

Les fonctions copules. Définition, exemples et propriétés.

I. Définition et Théorème de Sklar

(Ω, \mathcal{A}, P) est un espace de probabilité. Les variables aléatoires y sont définies et sont à valeurs réelles ($+\infty$ et $-\infty$ sont des valeurs exclues)

1) Rappel sur les fonctions de répartition des variables aléatoires

Si X est une variable aléatoire, sa fonction de répartition est la fonction F définie sur \mathbb{R} par $F(x) = P\{X \leq x\}$. F est croissante et continue à droite. De plus $\lim_{x \rightarrow -\infty} F(x) = 0$ et $\lim_{x \rightarrow +\infty} F(x) = 1$.

On prolonge F à $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$ par $F(-\infty) = 0$ et $F(+\infty) = 1$.

Si F est continue et strictement croissante sur \mathbb{R} , alors F admet un inverse F^{-1} défini sur $[0, 1]$ à valeurs dans $\overline{\mathbb{R}}$.

Dans les autres cas, on définit un pseudo-inverse par la formule :

$$F^{-1}(u) = \inf \{x : F(x) \geq u\}$$

pour $u \in [0, 1]$ et la convention $\inf \emptyset = +\infty$ et $\inf \mathbb{R} = \inf]-\infty, a] = -\infty$

Le pseudo-inverse possède les deux propriétés suivantes :

$$(i) \quad \{F(x) \leq u\} = \{x \leq F^{-1}(u)\} \quad \text{pour tout } u \in [0, 1] \text{ et } x \in \overline{\mathbb{R}}$$

$$(ii) \quad \forall x \in F(\overline{\mathbb{R}}) \quad F(F^{-1}(x)) = x.$$

On dit qu'une variable aléatoire X est continue si sa fonction de répartition l'est. Dans ce cas, $F(\overline{\mathbb{R}}) = [0, 1]$ et (ii) est vrai pour tout u .

Dans le cas d'un vecteur aléatoire (X_1, \dots, X_n) , on définit également sa fonction de répartition sur \mathbb{R}^n par

$$F(x_1, \dots, x_n) = P\{X_1 \leq x_1, \dots, X_n \leq x_n\}$$

et que l'on étend à $\overline{\mathbb{R}}^n$.

2) Définition et premières propriétés.

Prenons un vecteur aléatoire (X_1, \dots, X_n) et définissons

$$C: [0, 1]^n \rightarrow [0, 1]$$

$$(u_1, \dots, u_n) \mapsto F(F_1^{-1}(u_1), \dots, F_n^{-1}(u_n))$$

où F est la fonction de répartition du vecteur et les F_i sont celles des variables X_i , que l'on appelle les marginales du vecteur.

La fonction C est bien définie et possède les 3 propriétés suivantes:

$$\textcircled{1} \quad \forall (u_1, \dots, u_n) \in [0, 1]^n \quad C(u_1, \dots, u_{i-1}, 0, u_i, \dots, u_n) = 0$$

$$\textcircled{2} \quad \forall u_i \in F_i(\overline{\mathbb{R}}) \quad C(1, \dots, 1, u_i, 1, \dots, 1) = u_i$$

$$\textcircled{3} \quad \forall (u_1^{(1)}, \dots, u_n^{(1)}) \in [0, 1]^n$$

$$\forall (u_1^{(2)}, \dots, u_n^{(2)}) \in [0, 1]^n \quad \sum_{(i_1, \dots, i_n) \in \{1, 2\}^n} (-1)^{\sum_{j=1}^n i_j} C(u_1^{(i_1)}, u_2^{(i_2)}, \dots, u_n^{(i_n)}) \geq 0$$

avec $u_j^{(1)} \leq u_j^{(2)} \quad \forall j$

démonstration: $\textcircled{1}$ vient du fait que $F(x_1, \dots, x_{i-1}, -\infty, x_{i+1}, \dots, x_n) = 0$
pour $(x_1, \dots, x_n) \in \overline{\mathbb{R}}^n$.

$\textcircled{2}$ vient du fait que

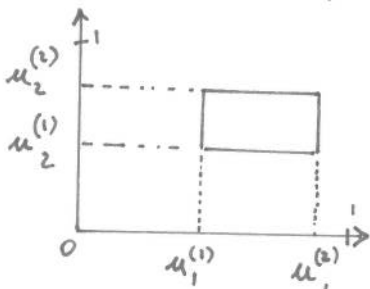
$$C(1, \dots, 1, u_i, 1, \dots, 1) = F(F_1^{-1}(1), F_2^{-1}(1), \dots, F_i^{-1}(u_i), F_{i+1}^{-1}(1), \dots, F_n^{-1}(1))$$

$$= F_i(F_i^{-1}(u_i)) = u_i \quad \text{si } u_i \in F_i(\overline{\mathbb{R}}).$$

$\textcircled{3}$ Dans le cas $n=2$, $\textcircled{3}$ signifie que pour

$$\begin{cases} 0 \leq u_1^{(1)} \leq u_1^{(2)} \leq 1 \\ 0 \leq u_2^{(1)} \leq u_2^{(2)} \leq 1 \end{cases},$$

$$C(u_2^{(2)}, u_2^{(2)}) - C(u_1^{(2)}, u_2^{(2)}) - C(u_1^{(1)}, u_2^{(2)}) + C(u_1^{(1)}, u_2^{(1)}) \geq 0$$



ce qui se réécrit: $0 \leq P\{F_1^{-1}(u_1^{(1)}) < X_1 \leq F_1^{-1}(u_1^{(2)}), F_2^{-1}(u_2^{(1)}) < X_2 \leq F_2^{-1}(u_2^{(2)})\}$

ce qui est vrai. Plus généralement dans le cas n , $\textcircled{3}$ exprime que

$$P\{F_1^{-1}(u_1^{(1)}) < X_1 \leq F_1^{-1}(u_1^{(2)}), \dots, F_n^{-1}(u_n^{(1)}) < X_n \leq F_n^{-1}(u_n^{(2)})\} \geq 0$$

Définition des fonctions copules:

On appelle fonction copule C (ou fonction copule d'ordre n) toute fonction définie sur $[0, 1]^n$ à valeurs dans $[0, 1]$ telle que

$$\textcircled{1} \quad \forall (u_1, \dots, u_n) \in [0, 1]^n \quad C(u_1, \dots, u_{i-1}, 0, u_{i+1}, \dots, u_n) = 0$$

$$\textcircled{2} \quad \forall (u_1, \dots, u_n) \in [0, 1]^n \quad C(1, \dots, 1, u_i, 1, \dots, 1) = u_i$$

$$\textcircled{3} \quad \forall (u_1^{(1)}, \dots, u_n^{(1)}) \in [0, 1]^n \text{ et } \forall (u_1^{(2)}, \dots, u_n^{(2)}) \in [0, 1]^n \quad u_j^{(1)} \leq u_j^{(2)}$$

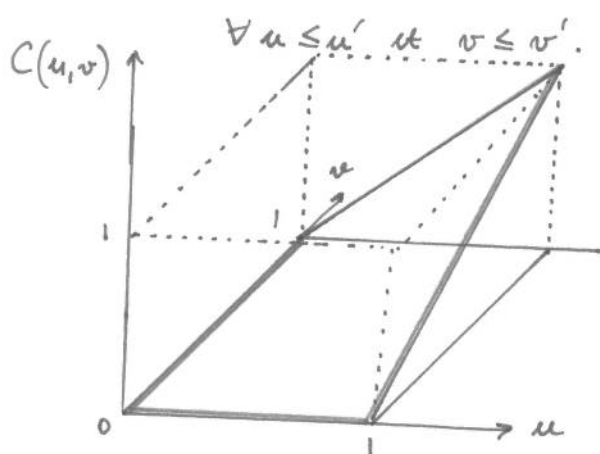
$$\sum_{(i_1, \dots, i_n) \in \{1, 2\}^n} (-1)^{\sum_{j=1}^n i_j} C(u_1^{(i_1)}, \dots, u_n^{(i_n)}) \geq 0.$$

Dans le cas $n = 2$, C est définie sur le carré unité $[0, 1]^2$

avec les propriétés $C(0, v) = C(u, 0) = 0 \quad \forall (u, v) \in [0, 1]^2$

$$C(1, v) = v \text{ et } C(u, 1) = u \quad \forall (u, v) \in [0, 1]^2$$

$$C(u', v') - C(u, v') - C(u', v) + C(u, v) \geq 0$$



Premières propriétés:

(i) $\forall (u_1, \dots, u_n) \in [0, 1]^n$ les fonctions $u \mapsto C(u_1, \dots, u_{i-1}, u, u_{i+1}, \dots, u_n)$ sont croissantes.

