section, nous insisterons davantage sur la construction stochastique multivariée associée à la copule dite de Clayton. Commençons par le cas $d = 2$ afin de simplifier la présentation.

Utilisation des odds ratio

En reprenant l'idée initiale de (20) développée par la suite dans (74), définissons la fonction suivante

$$\Theta(x_1, x_2) = \frac{C(x_1, x_2) C_{12}(x_1, x_2)}{C_1(x_1, x_2) C_2(x_1, x_2)}, x, y \in [0, 1],$$

où

$$C_1(x_1, x_2) = \frac{\partial C(x_1, x_2)}{\partial x_1} = \partial_1 C(x_1, x_2),$$

$$C_2(x_1, x_2) = \frac{\partial C(x_1, x_2)}{\partial x_2} = \partial_2 C(x_1, x_2)$$

et

$$C_{12}(x_1, x_2) = \frac{\partial^2 C(x_1, x_2)}{\partial x_1 \partial x_2} = \partial_{12} C(x_1, x_2).$$

(20) avait introduit cette fonction en essayant d'étudier la transmission de maladie des parents aux enfants. Si l'on considère le père et son fils par exemple, et que X_1 correspond à l'âge de décès du fils alors que X_2 correspond à l'âge de décès du père, les deux ne doivent probablement pas être indépendants (à cause de nombreux facteurs communs, ou de transmission de certaines pathologies). La fonction $\Theta(x_1, x_2)$ devrait permettre de mesurer le degré d'association entre X_1 et X_2 , le cas d'indépendance correspondant au cas où $\Theta(x_1, x_2) = 1$ (et réciproquement d'ailleurs), avec une forme de dépendance positive dès lors que $\Theta(x_1, x_2) > 1$, au sens où $\Theta(x_1, x_2) > 1$ si et seulement si (X_1, X_2) est croissant dans le coin gauche (left corner set increasing, LCSI). On note aussi que Θ tend vers l'infini dans le cas comonotone.

Remarque 6.13 Cette fonction Θ apparaît dans (33) sous le nom de « ratio de fonctions de hasard conditionnel ». En effet, si X et Y sont deux durées de vie et que l'on pose

$$h(x_1, x_2) = -\frac{\partial_{12} \mathbb{P}(X_1 > x_1, X_2 > x_2)}{\mathbb{P}(X_1 > x_1, X_2 > x_2)},$$

(correspondant au taux de hasard de (X_1, X_2)),

$$h_2(x_1, x_2) = -\frac{\partial_2 \mathbb{P}(X_1 > x_1, X_2 > x_2)}{\mathbb{P}(X_1 > x_1, X_2 > x_2)},$$

(le taux de hasard de X_2 sachant que X_1 a vécu au moins une durée x_1) et

$$h_{2|1}(X_1, X_2) = -\frac{\partial_{12}\mathbb{P}(X_1 > x_1, X_2 > x_2)}{\partial_1\mathbb{P}(X_1 > x_1, X_2 > x_2)},$$

(le taux de hasard de X_2 sachant que X_1 était décédé avant x_1 années), alors

$$\Theta(X_1, X_2) = \frac{h_{Y|X}(X_1, X_2)}{h_Y(X_1, X_2)} = \frac{h_{X|Y}(X_1, X_2)}{h_X(X_1, X_2)}.$$

En allant un peu plus loin, comme le notent (20) ou (74), on peut aussi écrire le ratio

$$\Theta(x_1, x_2) = \frac{h(x_1, x_2)}{h_1(x_1, x_2)h_2(x_1, x_2)}.$$

Cette fonction, définie dans (20) comme un odds-ratio, est symétrique et indépendant des lois marginales.

La copule de Pareto

Une autre approche a permis d'aboutir à la même copule. Pour cela, supposons que X_1 et X_2 soient conditionnellement indépendantes, sachant Θ , et que $\mathbb{P}(X_1 > x_1|\Theta = \theta) = \exp(-\theta x_1)$ et $\mathbb{P}(X_2 > x_2|\Theta = \theta) = \exp(-\theta x_2)$. Supposons de plus que Θ suive une loi Gamma, alors

$$\begin{aligned} \mathbb{P}(X_1 > x_1, X_2 > x_2) &= \int_0^\infty \exp(-\theta[x_1 + x_2]) \frac{\theta^{\gamma-1} \exp(-\theta/\beta)}{\beta^\gamma \Gamma(\gamma)} d\theta \\ &= (1 + \beta x_1 + \beta x_2)^{-\gamma}. \end{aligned} \quad (6.14)$$

On notera tout d'abord que les lois marginales (non conditionnelles) sont des lois de Pareto, $\mathbb{P}(X_1 > x_1) = (1 - \beta x_1)^{-\gamma}$ et $\mathbb{P}(X_2 > x_2) = (1 - \beta x_2)^{-\gamma}$. De plus, la copule de survie du couple (X_1, X_2) est

$$C^*(u_1, u_2) = (u_1^{-1/\gamma} + u_2^{-1/\gamma} - 1)^{-\gamma},$$

avec $\gamma > 0$, parfois appelée copule de Pareto.

L'approche par vieillissement (frailty model)

Ces modèles ont été introduits à la fin des années 70, et popularisée par (74). L'idée est d'étendre la modélisation de la partie précédente. On suppose que l'on veut modéliser deux durées de vie, X_1 et X_2 , indépendantes conditionnellement à un facteur exogène Θ . On assume aussi que les lois conditionnelles de X_1 et X_2 sachant Θ sont de la forme

$$\bar{F}_{1|\Theta}(x_1|\theta) = \mathbb{P}(X_1 > x_1|\Theta = \theta) = \bar{G}_1(x)^\theta,$$

pour une fonction de référence G_1 , pour tout θ , et de manière similaire pour la loi de X_2 sachant $\Theta = \theta$. Alors

$$\begin{aligned}\mathbb{P}(X_1 > x_1, X_2 > x_2) &= \mathbb{E}(\mathbb{P}(X_1 > x_1, X_2 > x_2 | \Theta)) \\ &= \mathbb{E}(\mathbb{P}(X_1 > x_1 | \Theta) \mathbb{P}(X_2 > x_2 | \Theta)),\end{aligned}$$

qui peut s'écrire

$$\mathbb{E}(\exp[-\Theta(-\log \mathbb{P}(X_1 > x_1))] \exp[-\Theta(-\log \mathbb{P}(X_2 > x_2))]),$$

d'où finalement, si ψ correspond à la transformée de Laplace de Θ , c'est-à-dire $\psi(t) = \mathbb{E}(\exp(-t\Theta))$, on peut écrire

$$\mathbb{P}(X_1 > x_1, X_2 > x_2) = \psi(-\log \mathbb{P}(X_1 > x_1) - \log \mathbb{P}(X_2 > x_2)).$$

Compte tenu de l'écriture des lois marginales, $\mathbb{P}(X_1 > x_1) = \psi(-\log \bar{G}_1(x_1))$ la coule de survie de (X_1, X_2) est alors donnée par

$$C^*(u_1, u_2) = \psi(\psi^{-1}(u_1) + \psi^{-1}(u_2)),$$

qui est une copule Archimédienne de générateur $\phi = \psi^{-1}$. Dans le cas particulier où le facteur exogène Θ suit une loi Gamma, $\psi(t) = (1+t)^{1/(\alpha)}$, alors C^* est la copule de Clayton.

La copule de Clayton est particulièrement intéressante dans les modèles de survie car elle vérifie une propriété de stabilité par troncature : si la copule de survie du couple (X_1, X_2) est une copule de Clayton (de paramètre α), il en sera de même pour (X_1, X_2) sachant $X_1 > t$ et $X_2 > t$, pour tout t (avec le même paramètre α). Comme démontré dans (?), il s'agit d'une caractérisation de la copule de Clayton au sein des copules Archimédiennes qui reste valide en dimension d quelconque.

Propriétés de la copule de Clayton, en dimension $d = 2$

Définition 6.22 Soit $\theta \geq -1$, la copule de Clayton de paramètre θ est définie sur $[0, 1] \times [0, 1]$ par

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

Si $\theta = 0$, on considèrera le cas limite $C^\perp(u_1, u_2) = u_1 u_2$. Plus généralement, notons que $\theta \rightarrow -1$, $\theta \rightarrow 0$ et $\theta \rightarrow \infty$, C correspond respectivement au cas anticomotone, au cas indépendant, et à la copule comotone. De plus, si $0 \leq \theta_1 \leq \theta_2$, notons que $C_{\theta_1}(u_1, u_2) \leq C_{\theta_2}(u_1, u_2)$ pour tout $u_1, u_2 \in [0, 1]$.

La densité de la Copule de Clayton est représentée sur la Figure 6.9.

Copule de Clayton en dimension $d > 2$

La copule de Clayton étant Archimédienne, de générateur $\phi(t) = t^{-\theta} - 1$, on peut aisément l'étendre à la dimension $d > 2$, à condition de se restreindre au cas $\theta \geq 0$.

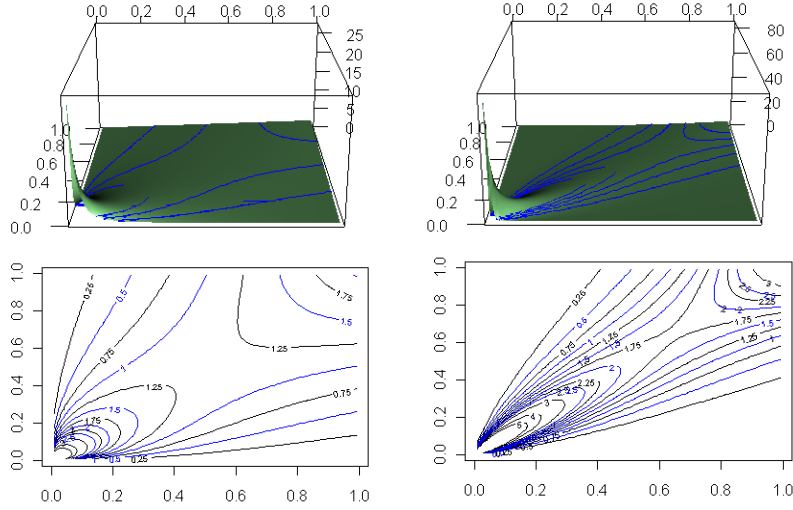


Figure 6.9 – Exemples de densités de la copule de Clayton.

Définition 6.23 Pour tout $\theta \geq 0$, la copule de Clayton de paramètre θ est définie sur $[0, 1] \times \dots \times [0, 1]$ par

$$C(u_1, \dots, u_d) = (u_1^{-1/\theta} + \dots + u_d^{-1/\theta} - (d-1))^{-\theta}. \quad (6.15)$$

Cette copule peut être obtenue de la manière suivante : posons

$$U_i = \left(1 + \frac{Y_i}{Z}\right)^{-\theta}, i = 1, \dots, d,$$

où les Y_i sont des variables exponentielles $\mathcal{E}(1)$ indépendantes, indépendante de $Z \sim \mathcal{G}(\theta, 1)$. Alors (U_1, \dots, U_n) admet pour loi jointe C donnée par 6.15.

6.4.7 Le modèle de Marshall et Olkin

La classe des copules dite de Marshall et Olkin est dérivée du modèle à choc commun introduit par (66).

Le modèle à choc commun

Soient X_1 et X_2 deux durées de vies, associés à deux composants (notés x_1 et x_2). Supposons que 3 chocs peuvent affecter ces composantes : deux associés aux composants x_1 et x_2 , indépendamment, et un dernier qui affecte les deux. Supposons que ces chocs sont modélisés par des processus de Poisson, avec une durée avant le premier choc Z_1 (exponentielle, de paramètre λ_1), qui affecte

x_1 , Z_2 (exponentielle, de paramètre λ_2), qui affecte x_2 , et Z_{12} (exponentielle, de paramètre λ_{12}), qui affecte x_1 et x_2 . Les dates de survenance des chocs sont supposés indépendantes.

Si l'on suppose que les chocs sont fatals pour les deux composants, la durée de vie des composants, (X_1, X_2) admet pour fonction de survie :

$$\begin{aligned}\bar{F}(X_1, X_2) &= \mathbb{P}(X_1 > x_1, X_2 > x_2) \\ &= \mathbb{P}(Z_1 > x_1) \cdot \mathbb{P}(Z_2 > x_2) \cdot \mathbb{P}(Z_{12} > \min\{x_1, x_2\}).\end{aligned}$$

Or les lois des Z_i sont exponentielles, et donc :

$$\mathbb{P}(X_1 > x_1, X_2 > x_2) = \exp(-\lambda_1 x_1 - \lambda_2 x_2 - \lambda_{12} \max\{x_1, x_2\}), x_1, x_2 > 0.$$

Proposition 6.7 \bar{F} satisfait une propriété faible d'absence de mémoire, au sens où

$$\bar{F}(x_1 + t, x_2 + t) = \bar{F}(x_1, x_2) \bar{F}(t, t).$$

En effet, soient $x_1, x_2, t > 0$, alors

$$\bar{F}(x_1 + t, x_2 + t) = \exp(-\lambda_1(x_1 + t) - \lambda_2(x_2 + t) - \lambda_{12} \max\{x_1 + t, x_2 + t\}),$$

c'est-à-dire

$$\begin{aligned}\bar{F}(x_1 + t, x_2 + t) &= \exp(-\lambda_1 x_1 - \lambda_2 x_2 - \lambda_{12} \max\{x_1, x_2\}) \\ &\quad \times \exp(-[\lambda_1 - \lambda_2 - \lambda_{12}] \cdot t),\end{aligned}$$

que l'on peut réécrire $\bar{F}(x_1 + t, x_2 + t) = \bar{F}(x_1, x_2) \bar{F}(t, t)$ pour tout $x_1, x_2, t > 0$.

Remarque 6.14 Il existe une propriété forte d'absence de mémoire (introduite dans (66)) à savoir

$$\bar{F}(x_1 + t_1, x_2 + t_2) = \bar{F}(x_1, x_2) \cdot \bar{F}(t_1, t_2),$$

pour tout $x_1, x_2, t_1, t_2 > 0$. (66) et (1) ont montré que les seuls vecteurs à composantes indépendantes pouvaient vérifier cette propriété. Notons également qu'une autre écriture possible de la propriété faible peut être que

$$\bar{F}((1 + \alpha)x_1, (1 + \alpha)x_2) = \bar{F}(x_1, x_2) \cdot \bar{F}(\alpha x_1, \alpha x_2), \text{ pour tout } \alpha > 0.$$

Cette équation fonctionnelle a été considérée en particulier dans (76). De telles fonctions sont liées aux problèmes d'extrêmes multivariés.

