

Étude de la structure de dépendance entre les indices boursiers S&P500 et Dow Jones Industrial Average

MAARAF Oumayma

Actuariat Finance

Résumé

Identifier la structure de dépendance entre deux indices boursiers est une étape cruciale dans la gestion des risques de marché. A l'aide de la théorie des copules, une étude de dépendance du S&P500 et du Dow J sera menée le long de cet article. Commenant par une mise à niveau des notions de bases des copules et s'achevant par une application numérique qui aboutira à l'identification de la copule optimale qui modélise la structure de dépendance de ces indices.

Introduction

Les copules représentent désormais un outil important pour gérer de manière flexible la co-évolution entre les marchés, les facteurs de risque et d'autres variables pertinentes étudiées en finance. Cet outil est emprunté en fait à la théorie des statistiques, mais il gagne de plus en plus de popularité tant parmi les universitaires que parmi les praticiens du domaine de la finance, et ceci à cause de l'augmentation de la volatilité et du comportement erratique des marchés financiers.

Dans cet article, on va étudier la structure de dépendance entre les log-rendements de deux indices boursiers américains, le S&P500 qui est basé sur 500 grandes sociétés cotées sur les bourses aux États-Unis et le Dow Jones Industrial Average qui est basé sur 30 plus grandes compagnies américaines.

En effet, on sera amené à explorer les données de ces deux indices afin de pouvoir identifier et estimer la copule qui modélise leur dépendance éventuelle.

Le code de la partie d'application numérique est disponible sur le lien suivant
<https://nbviewer.org/github/Oumayma-maaraf/Copula/blob/main/cop.ipynb>

1 Copules et modélisation de la dépendance

Définition 1.1. (*Copule*) : Une copule est une fonction $C : [0, 1]^2 \rightarrow [0, 1]$, qui vérifie les propriétés suivantes :

$$- \forall (u, v) \in [0, 1]^2,$$

$$C(u, 0) = 0 = C(0, v)$$

$$C(u, 1) = u \text{ et } C(1, v) = v$$

$$- \forall u_1, u_2, v_1, v_2 \in [0, 1] \text{ tels que } u_1 \leq u_2 \text{ et } v_1 \leq v_2$$

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

Théorème 1.1. (*Théorème de Sklar*) : Soit H une fonction de répartition jointe avec F et G ses marginales. Alors il existe une copule C telle que pour tout $x, y \in (\hat{R})$

$$H(x, y) = C(F(x), G(y)) \tag{1}$$

Si F et G sont continue alors C est unique.

Soit C une copule et F et G des fonctions de répartition, alors la fonction H définie par (1) est une fonction de répartition jointe avec F et G ses marginales.

Théorème 1.2. (*Tau de Kendall*) : Soient (X_1, Y_1) et (X_2, Y_2) deux vecteurs indépendants de variables aléatoires continues de loi jointe respectivement H_1 et H_2 , et de même loi marginale F (pour X_1 et X_2) et G (pour Y_1 et Y_2). Soient C_1 et C_2 les copules

respectives de (X_1, Y_1) et (X_2, Y_2) .
le tau de Kendall se définit comme suit

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

Et on a

$$\tau = 4 \iint_{[0,1]^2} C_2(u, v) dC_1(u, v) - 1$$

Démonstration. Comme nos variables sont continues on a

$$\begin{aligned} \tau &= 2\mathbb{P}[(X_1 - X_2)(Y_1 - Y_2) > 0] - 1 \\ &= 2(\mathbb{P}[X_1 < X_2, Y_1 < Y_2] + \mathbb{P}[X_1 > X_2, Y_1 > Y_2]) - 1 \end{aligned}$$

et d'après le théorème de Sklar $H_2(x, y) = C_2(F(x), G(y))$, donc

$$\begin{aligned} \mathbb{P}[X_1 > X_2, Y_1 > Y_2] &= \iint_{\mathbb{R}^2} \mathbb{P}[X_2 < x, Y_2 < y] dC_1(F(x), G(y)) \\ &= \iint_{\mathbb{R}^2} C_2(F(x), G(y)) dC_1(F(x), G(y)) \end{aligned}$$

En posant $u = F(x)$ et $v = G(y)$, on obtient

$$\mathbb{P}[X_1 > X_2, Y_1 > Y_2] = \iint_{\mathbb{R}^2} C_2(u, v) dC_1(u, v)$$

De manière similaire on obtient

$$\begin{aligned} \mathbb{P}[X_1 < X_2, Y_1 < Y_2] &= \iint_{\mathbb{R}^2} (1 - u - v + C_2(u, v)) dC_1(u, v) \\ &= 1 - \frac{1}{2} - \frac{1}{2} + \iint_{\mathbb{R}^2} C_2(u, v) dC_1(u, v) \\ &= \iint_{\mathbb{R}^2} C_2(u, v) dC_1(u, v) \end{aligned}$$

D'où le résultat

□

Définition 1.2. (*Dépendance des queues*) : Soient X et Y deux variables aléatoires continues de fonction de répartition respective F et G . On définit quand les limites suivantes existent :

- Le coefficient de dépendance de queue supérieure est :

$$\lambda_U = \lim_{t \rightarrow 1^-} \mathbb{P}[Y > G^{-1}(t) | X > F^{-1}(t)]$$

- Le coefficient de dépendance de queue inférieure est :

$$\lambda_L = \lim_{t \rightarrow 0^+} \mathbb{P}[Y \leq G^{-1}(t) | X \leq F^{-1}(t)]$$

Theorème 1.3. Soient X et Y deux variables aléatoires continues de fonction de répartition respective F et G , et C la copule associée à X et Y . On définit λ_U et λ_L de la même manière que la définition précédente. Si les limites de la définition précédente existent alors

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

et

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

Démonstration.