(ii) Les fonctions copules sont lipschitziennes: $\forall (u, v) \in [0, 1]^n \quad |C(u) - C(v)| \leq \sum_{i=1}^n |u_i - v_i|$

démonstration: (i) il suffit de prendre $u \leq v$ et $u^{(1)} = (u_1, \dots, u_{i-1}, u, u_{i+1}, \dots, u_n)$

$$u^{(2)} = (u_1, \dots, u_{i-1}, v, u_{i+1}, \dots, u_n)$$

dans $\textcircled{3}$ pour avoir $C(u_1, \dots, u_{i-1}, u, u_{i+1}, \dots, u_n) \leq C(u_1, \dots, u_{i-1}, v, u_{i+1}, \dots, u_n)$.

(ii) (dans le cas $n=2$) L'inégalité du triangle:

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

Supposons que $u_2 \leq v_2$, alors ③ donne

$$C(1, v_2) - C(1, u_2) - C(u_1, v_2) + C(u_1, u_2) \geq 0$$

$$C(u_1, v_2) - C(u_1, u_2) \leq v_2 - u_2$$

dans le cas $u_2 \geq v_2$ nous avons $C(u_1, u_2) - C(u_1, v_2) \leq u_2 - v_2$

$$\text{donc } |C(u_1, u_2) - C(u_1, v_2)| \leq |u_2 - v_2|$$

De même, on obtient, $|C(u_1, v_2) - C(v_1, v_2)| \leq |u_1 - v_1|$.

3) Théorème de Sklar

Étant donné un vecteur aléatoire de fonction de distribution F et dont les composantes ont les fonctions de répartition F_i $i=1, \dots, n$, il existe un copule C tel que pour tout $x_1, \dots, x_n \in \overline{\mathbb{R}}$,

$$F(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n))$$

Le copule est unique dans le cas où les variables aléatoires sont continues.

Réciproquement, étant données des fonctions de répartition F_1, \dots, F_n et une fonction copule C , on peut construire un vecteur aléatoire (X_1, \dots, X_n) sur un espace de probabilité (Ω, \mathcal{A}, P) de telle sorte que les F_i soient les fonctions de répartition des X_i et que la fonction de répartition du vecteur (X_1, \dots, X_n) soit donnée par

$$F(x_1, \dots, x_n) = C(F_1(x_1), \dots, F_n(x_n)).$$

Ce théorème est le résultat essentiel de toute la théorie.

démonstration : ① Dans le cas où les variables aléatoires sont continues.

On rappelle qu'alors, $F_i(\bar{\mathbb{R}}) = [0, 1]$ pour chaque i .

Note discussion en 2) montre alors l'existence d'une fonction copule.

L'unicité résulte du fait que si pour deux copules C et C' nous avons

$$\forall x_1, \dots, x_n \in \bar{\mathbb{R}} \quad C(F_1(x_1), \dots, F_n(x_n)) = C'(F_1(x_1), \dots, F_n(x_n))$$

$$\text{comme } F_i(\bar{\mathbb{R}}) = [0, 1]^n, \quad \forall u_i \in [0, 1] \quad \exists x_i \in \bar{\mathbb{R}} \quad F_i(x_i) = u_i.$$

$$\text{donc } \forall u_1, \dots, u_n \in [0, 1] \quad C(u_1, \dots, u_n) = C'(u_1, \dots, u_n).$$

② Dans le cas où les variables aléatoires peuvent être discontinues

c'est à dire $F_i(\bar{\mathbb{R}})$ peut être strictement inclus dans $[0, 1]$

La même discussion en 2) montre l'existence et l'unicité d'une

fonction C qui vérifie les axiomes des fonctions copules sur $\prod_i F_i(\bar{\mathbb{R}})$.

Nous allons donc prolonger cette fonction à tout $[0, 1]^n$ de telle sorte que les axiomes soient vérifiés sur $[0, 1]^n$. En général, il n'y a pas unicité.

Comme C est lipschitzienne sur $\prod_i F_i(\bar{\mathbb{R}})$, C se prolonge naturellement

en une fonction lipschitzienne sur $\prod_i \overline{F_i(\bar{\mathbb{R}})}$. Le complémentaire

dans $[0, 1]$ de chaque $\overline{F_i(\bar{\mathbb{R}})}$ est un ouvert qui s'écrit donc comme réunion dénombrable d'intervalles ouverts deux à deux disjoints.

Nous devons donc prolonger C sur chacun de ces parcs $\prod_{i=1}^n]a_i^{(k)}, b_i^{(k)}[$

pour $k=1, \dots$, il suffit de le faire sur le premier $\prod_{i=1}^n]a_i, b_i[$.

On montre que le prolongement par interpolation linéaire convient, c'est à dire:

$$\text{pour } x \in \prod_{i=1}^n]a_i, b_i[\quad x_i = \lambda_i a_i + (1-\lambda_i) b_i \quad \text{pour } \lambda_i \in]0, 1[$$

$$\text{on pose } C(x_1, \dots, x_n) = \lambda_1 \dots \lambda_n C(a_1, \dots, a_n) + (1-\lambda_1) \lambda_2 \dots \lambda_n C(b_1, a_2, \dots, a_n) \\ + \dots + (1-\lambda_1)(1-\lambda_2) \dots (1-\lambda_n) C(b_1, \dots, b_n).$$

il reste à montrer que ② et ③ sont vrais pour C . C'est facile pour ②, plus pénible pour ③.

③ Il suffit de montrer que la fonction d'ensemble μ définie sur les pavés $\prod_i]x_i^{(1)}, x_i^{(2)}]$ par

$$\mu\left(\prod_{i=1}^n]x_i^{(1)}, x_i^{(2)}]\right) = \sum_{(i_1, \dots, i_n) \in \{1, 2\}^n} (-1)^{\sum_{j=1}^n i_j} G(F_{i_1}(x_{i_1}^{(1)}), \dots, F_{i_n}(x_{i_n}^{(2)}))$$

peut bien s'étendre à la tribu borélienne de \mathbb{R}^n en une mesure de probabilité (Théorème de Carathéodory)

et que les marginales ont les fonctions de répartition F_i .

Comme dans le cas de la construction de la mesure de Lebesgue sur \mathbb{R} , la compacité de $[0, 1]^n$ est cruciale.

II Exemples et propriétés.

Copule d'indépendance $\Pi(u_1, \dots, u_n) = u_1 \cdots u_n.$

On peut vérifier qu'elle vérifie bien les axiomes de la définition.

On peut aussi utiliser la caractérisation suivante :

Proposition : Une fonction C définie sur $[0, 1]^n$ est une fonction copule si et seulement si elle est la restriction à $[0, 1]^n$ de la fonction de répartition d'un vecteur aléatoire (U_1, \dots, U_n) où les U_i ont la distribution uniforme sur $[0, 1]$.

démonstration : \Leftarrow La fonction de répartition du vecteur (U_1, \dots, U_n) notée C s'écrit de façon unique à l'aide d'une copule \tilde{C} :

$$C(u_1, \dots, u_n) = \tilde{C}(F_1(u_1), \dots, F_n(u_n)) \quad \text{pour tout } (u_1, \dots, u_n) \in [0, 1]^n.$$

comme $F_i(u_i) = u_i$, $C = \tilde{C}$ donc C est une copule.

\Rightarrow C'est la dernière partie du théorème de Sklar avec C et des $F_i(x_i) = x_i$.

Copule de comonotonie (ou borne supérieure de Fréchet) $M(u_1, \dots, u_n) = \min(u_1, \dots, u_n)$

Pour vérifier que c'est bien une copule on peut utiliser la proposition ci-dessus. M est la fonction de répartition du vecteur (U, \dots, U) où U est uniforme.

$$P\{U \leq u_1, \dots, U \leq u_n\} = P\{U \leq \min(u_1, \dots, u_n)\} = \min(u_1, \dots, u_n).$$

De même, si X est une v.a. continue alors la copule de (X, \dots, X) est M .

Plus généralement si f_1, \dots, f_n sont des fonctions croissantes strictement définies sur le support de X , alors les $f_i(X)$ sont des v.a. continues et leur copule est aussi M (d'où le nom de comonotonie).

$$P\{f_1(X) \leq x_1, \dots, f_n(X) \leq x_n\} = P\{X \leq f_1^{-1}(x_1), \dots, X \leq f_n^{-1}(x_n)\}$$
$$= \min(F_{f_1^{-1}(x_1)}, \dots, F_{f_n^{-1}(x_n)}) = \min\{F_{f_1(X)}(x_1), \dots, F_{f_n(X)}(x_n)\}$$

où $F_{f_i(X)}$ est la fonction de répartition de $f_i(X)$.

Plus généralement, nous avons:

Proposition: Si (X_1, \dots, X_n) est un vecteur de variables aléatoires continues
 On note C son copule. Si f_1, \dots, f_n sont des fonctions strictement
 croissantes définies sur les supports de X_1, \dots, X_n respectivement, alors
 C est aussi la copule du vecteur de variables aléatoires continues
 $(f_1(X_1), \dots, f_n(X_n))$.