Si on pose

$$\alpha = \frac{\lambda_{12}}{\lambda_1 + \lambda_{12}} \text{ et } \beta = \frac{\lambda_{12}}{\lambda_2 + \lambda_{12}},$$

en notant que les lois marginales X_1 et X_2 suivent des lois exponentielles de paramètre $\lambda_1 + \lambda_{12}$ et $\lambda_2 + \lambda_{12}$, respectivement, on en déduit aisément que la copule de survie du couple (X_1, X_2) s'écrit :

$$C^*(u_1, u_2) = u_1 u_2 \min\{u_1^{-\alpha}, u_2^{-\beta}\} = \min\{u_1^{1-\alpha} u_2, u_1 u_2^{1-\beta}\}.$$

Copule induite par le modèle de Marshall et Olkin

Définition 6.24 *Etant donnés $\alpha, \beta \in (0, 1)$, la copule de Marshall-Olkin de paramètre (α, β) est*

$$C(u_1, u_2) = \min\{u_1^{1-\alpha}u_2, u_1u_2^{1-\beta}\}.$$

Cette copule est parfois appelée copule de Cuadras-Augé dans le cas $\alpha = \beta$, qui apparaît dans (25).

Remarque 6.15 On peut remarquer que ces copules sont obtenues par mélange, et elles possèdent une composante singulière, de masse strictement positive suivant la courbe $\mathcal{C} = \{(u_1, u_2) \in [0, 1] \times [0, 1], u_1^\alpha = u_2^\beta\}$. De plus, la masse de cette courbe est

$$\begin{aligned} \mathbb{P}(U_1 = U_2) &= \mathbb{P}(\text{le premier choc affecte les deux composantes}) \\ &= \frac{\lambda_{12}}{\lambda_1 + \lambda_2 + \lambda_{12}} > 0. \end{aligned}$$

On notera que pour cette copule, le τ de Kendall et le ρ de Spearman (nous reviendrons en détail sur ces mesures de dépendance dans la section 6.5.2) sont respectivement

$$\tau = \frac{\alpha\beta}{\alpha + \beta - \alpha\beta} \text{ et } \rho = \frac{3\alpha\beta}{2\alpha + 2\beta - \alpha\beta},$$

et les indices de dépendance de queue (nous reviendrons en détail sur ces mesures de dépendance dans la section 6.5.4) sont donnés par

$$\lambda_L = \lim_{u \rightarrow 0} \frac{C(u, u)}{u} = \lim_{u \rightarrow 0} \frac{u^2 \min\{u^\alpha, u^\beta\}}{u} = \lim_{u \rightarrow 0} u \min\{u^\alpha, u^\beta\} = 0,$$

et

$$\lambda_U = \lim_{u \rightarrow 1} \frac{C^*(u, u)}{1 - u} = \lim_{u \rightarrow 1} \frac{1 - 2u + u^2 \min\{u^\alpha, u^\beta\}}{1 - u} = \lim_{u \rightarrow 1} \frac{1 - 2u + u^2 u^\alpha}{1 - u} = \alpha,$$

en supposant que $\alpha < \beta$. Plus généralement,

$$\lambda_L = 0 \text{ et } \lambda_U = \min\{\alpha, \beta\}.$$

Intérêts de la copule de Marshall et Olkin

Ce modèle à « choc commun » présente l'avantage de le rendre facilement interprétable, et pratique pour la programmation. Il a ainsi été abondamment utilisé en fiabilité (par exemple dans (50)), afin de modéliser les « risques compétitifs » (« competing risks » présentés dans (27)). En assurance vie, (42) et (12) ont montré son intérêt, et surtout la facilité avec laquelle on peut dériver des formules fermées en ne supposant plus les durées de vie entre époux comme

indépendantes. Si on note T_x et T_y les durées de vie résiduelles d'un mari et de son épouse à la signature d'un contrat d'assurance vie (ils sont alors respectivement l'âge x et y), on note classiquement ${}_k p_{\overline{xy}}$ la probabilité conditionnelle qu'au moins un des deux survive k années,

$${}_k p_{\overline{xy}} = 1 - \mathbb{P}(T_x \leq k, T_y \leq k).$$

Si l'on suppose ces durées de vie résiduelles exponentielles, alors

$${}_k p_{\overline{xy}} = {}_k p_x + {}_k p_y - \exp(-\lambda_{xy}k) {}_k p_x \cdot {}_k p_y.$$

L'annuité pour un contrat au dernier survivant s'écrit alors

$$a_{\overline{xy}} = \sum_{k=1}^{\infty} v^k \mathbb{P}(T_x > k \text{ or } T_y > k) = \sum_{k=1}^{\infty} v^k {}_k p_{\overline{xy}}.$$

En posant ${}_k p_x^* = \exp(\lambda_{xy}k) {}_k p_x$ et ${}_k p_y^* = \exp(\lambda_{xy}k) {}_k p_y$, les annuités s'écrivent alors as

$$a_{\overline{xy}} = \sum_{k=1}^{\infty} e^{-(\delta + \lambda_{xy}) \cdot k} ({}_k p_x^* + {}_k p_y^* - {}_k p_x^* \cdot {}_k p_y^*),$$

correspondant à un calcul fait en supposant l'indépendance entre les durées de vie, et en actualisant avec un facteur $\delta + \lambda_{xy}$. On note que plus la dépendance est forte, plus faible sera le montant de l'annuité.

Les généralisations du modèle de Marshall et Olkin

Le modèle de Marshall et Olkin est caractérisé par des lois marginales exponentielles, et une propriété faible d'absence de mémoire, c'est-à-dire

$$\overline{F}(x_1 + t, x_2 + t) = \overline{F}(x_1, x_2) \overline{F}(t, t), x_1, x_2, t > 0, \quad (6.16)$$

(70) a proposé une extension de la forme

$$\overline{F}(x_1 \star t, x_2 \star t) = \overline{F}(x_1, x_2) \overline{F}(t, t), x, y, t > 0, \quad (6.17)$$

où \star est un opérateur binaire associatif (c'est-à-dire $(x \star y) \star z = x \star (y \star z)$).

Comme le notaient (1) et (85), les opérateurs associatif continue (qui vérifient de plus la fonction ' $x \star y = x \star z$ ou $y \star x = z \star x$ implique $y = z$ ') peuvent s'écrire

$$x \star y = \phi^{-1}(\phi(x) + \phi(y)),$$

où ϕ est une fonction continue strictement monotone. Si on suppose de plus que \star admette un élément identité e (c'est-à-dire $x \star e = x$), alors, (70) a montré que la solution de la propriété d'absence de mémoire $\overline{F}(x \star t) = \overline{F}(x) \overline{F}(t)$ serait de la forme

$$\overline{F}(x) = \exp(-\lambda \phi(x)), \lambda > 0 \text{ et } e = \phi^{-1}(0) < t.$$

Aussi, la solution continue de la condition 6.17

$$\bar{F}(x_1 \star t, x_2 \star t) = \bar{F}(x_1, x_2) \bar{F}(t, t), \text{ pour } x_1, x_2, t > 0,$$

où les marges vérifient aussi la propriété d'absence de mémoire (qui s'écrit $\bar{F}_X(x \star t) = \bar{F}_X(x) \bar{F}_X(t)$) est de la forme

$$\bar{F}(x_1, x_2) = \exp(-\lambda_1 \phi(x_1) - \lambda_2 \phi(x_2) - \lambda_{12} \phi(\max\{x_1, x_2\})), \lambda_1, \lambda_2, \lambda_{12} > 0.$$

Exemple 6.30 Considérons les trois opérateurs suivants :

- si $x \star y = x + y$, alors ϕ est une fonction affine : on obtient des lois exponentielles ($\bar{F}(x) = \exp(-\lambda x)$), avec une propriété faible d'absence de mémoire, ce qui correspond au modèle proposé par Marshall et Olkin.
- si $x \star y = xy$, alors ϕ est la fonction log, et les lois marginales sont alors des lois de Pareto, $\bar{F}(x) = x^{-\lambda}$, $x > 1$. De plus

$$\bar{F}(x_1, x_2) = \frac{1}{x_1^{\lambda_1}} \frac{1}{x_2^{\lambda_2}} \frac{1}{\max\{x_1, x_2\}^{-\lambda_{12}}}, \quad (6.18)$$

qui est une loi de Pareto bivariée (mais pas celle présentée dans la section 6.4.6). Ce modèle s'interprète en notant que la modélisation par choc commun reste valide : (X_1, X_2) admet une loi de Pareto bivariée définie par l'Equation 6.18 si et seulement si il existe Z_1, Z_2 et Z_0 , suivant des loi de Pareto, indépendantes, telles que $X_1 = \min\{Z_1, Z_0\}$ et $X_2 = \min\{Z_2, Z_0\}$.

- si $x \star y = (x^\alpha + y^\alpha)^{1/\alpha}$, alors ϕ est une fonction puissance, et les lois marginales suivent des lois de Weibull, $\bar{F}(x) = \exp(-\lambda x^{-\alpha})$, $x > 0$. La fonction de survie jointe est alors

$$\bar{F}(x_1, x_2) = \exp(-\lambda_1 x_1^\alpha - \lambda_2 x_2^\alpha - \lambda_{12} \max\{x_1^\alpha, x_2^\alpha\}),$$

qui a été induit par (66). Cette fonction apparaît là aussi naturellement dans les modèles d'extrêmes multivariés.

6.4.8 Le modèle de Gumbel

La copule dite de Gumbel a été introduite dans (49). Elle est parfois appelée copule de « Gumbel-Hougaard », suite à son utilisation dans (53), ou encore la « distribution logistique extrême ».

Une loi logistique bivariée

Considérons la loi dont la fonction de répartition jointe est de la forme

$$F(x_1, x_2) = \exp\left(-(x_1^{-\theta} + x_2^{-\theta})^{1/\theta}\right), x_1, x_2 > 0,$$

où $\theta \geq 1$. Cette fonction peut se réécrire sous la forme

$$F(x_1, x_2) = \exp\left(-\left([- \log(e^{-1/x_1})]^\theta + [- \log(e^{-1/x_2})]^\theta\right)^{1/\theta}\right), x_1, x_2 > 0.$$

Les lois marginales de cette distribution sont des lois de Fréchet standards (de fonction de répartition $F_i(x_i) = e^{-1/x_i}$ pour $x_i > 0$). La copule associée à cette loi bivariée est alors

$$C(u_1, u_2) = \exp\left(-\left([-\log u_1]^\theta + [-\log u_2]^\theta\right)^{1/\theta}\right), u_1, u_2 \in [0, 1],$$

où $\theta \geq 1$.

Une distribution pour des défauts joints

Comme dans la section 6.4.6, considérons deux durées de vie résiduelles T_1 et T_2 , indépendantes conditionnellement à un facteur Θ (appelé en anglais « frailty factor ») de distribution stable, dont la transformée de Laplace est $\psi(t) = \exp(-t^\theta)$. On suppose que $T_1|\Theta$ et $T_2|\Theta$ suivent des lois exponentielles de paramètre Θ . La loi jointe est alors appelée distribution logistique extrême, et la copule associée est le copule de Gumbel. De plus, cette copule est Archimédienne, de générateur $\phi(t) = \psi^{-1}(t) = (-\log t)^\theta$, avec $\theta > 0$. On obtient également une extension de la copule en dimension $d \geq 2$.

Propriétés de la copule de Gumbel

Définition 6.25 Soit $\theta \geq 1$, la copule de Gumbel de paramètre θ est définie sur $[0, 1] \times [0, 1]$ par

$$C(u_1, u_2) = \exp\left(-\left([-\log u_1]^\theta + [-\log u_2]^\theta\right)^{1/\theta}\right), u_1, u_2 \in [0, 1].$$

La densité de la Copule de Gumbel est représentée sur la Figure 6.10.

Le τ de Kendall (nous reviendrons dans la section 6.5.2 sur cette mesure) d'une copule de Gumbel de paramètre $\theta \geq 1$ est

$$\tau = \frac{\theta - 1}{\theta},$$

comme le rappelle (73), et les indices de dépendance de queue (qui seront détaillés dans la section 6.5.4) sont $\lambda_L = 0$ et $\lambda_U = 2 - 2^{1/\theta}$.

Remarque 6.16 La copule de Gumbel est max-stable, au sens où pour tout $t > 0$,

$$C^t(u_1, u_2) = C(u_1^t, u_2^t), \text{ pour } u_1, u_2 \in [0, 1],$$

6.4.9 La copule Gaussienne

La copule est obtenue comme copule sous-jacente d'un vecteur Gaussien, par la relation

$$C(u_1, \dots, u_d) = F(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d)), \text{ pour } u_1, \dots, u_d \in [0, 1].$$

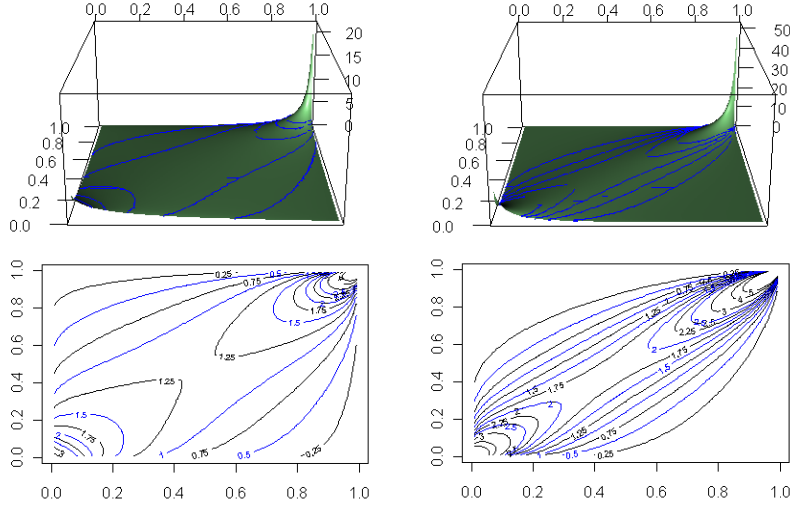


Figure 6.10 – Exemples de densités de la copule de Gumbel.

Définition 6.26 Pour tout $r \in [-1, 1]$, la copule Gaussienne est définie par

$$C(u_1, u_2) = \frac{1}{2\pi\sqrt{1-r^2}} \int_{-\infty}^{\Phi^{-1}(u_1)} \int_{-\infty}^{\Phi^{-1}(u_2)} \exp\left(\frac{x^2 - 2rxy + y^2}{2(1-r^2)}\right) dx dy,$$

si $r \in (-1, 1)$, avec respectivement C^- et C^+ si r vaut -1 ou $+1$.

Elle peut se définir en dimension $d \geq 2$, comme détaillé dans la section 6.4.2.

La densité de la copule Gaussienne peut se visualiser sur la Figure 6.11.

Pour une copule Gaussienne dont le vecteur Gaussien sous-jacent a pour corrélation $r \in [-1, 1]$, le τ de Kendall et le ρ de Spearman sont respectivement

$$\tau = \frac{2}{\pi} \sin^{-1}(r) \text{ et } \rho = \frac{6}{\pi} \sin^{-1}\left(\frac{r}{2}\right)$$

Ces deux mesures seront présentées plus en détails dans la section suivante.