$$\begin{aligned} \lambda_U &= \lim_{t \rightarrow 1^-} \mathbb{P} [Y > G^{-1}(t) | X > F^{-1}(t)] \\ &= \lim_{t \rightarrow 1^-} \mathbb{P} [G(Y) > t | F(X) > t] \\ &= \lim_{t \rightarrow 1^-} \frac{\mathbb{P} [G(Y) > t, F(X) > t]}{\mathbb{P} (F(X) > t)} \\ &= \lim_{t \rightarrow 1^-} \frac{1 - t - t + C(t, t)}{1 - t} \\ &= \lim_{t \rightarrow 1^-} \frac{1 - 2t + C(t, t)}{1 - t} \end{aligned}$$

Et

$$\begin{aligned} \lambda_L &= \lim_{t \rightarrow 0^+} \mathbb{P} [Y \leq G^{-1}(t) | X \leq F^{-1}(t)] \\ &= \lim_{t \rightarrow 0^+} \mathbb{P} [G(Y) \leq t | F(X) \leq t] \\ &= \lim_{t \rightarrow 0^+} \frac{\mathbb{P} [G(Y) \leq t, F(X) \leq t]}{\mathbb{P} (F(X) \leq t)} \\ &= \lim_{t \rightarrow 0^+} \frac{C(t, t)}{t} \end{aligned}$$

□

1.1 Copules Archimédiennes

Les copules archimédiennes peuvent être construites en utilisant une fonction $\phi : [0, 1] \rightarrow \mathbb{R}^+$, continue, décroissante, convexe et $\phi(1) = 0$. Cette fonction est appelée générateur. Si de plus $\phi(0) = +\infty$, elle est appelée générateur strict.

Définition 1.1.1. (*Pseudo-inverse*) : Soit ϕ une fonction vérifiant les conditions précédentes. Le pseudo-inverse de cette fonction est le suivant :

$$\phi^{[-1]}(v) = \begin{cases} \phi^{-1}(v) & 0 \leq v \leq \phi(0) \\ 0 & \phi(0) \leq v \leq +\infty \end{cases}$$

En particulier, on a pour tout $v \in [0, 1]$

$$\phi^{[-1]}(\phi(v)) = v$$

Définition 1.1.2. (Copule Archimédienne) : Soit ϕ un générateur et $\phi^{[-1]}$ son pseudo-inverse. La copule archimédienne C^ϕ est générée comme suit, pour tout $u, v \in [0, 1]$

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

Proposition 1.1.1. (Tau de Kendall) : Soit C une copule Archimédienne de générateur ϕ . Son tau de Kendall τ est donné par

$$\tau = 1 + 4 \int_0^1 \frac{\phi(u)}{\phi'(u)} du$$

Proposition 1.1.2. (Dépendance de queue supérieure) : Soit C une copule Archimédienne bivariée stricte, de générateur ϕ

- Si $\phi^{-1}(0)$ n'est pas défini, alors C ne possède pas de dépendance de queue supérieure
- Si $\phi^{-1}(0)$ est fini, le coefficient de dépendance de queue supérieure associé à C est donnée par :

$$\lambda_U = 2 - 2 \lim_{s \rightarrow 0^+} \left(\frac{\phi^{-1}(2s)}{\phi^{-1}(s)} \right)$$

Démonstration. On a

$$\begin{aligned} \lambda_U &= \lim_{t \rightarrow 1^-} \frac{1 - 2t + C(t, t)}{1 - t} \\ &= 2 - \lim_{t \rightarrow 1^-} \frac{1 - C(t, t)}{1 - t} \\ &= 2 - \lim_{t \rightarrow 1^-} \frac{1 - \phi^{-1}(2\phi(t))}{1 - t} \\ &= 2 - \lim_{s \rightarrow 0^+} \frac{1 - \phi^{-1}(2s)}{1 - \phi^{-1}(s)} \quad (s = \phi(t)) \end{aligned}$$

□

Proposition 1.1.3. (Dépendance de queue inférieure) : Soit C une copule Archimédienne bivariée stricte, de générateur ϕ :

Le coefficient de dépendance de queue inférieure associé à C est donnée par :

$$\lambda_L = 2 \lim_{s \rightarrow +\infty} \left(\frac{\phi^{-1}(2s)}{\phi^{-1}(s)} \right)$$

Démonstration. On part du Théorème 1.2 et on procède de la même manière que dans la démonstration précédente. □

1.1.1 Copule de Clayton

Soit $\theta > 0$, on considère la fonction $\phi : [0, 1] \rightarrow \bar{\mathbb{R}}^+$ définie comme suit :

$$\phi(u) = \frac{u^{-\theta} - 1}{\theta}$$

- ϕ est continue sur $[0, 1]$
- ϕ est au moins deux fois dérivable sur $]0, 1]$, sa dérivée est $\phi'(u) = -u^{-1-\theta}$, qui est négative strictement, donc ϕ est décroissante
- la dérivée seconde de ϕ est $\phi''(u) = (1 + \theta)u^{-2-\theta}$, qui est positive, donc ϕ est convexe
- $\phi(1) = 0$ et $\lim_{t \rightarrow 0^+} \frac{1}{u^\theta} = +\infty$

Donc ϕ est un générateur strict. On a $\phi(0) = +\infty$ donc $\phi^{[-1]} = \phi^{-1}$. Ainsi pour tout $v \in \bar{\mathbb{R}}^+$

$$\phi^{-1}(v) = (1 + v\theta)^{-\frac{1}{\theta}}$$

La Copule Archimédienne générée par la fonction ϕ est

$$\begin{aligned} C^\phi(u, v) &= (1 + (\phi(u) + \phi(v)) \times \theta)^{-\frac{1}{\theta}} \\ &= (1 + \left(\frac{u^{-\theta} - 1}{\theta} + \frac{v^{-\theta} - 1}{\theta} \right) \times \theta)^{-\frac{1}{\theta}} \\ &= (u^{-\theta} + v^{-\theta} - 1)^{-\frac{1}{\theta}} \end{aligned}$$

Cette copule est appelée **copule de Clayton**.