Copule gaussienne de matrice de corrélation R :

Il s'agit de la copule d'un vecteur gaussien de matrice de corrélation R
 de moyennes nulles et de variances 1. D'après la proposition précédente,
 il s'agit de la même copule si nous étions partis d'un vecteur gaussien
 de corrélation R et de moyennes et variances quelconques.

Si on note $\Phi(x) = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt$

et $\Phi_R^{(n)}(x_1, \dots, x_n) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_n} \frac{1}{(2\pi)^{n/2} \sqrt{\det R}} e^{-\frac{1}{2} y^T R^{-1} y} dy_1 \dots dy_n$

$C_{R, \text{gaussien}}^{(n)}(u_1, \dots, u_n) = \Phi_R^{(n)}(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_n))$

Copule de Student de matrice de corrélation R à d degrés de liberté.

$C_{R, d, \text{student}}^{(n)}(u_1, \dots, u_n) = T_{R, d}^{(n)}(T_d^{-1}(u_1), \dots, T_d^{-1}(u_n))$

avec $T_d(x) = \int_{-\infty}^x \frac{\Gamma(\frac{d+1}{2})}{\Gamma(\frac{d}{2})} \frac{1}{\sqrt{d\pi}} \frac{dt}{(1+t^2/d)^{\frac{d+1}{2}}}$

$T_{R, d}^{(n)}(x_1, \dots, x_n) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_n} \frac{\Gamma(\frac{d+n}{2})}{\Gamma(\frac{d}{2})} \frac{1}{(d\pi)^{n/2} \sqrt{\det R}} \frac{dt_1 \dots dt_n}{(1 + \frac{t^T R t}{d})^{\frac{d+n}{2}}}$

Copules archimédiennes: Soit Λ une loi portée sur \mathbb{R}_+ et φ sa transformée de Laplace $\varphi(t) = \int_0^\infty e^{-tx} \Lambda(dx)$ φ est décroissante strictement $\varphi(0)=1$ $\varphi(+\infty)=0$.

alors $C(u_1, \dots, u_n) = \varphi(\varphi^{-1}(u_1) + \dots + \varphi^{-1}(u_n))$ est une fonction copule.

démonstration: On va construire un vecteur aléatoire (X_1, \dots, X_n)

admettant C pour copule.

On part d'une variable aléatoire Y de loi Λ puis conditionnellement à Y on construit n variables aléatoires indépendantes de fonction de répartition $P\{X_i \leq x\} = e^{-Y\varphi^{-1}(x)}$

$$\begin{aligned} P\{X_1 \leq x_1, \dots, X_n \leq x_n\} &= E P\{X_1 \leq x_1, \dots, X_n \leq x_n \mid Y\} \\ &= E \prod_{i=1}^n e^{-Y\varphi^{-1}(x_i)} = E e^{-Y \sum_{i=1}^n \varphi^{-1}(x_i)} \\ &= \varphi(\varphi^{-1}(x_1) + \dots + \varphi^{-1}(x_n)) \end{aligned}$$

D'autre part les X_i sont uniformes: $P\{X_i \leq x\} = E P\{X_i \leq x \mid Y\}$

$$= E e^{-Y\varphi^{-1}(x)} = \varphi\varphi^{-1}(x) = x$$

exemple: famille de Gumbel-Hougaard avec $\varphi = e^{-x}$, $C = \exp(-[(\ln u_1)^p + \dots + (\ln u_n)^p]^{1/p})$

Proposition: Les Bornes de Fréchet. Pour toute copule C on a

$$\forall (u_1, \dots, u_n) \quad \left(\sum_{i=1}^n u_i - n + 1\right)^+ \leq C(u) \leq M(u) = \min(u_1, \dots, u_n)$$

$W = \left(\sum_{i=1}^n u_i - n + 1\right)^+$ est une fonction copule ssi $n=2$.

démonstration: Soit un vecteur de variables aléatoires uniformes de fonction de répartition C .

$$C(u_1, \dots, u_n) = P\{U_1 \leq u_1, \dots, U_n \leq u_n\} \leq P\{U_i \leq u_i\} = u_i \text{ pour chaque } i$$

donc a fortiori, $C(u_1, \dots, u_n) \leq \min(u_1, \dots, u_n)$.

pour l'autre inégalité:

$$\begin{aligned} C(u_1, \dots, u_n) &= P\{U_1 \leq u_1, \dots, U_n \leq u_n\} = 1 - P\{U_1 > u_1 \text{ ou } U_2 > u_2 \text{ ou } \dots \text{ ou } U_n > u_n\} \\ &\geq 1 - \sum_{i=1}^n P\{U_i > u_i\} = 1 - \sum_{i=1}^n (1 - P\{U_i \leq u_i\}) = \sum_{i=1}^n u_i - n + 1 \end{aligned}$$

d'autre part $C(u_1, \dots, u_n) \geq 0$ donc $C(u_1, \dots, u_n) \geq \left(\sum_{i=1}^n u_i - n + 1\right)^+$.

Quand $n=2$, $C(u, v) = (u+v-1)^+$ est la fonction de distribution de $(U, 1-U)$

Quand $n > 2$, $W\left(\left[\frac{1}{2}, 1\right]^n\right) = 1 - \frac{n}{2} < 0$ donc W ne peut pas être la fonction de répartition d'une

On peut définir un ordre partiel sur les copules $C < C'$ si $C(u_1, \dots, u_n) < C'(u_1, \dots, u_n)$ pour tout $(u_1, \dots, u_n) \in [0, 1]^n$.

Le copule de comonotonie est donc la borne supérieure des copules pour cette relation d'ordre.

La fonction W est la borne inférieure.

On peut en effet montrer que pour tout $(u_1, \dots, u_n) = u \in [0, 1]^n$ il existe une copule C dépendant de u telle que $C(u) = W(u)$ à ce point.

Il existe des familles de copules totalement ordonnées pour cette ordre.

Par exemple, en dimension 2, $C_p^{\text{gaussien}} < C_{p'}^{\text{gaussien}}$ dès que $p < p'$.

Définition : copule extrême. une fonction copule C est extrême si pour tout $r > 0$

$$C(u_1, \dots, u_n) = C(u_1^{1/r}, \dots, u_n^{1/r})^r.$$

exemple : la famille de Gumbel-Hougaard $C_\theta(u_1, \dots, u_n) = \exp\left(-\left(\sum_{i=1}^n (-\log u_i)^\theta\right)^{1/\theta}\right)$ pour $\theta > 0$. C'est une copule archimédienne.

La raison pour laquelle on appelle copule extrême une telle fonction copule est la suivante: prenons une famille indépendante de vecteurs $(X_i^{(k)}, \dots, X_n^{(k)})$ pour $k=1, \dots, p$. ces vecteurs ont même loi caractérisé par des marginals F_1, \dots, F_n et une copule C

Intéressons nous au vecteur des marginals maximaux

$$(M_1, \dots, M_n) = \left(\max_{1 \leq k \leq p} X_1^{(k)}, \dots, \max_{1 \leq k \leq p} X_n^{(k)} \right)$$

$$\begin{aligned} \text{La loi de } M_i \text{ est donnée par } P\{M_i \leq m\} &= P\left\{ \max_{1 \leq k \leq p} X_i^{(k)} \leq m \right\} \\ &= P\left\{ X_i^{(k)} \leq m \text{ pour } 1 \leq k \leq p \right\} \\ &= P\left\{ X_i^{(k)} \leq m \right\}^p = F_i(m)^p \end{aligned}$$

$$\text{donc } F_{M_i}(x) = F_i(x)^p.$$

Une copule du vecteur (M_1, \dots, M_n) satisfait

$$\begin{aligned} P\{M_1 \leq m_1, \dots, M_n \leq m_n\} &= P\left\{ X_i^{(k)} \leq m_i \text{ pour } 1 \leq k \leq p, 1 \leq i \leq n \right\} \\ &= P\left\{ X_i^{(k)} \leq m_i \text{ pour } 1 \leq i \leq n \right\}^p \\ &= C(F_1(m_1), \dots, F_n(m_n))^p = C\left(F_{M_1}(m_1)^{1/p}, \dots, F_{M_n}(m_n)^{1/p}\right)^p \end{aligned}$$

Ainsi le copule de valeurs maximaux est $C(u_1^{1/p}, \dots, u_n^{1/p})^p$.

S'il existe une copule $C'(u_1, \dots, u_n)$ telle que $\lim_{p \rightarrow \infty} C(u_1^{1/p}, \dots, u_n^{1/p})^p = C'(u_1, \dots, u_n)$

alors C' est une copule extrême.

Propriété de convexité: Soient $0 \leq \lambda_i$ $i=1, \dots, p$ avec $\sum_{i=1}^p \lambda_i = 1$ et C_1, \dots, C_p des fonctions copules, alors $\sum_{i=1}^p \lambda_i C_i$ est une fonction copule.