6.5 Mesurer et ordonner des dépendances

(82) a proposé une axiomatique sur les propriétés fondamentales que devrait satisfaire une mesure de concordance κ . Pour cela, il convient que la mesure soit cohérente avec une relation d'ordre sur les paires de variables aléatoires. L'ordre le plus naturel (il n'existe pas d'ordre « naturel » dans \mathbb{R}^d) est l'ordre \preceq_{PQD} (de dépendance positive par quadrant, parfois appelé aussi « ordre de

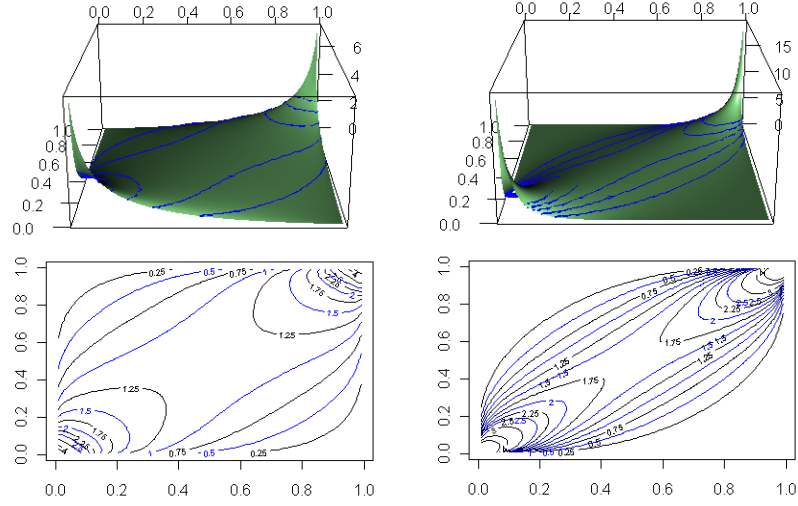


Figure 6.11 – Exemples de densités de la copule de Gaussienne.

concordance »), défini par $\mathbf{X} = (X_1, X_2) \preceq_{PQD} (Y_1, Y_2) = \mathbf{Y}$ si et seulement si

$$\mathbb{P}(F_{X_1}(X_1) \leq u_1, F_{X_2}(X_2) \leq u_2) \leq \mathbb{P}(F_{Y_1}(Y_1) \leq u_1, F_{Y_2}(Y_2) \leq u_2),$$

pour tout $0 \leq u_1, u_2 \leq 1$. Cet ordre fera penser à une extension de l'ordre \preceq_{VaR} évoqué dans le chapitre 3, en dimension supérieure. Les extensions en dimension $d \geq 2$ sont toutefois nombreuses, car si en dimension 1, comparer les fonctions de répartition et les fonctions de survie était équivalent, ça ne sera pas le cas en dimension plus grande.

Si on se restreint à une classe de Fréchet, de sorte que \mathbf{X} et \mathbf{Y} ont les mêmes lois marginales, de manière que $C_{\mathbf{X}}$ est la copule associée à \mathbf{X} , et $C_{\mathbf{Y}}$ est la copule associée à \mathbf{Y} , alors

$$\mathbf{X} = (X_1, X_2) \preceq_{PQD} (Y_1, Y_2) = \mathbf{Y} \text{ si et seulement si } C_{\mathbf{X}}(u_1, u_2) \leq C_{\mathbf{Y}}(u_1, u_2),$$

pour tout $u_1, u_2 \in [0, 1]$.

Définition 6.27 κ est une mesure de concordance si et seulement si κ vérifie les propriétés suivantes :

1. κ est définie pour toute paire de variables continues (X_1, X_2) ,
2. $-1 \leq \kappa(X_1, X_2) \leq +1$, $\kappa(X_1, X_1) = +1$ et $\kappa(X_1, -X_1) = -1$,
3. $\kappa(X_1, X_2) = \kappa(X_2, X_1)$,
4. si X_1 et X_2 sont indépendantes, alors $\kappa(X_1, X_2) = 0$,
5. $\kappa(-X_1, X_2) = \kappa(X_1, -X_2) = -\kappa(X_1, X_2)$,
6. si $(X_1, X_2) \preceq_{PQD} (Y_1, Y_2)$, alors $\kappa(X_1, X_2) \leq \kappa(Y_1, Y_2)$,

7. si $(X_1^1, X_2^1), (X_1^2, X_2^2), \dots$ est une suite de vecteurs qui converge en loi vers (X_1, X_2) alors $\kappa(X_1^n, X_2^n) \rightarrow \kappa(X_1, X_2)$ lorsque $n \rightarrow \infty$.

On notera que si ces axiomes semblent naturels, la corrélation par exemple ne satisfait pas le premier (cette dernière n'étant définie que pour des vecteurs dont les composantes sont dans L^2). A l'aide du second et du cinquième axiome, on voit que l'on pourra définir une notion de « dépendance positive », mais nous empêchera généralement d'avoir une équivalence dans le quatrième axiome. On remarquera également que les propriétés 2-7 correspondent à des propriétés de la corrélation dans le cas des vecteurs Gaussiens.

Remarque 6.17 (79) avait proposé de définir les mesures de dépendance sous la forme de distance à l'indépendance, ne permettant pas de distinguer dépendance positive et négative. (79) imposait en particulier que $\delta(X, \pm X) = +1$, et des propriétés de linéarité de la mesure, c'est-à-dire $\delta(aX+b, cY+d) = \delta(X_1, X_2)$ pour tout $a, c > 0$ et b, d . En supposant que δ est compris entre 0 et 1, (79) avait défini une distance à l'indépendance (au sens classique du terme, c'est à dire une fonction réelle positive). Mais selon (86) ces hypothèses sont probablement trop contraignantes.

Proposition 6.8 Si κ est une mesure de concordance, et si $f, g : \mathbb{R} \rightarrow \mathbb{R}$ sont deux fonctions croissantes alors $\kappa(f(X), g(Y)) = \kappa(X_1, X_2)$. De plus, $\kappa(X_1, X_2) = 1$ s'il existe f presque sûrement strictement croissante telle que $Y = f(X)$ with f ; et de manière analogue $\kappa(X_1, X_2) = -1$ si $Y = f(X)$ où f presque sûrement strictement décroissante.

La preuve est donnée dans (82).

On déduit de cette proposition que les mesures de concordance sont des fonctions de la copule uniquement, au sens où si (X_1, X_2) et (Y_1, Y_2) ont la même copule (notée C), alors $\kappa(X_1, Y_1) = \kappa(X_2, Y_2) = \kappa(C)$.

Exemple 6.31 Le β de Blomqvist, parfois appelé « coefficient de corrélation médiane » (et parfois noté q) est défini par

$$\begin{aligned} \beta &= \mathbb{P}((X_1 - \text{médiane}(X_1))(X_2 - \text{médiane}(X_2)) > 0) \\ &- \mathbb{P}((X_1 - \text{médiane}(X_1))(X_2 - \text{médiane}(X_2)) < 0), \end{aligned}$$

qui peut aussi s'écrire

$$\beta = 4C\left(\frac{1}{2}, \frac{1}{2}\right) - 1.$$

6.5.1 La corrélation de Pearson

Soient X_1 et X_2 deux variables aléatoires continues de variances finies, alors la corrélation linéaire (ou de Pearson) est définie par

$$\begin{aligned}\text{Corr}(X_1, X_2) &= \frac{\text{Cov}(X_1, X_2)}{\sqrt{\text{Var}(X) \cdot \text{Var}(Y)}} \\ &= \frac{\mathbb{E}([X - \mathbb{E}(X)][X - \mathbb{E}(Y)])}{\sqrt{\mathbb{E}([X - \mathbb{E}(X)]^2)\mathbb{E}([Y - \mathbb{E}(Y)]^2)}}.\end{aligned}$$

On parle de « corrélation linéaire » au sens où $\text{Corr}(X_1, X_2) = +1$ si et seulement s'il existe $a > 0$ et b tels que $X_2 = aX_1 + b$ presque sûrement. De plus, cet coefficient est invariant par transformations affines, c'est-à-dire

$$\text{Corr}(aX_1 + b, cX_2 + d) = \text{Corr}(X_1, X_2) \text{ si } a \text{ et } c \text{ sont de même signe.}$$

Exemple 6.32 Comme $\text{Corr}(X_1, X_2)$ peut s'écrire sous la forme $\mathbb{E}(\phi(X_1, X_2))$ où $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}$ supermodulaire, alors $\text{Corr}(X_1, X_2)$ vérifie l'inégalité de Tchen (obtenue par (91), cette relation sera détaillée davantage dans la section 6.6.1),

$$\begin{aligned}\text{Corr}(X_1^-, X_2^-) &= \text{Corr}(F_1^{-1}(U), F_2^{-1}(1 - U)) \leq \text{Corr}(X_1, X_2) \\ &\leq \text{Corr}(F_1^{-1}(U), F_2^{-1}(U)) = \text{Corr}(X_1^+, X_2^+),\end{aligned}$$

où U est uniformément distribué sur $[0, 1]$. En conséquence, le coefficient de corrélation ne décrit pas nécessairement l'intervalle $[-1, 1]$. Ainsi, si X_1 et X_2 suivent une loi lognormale, de paramètre de variance 1 et σ , respectivement,

$$\frac{e^\sigma - 1}{\sqrt{e^{\sigma^2} - 1}\sqrt{e - 1}} \leq \text{Corr}(X_1, X_2) \leq \frac{e^\sigma - 1}{\sqrt{e^{\sigma^2} - 1}\sqrt{e - 1}}. \quad (6.19)$$

comme le montrent (80) et (35). Ceci implique en particulier qu'un coefficient de corrélation proche de 0 peut parfaitement correspondre à un cas de comonotonie.

Exemple 6.33 Ce coefficient de corrélation apparaît naturellement dans les modèles de régression (linéaire). Si on considère la régression de Y sur X , les estimateurs des coefficients de la régression \hat{a} et \hat{b} qui minimisent la somme des carrés des erreurs $\mathbb{E}(Y - (aX + b))^2$, sont ici :

$$\hat{a} = \frac{\text{Cov}(X, Y)}{\text{Var}(X)} = \text{Corr}(X, Y) \sqrt{\frac{\text{Var}(Y)}{\text{Var}(X)}} \text{ et } \hat{b} = \mathbb{E}(Y) - \hat{a} \mathbb{E}(X).$$

6.5.2 Le tau de Kendall et le rho de Spearman

Le ρ de Spearman entre deux variables continues est la corrélation (au sens de Pearson) entre $U_1 = F_1(X_1)$ et $U_2 = F_2(X_2)$. Comme U_1 et U_2 sont uniformément réparties sur $[0, 1]$, $\mathbb{E}(U_1) = \mathbb{E}(U_2) = 1/2$, et $\text{Var}(U_1) = \text{Var}(U_2) = 1/12$, et donc

$$\rho(X_1, X_2) = \text{Corr}(U_1, U_2) = \frac{\mathbb{E}(U_1 U_2) - 1/4}{1/12} = 12 \mathbb{E}(U_1 U_2) - 3.$$

Définition 6.28 Soit (X_1, X_2) un couple de variables aléatoire de loi jointe $F \in \mathcal{F}(F_1, F_2)$, de copule C , et de lois marginales continues. Alors le ρ de Spearman est

$$\begin{aligned}\rho(X_1, X_2) &= 12 \int_0^1 \int_0^1 C(u_1, u_2) du_1 du_2 - 3 \\ &= 12 \int_{\mathbb{R}} \int_{\mathbb{R}} [F(x_1, x_2) - F_1(x_1)F_2(x_2)] dx_1 dx_2.\end{aligned}$$

Comme le note (33), l'expression ci-dessus est une distance moyenne entre F et F^\perp . Ce coefficient a été introduit sous sa forme empirique par Spearman en 1904, par la relation

$$\rho(X_1, X_2) = 3[\mathbb{P}((X_1 - Y_1)(Y_2 - Z_3) > 0) - \mathbb{P}((X_1 - Y_2)(X_2 - Z_3) < 0)],$$

où (X_1, X_2) , (Y_1, Y_2) et (Z_1, Z_2) sont trois versions indépendantes de (X_1, X_2) , tel que l'introduit également (73).

Définition 6.29 Soit (X_1, X_2) un couple de variables aléatoires continues de loi jointe F , de copule C , alors de τ de Kendall est défini par

$$\tau(X_1, X_2) = 4 \int_0^1 \int_0^1 C(u_1, u_2) dC(u_1, u_2) - 1 = 4\mathbb{E}(F(X_1, X_2)) - 1.$$

Là encore, initialement, le τ de Kendall n'a pas été défini, initialement, à l'aide des copules, mais comme une probabilité de concordance à laquelle on soustrait la probabilité de discordance du vecteur (X_1, X_2) , c'est-à-dire

$$\tau(X_1, X_2) = 3[\mathbb{P}((X_1 - Y_1)(X_2 - Y_2) > 0) - \mathbb{P}((X_1 - Y_1)(X_2 - Y_2) < 0)],$$

où (X_1, X_2) et (Y_1, Y_2) sont deux versions indépendantes de (X_1, X_2) , tel que démontré par (73).

Proposition 6.9 Si X_1 et X_2 sont continues, alors le τ de Kendall et le ρ de Spearman sont des mesures de concordance.

Une preuve de ce résultat est donnée dans (73).

En particulier pour des vecteurs comonotones, $\rho(X_1, X_2) = 1$ et $\tau(X_1, X_2) = 1$, alors que pour des vecteurs anticomonotones $\rho(X_1, X_2) = -1$ et $\tau(X_1, X_2) = -1$. Et la réciproque est vraie : les bornes ± 1 correspondent aux cas comonotones ou anti-comonotones. (ce résultat est démontré dans (35)).

Exemple 6.34 Pour la copule min-max présentée dans l'Exemple 6.7, on notera que

$$\tau(X_{1:n}, X_{n:n}) = \frac{1}{2n-1}.$$

Exemple 6.35 Si X_1 et X_2 forment un couple de copule Archimédienne, de générateur ϕ , alors le τ de Kendall s'écrit

$$\tau(X_1, X_2) = 1 + 4 \int_0^1 \frac{\phi(t)}{\phi'(t)} dt.$$

Exemple 6.36 Comme le montre (55), considérons pour $\theta \in [0, 1]$ un mélange entre la copule indépendante et la borne supérieure de Fréchet-Hoeffding (considéré dans l'exemple 6.9),

$$C(u_1, u_2) = (1 - \theta)C^\perp(u_1, u_2) + \theta C^+(u_1, u_2),$$

pour tout $(u_1, u_2) \in [0, 1] \times [0, 1]$. Alors si (X_1, X_2) a pour copule C , $\rho(X_1, X_2) = \theta$. Plus généralement, si l'on rajoute la borne inférieure de Fréchet-Hoeffding, comme suggéré par (41), avec comme paramètres de mélange $\alpha, \beta > 0$ tels que $\alpha + \beta \leq 1$,

$$C(u_1, u_2) = \alpha C^-(u_1, u_2) + (1 - \alpha - \beta)C^\perp(u_1, u_2) + \beta C^+(u_1, u_2)$$

pour tout $(u_1, u_2) \in [0, 1] \times [0, 1]$. Alors $\rho(X_1, X_2) = \beta - \alpha$.

Les Tableaux 6.2 et 6.3 montrent l'évolution du τ de Kendall et du ρ de Spearman en fonction du paramètre sous jacent (noté ici θ), pour quelques familles usuelles (définies dans (73) par exemple).