Le tau de Kendall de cette copule est donné par

$$\begin{aligned} \tau &= 1 - \frac{4}{\theta} \int_0^1 u^{1+\theta} (u^{-\theta} - 1) du \\ &= 1 - \frac{4}{\theta} \left(\left[\frac{1}{2} u^2 \right]_0^1 - \left[\frac{1}{\theta+2} u^{\theta+2} \right]_0^1 \right) \\ &= 1 - \frac{2}{\theta+2} \\ &= \frac{\theta}{\theta+2} \end{aligned}$$

Le coefficient de dépendance de queue supérieure associé à C^ϕ est donné par

$$\begin{aligned} \lambda_U &= 2 - 2 \lim_{s \rightarrow 0} \frac{(1 + 2s\theta)^{-\frac{1}{\theta}}}{(1 + s\theta)^{-\frac{1}{\theta}}} \\ &= 2 - 2 \lim_{s \rightarrow 0} \frac{1 - \frac{1}{\theta}(2s\theta) + o(s)}{1 - \frac{1}{\theta}(s\theta) + o(s)} \\ &= 2 - 2 \\ &= 0 \end{aligned}$$

Le coefficient de dépendance de queue inférieure associé à C^ϕ est donné par

$$\begin{aligned}\lambda_L &= 2 \lim_{s \rightarrow +\infty} \frac{(1 + 2s\theta)^{-\frac{1}{\theta}}}{(1 + s\theta)^{-\frac{1}{\theta}}} \\ &= \frac{2^{-\frac{1}{\theta}} \theta^{-\frac{1}{\theta}}}{\theta^{-\frac{1}{\theta}}} \\ &= 2^{-\frac{1}{\theta}}\end{aligned}$$

1.1.2 Copule de Gumbel

Soit $\theta \geq 1$, on considère la fonction $\phi : [0, 1] \rightarrow \bar{\mathbb{R}}^+$ définie comme suit :

$$\phi(u) = (-\ln(u))^\theta$$

- ϕ est continue sur $[0, 1]$
- ϕ est au moins deux fois sur $]0, 1]$, sa dérivée est $\phi'(u) = -\frac{1}{u}\theta(-\ln(u))^{\theta-1}$, qui est négative strictement, donc ϕ est décroissante
- la dérivée seconde de ϕ est

$$\phi''(u) = \frac{\theta((1 - \theta)\ln(u) - 1)(-\ln(u))^{\theta-1}}{u^2}$$

qui est positive, donc ϕ est convexe

- $\phi(1) = 0$ et $\lim_{t \rightarrow 0^+} (-\ln(u))^\theta = +\infty$

Donc ϕ est un générateur strict. On a $\phi(0) = +\infty$ donc $\phi^{[-1]} = \phi^{-1}$. Ainsi pour tout $v \in \bar{\mathbb{R}}^+$

$$\phi^{-1}(v) = \exp(-v^{\frac{1}{\theta}})$$

La Copule Archimédienne générée par ϕ est

$$\begin{aligned}C^\phi(u, v) &= \exp(-(\phi(u) + \phi(v))^{\frac{1}{\theta}}) \\ &= \exp(-[(-\ln(u))^\theta + (-\ln(v))^\theta]^{\frac{1}{\theta}})\end{aligned}$$

Cette copule est appelée **copule de Gumbel**.

Le tau de Kendall de cette copule est donné par

$$\begin{aligned}\tau &= 1 - \frac{4}{\theta} \int_0^1 u \frac{(-\ln(u))^\theta}{(-\ln(u))^{\theta-1}} du \\ &= 1 + \frac{4}{\theta} \int_0^1 u \ln(u) du \\ &= 1 - \frac{1}{\theta} \\ &= \frac{\theta - 1}{\theta}\end{aligned}$$

Le coefficient de dépendance de queue supérieure associé à C^ϕ est donné par

$$\begin{aligned}\lambda_U &= 2 - \lim_{s \rightarrow 0^+} \frac{1 - \exp(-(2s)^{\frac{1}{\theta}})}{1 - \exp(-s^{\frac{1}{\theta}})} \\ &= 2 - \lim_{s \rightarrow 0^+} \frac{1 - (1 - (2s)^{\frac{1}{\theta}}) + o(s^{\frac{1}{\theta}})}{1 - (1 - s^{\frac{1}{\theta}}) + o(s^{\frac{1}{\theta}})} \\ &= 2 - 2^{\frac{1}{\theta}}\end{aligned}$$

Le coefficient de dépendance de queue inférieure associé à C^ϕ est donné par

$$\begin{aligned}\lambda_L &= 2 \lim_{s \rightarrow +\infty} \frac{\exp(-(2s)^{\frac{1}{\theta}})}{\exp(-s^{\frac{1}{\theta}})} \\ &= 2 \lim_{s \rightarrow +\infty} \exp((1 - 2^{\frac{1}{\theta}})s^{\frac{1}{\theta}}) \\ &= 0\end{aligned}$$

1.1.3 Copule de Frank

Soit $\theta \neq 0$, on considère la fonction $\phi : [0, 1] \rightarrow \bar{\mathbb{R}}^+$ définie comme suit :

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

- ϕ est continue sur $[0, 1]$
- ϕ est au moins deux fois sur $]0, 1]$, sa dérivée est $\phi'(u) = \frac{\theta e^{-\theta u}}{e^{-\theta u} - 1}$, qui est négative strictement, donc ϕ est décroissante
- la dérivée seconde de ϕ est $\phi''(u) = \frac{\theta^2 e^{-\theta u}}{(e^{-\theta u} - 1)^2}$ qui est positive, donc ϕ est convexe
- $\phi(1) = 0$ et $\lim_{t \rightarrow 0^+} (-\ln(u))^\theta = +\infty$

Donc ϕ est un générateur strict. On a $\phi(0) = +\infty$ donc $\phi^{[-1]} = \phi^{-1}$. Ainsi pour tout $v \in \bar{\mathbb{R}}^+$

$$\phi^{-1}(v) = -\frac{1}{\theta} \ln(e^{-v}(e^{-\theta} - 1) + 1)$$

La Copule Archimédienne générée par ϕ est

$$C^\phi(u, v) = -\frac{1}{\theta} \ln\left(\frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{(e^{-\theta} - 1)} + 1\right)$$

Cette copule est appelée copule de Frank.