Plus généralement si $(C_\alpha)_{\alpha \in \mathbb{D}}$ est une famille de copule et Λ est une mesure de probabilité sur \mathbb{D} alors $\int_{\mathbb{D}} C_\alpha \Lambda(d\alpha)$ est une copule.

Restrictions de copules: Soit $C(u_1, \dots, u_n)$ une copule d'ordre n .

soit $2 \leq k \leq n$ et $\{i_1, \dots, i_k\} \subset \{1, \dots, n\}$ k indices distincts alors

$C(u_{i_1}, \dots, u_{i_k})$ est une copule d'ordre k .

En revanche, on ne connaît pas de méthode générale pour construire des copules d'ordre plus élevé à partir de copules d'ordre inférieur.

En général, $C_1(u, C_2(v, w))$ et $C_1(C_2(u, v), C_3(w, x))$ [où C_1, C_2 et C_3 sont des copules d'ordre 2] ne sont pas des copules.

Partie absolument continue et singulière d'une copule.

Comme toute mesure de probabilité sur $[0, 1]^n$, une fonction copule peut être décomposée en somme d'une mesure absolument continue par rapport à la mesure de Lebesgue sur $[0, 1]^n$ et d'une mesure singulière portée par un ensemble de mesure 0 pour cette mesure de Lebesgue (Théorème de Lebesgue-Radon-Nikodym).

De plus, la partie absolument continue peut s'écrire comme intégrale de la dérivée de Radon-Nikodym qui est notée $\frac{\partial^n}{\partial u_1 \dots \partial u_n} C(u_1, \dots, u_n)$, c'est-à-dire:

$$C(u_1, \dots, u_n) = \int_0^{u_1} \dots \int_0^{u_n} \frac{\partial^n}{\partial u_1 \dots \partial u_n} C(t_1, \dots, t_n) dt_1 \dots dt_n + S(u_1, \dots, u_n).$$

où S est la fonction de répartition de la mesure singulière. Cette mesure singulière n'a pas d'atomes puisque C n'en a pas et donc S est continue.

Lorsque la fonction C est telle que la dérivée n-ème $\frac{\partial^n}{\partial u_1 \dots \partial u_n} C$ existe partout.

sur $[0,1]^n$ et γ est intégrable, alors $S \equiv 0$ et cette dérivée (classique) et la densité de Radon-Nikodym sont égaux.

Dans ce cas, la définition de la fonction copule se simplifie en:

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

$$(2) \quad C(1, \dots, 1, u_i, 1, \dots, 1) = u_i$$

$$(3) \quad \frac{\partial^n}{\partial u_1 \dots \partial u_n} C(u_1, \dots, u_n) \geq 0 \quad \text{sur } [0,1]^n$$

Exemple de copule sans partie singulière: $\Pi = u_1 \dots u_n$

Exemple de copule sans partie absolument continue: $M = \min(u_1, \dots, u_n)$

La mesure de cette copule est portée par $\Delta = \{u_1 = u_2 = \dots = u_n\}$ qui est de mesure nulle pour Lebesgue sur $[0,1]^n$. D'autre part, $\frac{\partial^n}{\partial u_1 \dots \partial u_n} M = 0$ excepté sur $\{u_1 = u_2 = \dots = u_n\}$ où la dérivée n'existe pas.

Exemple de copule avec partie absolument continue ET singulière: $\theta \Pi + (1-\theta)M$ avec $0 < \theta < 1$

Calcul de probabilités conditionnelles.

Proposition: Soit C une fonction copule pour tout i $\frac{\partial C}{\partial u_i}(u_1, \dots, u_n)$ existe presque partout

$$\text{et } \frac{\partial C}{\partial u_i}(u_1, \dots, u_n) = P\{U_1 \leq u_1, \dots, U_{i-1} \leq u_{i-1}, U_{i+1} \leq u_{i+1}, \dots, U_n \leq u_n \mid U_i = u_i\} \quad \text{p.p.}$$

où (U_1, \dots, U_n) a pour loi C .

Sous l'hypothèse que C est telle que $\frac{\partial^p}{\partial u_{i_1} \dots \partial u_{i_p}} C(u_1, \dots, u_n)$ existe partout

et est intégrable sur $[0,1]^n$, alors:

$$\frac{\partial^p}{\partial u_{i_1} \dots \partial u_{i_p}} C(u_1, \dots, u_n) = P\{U_j \leq u_j \text{ pour } j \notin \{i_1, \dots, i_p\} \mid U_{i_k} = u_{i_k} \text{ pour } k \in \{1, \dots, p\}\}$$

preuve: (1) D'après (3) de la définition de copules $u_i \mapsto C(u_1, \dots, u_n)$ est

Lipschitz, donc, absolument continue sur $[0,1]$ donc dérivable p.p. et

$$C(u_1, \dots, u_n) = \int_0^{u_i} \frac{\partial C}{\partial u_i}(u_1, \dots, u_{i-1}, t, u_{i+1}, \dots, u_n) dt$$

$$\text{soit: } E\{1_{\{U_i \leq u_i\}} 1_{\{U_2 \leq u_2\}} \dots 1_{\{U_n \leq u_n\}}\} = E\{1_{\{U_i \leq u_i\}} \frac{\partial C}{\partial u_i}(u_1, \dots, u_{i-1}, U_i, u_{i+1}, \dots, u_n)\}$$

ce qui est la définition de l'espérance conditionnelle.

② L'hypothèse de dérivabilité partout + intégrabilité permet aussi d'écrire

$$C(u_1, \dots, u_n) = \int_0^{u_1} \dots \int_0^{u_n} \frac{\partial^n}{\partial u_1 \dots \partial u_n} C(u) dt_1 \dots dt_n.$$

ce qui est aussi équivalent à la définition de l'espérance conditionnelle.

Simulation de variables aléatoires de marginales données et de copule donnée.

Soient F_1, \dots, F_n les fonctions de répartition des marginales et C la copule.

Supposons tout d'abord que nous avons simulé un vecteur (U_1, \dots, U_n) de variables aléatoires uniformes sur $[0, 1]$ suivant la loi donnée par la copule C .

Définissons les inverses $F_i^{-1}(u) = \inf \{x : F_i(x) \geq u\}$, alors le vecteur $(F_1^{-1}(U_1), \dots, F_n^{-1}(U_n))$ a la bonne loi.

En effet: $F_i^{-1}(U_i)$ a pour fonction de répartition F_i :

$$P \{ F_i^{-1}(U_i) \leq x \} = P \{ U_i \leq F_i(x) \} = F_i(x)$$

D'autre part

$$\begin{aligned} P \{ F_1^{-1}(U_1) \leq x_1, \dots, F_n^{-1}(U_n) \leq x_n \} &= P \{ U_1 \leq F_1(x_1), \dots, U_n \leq F_n(x_n) \} \\ &= C(F_1(x_1), \dots, F_n(x_n)). \end{aligned}$$

Il reste à expliquer comment simuler (U_1, \dots, U_n) .

On commence par simuler $U_1 \sim \text{Unif}(0, 1)$. Puis, comme,

$$P \{ U_2 \leq u \mid U_1 \} = P \{ U_2 \leq u, U_3 \leq 1, \dots, U_n \leq 1 \mid U_1 \} = \frac{\partial C}{\partial u_1}(u, u, 1, \dots, 1)$$

on a donc la fonction de répartition conditionnelle et on peut utiliser son inverse pour simuler des valeurs de $U_2 \mid U_1$.

Ainsi de suite si la copule est suffisamment dérivable

$$P \{ U_3 \leq u \mid U_1 \text{ et } U_2 \} = \frac{\partial^2 C}{\partial u_1 \partial u_2}(U_1, U_2, u, 1, \dots, 1)$$

etc.

Applications de copules à la finance

I Calcul du prix d'options multi-sous-jacents.

1) Rappel du résultat de Breeden-Litzenberger

On suppose que les taux sont constants (pour simplifier mais ce n'est pas nécessaire) S est le prix d'un sous-jacent. Les prix des calls et des puts à l'instant 0 sont donnés par la règle d'évaluation risque neutre :

$$\text{Call}(T, K) = E \left\{ e^{-rT} (S_T - K)^+ \right\}$$

$$\text{Put}(T, K) = E \left\{ e^{-rT} (K - S_T)^+ \right\}.$$

Notons que $(S_T - K)^+ = \int_{\mathbb{R}} \mathbb{1}_{\{K \leq x \leq S_T\}} dx = \int_K^{\infty} \mathbb{1}_{\{S_T \geq x\}} dx$

d'où $\text{Call}(T, K) = e^{-rT} \int_K^{\infty} P\{S_T \geq x\} dx = e^{-rT} \int_K^{\infty} 1 - P\{S_T \leq x\} dx.$

et donc
$$F_S(x) = 1 - P\{S_T \leq x\} = -e^{rT} \frac{\partial \text{Call}(T, x)}{\partial K}$$

De même,
$$F_S(x) = e^{rT} \frac{\partial \text{Put}(T, x)}{\partial K}$$

En pratique, on se trouve souvent dans la situation suivante :

on observe suffisamment de prix d'options sur chacun des deux sous-jacents S^1 et S^2 et on souhaite calculer le prix d'une option dont le payoff h dépend des deux valeurs terminales S_T^1 et S_T^2 . Pour calculer le prix, on regarde

$$e^{-rT} E \left\{ h(S_T^1, S_T^2) \right\}.$$

L'approche par les copules est très naturelle puisqu'on connaît les lois marginales par le marché et la copule permet de construire la distribution jointe. Évidemment, on peut considérer de la même façon des options sur n sous-jacents.

