Copules Gaussiennes

Kidiomi Oussou
Amine Meddour
Daouda Cisse
Mehdi Jacqueline

ISFA

Décembre 2023

Sommaire

Sommaire

Introduction

Présentation de la copule gaussienne
- Définition d'une copule
- Définition de la copule gaussienne
- Mesure de dépendance
- Mesure de dépendance
- Dépendance des extrêmes
- Utilisation de la copule en pratique

Représentations graphiques
- Graphiques
- Graphiques

Application 1

Application 2

Conclusion

Introduction

Modélisation des dépendances entre variables aléatoires.

- Finance, météorologie, gestion des risques, ...,
- Capture des dépendances linéaires, symétriques et continues,
- Facilité de compréhension, d'utilisation et d'estimation.

Présentation de la copule gaussienne - Définition d'une copule

Rappel :

Soit $U_1, U_2, \ldots, U_n$ des variables aléatoires continues avec des fonctions de répartition marginales $F_1, F_2, \ldots, F_n$. Une copule C est une fonction $C : [0,1]^n \to [0,1]$ telle que pour tout $u_1, u_2, \ldots, u_n$ dans l'intervalle [0, 1], la copule C satisfait :

$$C(u_1, u_2, \ldots, u_n) = P(U_1 \leq F_1^{-1}(u_1), U_2 \leq F_2^{-1}(u_2), \ldots, U_n \leq F_n^{-1}(u_n))$$

Théoreme de Sklar :

Soit H la fonction de répartition cumulative conjointe de variables aléatoires $X_1, X_2, \ldots, X_n$, avec des marginales $F_1, F_2, \ldots, F_n$. Alors, il existe une copule C telle que, pour tout $(x_1, x_2, \ldots, x_n)$ dans l'espace des réels, la relation suivante est vérifiée :

$$H(x_1, x_2, \ldots, x_n) = C(F_1(x_1), F_2(x_2), \ldots, F_n(x_n))$$

Présentation de la copule gaussienne - Définition de la copule gaussienne

La copule gaussienne en dimension D est définie comme suit :

$$C(u_1, u_2, \ldots, u_D) = F_D\left(F^{-1}(u_1), F^{-1}(u_2), \ldots, F^{-1}(u_D)\right)$$

où :

- $C(u_1, u_2, \ldots, u_D)$ est la copule gaussienne en dimension D.
- F_D est la fonction de répartition conjointe de la distribution normale multivariée.
- F^{-1} est la fonction quantile de la distribution normale.

Dans notre cas, on s'intéressera au cas en dimension 2. On a alors la formule :

$$C(u, v) = F_2\left(F^{-1}(u), F^{-1}(v)\right)$$

La copule gaussienne $C(u, v)$ est donnée par :

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

où :

- $C(u, v)$ est la répartition de la copule gaussienne,
- r est le coefficient de corrélation de la copule gaussienne
- $F(u)$ est la fonction de répartition d'une loi normale centrée réduite, $F(u) = \int_{-\infty}^{u} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) dz$

Présentation de la copule gaussienne - Mesure de dépendance

La relation entre le coefficient de corrélation de la copule gaussienne (r) et le coefficient de concordance de Kendall (τ) est donnée par :

$$r = \sin\left(\frac{\pi\tau}{2}\right)$$

Inversement, on a alors :

$$\tau = \frac{2}{\pi}\sin^{-1}(r)$$

La relation entre le coefficient de corrélation de la copule gaussienne (r) et le coefficient de corrélation de Spearman (ρ) est donnée par la formule suivante :

$$r = 2 \sin\left(\frac{\rho\pi}{6}\right)$$

On a aussi :

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

Présentation de la copule gaussienne - Densité

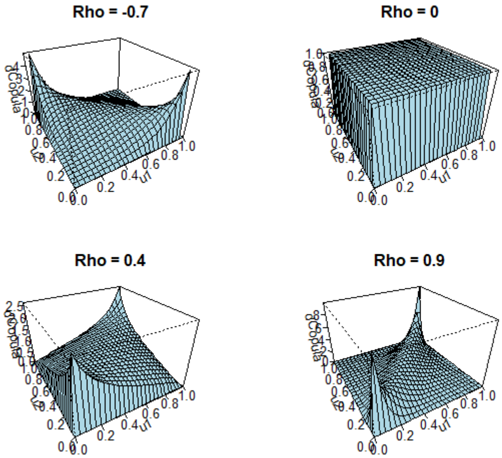


Figure: Representation des densités

Présentation de la copule gaussienne - Dépendance des extrêmes

Dans le cadre de la gestion des risques, on s'interesse à la dépendance entre des événements risqués c'est à dire rares. On définit alors un outil permettant de quantifier une dépendance forte des extrêmes entre deux variables aléatoires.
On définit le paramètre de dépendance forte des extrêmes pour la queue inférieure.