Table 6.2 – τ de Kendall en fonction du paramètre θ de la copule sous-jacente.

| τ de Kendall | 0.0 | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 | 1.0 |
|----------------------|------|------|------|------|------|------|------|------|------|------|-----------|
| Gaussian θ | 0.00 | 0.16 | 0.31 | 0.45 | 0.59 | 0.71 | 0.81 | 0.89 | 0.95 | 0.99 | 1.00 |
| Gumbel θ | 1.00 | 1.11 | 1.25 | 1.43 | 1.67 | 2.00 | 2.50 | 3.33 | 5.00 | 10.0 | $+\infty$ |
| Plackett θ | 1.00 | 1.57 | 2.48 | 4.00 | 6.60 | 11.4 | 21.1 | 44.1 | 115 | 530 | $+\infty$ |
| Clayton θ | 0.00 | 0.22 | 0.50 | 0.86 | 1.33 | 2.00 | 3.00 | 4.67 | 8.00 | 18.0 | $+\infty$ |
| Frank θ | 0.00 | 0.91 | 1.86 | 2.92 | 4.16 | 5.74 | 7.93 | 11.4 | 18.2 | 20.9 | $+\infty$ |
| Joe θ | 1.00 | 1.19 | 1.44 | 1.77 | 2.21 | 2.86 | 3.83 | 4.56 | 8.77 | 14.4 | $+\infty$ |
| Galambos θ | 0.00 | 0.34 | 0.51 | 0.70 | 0.95 | 1.28 | 1.79 | 2.62 | 4.29 | 9.30 | $+\infty$ |
| Morgenstein θ | 0.00 | 0.45 | 0.90 | - | - | - | - | - | - | - | - |

Remarque 6.18 En analyse canonique nonlinéaire, on cherche les transformations (nonlinéaires) qui maximisent la corrélation entre deux variables, comme dans (79)

$$r^*(X_1, X_2) = \operatorname{argmax}\{\operatorname{Corr}(g_1(X_1), g_2(X_2))\}$$

où $a, b : \mathbb{R} \rightarrow \mathbb{R}$ sont des fonctions mesurables telles que $\operatorname{Var}(g_1(X_1))$ et $\operatorname{Var}(g_2(X_2))$ soient finies. On parle alors de de corrélation maximale. On notera que cette

Table 6.3 – ρ de Spearman en fonction du paramètre θ de la copule sous-jacente.

| ρ de Spearman | 0.0 | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 | 1.0 |
|----------------------|------|------|------|------|------|------|------|------|------|------|-----------|
| Gaussian θ | 0.00 | 0.10 | 0.21 | 0.31 | 0.42 | 0.52 | 0.62 | 0.72 | 0.81 | 0.91 | 1.00 |
| Gumbel θ | 1.00 | 1.07 | 1.16 | 1.26 | 1.38 | 1.54 | 1.75 | 2.07 | 2.58 | 3.73 | $+\infty$ |
| Plackett θ | 1.00 | 1.35 | 1.84 | 2.52 | 3.54 | 5.12 | 7.76 | 12.7 | 24.2 | 66.1 | $+\infty$ |
| Clayton θ | 0.00 | 0.14 | 0.31 | 0.51 | 0.76 | 1.06 | 1.51 | 2.14 | 3.19 | 5.56 | $+\infty$ |
| Frank θ | 0.00 | 0.60 | 1.22 | 1.88 | 2.61 | 3.45 | 4.47 | 5.82 | 7.90 | 12.2 | $+\infty$ |
| Joe θ | 1.00 | 1.12 | 1.27 | 1.46 | 1.69 | 1.99 | 2.39 | 3.00 | 4.03 | 6.37 | $+\infty$ |
| Galambos θ | 0.00 | 0.28 | 0.40 | 0.51 | 0.65 | 0.81 | 1.03 | 1.34 | 1.86 | 3.01 | $+\infty$ |
| Morgenstein θ | 0.00 | 0.30 | 0.60 | 0.90 | - | - | - | - | - | - | - |

mesure peut être intéressante, en particulier car $r^*(X_1, X_2) = 0$ si et seulement si les variables X_1 et X_2 sont indépendantes. De plus, dans le cas Gaussien, comme l'a montré (58), $r^*(X_1, X_2) = r(X_1, X_2)$, autrement dit la corrélation est maximale avec des transformations affines. (56) avaient suggéré de se limiter aux fonctions g_1 et g_2 monotones, introduisant ainsi la corrélation monotone.

6.5.3 Autres mesures de corrélation

Parmi les autres mesures intéressantes, car ne dépendant que des rangs, on pourra introduire l'indice γ de Gini, le β de Blomqvist, ou encore la classe de Schweizer et Wolff.

Définition 6.30 Soit (X_1, X_2) un couple de copule C . L'indice γ de Gini est défini par

$$\gamma(X_1, X_2) = 4 \left(\int_0^1 C(s, 1-s) ds - \int_0^1 |s - C(s, s)| ds \right).$$

(10) a proposé un test dit de quadrant, consistant évaluer la probabilité que X et Y dépassent conjointement la valeur médiane, en introduisant la mesure suivante :

Définition 6.31 Soit (X_1, X_2) un couple de copule C . L'indice β de Blomqvist est défini par

$$\begin{aligned} \beta(X_1, X_2) &= 2 \mathbb{P} \left(\left[X_1 > F_1^{-1} \left(\frac{1}{2} \right) \right] \left[X_2 > F_2^{-1} \left(\frac{1}{2} \right) \right] > 0 \right) - 1 \\ &= C \left(\frac{1}{2}, \frac{1}{2} \right) - 1. \end{aligned}$$

Enfin, (86) et (87) ont proposé de réécrire le ρ de Spearman,

$$\rho(X_1, X_2) = 12 \int_0^1 \int_0^1 (C(u_1, u_2) - C^\perp(u_1, u_2)) du_1 du_2$$

en considérant

$$\int_0^1 \int_0^1 |C(u_1, u_2) - C^\perp(u_1, u_2)| du_1 du_2$$

puis en changeant la norme utilisée pour mesurer la distance entre C et C^\perp ,

$$k(X_1, X_2) \propto \int_0^1 \int_0^1 \|C(u_1, u_2) - C^\perp(u_1, u_2)\|_* du_1 du_2,$$

où $\|\cdot\|_*$ désigne une norme (pas forcément la norme L_1), à une constante de normalisation près (de manière à avoir $k(X_1, X_1) = 1$).

Définition 6.32 Soit (X_1, X_2) un couple de copule C . L'indice Φ^2 de Hoeffding est défini à partir de la norme L_2 , par

$$k(X_1, X_2) = \sqrt{90 \int_0^1 \int_0^1 [C(u_1, u_2) - C^\perp(u_1, u_2)]^2 du_1 du_2}.$$

Définition 6.33 Soit (X_1, X_2) un couple de copule C . L'indice κ est défini à partir de la norme L_∞ par

$$\kappa(X_1, X_2) = 4 \sup_{[0,1]^2} \{|C(u_1, u_2) - C^\perp(u_1, u_2)|\}.$$

Cette dernière mesure peut se rapprocher d'une distance de Kolmogorov-Smirnov par rapport à l'indépendance.

6.5.4 Dépendance locale, dans les queues

Pour étudier la dépendance dans les queues de distribution (92) avait suggéré d'introduire des fonctions de concentrations dans les queues.

Définition 6.34 Pour la queue inférieure,

$$\begin{aligned} L(z) &= \frac{\mathbb{P}(X_1 < F_1^{-1}(z), X_2 < F_2^{-1}(z))}{z} = \frac{C(z, z)}{z} \\ &= \mathbb{P}(X_1 < F_1^{-1}(z) | X_2 < F_2^{-1}(z)) = \mathbb{P}(X_2 < F_2^{-1}(z) | X_1 < F_1^{-1}(z), X_2 < F_2^{-1}(z)), \end{aligned}$$

et pour la queue supérieure,

$$R(z) = \frac{\mathbb{P}(X_1 > F_1^{-1}(z), X_2 > F_2^{-1}(z))}{(1-z)} = \mathbb{P}(U > z | V > z).$$

(54) avait définie le paramètre de dépendance de queue supérieure et de queue inférieure, respectivement, en posant

$$\lambda_U = R(1) = \lim_{z \rightarrow 1} R(z) \text{ et } \lambda_L = L(0) = \lim_{z \rightarrow 0} L(z).$$

Définition 6.35 Soit (X_1, X_2) un couple aléatoire dans \mathbb{R}^2 . Les indices de dépendance de queue inférieure (L) et supérieure (U) sont définis respectivement, dès lors que les limites existent, par

$$\lambda_L = \lim_{u \rightarrow 0} \mathbb{P}(X_1 \leq F_1^{-1}(u) | X_2 \leq F_2^{-1}(u)),$$

et

$$\lambda_U = \lim_{u \rightarrow 1} \mathbb{P}(X_1 > F_1^{-1}(u) | X_2 > F_2^{-1}(u)).$$

Proposition 6.10 Soit (X_1, X_2) un couple de copule C et de copule duale C^* , alors les indices de dépendance de queue, s'ils existent, sont définis par

$$\lambda_L = \lim_{u \rightarrow 0} \frac{C(u, u)}{u} \text{ et } \lambda_U = \lim_{u \rightarrow 0} \frac{C^*(u, u)}{u}.$$

Exemple 6.37 Dans le cas de copules Archimédiennes, (73), (16) ou (18) montrent que

$$\lambda_U = 2 - \lim_{x \rightarrow 0} \frac{1 - \phi^{-1}(2x)}{1 - \phi^{-1}(x)} \text{ et } \lambda_L = \lim_{x \rightarrow 0} \frac{\phi^{-1}(2\phi(x))}{x} = \lim_{x \rightarrow \infty} \frac{\phi^{-1}(2x)}{\phi^{-1}(x)}.$$

De plus, des propriétés peuvent être obtenues dans le cas de distortion de générateurs, $\phi_{\alpha, \beta}(\cdot) = \phi(\cdot^\alpha)^\beta$. Dans ce cas, les coefficients de dépendance de queue supérieure et inférieure sont respectivement

$$\lambda_U \text{ et } \lambda_L^{1/\alpha} \text{ pour } \phi_{\alpha, 1}(\cdot) = \phi(\cdot^\alpha)$$

et

$$2 - (2 - \lambda_U)^{1/\beta} \text{ et } \lambda_L^{1/\beta} \text{ for } \phi_{1, \beta}(\cdot) = \phi(\cdot)^\beta$$

(60) a proposé une approche alternative pour quantifier la dépendance dans les queues. Considérons un couple de variables aléatoires de même loi, $X_1 \stackrel{\mathcal{L}}{=} X_2$. Alors :

– sous hypothèse d'indépendance,

$$\mathbb{P}(X_1 > t, X_2 > t) = \mathbb{P}(X_1 > t) \times \mathbb{P}(X_2 > t) = \mathbb{P}(X_1 > t)^2,$$

– sous hypothèse de comonotonie,

$$\mathbb{P}(X_1 > t, X_2 > t) = \mathbb{P}(X_1 > t) = \mathbb{P}(X_1 > t)^1.$$

On peut alors supposer que $\mathbb{P}(X_1 > t, X_2 > t) \sim \mathbb{P}(X_1 > t)^{1/\eta}$ lorsque $t \rightarrow \infty$, où $\eta \in (0, 1]$ sera appelée indice de dépendance de queue. Aussi, suivant l'idée de (21) on utilise la définition suivante d'indice de queue inférieure et de queue supérieure, respectivement notés η_U et η_L :

Définition 6.36 Soient

$$\bar{\chi}_U(z) = \frac{2 \log(1-z)}{\log C^\star(z, z)} - 1 \text{ et } \bar{\chi}_L(z) = \frac{2 \log(1-z)}{\log C(z, z)} - 1$$

Alors

$$\eta_U = \frac{1}{2} \left(1 + \lim_{z \rightarrow 0} \bar{\chi}_U(z) \right) \text{ et } \eta_L = \frac{1}{2} \left(1 + \lim_{z \rightarrow 0} \bar{\chi}_L(z) \right)$$

sont appelés indices de queue supérieure et inférieure, respectivement.

On notera que le cas indépendant correspondra au cas où η vaut $1/2$.

Exemple 6.38 Si (X_1, X_2) a une copule de Gumbel, et des marges Fréchet de paramètre 1,

$$\mathbb{P}(X \leq x, Y \leq y) = \exp(-(x^{-\alpha} + y^{-\alpha})^{1/\alpha}), \text{ où } \alpha \geq 0,$$

alors $\eta_U = 1$ alors que $\eta_L = 1/2^\alpha$. On peut montrer que dans le cas d'une copule de Clayton, $\eta_U = 1/2$ et $\eta_L = 1$. Dans le cas d'une copule Gaussienne de corrélation $r \in [-1, +1]$, $\eta_U = \eta_L = (1+r)/2$.

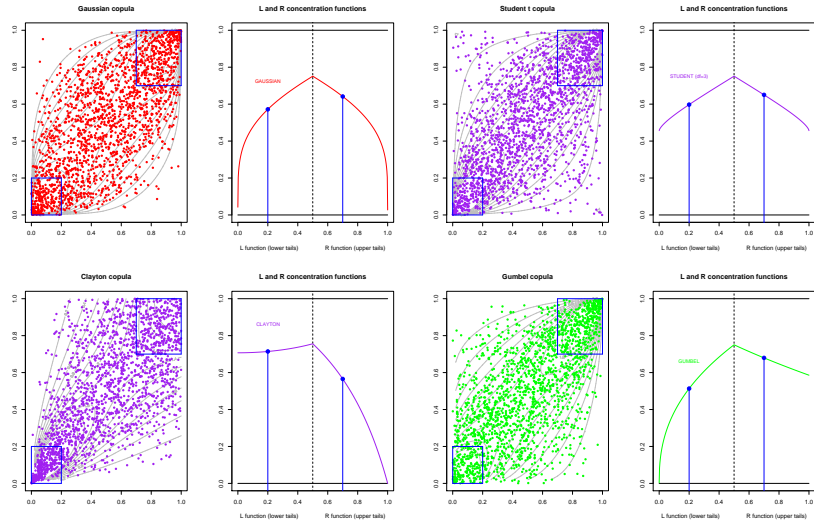


Figure 6.12 – Fonctions cumulatives L (sur $[0, 1/2]$) et R (sur $[1/2, 1]$) avec dans la partie supérieure, le cas Gaussien et le cas Student t à droite, et dans la partie inférieure le cas Clayton et Gumbel.

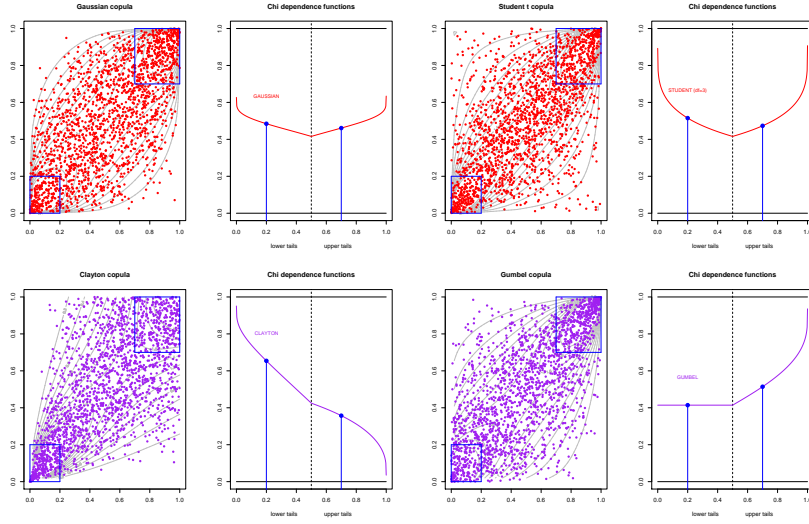


Figure 6.13 – Fonctions cumulatives $\bar{\chi}_L$ (sur $[0, 1/2]$) et $\bar{\chi}_U$ (sur $[1/2, 1]$) avec dans la partie supérieure, le cas Gaussien et le cas Student t à droite, et dans la partie inférieure le cas Clayton et Gumbel.