Le tau de Kendall de cette copule est donné par

$$\tau = 1 - 4 \frac{\frac{1}{\theta} \int_0^\theta \frac{t}{e^t - 1} dt - 1}{\theta}$$

Le coefficient de dépendance de queue supérieure associé à C^ϕ est donné par

$$\begin{aligned}\lambda_U &= 2 - 2 \lim_{s \rightarrow 0^+} \frac{\ln(e^{-2s}(e^{-\theta} - 1) + 1)}{\ln(e^{-s}(e^{-\theta} - 1) + 1)} \\ &= 2 - 2 \lim_{s \rightarrow 0^+} \frac{\ln(1 + (e^{-\theta} - 1)(1 - 2s) + o(s))}{\ln(1 + (e^{-\theta} - 1)(1 - s) + o(s))} \\ &= 2 - 2 \\ &= 0\end{aligned}$$

Le coefficient de dépendance de queue inférieure associé à C^ϕ est donné par

$$\lambda_L = 0$$

2 Inférence statistique de copules

2.1 Dépendance

2.1.1 Khi-Plot

Il s'agit d'une méthode qui permet la détection de la présence d'association entre deux variables à partir d'un échantillon aléatoire. Elle s'appuie exclusivement sur les couples de rangs des observations.

Soit la paire (X_i, Y_i) , $1 \leq i \leq n$, on pose

$$F_i = \frac{1}{n-1} \sum_{j=1}^n \mathbb{1}_{\{j \neq i: X_j \leq X_i\}}$$

$$G_i = \frac{1}{n-1} \sum_{j=1}^n \mathbb{1}_{\{j \neq i: Y_j \leq Y_i\}}$$

$$H_i = \frac{1}{n-1} \sum_{j=1}^n \mathbb{1}_{\{j \neq i: X_j \leq X_i, Y_j \leq Y_i\}}$$

Le graph de Khi-plot se base sur deux statistiques :

- la statistique χ : telle que pour tout $1 \leq i \leq n$

$$\chi_i = \frac{H_i - F_i G_i}{\sqrt{F_i(1 - F_i)G_i(1 - G_i)}}$$

On a donc $\chi \in [-1, 1]$.

- la statistique λ : telle que pour tout $1 \leq i \leq n$

$$\lambda_i = 4 \operatorname{sgn}((F_i - 0.5)(G_i - 0.5)) \max((F_i - 0.5)^2, (G_i - 0.5)^2)$$

Avec

$$\operatorname{sgn}(x) = \begin{cases} 1 & x \geq 0 \\ -1 & x < 0 \end{cases}$$

On a donc $\lambda_i \in [-1, 1]$. Pour éliminer les valeurs aberrantes, on trace uniquement les paires qui vérifient

$$|\lambda_i| < 4 \left(\frac{1}{n-1} - 0.5 \right)^2$$

2.1.2 K-Plot

Pour construire ce graphique :

- on calcule les H_i
- on ordonne les H_i : $(H_{(i)})_{i \in \{1, \dots, n\}}$
- on calcule l'espérance de la i -ème statistique d'ordre d'un échantillon aléatoire de taille n de $C(U, V)$

$$W_{i:n} = n C_{n-1}^{i-1} \int_0^1 w(K_0(w))^{i-1} (1 - K_0(w))^{n-i} dK_0(w)$$

où K_0 est la fonction de répartition de $C(U; V)$ sous l'hypothèse de l'indépendance.

- on trace les paires $(W_{i:n}; H_{(i)})$

2.2 Estimation des copules

2.2.1 Méthode des moments

Cette méthode est utilisée pour les mesures de dépendance. Si la mesure de dépendance s'exprime en fonction du paramètre de copule

$$\kappa = f(\theta)$$

Alors pour une fonction f qui admet un inverse, et pour un estimateur non-paramétrique de κ , on a

$$\hat{\theta} = f^{-1}(\hat{\kappa})$$

En particulier, le tau de Kendall peut être estimé par (n nombre d'observations)

$$\hat{\tau} = \frac{2}{n(n-1)} \sum_{j=2}^n \sum_{i=1}^{j-1} \text{sgn}[(X_j - X_i)(Y_j - Y_i)]$$

Et donc on peut appliquer cette méthode pour les copules dont le tau de Kendall s'écrit en fonction de leurs paramètre.

Démonstration. (Estimateur du tau de Kendall) : La somme $\sum_{j=2}^n \sum_{i=1}^{j-1} \text{sgn}[(X_j - X_i)(Y_j - Y_i)]$ est finie, et on a

$$\begin{aligned} \mathbb{E}(\hat{\tau}) &= \frac{2}{n(n-1)} \sum_{j=2}^n \sum_{i=1}^{j-1} \mathbb{E}(\text{sgn}[(X_j - X_i)(Y_j - Y_i)]) \\ &= \frac{2}{n(n-1)} \sum_{j=2}^n \sum_{i=1}^{j-1} [\mathbb{P}[(X_j - X_i)(Y_j - Y_i) \geq 0] - \mathbb{P}[(X_j - X_i)(Y_j - Y_i) < 0]] \end{aligned}$$

Ce qui correspond bien à la définition du tau de Kendall pour $((X_i, Y_i))_{i \in \{1, \dots, n\}}$ □