2) Quelques exemples d'options sur 2 sous-jacents et comment exprimer leurs prix en terme de copule.

a) Options binaires: $h(S_T^1, S_T^2) = \mathbb{1}_{\{S_T^1 > K, S_T^2 > L\}}$

le prix p est donné par:

$$\begin{aligned} p &= e^{-rT} P\{S_T^1 > K, S_T^2 > L\} = e^{-rT} (1 - P\{S_T^1 \leq K \text{ ou } S_T^2 \leq L\}) \\ &= e^{-rT} (1 - P\{S_T^1 \leq K\} - P\{S_T^2 \leq L\} + P\{S_T^1 \leq K \text{ et } S_T^2 \leq L\}) \\ &= e^{-rT} - \frac{\partial \text{put}^1}{\partial K}(T, K) - \frac{\partial \text{put}^2}{\partial K}(T, L) + e^{-rT} C\left(\frac{\partial \text{put}^1}{\partial K}(T, K)e^{rT}, e^{rT} \frac{\partial \text{put}^2}{\partial K}(T, L)\right) \end{aligned}$$

b) Options sur le maximum: $h(S_T^1, S_T^2) = (\max(S_T^1, S_T^2) - K)^+$

$$\begin{aligned} p &= e^{-rT} E\left\{(\max(S_T^1, S_T^2) - K)^+\right\} = e^{-rT} E \int_{\mathbb{R}} \mathbb{1}_{\{K \leq x \leq \max(S_T^1, S_T^2)\}} dx \\ &= e^{-rT} E \int_K^\infty 1 - \mathbb{1}_{\{\max(S_T^1, S_T^2) \leq x\}} dx \\ &= e^{-rT} \int_K^\infty 1 - P\{S_T^1 \leq x, S_T^2 \leq x\} dx \\ &= e^{-rT} \int_K^\infty 1 - C\left(e^{rT} \frac{\partial \text{put}^1}{\partial K}(T, x), e^{rT} \frac{\partial \text{put}^2}{\partial K}(T, x)\right) dx \end{aligned}$$

c) Options Spreads: $h(S_T^1, S_T^2) = (S_T^2 - S_T^1 - K)^+$

$$\begin{aligned} p &= e^{-rT} E\left\{(S_T^2 - S_T^1 - K)^+\right\} = e^{-rT} E \int_{\mathbb{R}} \mathbb{1}_{\{K + S_T^1 \leq x \leq S_T^2\}} dx \\ &= e^{-rT} E \int_{\mathbb{R}} \mathbb{1}_{\{K + S_T^1 \leq x\}} - \mathbb{1}_{\{K + S_T^1 \leq x \text{ et } S_T^2 \leq x\}} dx \\ &= e^{-rT} \int_{\mathbb{R}} e^{rT} \frac{\partial \text{put}^1}{\partial K}(T, x - K) - C\left(e^{rT} \frac{\partial \text{put}^1}{\partial K}(T, x - K), e^{rT} \frac{\partial \text{put}^2}{\partial K}(T, x)\right) dx \\ &= \int_{\mathbb{R}} \frac{\partial \text{put}^1}{\partial K}(T, x - K) - e^{-rT} C\left(e^{rT} \frac{\partial \text{put}^1}{\partial K}(T, x - K), e^{rT} \frac{\partial \text{put}^2}{\partial K}(T, x)\right) dx \end{aligned}$$

3) Quelle copule choisir? Bornes d'arbitrage sur les prix - Sous/Sur Réplication.

Pour calculer le prix et la couverture de ces options multi sous-jacents il convient de connaître la copule C . Nous avons assez peu d'information comparé à la "taille" de l'ensemble des copules possibles.

a) On peut estimer une copule "historique" à partir de observations simultanées passés de S^1 et S^2 . C'est l'approche statistique.

Il y a plusieurs méthodes paramétrique, non paramétrique, etc. Malheureusement, les copules historique et risque-neutre ne sont pas nécessairement identiques.

b) On peut avoir une opinion a priori et trader en suivant cette opinion.
par exemple: copule gaussien avec corrélation $\rho = 0,5$.

c) Les représentations des prix avec les copules donnent des bornes inférieure et supérieure sur les prix ainsi que les stratégies de sous et sur réplication correspondante. C'est tout ce que l'on peut dire si on ne connaît rien sur C .

par exemple: option spread. on sait que tout copule C satisfait à $(u+v-1)^+ \leq C(u,v) \leq \min(u,v)$ donc le prix est encadré par:

$$\inf_C e^{-rT} E \left\{ (S_T^2 - S_T^1 - K)^+ \right\} = \int_{\mathbb{R}} \left(\frac{\partial \text{put}^1}{\partial K}(T,x) - \frac{\partial \text{put}^2}{\partial K}(T,x+K) \right)^+ dx$$

$$\sup_C e^{-rT} E \left\{ (S_T^2 - S_T^1 - K)^+ \right\} = \int_{\mathbb{R}} \frac{\partial \text{put}^1}{\partial K}(T,x) - \left(\frac{\partial \text{put}^1}{\partial K}(T,x) + \frac{\partial \text{put}^2}{\partial K}(T,x+K) - \overrightarrow{E} \right)^+ dx$$

La stratégie statistique de sous réplication est donnée par la formule suivante $\left\{ x : \frac{\partial \text{put}^1}{\partial K}(T,x) > \frac{\partial \text{put}^2}{\partial K}(T,x+K) \right\} = \bigcup_i]a_i, b_i[$ [disjoints

$$\text{borne inf} = \sum_i \text{put}^1(T, b_i) - \text{put}^2(T, K+b_i) - \text{put}^1(T, a_i) + \text{put}^1(T, K+a_i)$$

d) Ces représentations permettent aussi de comprendre comment le prix est affecté par "la corrélation".

Par exemple, pour l'option spread, pour les mêmes marginales, si $C < C'$ (code partiel des coupons), alors $P_{C'} < P_C$.

Si on choisit de calculer le prix avec une copule gaussienne on a $P_{C'} < P_C$ si $\rho < \rho'$ et ce quelque soient les lois marginales.

II Application au calcul des prix de tranches de CDO.

Nous allons nous intéresser à deux produits dérivés de crédit :

le Credit Default Swap (CDS) sur indice et le Collateralized Debt Obligation (CDO).

1) Notations: Aujourd'hui est l'instant 0. On a un échéancier

$$0 = T_0 < T_1 < \dots < T_M. \quad \text{On suppose } T_i - T_{i-1} = \delta$$

Nous considérons un panier de n obligations de valeur nominale $\frac{1}{n}$ chacune. Ces obligations peuvent faire défaut et le détenteur de l'obligation i subit alors une perte que l'on écrit sous la forme $\frac{1}{n}(1 - R_i)$ où R_i est appelé le taux de recouvrement (il peut être aléatoire et différent pour chaque obligation).

N_t = nombre d'obligations ayant ^{déjà} fait défaut à l'instant t

L_t = montant cumulé des pertes du portefeuille à l'instant t .

Si les taux de recouvrement R_i sont déterministes et égaux alors :

$$L_t = (1 - R) \frac{1}{n} N_t.$$

L'instant aléatoire où l'obligation i fait défaut est noté T_i .

2) Valorisation d'un CDS sur indice. (indice = panier de n obligations)

a) L'acheteur de protection paye aux dates T_0, \dots, T_{M-1} une prime S proportionnelle au nominal des obligations qui n'ont pas encore fait défaut à chacune de ces dates.