6.5.5 Quantifier la dépendance en dimension $d > 2$

En dimension $d > 2$, le cas Gaussien nous invite à étudier la dépendance par paires, en regardant des matrices de mesures de dépendance. On posera ainsi

$$\mathbf{R} = [\text{Corr}(X_i, X_j)], \text{ pour } i, j = 1, \dots, d.$$

Mais il est possible de résumer l'information en une mesure unique. Nous avons noté que les mesures de dépendance naturelles pouvaient parfois se voir comme des « distances » entre la copule sous-jacente C et la copule indépendante C^\perp , normalisé de manière à avoir 1 dans le cas comonotone. En particulier,

$$\rho(X_1, X_2) = \frac{\int_{[0,1] \times [0,1]} C(u_1, u_2) - C^\perp(u_1, u_2) du dv}{\int_{[0,1] \times [0,1]} C^+(u_1, u_2) - C^\perp(u_1, u_2) du dv} \quad (6.20)$$

En suivant les idées de (96) ou (54), il est possible d'étendre cette définition en dimension $d > 2$,

$$\begin{aligned} \rho(\mathbf{X}) &= \frac{\int_{[0,1]^d} C(\mathbf{u}) - C^\perp(\mathbf{u}) d\mathbf{u}}{\int_{[0,1] \times [0,1]} C^+(\mathbf{u}) - C^\perp(\mathbf{u}) d\mathbf{u}} \\ &= \frac{d+1}{2^d - (d+1)} \left(2^d \int_{[0,1]^d} C(\mathbf{u}) d\mathbf{u} - 1 \right) \end{aligned} \quad (6.21)$$

pour le ρ de Spearman, et pour le τ de Kendall

$$\tau(\mathbf{X}) = \frac{1}{2^{d-1} - 1} \left(2^d \int_{[0,1]^d} C(\mathbf{u}) dC(\mathbf{u}) - 1 \right) \quad (6.22)$$

On se doute toutefois que la borne inférieure ne devrait pas être -1 (C^- n'était plus une copule en dimension $d > 2$), mais on peut montrer que

$$\tau \geq \frac{-1}{2^{d-1} - 1} \text{ et } \rho \geq \frac{2^d - (d+1)!}{d!(2^d - (d+1))}.$$

Exemple 6.39 En dimension 3, le τ de Kendall est la moyenne des τ de Kendall par paires,

$$\tau(X_1, X_2, X_3) = \frac{1}{3} (\tau(X_1, X_2) + \tau(X_1, X_3) + \tau(X_2, X_3)).$$

On peut également définir un β de Blomqvist multivarié,

$$\beta = \frac{2^{d-1}}{2^{d-1} - 1} \left[C\left(\frac{1}{2}, \dots, \frac{1}{2}\right) + C^*\left(\frac{1}{2}, \dots, \frac{1}{2}\right) - 2^{1-d} \right]$$

Exemple 6.40 Dans le cas des copules Archimédiennes, de générateur ϕ , notons que

$$\beta = \frac{2^{d-1}}{2^{d-1} - 1} \cdot \phi^{-1} \left[d\phi\left(\frac{1}{2}\right) + \sum_{i=1}^d (-1)^i \binom{d}{i} \phi^{-1} \left(i\phi\left(\frac{1}{2}\right) \right) - 2^{1-d} \right].$$

6.6 De l'aggrégation des risques multiples

Dans cette section, nous aborderons la distribution d'une agrégation de plusieurs risques. Il peut s'agir de la somme des risques $X_1 + \dots + X_d$ (ou leur moyenne) mais de manière générale, une fonction d'aggrégation est une fonction $h : \mathbb{R}^d \rightarrow \mathbb{R}$.

Remarque 6.19 Nous allons considérer des fonctions un peu différentes de celles étudiées dans (2). Dans cette littérature, une fonction d'aggrégation est une fonction $h : \mathbb{R}^d \rightarrow \mathbb{R}$ telle que

- $h(\mathbf{0}) = h(0, \dots, 0) = 0$,
- $h(\mathbf{1}) = h(1, \dots, 1) = 1$,
- si $\mathbf{x} \leq \mathbf{y}$ (au sens $x_i \leq y_i$ pour tout i) $h(\mathbf{x}) \leq h(\mathbf{y})$.

On note que sous cette forme, les fonctions d'aggrégation sont proches des semi-copules introduites dans la section 6.3. Ici, on ne supposera pas forcément que $h(\mathbf{1}) = 1$ par exemple.

Exemple 6.41 En hydrologie, si X_i désigne le niveau d'un fleuve à un endroit, il n'est pas rare de chercher à modéliser les « low-flow », c'est à dire chercher la loi de $\min\{X_i\}$ (voire d'une version normalisée). En réassurance, il est commun de considérer des traités qui considèrent des remboursements qui partagent prorata capita les dépenses associées, de telle sorte que le remboursement s'écrit alors

$$S = h(X_1, X_2) = \begin{cases} 0 & \text{si } X_1 \leq R \\ X_1 - R + \frac{X_1 - R}{L - R} \cdot X_2 & \text{si } X_1 \in [R, L] \\ L - R + \frac{L - R}{L} X_2 & \text{si } X_1 > L \end{cases}$$

pour un niveau de rétention R et une limite L , où X_1 est le coût du sinistres (indemnité versée à l'assuré) et X_2 désigne les frais associés (expertise, avocat, etc.)

Si l'on spécifie les lois marginales et la copule, on connaît la loi du vecteur (X_1, \dots, X_d) , et on peut en déduire la loi de $S = h(X_1, \dots, X_d)$, quelle que soit la fonction d'aggrégation considérée. On peut aussi en déduire $\mathcal{R}(h(X_1, \dots, X_d))$ pour toute mesure de risque \mathcal{R} telle que définie dans le Chapitre 3.

Exemple 6.42 En finance, les options multisupports sont des exemples classiques où l'on a besoin d'agréger des risques. Rappelons que le prix d'un call européen est de la forme $\Gamma(T, K) = e^{-rT} \mathbb{E}_{\mathbb{Q}}((S_T - K)_+)$, où (S_t) est le prix du sous-jacent, et que l'on considère une option de maturité T , et de strike K . \mathbb{Q} est une probabilité risque neutre (nous supposons ici qu'elle existe, et qu'elle est unique). En fait, (13) montre comment extraire la probabilité risque neutre des prix des options. Notons que $(S_T - K)_+ = \int_K^\infty \mathbf{1}(S_T > x) dx$, et donc

$$\Gamma(T, K) = e^{-rT} \int_K^\infty \mathbb{Q}(S_T > x) dx,$$

ce qui donne

$$\mathbb{Q}(S_T \leq x) = -e^{-rT} \frac{\partial \Gamma}{\partial K}(T, x).$$

Considérons une option sur deux actifs, dont le payoff serait basé sur le maximum des prix atteints à la date T , $h(S_T^1, S_T^2) = (\max\{S_T^1, S_T^2\} - K)_+$. Le prix du call offrant ce payoff est

$$\begin{aligned} \Gamma(T, K) &= e^{-rT} \mathbb{E}_{\mathbb{Q}}((\max\{S_T^1, S_T^2\} - K)_+) \\ &= e^{-rT} \mathbb{E}_{\mathbb{Q}} \left(\int_K^\infty 1 - \mathbf{1}(\max\{S_T^1, S_T^2\} \leq x) dx \right) \\ &= e^{-rT} \int_K^\infty 1 - \underbrace{\mathbb{Q}(\max\{S_T^1, S_T^2\} \leq x)}_{\mathbb{Q}(S_T^1 \leq x, S_T^2 \leq x)} dx, \end{aligned}$$

aussi, si (S_T^1, S_T^2) a pour copule C (sous la probabilité \mathbb{Q}), alors Γ vérifie une relation de la forme

$$\Gamma(T, K) = e^{-rT} \int_K^\infty 1 - C \left(e^{rT} \frac{\partial \Gamma^1}{\partial K}(T, x), e^{rT} \frac{\partial \Gamma^2}{\partial K}(T, x) \right) dx.$$

où Γ^i est le prix du call européen associé au sous-jacent i . Si le payoff était $h(S_T^1, S_T^2) = ([S_T^1 - S_T^2] - K)_+$, le prix d'un tel call serait

$$\begin{aligned} C(T, K) &= e^{-rT} \mathbb{E}_{\mathbb{Q}}((S_T^1 - S_T^2 - K)_+) \\ &= e^{-rT} \mathbb{E}_{\mathbb{Q}} \left(\int_{-\infty}^{\infty} \mathbf{1}(S_T^2 + K \leq x \leq S_T^1) dx \right) \\ &= e^{-rT} \int_{-\infty}^{\infty} \mathbb{Q}(K + S_T^2 \leq x) - \underbrace{\mathbb{Q}(S_T^2 + K \leq x, S_T^1 \leq x)}_{\mathbb{Q}(S_T^1 \leq x, S_T^2 \leq x+K)} dx, \end{aligned}$$

et là encore, si (S_T^1, S_T^2) a pour copule C (sous \mathbb{Q}), on peut réécrire $C(T, K)$

$$e^{-rT} \int_{-\infty}^{\infty} e^{rT} \frac{\partial \Gamma^2}{\partial K}(T, x - K) - C \left(e^{rT} \frac{\partial \Gamma^1}{\partial K}(T, x), e^{rT} \frac{\partial P^2}{\partial K}(T, x - K) \right) dx.$$

Par la suite, nous supposons que nous ne connaissons que les lois marginales, c'est-à-dire $\mathbf{X} \in \mathcal{F}(F_1, \dots, F_d)$, mais nous essayerons de savoir ce que cela signifie pour $h(\mathbf{X})$.

Remarque 6.20 (95) avait appelé l'étude des valeurs possibles de $h(X_1, X_2)$, quand seules les lois marginales sont spécifiées, de l'arithmétique probabiliste, en reliant les lois de $S = h(X_1, X_2)$ aux lois de X_1 et X_2 . En particulier, l'obtention de bornes pour la loi (la fonction de répartition) de $h(X_1, X_2)$ y est détaillé, pour des fonctions h particulières, comme la somme $+$, le minimum \wedge ou le maximum \vee .

6.6.1 Espérance de fonctions nonlinéaires

Rappelons tout d'abord que, peu importe la structure de dépendance entre X_1 et X_2 , deux variables d'espérance finie, par linéarité de l'espérance,

$$\mathbb{E}(h(X_1, X_2)) = h(\mathbb{E}(X_1), \mathbb{E}(X_2)) \text{ si } h \text{ est linéaire.}$$

De manière générale, on a le résultat suivant :

Proposition 6.11 Soit $h : \mathbb{R}^2 \rightarrow \mathbb{R}$ une fonction 2-croissante, c'est-à-dire

$$h(y_1, y_2) + h(x_1, x_2) - h(x_1, y_2) - h(y_1, x_2) \geq 0,$$

pour tout $x_1 \leq y_1$ et $x_2 \leq y_2$. Alors

$$\mathbb{E}(h(X_1^-, X_2^-)) \leq \mathbb{E}(h(X_1, X_2)) \leq \mathbb{E}(h(X_1^+, X_2^+)),$$

où respectivement (X_1^-, X_2^-) et (X_1^+, X_2^+) sont des versions anti-comonotones et comonotones de (X_1, X_2) .

La preuve de ce résultat se trouve dans (91). Cette proposition peut se généraliser en dimension plus grand, mais la notion de croissance n'est alors plus la d -croissance, mais la supermodularité,

Proposition 6.12 *Soit $\phi : \mathbb{R}^d \rightarrow \mathbb{R}$ une fonction supermodulaire, c'est-à-dire*

$$\begin{aligned} & \phi(\max\{x_1, y_1\}, \dots, \max\{x_d, y_d\}) \\ & + \phi(\min\{x_1, y_1\}, \dots, \min\{x_d, y_d\}) \\ & - \phi(x_1, \dots, x_d) - \phi(y_1, \dots, y_d) \geq 0, \end{aligned}$$

pour tout $\mathbf{x} = (x_1, \dots, x_d)$ et $\mathbf{y} = (y_1, \dots, y_d)$. Alors pour tout $\mathbf{X} = (X_1, \dots, X_d) \in \mathcal{F}(F_1, \dots, F_d)$,

$$\mathbb{E}(\phi(X_1, \dots, X_d)) \leq \mathbb{E}(\phi(X_1^+, \dots, X_d^+)).$$

où \mathbf{X}^+ est une version comonotone de \mathbf{X} .

Remarque 6.21 Si ϕ est suffisamment dérivable, la d croissance est équivalente à $\partial_{1,\dots,d}\phi$ positive partout, alors que la supermodularité est équivalente à $\partial_{i,j}\phi$ positive pour tout $i \neq j$.

Exemple 6.43 Le prime (pure) stop-loss d'un traité de réassurance s'écrit comme l'espérance d'une fonction supermodulaire dès lors que l'on somme des risques : en effet, $\phi(x_1, \dots, x_d) = (x_1 + \dots + x_d - k)_+$ est une fonction supermodulaire.

Exemple 6.44 Pour une assurance vie jointe sur n années, l'annuité s'écrit

$$a_{xy:\overline{n}} = \sum_{k=1}^n v^k \mathbb{P}(T_x > k \text{ et } T_y > k) = \sum_{k=1}^n v^k {}_k p_{xy},$$

où v est le facteur d'actualisation, et (T_x, T_y) les durées de vie résiduelles des deux assurés. Alors

$$a_{xy:\overline{n}}^- \leq a_{xy:\overline{n}} \leq a_{xy:\overline{n}}^+,$$

où

$$\begin{aligned} a_{xy:\overline{n}}^- &= \sum_{k=1}^n v^k \max\{{}_k p_x + {}_k p_y - 1, 0\} \text{ (cas anti-comonotone),} \\ a_{xy:\overline{n}}^+ &= \sum_{k=1}^n v^k \min\{{}_k p_x, {}_k p_y\} \text{ (cas comonotone).} \end{aligned}$$

Exemple 6.45 Dans le cas d'une assurance au dernier survivant, sur n années, l'annuité s'écrit

$$a_{\overline{xy}:\overline{n}} = \sum_{k=1}^n v^k \mathbb{P}(T_x > k \text{ or } T_y > k) = \sum_{k=1}^n v^k {}_k p_{\overline{xy}},$$

où ${}_k p_{\overline{xy}} = \mathbb{P}(T_x > k \text{ or } T_y > k) = {}_k p_x + {}_k p_y - {}_k p_{xy}$. Alors

$$a_{\overline{xy}:\overline{n}}^- \leq a_{\overline{xy}:\overline{n}} \leq a_{\overline{xy}:\overline{n}}^+,$$

où

$$a_{\overline{xy}:\overline{n}}^- = \sum_{k=1}^n v^k (1 - \min\{{}_k q_x, {}_k q_y\}) \quad (\text{cas comonotone}),$$

$$a_{\overline{xy}:\overline{n}}^+ = \sum_{k=1}^n v^k (1 - \max\{{}_k q_x + {}_k q_y - 1, 0\}) \quad (\text{cas anti-comonotone}).$$

Exemple 6.46 L'annuité d'une pension de veuvage s'écrit

$$a_{x|y} = a_y - a_{xy} = \sum_{k=1}^{\infty} v^k {}_k p_y - \sum_{k=1}^{\infty} v^k {}_k p_{xy}.$$

Aussi,

$$a_{x|y}^- \leq a_{x|y} \leq a_{x|y}^+,$$

où

$$a_{x|y}^- = a_y - a_{xy} = \sum_{k=1}^{\infty} v^k {}_k p_y - \sum_{k=1}^{\infty} v^k \min\{{}_k p_x, {}_k p_y\}. \quad (\text{cas comonotone}),$$

$$a_{x|y}^+ = a_y - a_{xy} = \sum_{k=1}^{\infty} v^k {}_k p_y - \sum_{k=1}^{\infty} v^k \max\{{}_k p_x + {}_k p_y - 1, 0\}. \quad (\text{cas anti-comonotone}).$$

Sur l'importance de la comonotonie comme majorant de certaines quantités en actuariat, on pourra consulter (32), (31) et (93).