2.3 Tests d'adéquation pour les modèles de copules

Soient X et Y deux variables aléatoires continues de fonction de répartition respective F et G . Soient $(X_i)_{i \in \{1, \dots, n\}}$ et $(Y_i)_{i \in \{1, \dots, n\}}$ n observation de X et Y . On estime F et G par

$$\begin{aligned} F_n(x) &= \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{X_i \leq x\}} \\ G_n(y) &= \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{Y_i \leq y\}} \end{aligned}$$

Et pour tout $1 \leq i \leq n$, on pose

$$U_i = F_n(X_i) \text{ et } V_i = G_n(Y_i)$$

Et

$$U_i = \frac{R_i}{n} \text{ et } V_i = \frac{S_i}{n}$$

Définition 2.3.1. (*Copule empirique*) : La copule empirique a été introduite par Deheuvels (1979) et connue sous l'appellation *Fonction empirique de dépendance* : pour tout $(u, v) \in [0, 1]^2$

$$\hat{C}_n(u, v) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{\frac{R_i}{n} \leq u, \frac{S_i}{n} \leq v\}}$$

2.3.1 Principe

Après l'estimation des copules(e) C_θ , on teste si la vraie copule appartient à la famille qu'on a choisit, c'est-à-dire on teste

$$H_0 : C \in C_\theta$$

D'après la méthode des moments, on a déjà estimé θ sous l'hypothèse H_0 . On calcule alors la distance $\|C_{\hat{\theta}} - \hat{C}_n\|$.

On mesure la distance par

$$D_n = n \iint_{[0,1]^2} [C_{\hat{\theta}}(u, v) - \hat{C}_n(u, v)]^2 d\hat{C}_n(u, v)$$

Pour avoir avec une formule simple de D_n , on travaille avec la **statistique de Cramér-von Mises**

$$D_n = \sum_{i=1}^n \left[C_{\hat{\theta}} \left(\frac{R_i}{n+1}, \frac{S_i}{n+1} \right) - \hat{C}_n \left(\frac{R_i}{n+1}, \frac{S_i}{n+1} \right) \right]^2$$

Et aussi la **statistique de Kolmogorov-Smirnov**

$$D_n = \max_{1 \leq i \leq n} \left| C_{\hat{\theta}} \left(\frac{R_i}{n+1}, \frac{S_i}{n+1} \right) - \hat{C}_n \left(\frac{R_i}{n+1}, \frac{S_i}{n+1} \right) \right|$$

3 Application numérique

On travaille avec les log rendements des deux indices boursiers *S&P500* et *DowJ*, depuis le 31/12/2007 jusqu'au 23/12/2021.

3.1 Description des données

	Nombre d'observation	moyenne	écart-type	minimum	premier quantile	médiane	troisième quantile	maxium
S&P500	3521	0.000332	0.013050	-0.1276	-0.0039	0.000728	0.005800	0.109572
Dow J	3521	0.000283	0.012530	-0.1384	-0.0040	0.000597	0.005534	0.107643

TABLE 1 – Données descriptives de S&P 500 et Dow J

On constate que l'estimation de la moyenne des log-rendement des deux indices est très proches de zéro. Les coefficients d'aplatissement et d'asymétrie empiriques sont les suivants

	Coef. Asymétrie	Coef. Aplatissement
S&P500	-0.5628	13.29
Dow J	-0.4864	15.682

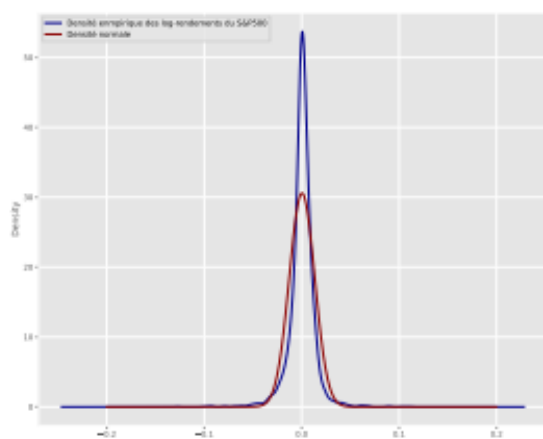
TABLE 2 – Coefficient d'aplatissement et d'asymétrie de S&P 500 et Dow J

On constate que le coefficient d'asymétrie empirique est négatif pour les deux indices, ce qui veut dire que les distribution des log-rendements (des deux indices) est plus étalée à gauche. On remarque aussi que le coefficient d'aplatissement plus grand que 3 pour les deux indices, ce qui veut dire que les queues de la distribution des log-rendements des deux indices sont beaucoup plus épaisses que celles de la loi normale et que la distribution est leptokurtique.

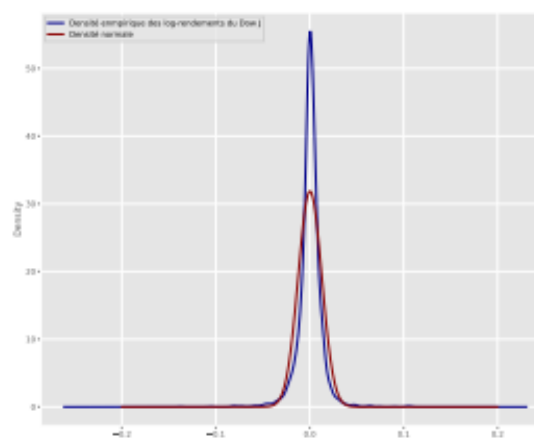
3.2 Lois marginales

D'après les graphiques de la densité empirique des log rendements des deux indices et de la loi normale, on constate que la densité empirique est vraiment leptokurtique. On remarque aussi, d'après les QQ-plot que les quantiles empiriques sont différents des quantiles théoriques de la loi normale.