$$\begin{aligned} \lambda_L &= \lim_{u \to 0} P\left(X \leq F_X^{-1}(u) \,|\, Y \leq F_Y^{-1}(u)\right) \\ &= \lim_{u \to 0} \frac{C(u,u)}{u} \end{aligned} \tag{1}$$

De la même manière, on a le paramètre de dépendance forte des extrêmes pour la queue supérieure :

$$\begin{aligned} \lambda_U &= \lim_{u \to 1} P\left(X \geq F_X^{-1}(u) \,|\, Y \geq F_Y^{-1}(u)\right) \\ &= \lim_{u \to 1} \frac{1 - 2u + C(u,u)}{1-u} \end{aligned} \tag{2}$$

Dans le cas de la copule gaussienne , $\lambda_L = 0$ et $\lambda_L = 0$

On définit le coefficient de dépendance faible des extrêmes à gauche par

$$\tilde{\lambda}_L = \lim_{u \to 0} \frac{\log(u)}{\log(C(u,u))} \tag{3}$$

et le coefficient de dépendance faible des extrêmes à droite par

$$\tilde{\lambda}_U = \lim_{u \to 1} \frac{2\log(1-u)}{\log(1-2u+C(u,u))-1}.$$

On dit que la copule C (ou le couple de variables aléatoires (X,Y) qui admet la copule C) présente de la dépendance faible des extrêmes à gauche (respectivement à droite) si $\tilde{\lambda}_L > 0$ (respectivement $\tilde{\lambda}_U > 0$). La copule Gaussienne présente de la dépendance faible des extrêmes puisque $\tilde{\lambda}_L > 0$ et $\tilde{\lambda}_U > 0$.

Présentation de la copule gaussienne - Utilisation de la copule en pratique

Question : Le copule gaussienne est elle adaptée à mon jeu de donnée ?
1) Visualiser les données (Rank Rank Plot ...)
2) Effectuer un test statistique
2.a) On suppose que $H_0 : C \in (C_r)$ est vraie
2.b) On estime r_n :
2.b.1) Par maximum de vraisemblance en maximisant :

$$\sum_{i=1}^{n} log[c_r(F_n(X_i), G_n(Y_i))]$$

avec:

$$F_n(X_i) = \frac{R_i}{n+1}, \, G_n(X_i) = \frac{S_i}{n+1}$$

avec R_i et S_i les rangs des observations des variables X et Y

2.b.2) Par la méthode des moments via le ρ de Spearman empirique :

$$\rho_n = \frac{12}{n(n^2-1)} \sum_{i=1}^{n} R_i S_i - 3\frac{n+1}{n-1}$$

2.c) On compare ensuite la copule estimée C_{r_n} à un estimateur de la copule $\hat{C}_n$ qui vaut :

$$\hat{C}_n(u,v) = \frac{1}{n} \sum_{i=1}^{n} 1_{(\hat{U}_i \leq u, \hat{V} \leq v)},$$

avec $\hat{U} = \frac{R_i}{n+1}$ et $\hat{V} = \frac{S_i}{n+1}$, cela via un test d'ajustement tel que le test de Kolmogorov-Smirnov