6.6.2 Comparer des sommes de risques

Dans un certain nombre de cas, s'il existe un préordre $\preceq_{\mathbb{R}^d}$ tel que $\mathbf{X} \preceq_{\mathbb{R}^d} \mathbf{Y}$, alors, il existe un autre préordre $\preceq_{\mathbb{R}}$ tel que

$$X_1 + \cdots + X_d \preceq_{\mathbb{R}} Y_1 + \cdots + Y_d.$$

Mais il faut commencer par définir des relations d'ordre pour les vecteurs aléatoires de \mathbb{R}^d . L'ordre simple à utiliser en dimension 2 est probablement l'ordre de corrélation :

Définition 6.37 Soient $\mathbf{X} = (X_1, X_2)$ et $\mathbf{Y} = (Y_1, Y_2)$, dans la même classe de Fréchet $\mathcal{F}(F_1, F_2)$, de lois marginales continues, alors \mathbf{X} sera moins corrélé que \mathbf{Y} , noté $\mathbf{X} \preceq_{\text{corr}} \mathbf{Y}$ si l'une des condition suivante est vérifiée

- pour toutes fonctions croissantes φ_1 et φ_2 telles que les covariances existe,

$$\text{Corr}(\varphi_1(X_1), \varphi_2(X_2)) \leq \text{Corr}(\varphi_1(Y_1), \varphi_2(Y_2))$$

- pour tout $\mathbf{z} \in \mathbb{R}^2$, $F_X(\mathbf{z}) \leq F_Y(\mathbf{z})$
- pour tout $\mathbf{u} \in [0, 1]^2$, $C_X(\mathbf{u}) \leq C_Y(\mathbf{u})$

En dimension 2, on a le résultat suivant :

Proposition 6.13 *Soient $\mathbf{X} = (X_1, X_2), \mathbf{Y} = (Y_1, Y_2) \in \mathcal{F}(F_1, F_2)$. si $\mathbf{X} \preceq_{corr} \mathbf{Y}$, alors pour toute fonction $h : \mathbb{R}^2 \rightarrow \mathbb{R}$ 2-croissante, $h(X_1, X_2) \preceq_{TVaR} h(Y_1, Y_2)$, en utilisant l'ordre induit par les TVaR dans le chapitre 3.*

En dimension plus grande, un résultat similaire peut-être obtenu en introduisant l'ordre supermodulaire.

6.6.3 Mesures de risques pour la somme de risques

En dimension $d = 2$, si des versions comonotones et anticomontones permettent d'obtenir des bornes à certains quantités, cela n'est en général pas le cas pour une mesure de risque quelconque, \mathcal{R} , c'est-à-dire

$$\mathcal{R}^- \leq \mathcal{R}(X_1^- + X_2^-) \not\leq \mathcal{R}(X_1 + X_2) \not\leq \mathcal{R}(X_1^+ + X_2^+) \leq \mathcal{R}^+,$$

où la borne supérieure \mathcal{R}^+ peut excéder le cas comonotone, par exemple.

Dans le cas où \mathcal{R} désigne la Value-at-Risk (telle que définie dans le chapitre 3) pour un seuil $q \in (0, 1)$, rappelons que

$$\mathcal{R}(X_1 + X_2) = \text{VaR}[X_1 + X_2; q] = F_{X_1 + X_2}^{-1}(q) = \inf\{x \in \mathbb{R} | F_{X_1 + X_2}(x) \geq q\}.$$

Exemple 6.47 Si $X_1 \sim \mathcal{E}(\alpha)$ et $Y_1 \sim \mathcal{E}(\beta)$, de telle sorte que $\mathbb{P}(X_2 > x_1) = \exp(-x_1/\alpha)$ et $\mathbb{P}(X_2 > x_2) = \exp(-x_2/\beta)$ pour tout $x_1, x_2 \in \mathbb{R}^+$, on a les inégalités

$$\exp(-x/\max\{\alpha, \beta\}) \leq \mathbb{P}[X_1 + X_2 > x] \leq \exp(-(x - \xi)_+/(\alpha + \beta))$$

sont alors valides pour tout $x \in \mathbb{R}^+$, quelle que soit la dépendance entre X_1 et X_2 , où

$$\xi = (\alpha + \beta) \log(\alpha + \beta) - \alpha \log \alpha - \beta \log \beta.$$

De plus, on a

$$-\max\{\alpha, \beta\} \log(1 - q) \leq \text{VaR}[X_1 + X_2; q] \leq \xi - (\alpha + \beta) \log(1 - q)$$

pour tout niveau $q \in (0, 1)$.

Si $\alpha = \beta = 1$, rappelons que sous hypothèse d'indépendance $X_1 + X_2 \sim \mathcal{G}(2, 1)$ alors que sous hypothèse de comonotonie, $X_1 + X_2 \sim \mathcal{E}(2)$.

Exemple 6.48 De manière générale, quelles que soient les lois F_1 et F_2 , il est possible (au moins numériquement) de calculer les bornes inférieures et supérieures. La Figure 6.14 montre la Value-at-Risk pour la somme de deux risques Gaussien, alors que la Figure 6.15 montre la Value-at-Risk pour la somme de deux risques suivant des lois Gamma.

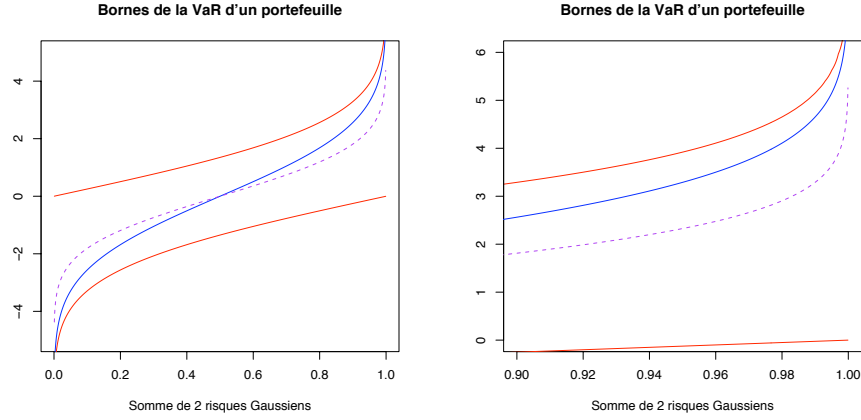


Figure 6.14 – Value-at-Risk pour la somme de 2 variables gaussiennes $\mathcal{N}(0, 1)$, avec le cas indépendant en pointillé, et le cas comonotone en trait plein. Les courbes en bas et en haut étant les bornes inférieures et supérieures. Le graphique de droite correspond à un agrandissement pour les quantiles excédant le niveau 90%.

Dans un cadre général, et plus théorique, (85) étudiait les bornes possibles pour la Value-at-Risk (ou plutôt la loi jointe) de $h(X_1, X_2)$ où $h : \mathbb{R}^2 \rightarrow \mathbb{R}$ et $(X_1, X_2) \in \mathcal{F}(F_1, F_2)$, introduisant le concept de convolutions supremal et infimal,

$$F_{\sup}(F_1, F_2)(z) = \sup \{C(F_1(x_1), F_2(x_2)), h(x_1, x_2) = z\} \quad (6.23)$$

$$F_{\inf}(F_1, F_2)(z) = \inf \{C(F_1(x_1), F_2(x_2)), h(x_1, x_2) = z\} \quad (6.24)$$

(95) a proposé des algorithmes numériques pour calculer ces bornes. Dans le cas de la somme, l'idée est de noter que la distribution des bornes correspond à la distribution de S_{\min} and S_{\max} , où

$$\mathbb{P}(S_{\max} < s) = \sup_{x \in \mathbb{R}} \max\{\mathbb{P}(X_1 < x) + \mathbb{P}(X_2 < s - x) - 1, 0\}$$

et

$$\mathbb{P}(S_{\min} \leq s) = \inf_{x \in \mathbb{R}} \min\{\mathbb{P}(X_1 \leq x) + \mathbb{P}(X_2 \leq s - x), 1\}.$$

On obtient alors le résultat suivant :

Proposition 6.14 Soit $\mathbf{X} = (X_1, X_2) \in \mathcal{F}(F_1, F_2)$ alors pour tout $s \in \mathbb{R}$,

$$\tau_{C-}(F_1, F_2)(s) \leq \mathbb{P}(X_1 + X_2 \leq s) \leq \rho_{C-}(F_1, F_2)(s),$$

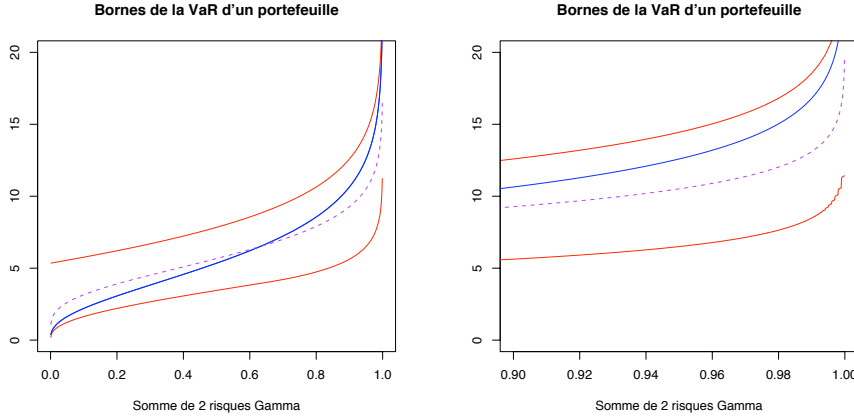


Figure 6.15 – Value-at-Risk pour la somme de 2 variables gaussiennes $\mathcal{G}(3, 1)$, avec le cas indépendant en pointillé, et le cas comontone en trait plein. Les courbes en bas et en haut étant les bornes inférieures et supérieures. Le graphique de droite correspond à un agrandissement pour les quantiles excédant le niveau 90%.

où

$$\tau_C(F_1, F_2)(s) = \sup_{x_1, x_2 \in \mathbb{R}} \{C(F_1(x_1), F_2(x_2)), x_1 + x_2 = s\}$$

et si $\tilde{C}(u_1, u_2) = u_1 + u_2 - C(u_1, u_2)$,

$$\rho_C(F_1, F_2)(s) = \inf_{x_1, x_2 \in \mathbb{R}} \{\tilde{C}(F_1(x_1), F_2(x_2)), x_1 + x_2 = s\}.$$

Ces bornes, obtenues en dimension $d = 2$ peuvent être étendues en dimension supérieures, mais elles ne sont pas nécessairement atteintes,

$$\text{VaR}(X_1 + \dots + X_d, \alpha) \leq G^{-1}(\alpha)$$

où

$$G(s) = \sup_{(x_1, \dots, x_{d-1})} \{C^-(F_1(x_1) + \dots + F_{d-1}(x_{d-1}) + F_d(s - x_1 - \dots - x_{d-1}))\}.$$

Remarque 6.22 (36) ont suggéré d'abord le problème comme un programme d'optimisation, sous contrainte, et la résolution du programme dual donne

$$\text{VaR}(X_1 + \dots + X_d, \alpha) \leq H^{-1}(\alpha)$$

où

$$H(s) = 1 - d \inf_{r \in [0, s/d)} \left\{ \frac{1}{s - rd} \int_r^{s - (d-1)r} [1 - F(x)] dx \right\},$$

dans le cas particulier où les variables X_1, \dots, X_d ont la même loi (notée F).

Dans le cas de trois risques lognormaux, (36) obtient les bornes de la Table 6.4. Pour rappel, les accords de Bâle pour les banques suggèrent de retenir la somme des Value-at-Risk, ce qui revient certes à supposer les risques comonotones, mais ne constitue pas vraiment un montant prudent.

Table 6.4 – Borne supérieure de $\text{VaR}(X_1 + X_2 + X_3; \alpha)$ pour des sommes de variables $LN(-0.2, 1)$

| α | indépendance | comonotonie | borne duale | borne |
|----------|--------------|-------------|----------------|-------|
| 90,0% | 7.54 | 8.85 | 14.44 | 15.38 |
| 95,0% | 9.71 | 12.73 | 19.50 | 20.63 |
| 99,0% | 16.06 | 25.16 | 35.31 | 37.03 |
| 99,9% | 29.78 | 53.99 | 69.98 | 73.81 |

6.7 Inférence statistique

Dans cette dernière section, nous allons évoquer brièvement l'utilisation des copules en statistique, sous l'hypothèse que l'on dispose d'un échantillon \mathbf{X}^i i.i.d.

6.7.1 Méthodes paramétriques

Considérons un vecteur aléatoire \mathbf{X} , absolument continu, telle que la densité jointe s'écrive

$$f(x_1, \dots, x_d) = c(F_1(x_1), \dots, F_d(x_d)) \cdot \prod_{i=1}^d f_i(x_i)$$

où $c : [0, 1]^d \rightarrow \mathbb{R}^+$ est la densité de la copule associée à \mathbf{X} , et où f_i est la densité de la variable X_i .

La log-vraisemblance $\log \mathcal{L}$ associée à un échantillon $\mathbf{X}^1, \dots, \mathbf{X}^n$ i.i.d. s'écrit

$$\log \mathcal{L}_n = \sum_{k=1}^n \log c(F_1(x_1^k), \dots, F_d(x_d^k)) + \sum_{i=1}^d \sum_{k=1}^n \log f_i(x_i^k)$$

et peut se décomposer en deux termes : celui de gauche est associé à la structure de dépendance, et le second aux lois marginales. Notons que le second terme est le seul qui apparaît si l'on suppose que les composantes du vecteur \mathbf{X} sont indépendantes.