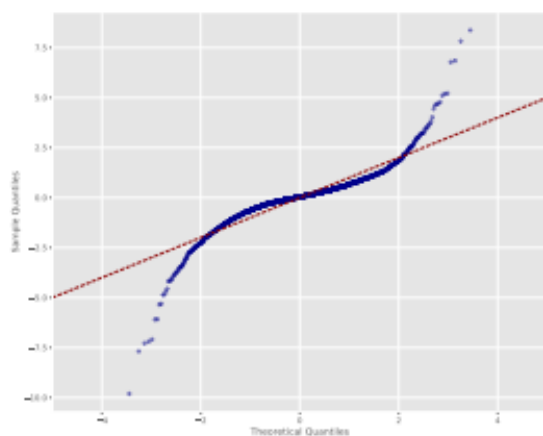
Toutes ces constations nous mènent à questionner la normalité des log-rendements des deux indices S&P500 et DowJ. Il faut donc faire un test statistique.



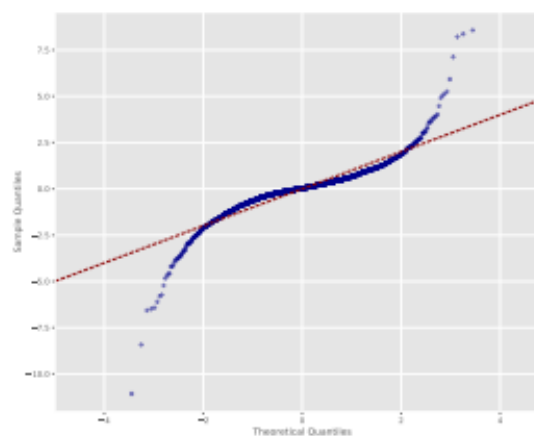
(a) Densité empirique des log-rendements du S& P500 et densité normale



(b) Densité empirique des log-rendements du Dow J et densité normale



(a) qqplot des log-rendements S& P500 et loi normal



(b) qqplot des log-rendements Dow J et loi normal

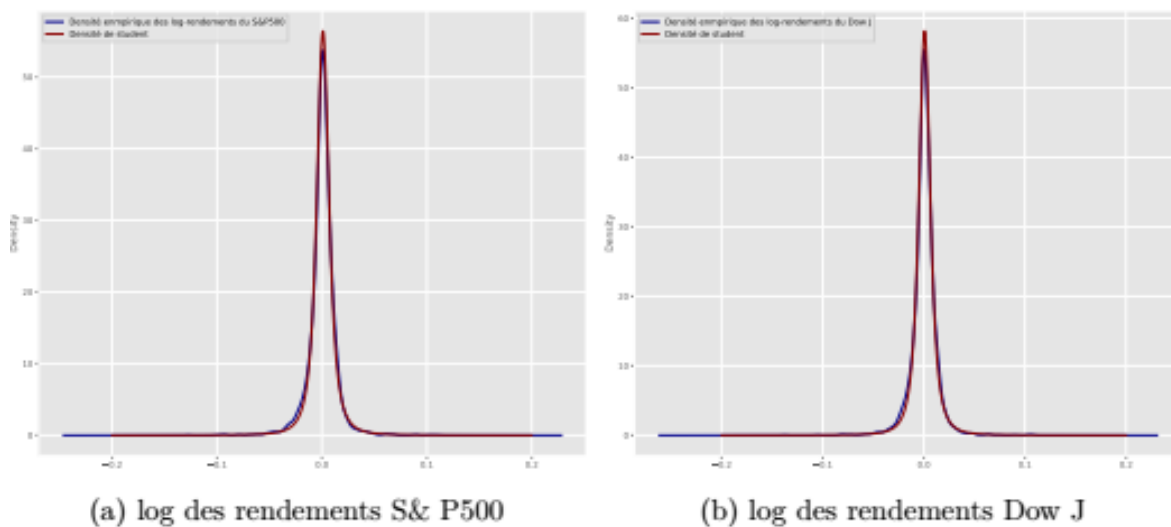
On utilise le test de Kolmogorov-Smirnov pour tester la normalité des log-rendements des deux indices

	Statistique	P-value
S& P500	0.47715	0.000
Dow J	0.47848	0.000

TABLE 3 – Résultat du test de normalité Kolmogorov-Smirnov

Il faut donc rejeter l'hypothèse de normalité.

On remarque que la densité de Student s'ajuste bien aux densités empiriques des deux indices. On teste alors si les log-rendements suivent une loi de Student, ceci en utilisant le test de Kolmogorov-Smirnov.



Les résultats du test nous confirme que les log-rendements des deux indices suivent une loi de Student de degré de liberté figurant dans le tableau 5.

	Statistique	P-value
S& P500	0.01946	0.1368
Dow J	0.01955	0.1335

TABLE 4 – Résultat du test de Kolmogorov-Smirnov (Student)

	Degré de liberté
S& P 500	2.12
Dow J	2.17

TABLE 5 – Degré de liberté de loi de Student des deux indices

3.3 Étude de la dépendance

3.3.1 Matrice de corrélation

La matrice de corrélation (empirique) nous donne une idée sur la dépendance (corrélation linéaire) entre les log rendements des deux indices.

	S& P 500	Dow J
S& P 500	1	0.977
Dow J	0.977	1

TABLE 6 – Matrice de corrélation

On constate que la corrélation linéaire entre les log rendements des deux indices s'élève à 0.976.