La valorisation risque-neutre (qui est à justifier mais que nous ne ferons pas)

donne:

$$V_0^{\text{prime}} = \sum_{i=0}^{M-1} S e^{-rT_i} \delta E \left\{ 1 - \frac{N_{T_i}}{n} \right\}$$

b) En échange, il reçoit du vendeur de protection, aux instants de défaut qui ont lieu avant T_M le montant des pertes. Par définition,

$$L_t = \sum_{j=1}^n (1-R_j) \frac{1}{n} \mathbb{1}_{\{\tau_j \leq t\}}$$

La valorisation risque-neutre donne:

$$\begin{aligned} V_0^{\text{défauts}} &= \sum_{j=1}^n \frac{1}{n} E \left\{ (1-R_j) e^{-r\tau_j} \mathbb{1}_{\{\tau_j \leq t\}} \right\} \\ &= E \int_0^{T_M} e^{-rt} dL_t \end{aligned}$$

en intégrant par parties:

$$V_0^{\text{défauts}} = E \left\{ e^{-rT_M} L_{T_M} \right\} + \int_0^{T_M} r e^{-rt} E \left\{ L_t \right\} dt$$

c) La prime S est calculée en faisant $V_0^{\text{défauts}} = V_0^{\text{prime}}$.
Pour ce faire, il suffit de connaître $E\{L_t\}$ et $E\{N_t\}$ $0 \leq t \leq T_M$.

3) Valorisation d'une tranche de CDO.

Le fonctionnement est à peu près le même sauf que la protection ne porte que sur une partie des pertes du portefeuille (= la tranche) c'est à dire l'acheteur est protégé contre les pertes telles que le montant cumulé est supérieur à \underline{K} (attachment point) et inférieur à \bar{K} (dettachment point). En échange, il paie une prime à T_0, \dots, T_{M-1} proportionnelle au nominal de la tranche moins les pertes (c'est un peu différent du CDS sur indice).

Le nominal de la tranche est $\bar{K} - \underline{K}$. Les pertes cumulées de la tranche sont: $(L_t - \underline{K})^+ - (L_t - \bar{K})^+ = \int_0^t \mathbb{1}_{\{\underline{K} \leq L_t \leq \bar{K}\}} dL_t$.

d'où la valorisation risque neutre de la prime:

$$V_0^{\text{prime}} = \sum_{i=0}^{M-1} S e^{-rT_i} \delta \left(\bar{K} - \underline{K} - E \left[(L_{T_i} - \underline{K})^+ \right] + E \left[(L_{T_i} - \bar{K})^+ \right] \right)$$

La valorisation risque-neutre donne

$$V_0^{\text{defauts}} = E \int_0^{T_M} e^{-rt} d((L_t - \underline{K})^+ - (L_t - \bar{K})^+)$$

$$V_0^{\text{defauts}} = e^{-rT_M} E \left\{ (L_{T_M} - \underline{K})^+ - (L_{T_M} - \bar{K})^+ \right\} + \int_0^{T_M} re^{-rt} E \left\{ (L_t - \underline{K})^+ - (L_t - \bar{K})^+ \right\} dt$$

Dans ce cas, nous avons besoin de connaître la loi de L_t pour pouvoir calculer la prime de la tranche. Rappelons que

$$L_t = \sum_{j=1}^n (1 - R_j) \frac{1}{n} \mathbb{1}_{\{\tau_j \leq t\}}$$

pour connaître la loi de L_t , nous avons besoin de connaître la loi jointe des τ_j . Nous supposons que $R_j = R$ déterministe et que les τ_j ont

même loi exponentielle de paramètre λ . Cette dernière hypothèse est justifiée par le fait que la valeur des tranches dépend peu de cette loi.

Ainsi: $E\{L_t\} = (1-R)(1 - e^{-\lambda t})$ et λ peut être calibré à partir de la prime du CDS sur indice puisque celle-ci ne dépend que des $E\{L_t\}$.

Les marginaux sont connus ($\text{Exp}(\lambda)$), il reste à définir une copule.

Le marché utilise une copule Gaussienne de matrice de corrélation

$$\begin{bmatrix} 1 & \rho & \rho & \rho \\ \rho & 1 & \rho & \rho \\ \rho & \rho & 1 & \rho \\ \rho & \rho & \rho & 1 \end{bmatrix} \quad \rho \geq 0$$

Il reste ensuite à faire des simulations pour calculer $E\{(L_t - K)^+\}$.

Note: dans la littérature du risque de crédit, l'utilisation de la copule gaussienne est introduite en disant que

$$\tau_i = Q^{-1} \left(\Phi \left(\sqrt{\rho} Z + \sqrt{1-\rho} Z'_i \right) \right)$$

où $Q(t) = P\{\tau_i < t\}$ est la fonction de répartition commune à tous les τ_i (typiquement exponentielle) Z est un aléa ^{gaussien} qui représente l'état général de l'économie et les Z'_i qui sont tous indépendants entre eux et avec Z représente l'état de la i -ième obligation. Il est immédiat de voir que la copule de (τ_1, \dots, τ_n) est la copule gaussienne avec la matrice de corrélation donnée ci-dessus.

III Application à l'agrégation des mesures de risques.

Le cadre de ce paragraphe est le suivant : on a mesuré le risque associé à chacun des actifs du portefeuille par une des méthodes décrites dans la première partie du cours. Par exemple si X_i sont les variables aléatoires qui modélisent les variations de chacun de ces actifs $i=1, \dots, n$. sur la durée considérée (1 à 10 jours), nous avons mesuré la $\text{VaR}_{i,\alpha}(X_i) = -F_i^{-1}(\alpha)$ où $F_i(x) = P\{X_i \leq x\}$.

ou l'expected short fall $\text{ESF}_{i,\alpha}(X_i) = -E\{X_i \mid X_i \leq -\text{VaR}_{i,\alpha}(X_i)\}$.

La question est donc de calculer la VaR_{α} ou l' ESF_{α} de $\sum_{i=1}^n X_i$.

Ces quantités dépendent fortement de la structure de dépendance et ce calcul peut se faire avec des copules. Là encore, deux approches sont possibles (a) estimer ou choisir une copule (b) utiliser la théorie des copules pour avoir des bornes sur les mesures de risques agrégées. Nous nous concentrons sur (b) dans ce paragraphe. Le point (a) sera abordé au prochain chapitre.

1) Théorème de Makarov.

Ce théorème donne des bornes sur la fonction de répartition de la somme $\sum_{i=1}^n X_i$ sachant les lois marginales F_1, \dots, F_n de X_1, \dots, X_n .

$$\sup_{\sum_{i=1}^n z_i = z} W(F_1(z_1), \dots, F_n(z_n)) \leq P\left\{\sum_{i=1}^n X_i \leq z\right\} \leq \inf_{\sum_{i=1}^n z_i = z} \tilde{W}(F_1(z_1), \dots, F_n(z_n))$$

$$\text{où } W(u_1, \dots, u_n) = \left(\sum_{i=1}^n u_i - n + 1\right)^+ \text{ et } \tilde{W}(u_1, \dots, u_n) = \min\left(1, \sum_{i=1}^n u_i\right).$$

preuve: Soient z_1, \dots, z_n tels que $\sum_{i=1}^n z_i = z$. alors

$$\forall i=1, \dots, n \quad X_i \leq z_i \Rightarrow \sum X_i \leq z$$

$$\text{donc } P\left\{\sum X_i \leq z\right\} \geq P\left\{\forall i \quad X_i \leq z_i\right\} = C(F_1(z_1), \dots, F_n(z_n))$$

où C est une copule de (X_1, \dots, X_n) comme $C > W$ pour tout couple $($

nous obtenons l'inégalité de gauche en prenant le supremum sur toutes les minoration.

Du côté droit : on remarque de même que : $\forall i \ x_i > z_i \Rightarrow \sum x_i > z$

$$\text{donc } P \left\{ \sum X_i > z \right\} \geq P \left\{ \forall i \ x_i > z_i \right\} = P \left\{ \forall i \ -x_i < -z_i \right\}$$

Notons $\hat{F}_i(x_i) = P \left\{ -X_i \leq x_i \right\}$ les fonctions de répartition de $-X_1, \dots, -X_n$ et \hat{C} un des copules. Comme \hat{C} est continue de (u_1, \dots, u_n) ,

$$P \left\{ \sum X_i > z \right\} \geq \hat{C} \left(\hat{F}_1((-z_1)^-), \dots, \hat{F}_n((-z_n)^-) \right)$$

$$\text{or } \hat{F}_i(x_i^-) = P \left\{ -X_i < x_i \right\} = 1 - P \left\{ X_i \leq -x_i \right\} = 1 - F_i(-x_i)$$

$$\text{donc } 1 - P \left\{ \sum X_i \leq z \right\} \geq \hat{C} \left(1 - F_1(z_1), \dots, 1 - F_n(z_n) \right)$$

$$P \left\{ \sum X_i \leq z \right\} \leq 1 - \hat{C} \left(1 - F_1(z_1), \dots, 1 - F_n(z_n) \right).$$

Comme \hat{C} est une copule, $\hat{C} < W$ et on remarque que

$$1 - W(1 - u_1, \dots, 1 - u_n) = 1 - \left(1 - \sum_{i=1}^n u_i \right)^+ = \min \left(1, \sum_{i=1}^n u_i \right) = \tilde{W}(u_1, \dots, u_n).$$

On obtient l'inégalité de droite en prenant l'infimum des ces majorations.