On supposera que la copule C appartient à une famille paramétrique $\mathcal{C} = \{C_{\boldsymbol{\theta}}, \boldsymbol{\theta} \in \boldsymbol{\Theta}\}$ et que les lois marginales sont également dans des familles paramétriques, $F_i \in \mathcal{F}_i = \{F_{\alpha_i}, \alpha_i \in A_i\}$.

Sous les conditions usuelles de régularités, l'estimateur du maximum de vraisemblance $(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\alpha}})$, solution de

$$(\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\alpha}}) = \operatorname{argmax}\{\log \mathcal{L}_n(\boldsymbol{\theta}, \boldsymbol{\alpha})\},$$

est consistant et asymptotiquement Gaussien, au sens où

$$\sqrt{n} \left((\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\alpha}}) - (\boldsymbol{\theta}, \boldsymbol{\alpha}) \right) \xrightarrow{\mathcal{L}} \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}),$$

avec

$$\boldsymbol{\Sigma} = - \left(\lim_{n \rightarrow \infty} \frac{\partial^2 \log \mathcal{L}_n}{\partial(\boldsymbol{\theta}, \boldsymbol{\alpha}) \partial(\boldsymbol{\theta}, \boldsymbol{\alpha})'} \Big|_{(\boldsymbol{\theta}, \boldsymbol{\alpha})} \right)^{-1}.$$

6.7.2 Méthodes semi-paramétriques

Ici, seule la copule est paramétrique, et l'on utilise les fonction de répartition empiriques des lois marginales pour estimer le paramètre de la copule. En l'occurrence,

$$\hat{\boldsymbol{\theta}} = \operatorname{argmin} \left\{ \sum_{k=1}^n \log c(\hat{F}_1(x_1^k), \dots, \hat{F}_d(x_d^k)) \right\}.$$

Des propriétés de normalités asymptotiques de cet estimateur peuvent être obtenues.

6.7.3 Méthodes non-paramétriques d'estimation de copule

(15) a proposé une revue de la littérature sur l'estimation non-paramétrique de densités de copules, en insistant sur l'estimation à noyau. Mais l'estimation à noyau étant biaisé (multiplicativement) sur les bords, il peut être intéressant de l'adapter pour obtenir un estimateur sans biais partout sur $[0, 1]^d$. Par la suite, on se limitera au cas $d = 2$ pour la simplicité de l'exposé. On suppose disposer d'observations $\mathbf{U}^i = (U_1^i, U_2^i)$, i.i.d., distribuées suivant C , de densité c .

Une première piste est de transformer les variables, en considérant $(X_1^i, X_2^i) = (G^{-1}(U_1^i), G^{-1}(U_2^i))$, où G est une fonction strictement croissante $\mathbb{R} \rightarrow [0, 1]$, tel que le couple (X_1, X_2) admette une densité. Pour tout $(x_1, x_2) \in \mathbb{R}^2$, posons

$$\hat{f}(x_1, x_2) = \frac{1}{nh^2} \sum_{i=1}^n K\left(\frac{x_1 - X_1^i}{h}\right) K\left(\frac{x_2 - X_2^i}{h}\right),$$

correspondant à l'estimateur à noyau usuel, en dimension 2. Or comme on peut écrire, en notant g la dérivée de G :

$$f(x_1, x_2) = g(x_1)g(x_2)c[G(x_1), G(x_2)], \quad (6.25)$$

on obtient

$$c(u_1, u_2) = \frac{f(G^{-1}(u_1), G^{-1}(u_2))}{g(G^{-1}(u_1))g(G^{-1}(u_2))}, \text{ pour } (u_1, u_2) \in [0, 1] \times [0, 1], \quad (6.26)$$

ce qui donne, en substituant \hat{f} dans (6.26),

$$\begin{aligned} \hat{c}(u_1, u_2) &= \frac{1}{nh \cdot g(G^{-1}(u_1)) \cdot g(G^{-1}(u_2))} \\ &\quad \sum_{i=1}^n K\left(\frac{G^{-1}(u_1) - G^{-1}(U_1^i)}{h}, \frac{G^{-1}(u_2) - G^{-1}(U_2^i)}{h}\right) \end{aligned}$$

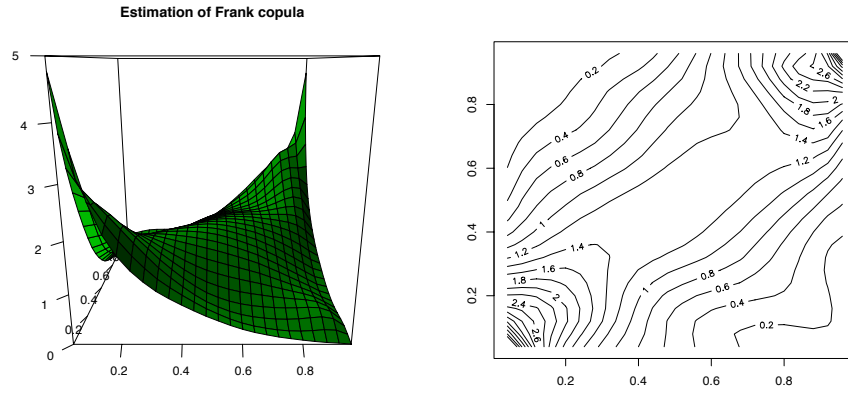


Figure 6.16 – Estimation d’une densité de copule de Frank, à partir de $n = 250$ simulation d’une copule de Frank, à l’aide d’une tranformation Gaussienne ($G = \Phi$), et un noyau Gaussien bivarié.

Un autre estimateur classique est l’estimateur par noyau Beta de la densité de la copule au point (u_1, u_2) , obtenu à l’aide de produits de noyaux Beta,

$$\hat{c}(u_1, u_2) = \frac{1}{n} \sum_{i=1}^n K\left(X_i, \frac{u_1}{b} + 1, \frac{1 - u_1}{b} + 1\right) \cdot K\left(Y_i, \frac{u_2}{b} + 1, \frac{1 - u_2}{b} + 1\right),$$

où $K(\cdot, \alpha, \beta)$ est la densité de la loi Beta de paramètres α et β .

Dans le cas Archimédien, nous avons noté dans la Remarque 6.11 qu’on pouvait caractériser une copule Archimédienne à l’aide de la fonction de Kendall K . L’estimateur nonparamétrique simple de cette fonction est

$$K_n(t) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}(Z_i \leq t)$$

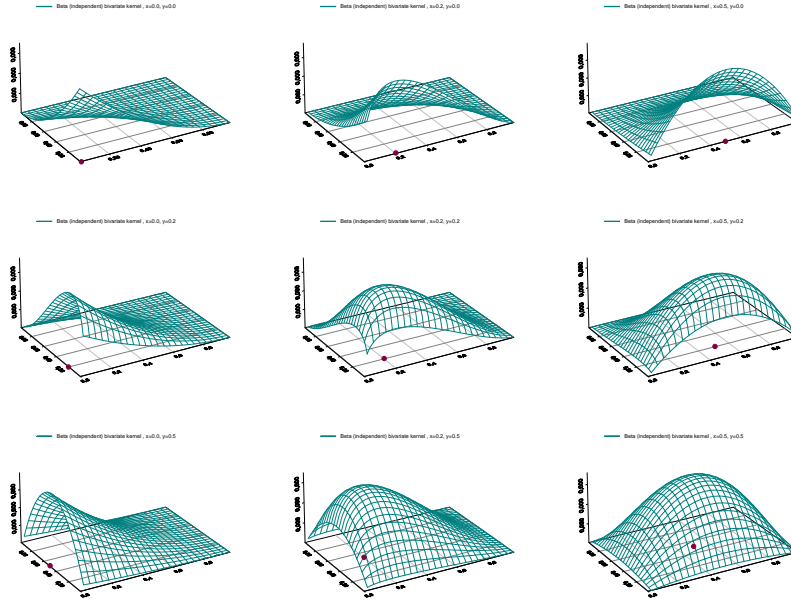


Figure 6.17 – Forme des noyaux Beta $K(\cdot, x/b + 1, (1 - x)/b + 1) \times K(\cdot, y/b + 1, (1 - y)/b + 1)$ pour $b = 0.2$, et différentes valeurs de x et y .

où

$$Z_i = \frac{1}{n-1} \sum_{j \neq i} \mathbf{1}(X_1^j < X_1^i, X_2^j < X_2^i).$$

L'estimateur du générateur associé est alors

$$\phi_n(t) = \exp \left(\int_{t_0}^t \frac{ds}{s - K_n(s)} \right).$$

6.7.4 Tests d'ajustement

Supposons que l'on cherche à tester une hypothèse de la forme $C \in \mathcal{C}$, où \mathcal{C} est une famille de copules (l'hypothèse alternative étant $C \notin \mathcal{C}$). Dans le cas univarié, on pense au test d'Anderson-Darling, ou à l'approche graphique du QQ-plot. Mais en dimension supérieure, c'est plus compliqué.

Si la famille \mathcal{C} est une famille paramétrique, (38) ou (45) ont suggéré d'utiliser le test de Cramér-von Mises, avec comme statistique

$$T = n \int_{[0,1]^d} [C_n(\mathbf{u}) - C_{\hat{\theta}}(\mathbf{u})]^2 dC_n(\mathbf{u}),$$

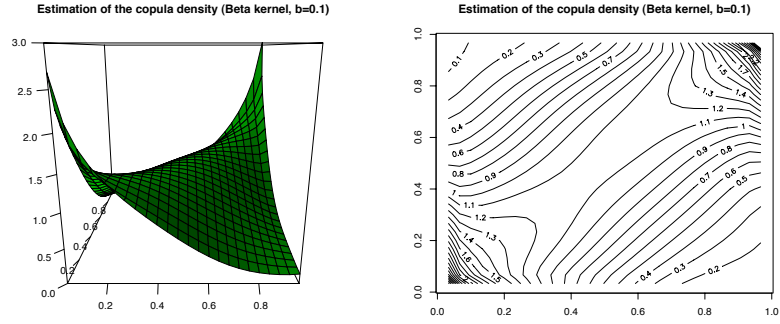


Figure 6.18 – Estimation de la densité de copule par noyaux Beta, $b = 0.1$ (simulation suivant une copule de Frank).

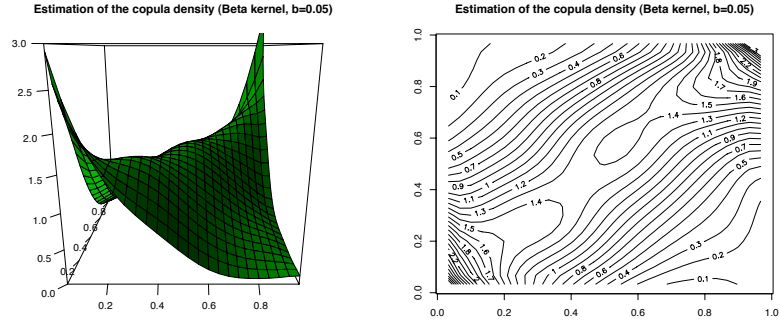


Figure 6.19 – Estimation de la densité de copule par noyaux Beta, $b = 0.05$

où C_n est la copule empirique, c'est-à-dire

$$C_n(u_1, \dots, u_d) = \frac{1}{n+1} \sum_{i=1}^n \mathbf{1}(U_1^i \leq u_1, \dots, U_d^i \leq u_d).$$

Une approche un peu plus simple est d'utiliser la fonction de Kendall $K(t) = \mathbb{P}(C(\mathbf{U}) \leq t)$. Si on suppose que cette fonction appartient à une famille paramétrique, alors la statistique précédente peut s'écrire

$$T = n \int_{[0,1]} \left[K_n(t) - K_{\hat{\theta}}(t) \right]^2 dK_n(t),$$

comme suggéré par (47).

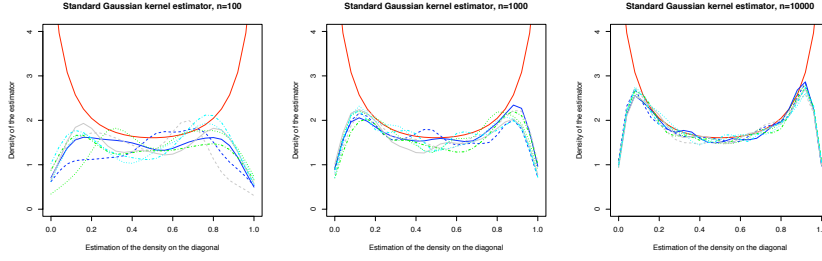


Figure 6.20 – Estimation de la densité sur la diagonale, par noyaux Gaussiens.

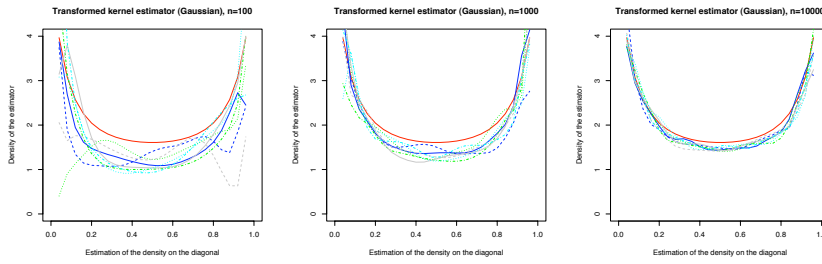


Figure 6.21 – Estimation de la densité sur la diagonale, transformation puis transformation inverse.

Une autre idée peut être de revenir à la transformation de (81). Supposons que \mathbf{U} ait pour copule C , alors \mathbf{V} défini par

$$\begin{cases} V_1 = U_1 \\ V_2 = C_{2|1}(U_2|U_1) \\ V_3 = C_{3|2,1}(U_3|U_2, U_1) \\ \dots \\ V_d = C_{d|d-1, \dots, 2,1}(U_d|U_{d-1}, \dots, U_2, U_1) \end{cases}$$

est un vecteur dont la loi est C^\perp (on utilise ici la méthode de simulation évoquée dans la section 6.3.5, à l'envers). Il suffit alors de faire des tests d'indépendance. Toutefois, il convient de faire plusieurs tests, en testant toutes les permutations possibles d'indices. (46) a adapté le test présenté auparavant dans ce cas, à l'aide de la statistique de Cramér-von Mises,

$$T = n \int_{[0,1]^d} [C_n(\mathbf{v}) - C^\perp(\mathbf{v})]^2 dC_n(\mathbf{v}),$$

où C_n est ici la copule empirique associée à \mathbf{V} .

Pour certaines familles de lois, il existe des tests spécifiques (en particulier

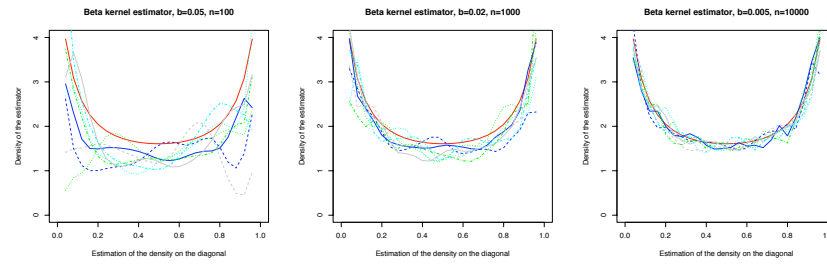


Figure 6.22 – Estimation de la densité sur la diagonale, par noyaux Beta.

des tests de normalité multivariés peuvent être utilisés pour tester l'ajustement d'une copule Gaussienne).