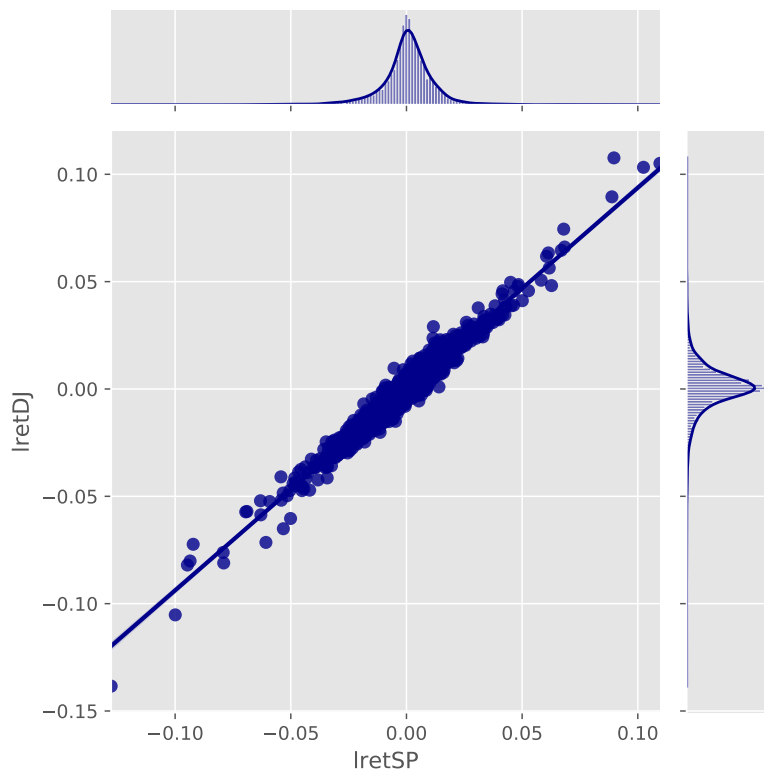


FIGURE 1 – Graph de dispersion des log-rendements des deux indices

3.3.2 Diagramme du Rank Rank

D'après le Rank Rank, il y'a une forte dépendance positive entre les log-rendements des deux indices, ainsi qu'une dépendance des queues supérieure et inférieure.

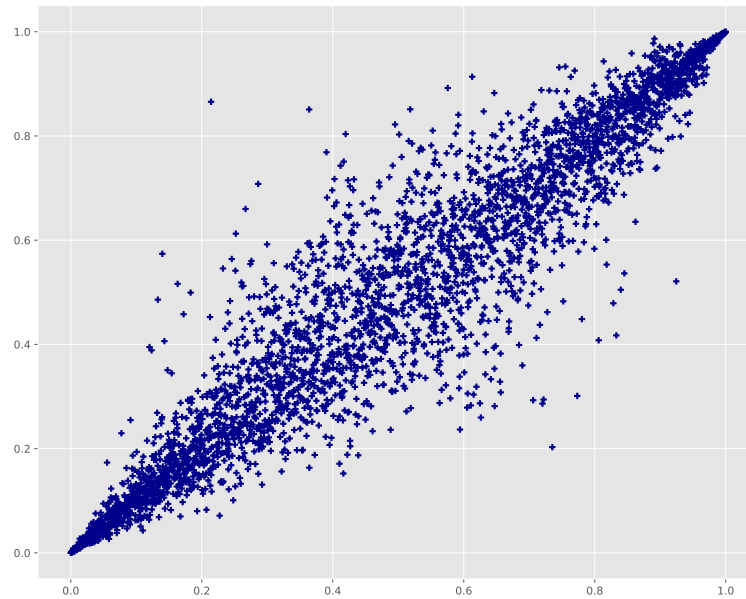


FIGURE 2 – Diagramme de Rank Rank

3.3.3 Khi plot

D'après le Khi-plot, il y'a dépendance positive entre les log-rendements des deux indices (Le diagramme se situe au-dessus de l'intervalle de l'indépendance)

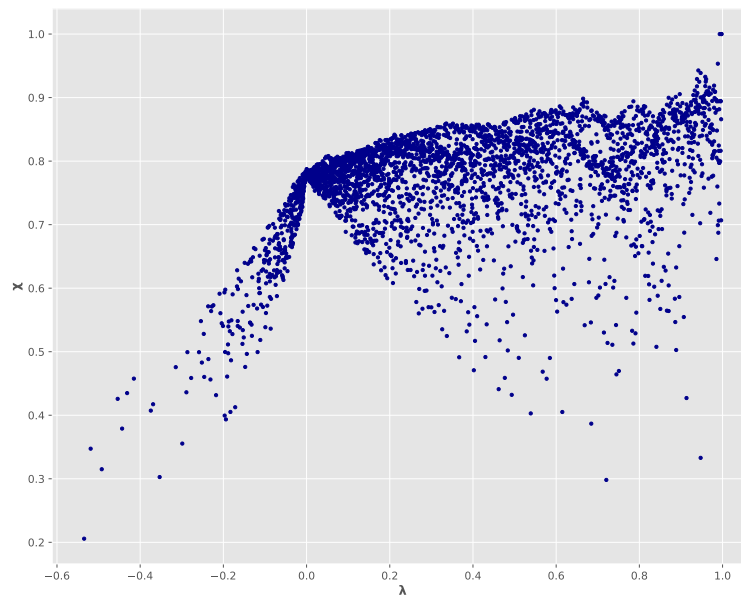


FIGURE 3 – khi plot des log rendements de S&P500 et DowJ

3.3.4 K plot

Le K plot nous confirme la dépendance positive (ici comonocité) entre les log-rendements des deux indices.