Remarque: On peut montrer que ces bornes ne peuvent pas être améliorées.

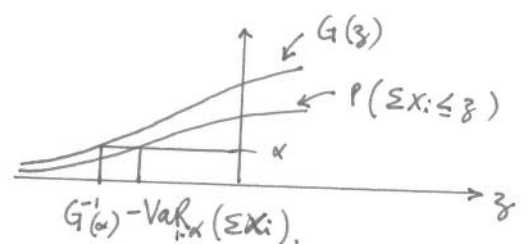
c'est à dire pour n'importe quelle famille de marginaux F_1, \dots, F_n , il existe un copule telle que $\sum X_i$ ait pour fonction de répartition l'une des deux bornes du théorème.

2) Application à la $\text{Var}_{i-\alpha}(\sum X_i)$ (On suppose les F_i continues)

$$\text{Notons } G(z) = \inf_{\sum_{i=1}^n z_i = z} \tilde{W}(F_1(z_1), \dots, F_n(z_n)).$$

$$\text{Comme } P \left\{ \sum_{i=1}^n X_i \leq z \right\} \leq G(z)$$

$$\boxed{\text{Var}_{i-\alpha}(\sum X_i) \leq -G^{-1}(\alpha)}$$



$$G^{-1}(\alpha) = \inf \left\{ z : \inf_{\sum z_i = z} \tilde{W}(F_1(z_1), \dots, F_n(z_n)) \geq \alpha \right\}$$

$$= \inf \left\{ \sum_{i=1}^n F_i^{-1}(q_i) : \tilde{W}(u_1, \dots, u_n) \geq \alpha \right\} = \inf_{\tilde{W}(u_1, \dots, u_n) = \alpha} \sum_{i=1}^n F_i^{-1}(u_i).$$

3) Application à l'expected shortfall $ESF_{\alpha}(\sum X_i)$.

On veut calculer la pire des expected shortfall :

$$-\inf_C + E \left\{ \sum X_i \mid \sum X_i \leq -\text{VaR}_{\alpha}(\sum X_i) \right\}$$

Pour faciliter les calculs, on ne s'intéresse qu'aux cas où $-\text{VaR}_{\alpha}(\sum X_i) \leq 0$ (ce qui n'est pas vraiment une contrainte en pratique!)

$$E \left\{ \sum X_i \mid \sum X_i \leq -\text{VaR}_{\alpha}(\sum X_i) \right\} = \frac{E \left\{ \sum X_i \mathbb{1}_{\left\{ \sum X_i \leq -\text{VaR}_{\alpha}(\sum X_i) \right\}} \right\}}{P \left\{ \sum X_i \leq -\text{VaR}_{\alpha}(\sum X_i) \right\}}$$

Par définition de la VaR_{α} : $P \left\{ \sum X_i \leq -\text{VaR}_{\alpha}(\sum X_i) \right\} \leq \alpha$.

$$\begin{aligned} \text{d'où } ESF_{\alpha}(\sum X_i) &\geq \frac{1}{\alpha} E \left\{ \sum X_i \mathbb{1}_{\left\{ \sum X_i \leq -\text{VaR}_{\alpha}(\sum X_i) \right\}} \right\} \\ &= \frac{1}{\alpha} E \int_{-\infty}^0 -\mathbb{1}_{\left\{ \sum X_i \leq z \right\}} \mathbb{1}_{\left\{ \sum X_i \leq -\text{VaR}_{\alpha}(\sum X_i) \right\}} dz \\ &= -\frac{1}{\alpha} \int_{-\infty}^0 P \left\{ \sum X_i \leq z \wedge -\text{VaR}_{\alpha}(\sum X_i) \right\} dz \end{aligned}$$

$$\begin{aligned} &= -\frac{1}{\alpha} \int_{-\infty}^{-\text{VaR}_{\alpha}(\sum X_i)} P \left\{ \sum X_i \leq z \right\} dz - \frac{1}{\alpha} \int_{-\text{VaR}_{\alpha}(\sum X_i)}^0 P \left\{ \sum X_i \leq \text{VaR}_{\alpha}(\sum X_i) \right\} dz \\ &\geq -\frac{1}{\alpha} \int_{-\infty}^{-\text{VaR}_{\alpha}(\sum X_i)} P \left\{ \sum X_i \leq z \right\} dz - \text{VaR}_{\alpha}(\sum X_i). \end{aligned}$$

$$\text{d'où } ESF_{\alpha}(\sum X_i) \leq -G^{-1}(\alpha) + \frac{1}{\alpha} \int_{-\infty}^{\hat{G}^{-1}(\alpha)} G(z) dz$$

avec $G(z) = \inf_{\sum z_i = z} \tilde{W}(F_1(z_1), \dots, F_n(z_n))$

$$\hat{G}^{-1}(\alpha) = \inf_{W(u_1, \dots, u_n) = \alpha} \sum_{i=1}^n F_i^{-1}(u_i)$$

Comment choisir une copule ?

I. Méthodes statistiques.

1) Méthodes paramétriques Maximum de vraisemblance.

a) Maximum de vraisemblance.

Supposons que nous observons un échantillon de taille T de vecteurs (x_1^t, \dots, x_n^t) pour $t=1, \dots, T$, réalisations indep de (X_1, \dots, X_n)

Si nous faisons l'hypothèse que la loi du vecteur (X_1, \dots, X_n) est donnée par $F_\theta(x_1, \dots, x_n)$ (fonction de répartition dépendant d'un paramètre θ), alors on choisit θ comme le paramètre qui maximise la probabilité d'avoir observé (x_1^t, \dots, x_n^t)

la vraisemblance peut s'écrire avec la copule :

$$C_\theta(F_{\theta,1}(x_1), \dots, F_{\theta,n}(x_n)) = F_\theta(x_1, \dots, x_n)$$

d'où la vraisemblance d'avoir observé ces (x_1^t, \dots, x_n^t) $t=1, \dots, T$.

$$\prod_{t=1}^T \frac{\partial^n C_\theta}{\partial x_1 \dots \partial x_n} (F_{\theta,1}(x_1^t), \dots, F_{\theta,n}(x_n^t)) \cdot \prod_{i=1}^n f_{\theta,i}(x_i^t)$$

où $f_{\theta,i}(x_i)$ est la densité de x_i .

La log-vraisemblance s'écrit

$$l(\theta) = \sum_{t=1}^T \log \frac{\partial^n C_\theta}{\partial x_1 \dots \partial x_n} (F_{\theta,1}(x_1^t), \dots, F_{\theta,n}(x_n^t)) + \sum_{t=1}^T \sum_{i=1}^n \log f_{\theta,i}(x_i^t)$$

que l'on doit maximiser par rapport à θ .

Le problème est que la dimension de l'espace des paramètres est généralement très grand et le problème est difficile numériquement.

b) Méthode IFM (Inference Functions for Margins)

Pour remédier à ce problème, on peut estimer la loi jointe en deux temps : d'abord les marginales puis la copule.

On suppose que chaque loi marginale est paramétré par θ_i et que la copule est paramétré par α . Le log vraisemblance s'écrit toujours:

$$l(\theta_1, \dots, \theta_n, \alpha) = \sum_{t=1}^T \log \frac{\partial^4}{\partial u_1 \dots \partial u_n} C_\alpha(F_{\theta_{1,1}}(x_1^t), \dots, F_{\theta_{n,n}}(x_n^t)) + \sum_{i=1}^n \log f_{\theta_{i,i}}(x_i^t)$$

On estime alors $\hat{\theta}_i = \arg \max \sum_{t=1}^T \log f_{\theta_{i,i}}(x_i^t)$.

puis dans un second temps

$$\hat{\alpha} = \arg \max \sum_{t=1}^T \log \frac{\partial^n}{\partial u_1 \dots \partial u_n} C_\alpha(F_{\hat{\theta}_{1,1}}(x_1^t), \dots, F_{\hat{\theta}_{n,n}}(x_n^t))$$

2) Méthode non paramétrique : la copule empirique. (Deheuvels 1979)

On ordonne les observations suivant chaque composante :

$$\begin{aligned} x_1^{(1)} &\leq x_1^{(2)} \leq \dots \leq x_1^{(T)} \\ x_2^{(1)} &\leq x_2^{(2)} \leq \dots \leq x_2^{(T)} \\ &\vdots \end{aligned}$$

La copule empirique est donné par :

$$\hat{C}\left(\frac{t_1}{T}, \dots, \frac{t_n}{T}\right) = \frac{1}{T} \sum_{t=1}^T \mathbb{1}_{\{x_1^t \leq x_1^{(t_1)}, \dots, x_n^t \leq x_n^{(t_n)}\}}$$

ce qui s'écrit aussi : $\frac{1}{T} \sum_{i_1=1}^T \dots \sum_{i_n=1}^T \mathbb{1}_{\{(x_1^{(i_1)}, \dots, x_n^{(i_n)}) \in \text{échantillon}\}}$

L'idée derrière cette définition est que les transformations croissantes des données conservent la copule. Ici on choisit le rang comme telle transformation

II Mesures de concordance.