Bibliographie

- [1] J. Aczél. *Lectures on functional equations and their applications*. Academic Press, 1966.
- [2] C. Alsina, M. J. Frank, and B. Schweizer. *Associative Functions : Triangular Norms And Copulas*. World Scientific Publishing Company, 2006.
- [3] C. Alsina, R.B. Nelsen, and B. Schweizer. On the characteristic function of a class of binary operations on distribution functions. *Statistic and Probability Letters*, 7 :85–89, 1993.
- [4] F. Anscombe. *Computing in statistical science through APL*. Springer Verlag, 1981.
- [5] R.E. Barlow and F. Spizzichino. Schur-concave survival functions and survival analysis. *Journal of Computational and Applied Mathematics*, 46 :437–447, 1993.
- [6] B. Bassan and F. Spizzichino. Stochastic comparison for residual lifetimes and Bayesian notions of multivariate aging. *Advances in Applied Probability*, 31 :1078–1094, 1993.
- [7] B. Bassan and F. Spizzichino. Relations among univariate aging, bivariate aging and dependence for exchangeable lifetimes. *Journal of Multivariate Analysis*, 93(2) :313 – 339, 2005.
- [8] N. Bäuerle and A. Müller. Modelling and comparing dependencies in multivariate risk portfolios. *Astin Bulletin*, 28 :59–76, 1998.
- [9] T. Bayes. An essay towards solving a problem in the doctrine of chances. *Philosophical Transactions of the Royal Society of London*, 54, 1763.
- [10] N. Blomqvist. On a measure of dependence between two random variables. *The Annals of Mathematical Statistics*, 21 :593–600, 1950.
- [11] G. Boole. *The Laws of Thought*. Dover, New York, 1954.
- [12] N.L. Bowers, H.U. Gerber, J.C. Hickman, D.A. Jones, and C.J. Nesbitt. *Actuarial mathematics*. The Society of Actuaries, 1997.

- [13] D.T. Breeden and R.H. Litzenberger. Prices of State-Contingent Claims Implicit in Option Prices. *The Journal of Business*, 51 :621–651, 1978.
- [14] A. Charpentier. Dynamic dependence ordering for archimedean copulas and distorted copulas. *Kybernetika*, 44 :777–794, 2008.
- [15] A. Charpentier, J.-D. Fermanian, and O. Scaillet. The Estimation of Copulas : Theory and Practice. In J. Rank, editor, *Copulas : From theory to application in finance*, pages 35–62. Risk Books, 2007.
- [16] A. Charpentier and J. Segers. Lower tail dependence for archimedean copulas : Characterizations and pitfalls. *Insurance : Mathematics and Economics*, 40(3) :525 – 532, 2007.
- [17] A. Charpentier and J. Segers. Convergence of archimedean copulas. *Statistics & Probability Letters*, 78(4) :412 – 419, 2008.
- [18] A. Charpentier and J. Segers. Tails of multivariate archimedean copulas. *Journal of Multivariate Analysis*, 100(7) :1521 – 1537, 2009.
- [19] Y.S. Chow and H. Teicher. *Probability theory : independence, interchangeability, martingales*. Springer Verlag, 1997.
- [20] D. G. Clayton. A model for association in bivariate life tables and its application in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika*, 65(1) :141–151, 1978.
- [21] J. Coles, S. Heffernan and J.A. Tawn. Dependence measures for multivariate extremes. *Extremes*, 2 :339–365, 1999.
- [22] R.D. Cook and M.E. Johnson. A family of distributions for modelling non-elliptically symmetric multivariate data. *Journal of the Royal Statistical Society. Series B (Methodological)*, 43(43) :210–218, 1981.
- [23] KMV Corporation. Modelling default risk. *KMV Corporation Technical document*, 1997.
- [24] H. Cossette, T. Duchesne, and E. Marceau. Modeling catastrophes and their impact on insurance portfolios. *North American Actuarial Journal*, 7 :1–22, 2003.
- [25] C. M. Cuadras and J. Augé. A continuous general multivariate distribution and its properties. *Communications in Statistics-Theory and Methods*, 10(4) :339–353, 1981.
- [26] W.F. Darsow, B. Nguyen, and E.T. Olsen. Sur les fonctions absolument monotones. *Illinois Journal of Mathematics*, 36 :600–642, 1992.
- [27] H.A. David and M.L. Moeschberger. *Theory of Competing Risks*. Griffin’s Statistical Monographs, 1978.

- [28] B. de Finetti. La prévision : ses lois logiques, ses sources subjectives. *Annales de l'Institut Henri Poincaré*, 7 :1–68, 1937.
- [29] P. Deheuvels. La fonction de dépendance empirique et ses propriétés . *Académie Royale de Belgique (Bulletin de la Classe des Sciences)*, 65 :274–292, 1979.
- [30] L. Devroye. *Non-Uniform Random Variate Generation*. Springer, 1986.
- [31] J. Dhaene, M. Denuit, M.J. Goovaerts, R. Kaas, and D. Vyncke. The concept of comonotonicity in actuarial science and finance : applications. *Insurance : Mathematics and Economics*, 31 :133–161, 2002.
- [32] J. Dhaene, M. Denuit, M.J. Goovaerts, R. Kaas, and D. Vyncke. The concept of comonotonicity in actuarial science and finance : theory. *Insurance : Mathematics and Economics*, 31 :3–33, 2002.
- [33] D. Drouot-Mari and S. Kotz. *Correlation and dependence*. World Scientific Publishing Company Inc., 2001.
- [34] F. Durante, J. Quesada-Molina, and C. Sempi. A generalization of the archimedean class of bivariate copulas. *Annals of the Institute of Statistical Mathematics*, 59 :487–498, 2007.
- [35] P. Embrechts, A. J. McNeil, and D. Straumann. Correlation and Dependence in Risk Management : Properties and Pitfalls. In M. Dempster, editor, *Risk Management : Value at Risk and Beyond*, pages 176–223. Cambridge University Press, 2002.
- [36] P. Embrechts and G. Puccetti. Bounds for the sum of dependent risks having overlapping marginals. *Journal of Multivariate Analysis*, 101 :177–190, 2009.
- [37] K.-T. Fang, S. Kotz, and K.-W. Ng. *Symmetric Multivariate and Related Distributions*. Chapman & Hall/CRC, 1989.
- [38] J.-D. Fermanian. Goodness of fit tests for copulas. *Journal of Multivariate Analysis*, 95(1) :119–152, 2005.
- [39] N. I. Fisher. Copulas. *Encyclopedia of Statistical Sciences*, 1 :159–163, 1997.
- [40] M. J. Frank. On the simultaneous associativity of $F(x, y)$ and $x+y-F(x, y)$. *Aequationes Mathematicae*, 19 :194–226, 1979.
- [41] M. Fréchet. Remarques au sujet de la note précédente. *Comptes Rendus de l'Académie des Sciences de Paris*, 249 :2719–2720, 1958.
- [42] W.E. Frees and E.A. Valdez. Understanding relationships using copulas. *North American Actuarial Journal*, 2 :1–25, 1998.

- [43] C. Genest and R. J. MacKay. Copules archimédiennes et familles de lois bidimensionnelles dont les marges sont données. *The Canadian Journal of Statistics*, 14 :145–159, 1986.
- [44] C. Genest, J.J. Quesada Molina, and J.A. Rodríguez Lallena. De l'impossibilité de construire des lois à marges multidimensionnelles données à partir de copules. *Comptes rendus de l'Académie des sciences de Paris*, 320 :723–726, 1995.
- [45] C. Genest and B. Rémillard. Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models. *Annales de l'Institut Henri Poincaré : Probabilités et statistiques*, 44 :1096–1127, 2008.
- [46] C. Genest, B. Rémillard, and D. Beaudoin. Goodness-of-fit tests for copulas : A review and a power study. *Insurance : Mathematics and Economics*, 44 :199–213, 2009.
- [47] C. Genest and L.-P. Rivest. Statistical Inference Procedures for Bivariate Archimedean Copulas. *Journal of the American Statistical Association*, 88(423) :1034–1043, 1993.
- [48] CreditMetrics Group. Creditmetrics technical document. *JP Morgan*, 1997.
- [49] E. J. Gumbel. Bivariate exponential distributions. *Journal of the American Statistical Association*, 55 :698–707, 1960.
- [50] R.A. Harris. Reliability applications of a bivariate exponential distribution. *Operations Research*, 16 :16–27, 1970.
- [51] E. Hewitt and L.J. Savage. Symmetric measures on Cartesian products. *Transactions of the American Mathematical Society*, 80 :470–501, 1955.
- [52] W. Hoeffding. Masstabinvariante Korrelationstheorie. *Schriften des Mathematischen Instituts und des Instituts für Angewandte Mathematik de Universität Berlin*, 5 :179–233, 1940.
- [53] P. Hougaard. Survival models for heterogeneous populations derived from stable distributions. *Biometrika*, 73(2) :387–396, 1986.
- [54] H. Joe. Families of min-stable multivariate exponential and multivariate extreme value distributions. *Statistical and Probability Letters*, 9 :75–81, 1990.
- [55] H. Joe. *Multivariate Models and Dependence Concepts*. Chapman & Hall/CRC, 1997.
- [56] G. Kimeldorf and A. Sampson. One-parameter families of bivariate distributions with fixed marginals. *Communications in Statistics*, 4 :293–301, 1975.

- [57] E.P. Klement, R. Mesiar, and E. Pap. Transformations of copulas. *Kybernetika*, 41 :425–434, 2005.
- [58] H.O Lancaster. The structure of bivariate distributions. *Annals of Mathematical Statistics*, 29 :719–736, 1958.
- [59] P.-S. Laplace. *Théorie Analytique des Probabilités*. Paris, Ve. Courcier, Paris, 1812.
- [60] A.W. Ledford and J.A. Tawn. Statistics for Near Independence in Multivariate Extreme Values. *Biometrika*, 83 :169–187, 1996.
- [61] C. Lefèvre and S. Utev. Mixed Poisson approximation in the collective epidemic model. *Stochastic Processes and their Applications*, 69 :217–246, 1997.
- [62] D.X. Li. On default correlation : a copula function approach. *RiskMetrics Group, Working Paper*, 1999.
- [63] S Loisel. A trivariate non-Gaussian copula having 2-dimensional Gaussian copulas as margins. 2009.
- [64] R.W. Madsen. Generalized binomial distribution. *Communications in statistics. Theory and methods*, 22 :3065–3086, 1993.
- [65] K. V. Mardia. Mardia’s test of multinormality. In S. Kotz & N. L. Johnson, editor, *Encyclopedia of statistical sciences (Vol. 5)*, pages 217–221. Wiley, 1985.
- [66] A. W. Marshall and I. Olkin. A generalized bivariate exponential distribution. *Journal of Applied Probability*, 4 :291–302, 1967.
- [67] A. J. McNeil and J. Nešlehová. Multivariate Archimedean copulas, d -monotone functions and l_1 -norm symmetric distributions. *The Annals of Statistics*, 37(5b) :3059–3097, 2009.
- [68] M. Merton. On the pricing of corporate debt : the risk structure of interest rates. *Journal of Finance*, 28 :449–470, 1974.
- [69] Patricia Mariela Morillas. A method to obtain new copulas from a given one. *Metrika*, 61 :169–184, 2005. 10.1007/s001840400330.
- [70] P. Muliere and M. Scarsini. Characterization of Marshall-Olkin type class of distributions. *Annals of the Institute of Statistical Mathematics*, 39 :429–441, 1987.
- [71] A. Müller and M. Scarsini. Archimedean copulae and positive dependence. *Journal of Multivariate Analysis*, 93 :434–445, 2004.
- [72] R. B. Nelsen. Dependence Modeling with Archimedean Copulas. [http ://www.lclark.edu/ mathsci/brazil2.pdf](http://www.lclark.edu/mathsci/brazil2.pdf) (2010-03-16), 2005.

- [73] R. B. Nelsen. *An Introduction to Copulas*. Springer, 2007.
- [74] D. Oakes. Bivariate survival models induced by frailties. *Journal of the American Statistical Association*, 84,406 :487–493, 1989.
- [75] I. Olkin and A.W. Marshall. *Inequalities : Theory of Majorization and Its Applications*. Academic Press, New York, 1979.
- [76] J. Pickands. Multivariate extreme value distributions. Preprint, University of Pennsylvania, 1976.
- [77] J. Pickands. Multivariate extreme value distributions. *M. Bull. Internat. Statist. Inst. Proc. 43d. Session (Buenos Aires)*, pages 859–878, 1981.
- [78] J. Pickands. Multivariate negative exponential and extreme value distributions. In J. Häslér and R.D. Reiss, editors, *Extreme Value Theory*. Springer Verlag, 1989.
- [79] A. Rényi. On measures of dependence. *Acta Mathematica Hungarica*, 10 :441–451, 1959.
- [80] J.P. Romano and A.F. Siegel. *Counterexamples in Probability and Statistics*. Wadsworth and Brooks, 1996.
- [81] M. Rosenblatt. Remarks on a Multivariate Transformation. *The Annals of Mathematical Statistics*, 23(3) :470–472, 1952.
- [82] M. Scarsini. On measures of concordance. *Stochastica*, 8(3) :201–218, 1984.
- [83] H. Schlesinger. Decomposing catastrophic risk. *Insurance : Mathematics and Economics*, 24 :95–101, 1999.
- [84] V. Schmitz. Revealing the dependence structure between $x(1)$ and $x(n)$. *Journal of Statistical Planning and Inference*, 123(1) :41 – 47, 2004.
- [85] B. Schweizer and A. Sklar. *Probabilistic Metric Spaces*. North-Holland, New York, 1983.
- [86] B. Schweizer and E. F. Wolff. Sur une mesure de dépendance pour les variables aléatoires. *Comptes Rendus de l'Académie des Sciences de Paris*, 283 :659–661, 1976.
- [87] B. Schweizer and E. F. Wolff. On nonparametric measures of dependence for random variables. *The Annals of Statistics*, 9 :879–885, 1981.
- [88] A. Sklar. Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut de Statistique de l'Université de Paris*, 8 :229–231, 1959.
- [89] A. Sklar. Random variables, distribution functions, and copulas—a personal look backward and forward. *Distributions with Fixed Marginals and Related Topics*, 28 :1–14, 1996.

- [90] F. Spizzichino. *Subjective probability models for lifetimes*. Chapman & Hall/CRC, 2001.
- [91] A.H. Tchen. Inequalities for distributions with given marginals. *Annals of Probability*, 8 :814–827, 1980.
- [92] G. Venter. Tails of copulas. *Proceedings of the Casualty Actuarial Society*, 2002.
- [93] D. Vyncke. Comonotonicity : the perfect dependence, 2003. PhD thesis.
- [94] T. Wilde. Creditrisk+, a credit risk management framework. *Credit Suisse First Boston*, 1999.
- [95] R.C. Williamson. Probabilistic arithmetic, 1989. PhD thesis.
- [96] E.F. Wolff. n -dimensional measures of dependence,. *Stochastica*, 4 :175–188, 1980.