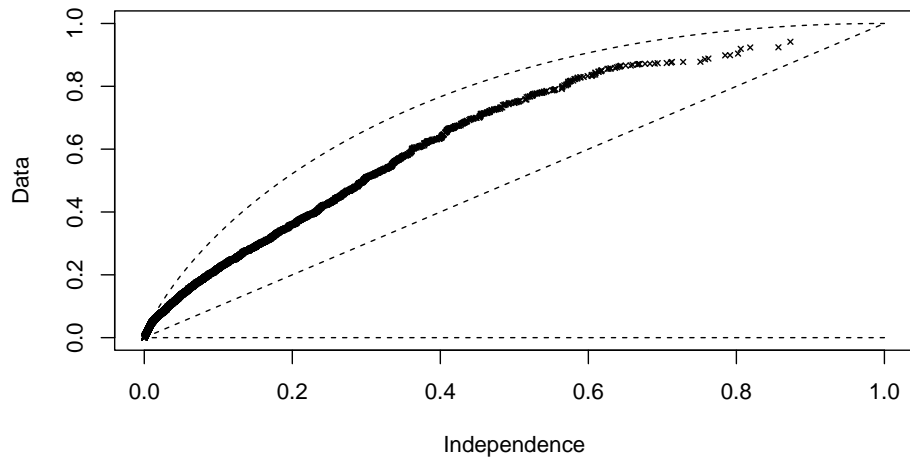
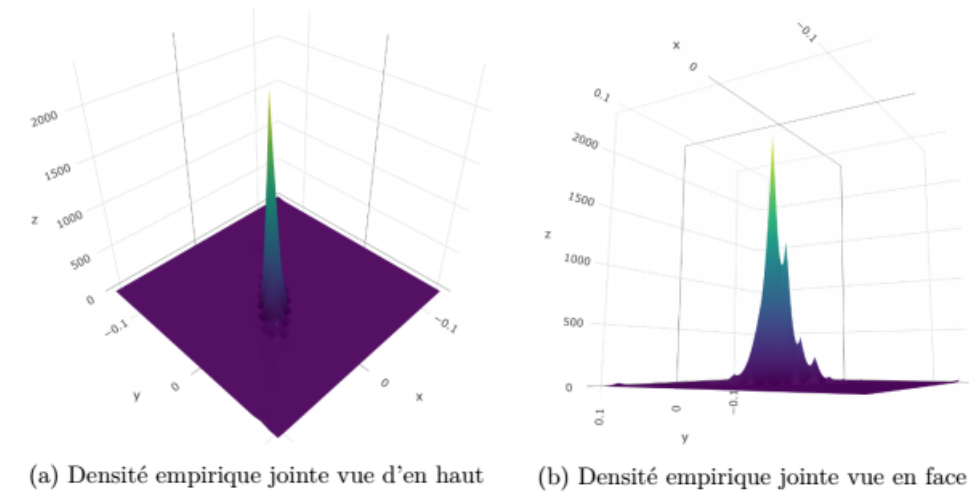


FIGURE 4 – K plot des log rendements de S&P500 et DowJ

3.4 Estimation de la Copule

La densité empirique jointe des log-rendement des deux indices est la suivante. Elle illustre toutes les constations précédentes (dépendance positive, dépendance des queues supérieure et inférieure)



D'après tout ce qu'on a constaté grâce aux différents diagrammes et graphs, on peut examiner les copules de Gumbel de Clayton et de Frank comme des candidates potentielles. En effet la copule de Gumbel permet la dépendance de queue supérieure entre les variables individuelles, celle de Clayton autorise tout niveau non nul de dépendance de queue inférieure entre les variables individuelles.

3.4.1 Estimation non-paramétrique du tau de Kendall

En utilisant la formule de l'estimateur non-paramétrique du tau de Kendall, on trouve

$$\hat{\tau} = 0.81983$$

3.4.2 Estimation des paramètres de copule et des coefficients de dépendance de queues

On se base sur la méthode des moments pour estimer le paramètre de copule (Gumbel, Clayton), et puis avec cette estimation calculer les coefficients de dépendance supérieure et inférieure de queue.

Pour la copule de Frank, la méthode de moments est difficile à implémenter, vu qu'il faut inverser la fonction de Debye. On se contente alors des estimations données par des fonctions prédéfinies sur le logiciel R.

paramètre/Coef.	Formule	Estimation
$\hat{\theta}$	$\frac{1}{1 - \hat{\tau}}$	5.55
$\hat{\lambda}_U$	$2 - 2^{\frac{1}{\hat{\theta}}}$	0.86698
$\hat{\lambda}_L$	0	0

TABLE 7 – Estimation de paramètre et de coefficients de dépendance supérieure de queue pour la copule de Gumbel

paramètre/Coef.	Formule	Estimation
$\hat{\theta}$	$\frac{2\hat{\tau}}{1 - \hat{\tau}}$	9.10
$\hat{\lambda}_U$	0	0
$\hat{\lambda}_L$	$2^{\frac{-1}{\hat{\theta}}}$	0.92666

TABLE 8 – Estimation de paramètre et de coefficients de dépendance inférieure de queue pour la copule de Clayton

paramètre/Coef.	Formule	Estimation
$\hat{\theta}$	-	19.73
$\hat{\lambda}_U$	0	0
$\hat{\lambda}_L$	0	0

TABLE 9 – Estimation de paramètre de la copule de Frank

3.5 Test d'adéquation pour les modèles de copules

On se base sur les statistiques de Cremér-von-Mises et de Kolmogorov-Smirnov pour choisir la copule optimale.

	Gumbel	Clayton	Frank
D_{CVM}	0.1905	0.6766	0.4164
D_{KS}	0.0174	0.0336	0.0221

TABLE 10 – Distance de Cramer-von Mises et Distance de Kolmogorov-Smirnov pour les copules de Gumbel de Clayton et de Frank

On trouve que c'est la Copule de Gumbel de paramètre cité au dessus qui est optimale parmi ces trois copules.

4 Conclusion

Les log-rendements des deux indices boursiers S&P500 et DowJ sont dépendants positivement. L'estimation et le test d'adéquation de la copule associée à montrer que parmi les copules de Gumbel de Clayton et de Frank, la copule de Gumbel est optimale. Néanmoins, en utilisant une fonction prédéfinie sur le logiciel R, on trouve que la copule optimale (associée à ces deux indices) est la copule BB7 (Joe-Clayton), qui est une copule à deux paramètres. Donc une étude de cette copule serait intéressante.

Références

- [1] Umberto Cherubini, Elisa Luciano, Walter Vecchiato (2004) Copula Methods in Finance, Wiley Finance Series.
- [2] Roger B. Nelsen (2006) An Introduction to Copulas, Springer Series in Statistics, 2nd ed.
- [3] Cours de Thierry RONCALLI : Copules et Aspects Multidimensionnels du Risque