Nous avons vu que les copules décrivent exactement la structure de dépendance d'un vecteur aléatoire. Les différentes structures possibles sont évidemment très nombreuses et on se propose d'introduire quelques paramètres qui résument cette structure. Ce sont le rho de Spearman et le tau de Kendall. Ils peuvent être estimés sur les données et permettent donc d'aider à choisir la copule. Nous nous plaçons dans le cas de deux variables aléatoires, (X, Y) continues.

1) La corrélation linéaire de (X, Y)

Si $\text{Var } X \text{ Var } Y > 0$, la corrélation linéaire (usuelle) est définie

$$\text{par } \text{Corr}(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var } X \text{ Var } Y}}$$

On sait que $\text{Corr}(X, Y) = +1$ si $Y = aX + b$ avec $a > 0, b \in \mathbb{R}$

$\text{Corr}(X, Y) = -1$ si $Y = aX + b$ avec $a < 0, b \in \mathbb{R}$.

La corrélation linéaire ne mesure donc que la dépendance linéaire.

On peut avoir $Y = f(X)$ où f est non-linéaire et $\text{Corr}(X, Y) \neq 1$.

par exemple: $X = \exp\left(-\frac{\sigma^2}{2} + \sigma Z\right)$ $Y = \exp\left(-\frac{\sigma'^2}{2} + \sigma' Z\right)$ où $Z \sim \mathcal{Gm}(0, 1)$

$$\text{Corr}(X, Y) = e^{\sigma\sigma'} - 1 \quad \text{Var } X = e^{\sigma^2} - 1 \quad \text{d'où} \quad \text{Corr}(X, Y) = \frac{e^{\sigma\sigma'} - 1}{\sqrt{e^{\sigma^2} - 1} \sqrt{e^{\sigma'^2} - 1}} < 1$$

2) Le rho de Spearman.

Pour avoir une mesure de corrélation qui soit indépendante par transformations monotones des variables aléatoires, on définit le rho de Spearman ρ_S qui est simplement la corrélation linéaire entre $F(X)$ et $G(Y)$ où F et G sont les fonctions de répartition de X et de Y respectivement.

$$\rho_S(X, Y) = \text{Corr}(F(X), G(Y))$$

Comme $F(x)$ et $G(y)$ sont uniformes sur $(0,1)$

$$\rho_S(x, y) = 12 E\{F(x)G(y)\} - 3.$$

Il est bien entendu que si u et v sont deux fonctions croissantes sur \mathbb{R}

alors $\rho_S(u(x), v(y)) = \rho_S(x, y)$. De même, si $Y = f(X)$ où f monotone, alors $\rho_S(x, Y) = \rho_S(x, x) = 1$.

Introduisons une copule C du couple (X, Y) , qui est aussi une copule du couple $(F(x), G(y)) = (U, V)$ où U et V sont uniformes sur $(0,1)$.

$$\begin{aligned} \rho_S(x, y) &= 12 E\{UV\} - 3 = 12 E\{(1-U)(1-V)\} - 3 \\ &= 12 E \int_0^{1-u} \int_0^{1-v} du dv - 3 = 12 E \int_0^1 \int_0^1 \mathbb{1}_{\{1-u \geq u\}} \mathbb{1}_{\{1-v \geq v\}} du dv - 3 \\ &= 12 \int_0^1 \int_0^1 C(1-u, 1-v) du dv - 3 = 12 \int_0^1 \int_0^1 C(u, v) du dv - 3. \end{aligned}$$

$$\rho_S(x, y) = 12 \int_0^1 \int_0^1 C(u, v) du dv - 3.$$

3) Le tau de Kendall.

Soient (X, Y) deux variables aléatoires avec une copule C . Le tau de Kendall τ est défini comme

$$\tau(X, Y) = 4 E\{C(U, V)\} - 1$$

où U, V sont uniformes $(0,1)$ de copule C .

Propriétés : (a) $\tau(X, Y) = 1 - 4 \int_0^1 \int_0^1 \frac{\partial C}{\partial u}(u, v) \frac{\partial C}{\partial v}(u, v) du dv$

$$(b) \tau(X, Y) = P\{(X_1 - X_2)(Y_1 - Y_2) > 0\} - P\{(X_1 - X_2)(Y_1 - Y_2) < 0\}$$

où (X_1, Y_1) et (X_2, Y_2) sont deux vecteurs indépendants de même loi que (X, Y) et X et Y sont continus

preuve : (a) en supposant que C soit continûment différentiable :

$$\begin{aligned} E\{C(U, V)\} &= E \int_0^U \int_0^V \frac{\partial^2 C}{\partial u \partial v}(u, v) du dv = E \int_0^1 \int_0^1 \mathbb{1}_{\{U \geq u, V \geq v\}} \frac{\partial^2 C(u, v)}{\partial u \partial v} du dv \\ &= E \int_0^1 \int_0^1 1 - \mathbb{1}_{\{U \leq u\}} - \mathbb{1}_{\{V \leq v\}} + \mathbb{1}_{\{U \leq u, V \leq v\}} \frac{\partial^2 C}{\partial u \partial v}(u, v) du dv \end{aligned}$$

$$\begin{aligned}
&= 1 - \int_0^1 \int_0^1 u \frac{\partial^2 C}{\partial u \partial v}(u,v) du dv - \int_0^1 \int_0^1 v \frac{\partial^2 C}{\partial u \partial v}(u,v) du dv + \int_0^1 \int_0^1 C(u,v) \frac{\partial^2 C}{\partial u \partial v}(u,v) du dv \\
&= 1 - \frac{1}{2} - \frac{1}{2} + \int_0^1 \left[C(u,v) \frac{\partial C}{\partial v}(u,v) \right]_{v=0}^{v=1} dv - \int_0^1 \int_0^1 \frac{\partial C}{\partial u}(u,v) \frac{\partial C}{\partial v}(u,v) du dv. \\
&= \frac{1}{2} - \int_0^1 \int_0^1 \frac{\partial C}{\partial u}(u,v) \frac{\partial C}{\partial v}(u,v) du dv.
\end{aligned}$$

(b) Comme (X, Y) sont continues,

$$P\{(X_1 - X_2)(Y_1 - Y_2) > 0\} - P\{(X_1 - X_2)(Y_1 - Y_2) < 0\} = 2P\{(X_1 - X_2)(Y_1 - Y_2) > 0\} - 1$$

d'autre part, pour toute fonctions strictement croissantes F et G :

$$P\{(X_1 - X_2)(Y_1 - Y_2) > 0\} = P\{(F(X_1) - F(X_2))(G(Y_1) - G(Y_2)) > 0\}$$

et on peut donc supposer que X, Y sont uniformes.

$$\begin{aligned}
P\{(U_1 - U_2)(V_1 - V_2) > 0\} &= P\{U_1 > U_2, V_1 > V_2\} + P\{U_1 < U_2, V_1 < V_2\} \\
&= E\{C(U, V)\} + 1 - \frac{1}{2} - \frac{1}{2} + E\{C(U, V)\} = 2E\{C(U, V)\}.
\end{aligned}$$

4) Indice de dépendance de queues.

Les rho de Spearman et tau de Kendall sont des mesures de concordance globale. Nous allons maintenant définir à partir d'une copule une mesure de dépendance des valeurs extrêmes.

Pour une copule C donnée on définit l'indice supérieur de dépendance de queue λ_U par:

$$\lambda_U = \lim_{u \rightarrow 1} \frac{1 - 2u + C(u, u)}{1 - u} \quad \text{si cette limite existe}$$

et l'indice inférieur λ_L :

$$\lambda_L = \lim_{u \rightarrow 0} \frac{C(u, u)}{u}$$

L'idée est la suivante si U_1 et U_2 sont deux uniformes de copule C ,

$$\text{alors: } P\{U_1 > u \mid U_2 > u\} = \frac{1 - 2u + C(u, u)}{1 - u} = P\{U_2 > u \mid U_1 > u\}.$$

lorsque λ_U existe, il décrit la limite de la probabilité d'avoir une

de même, $P\{U_1 < u \mid U_2 < u\} = P\{U_2 < u \mid U_1 < u\} = \frac{C(u, u)}{u}$.

On voit donc que $\lambda_U, \lambda_L \in [0, 1]$.

Quand λ_U (ou λ_L) est 0, on dit que la copule ne présente pas de dépendance de queues supérieure (ou inférieure). Les extrêmes sont "asymptotiquement indépendants".

Dans le cadre du problème d'aggrégation de VaR, ces indices sont très importants puisqu'ils vont mesurer les effets de extrêmes corrélés.

On remarque que les indices λ_U et λ_L sont nuls pour la copule gaussienne multivariée dans le cas où $|\rho| \neq 1$.