Représentation graphique & Application - Graphiques

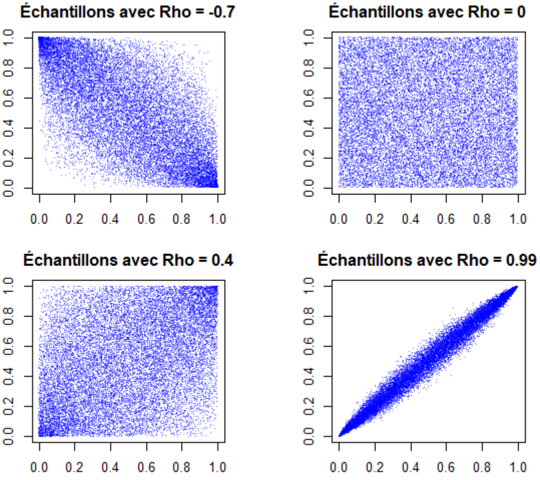


Figure: Graphe des points corrélés

Représentation graphique & Application - Graphiques

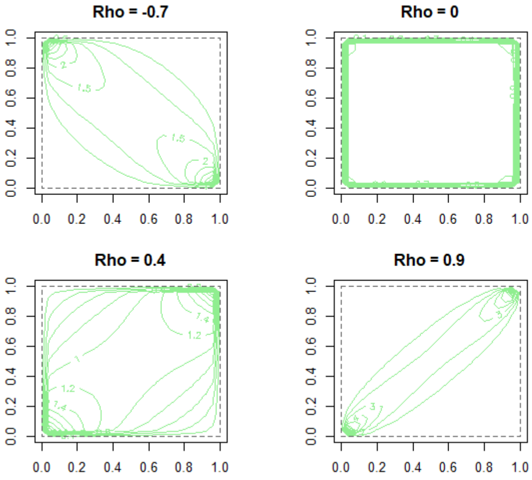
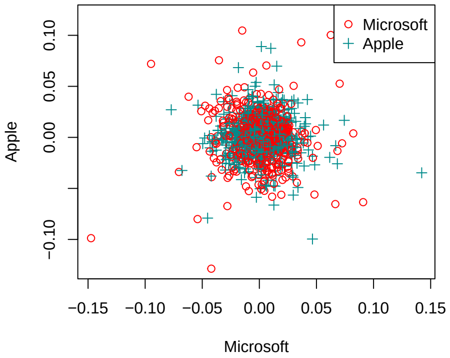


Figure: Lignes de niveau

Application 1 : Rendements Apple/Microsoft

Nous souhaitons étudier la corrélation entre les rendements journaliers des actions Apple et Microsoft. Nous utilisons les rendements journaliers des cinq dernières années.



Application 1 : Rendements Apple/Microsoft

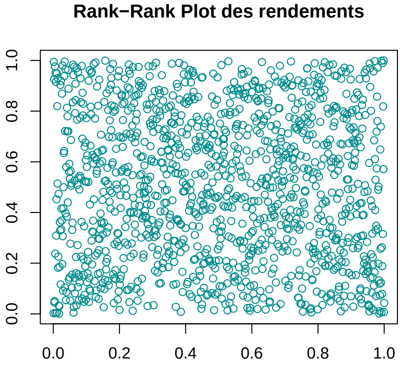


Figure: Rank-Rank Plot des rendements

Application 1 : Rendements Apple/Microsoft

- Estimation du paramètre de la copule par méthode basée sur les moments et la formule du τ de Kendall. (ici $r = -0.026$)

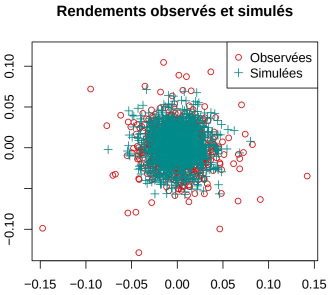


Figure: Scatter Plot Observations/Simulations

Application 2 : Rendements ETF SP/DJ

Nous souhaitons étudier la corrélation entre les rendements journaliers des ETF (Exchange Traded Fund) SP500 et DJ30. Nous utilisons les rendements journaliers des cours de clôtures entre 15/12/2011 et 31/12/2018.

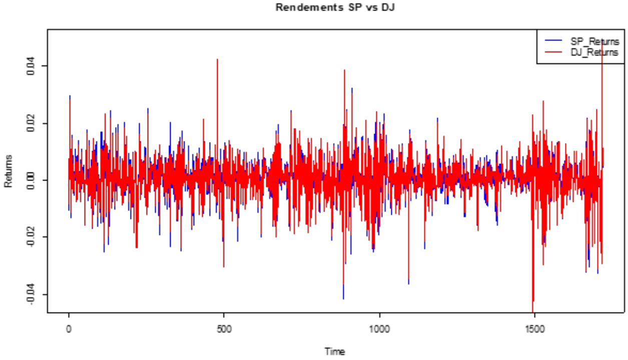


Figure: Plot rendements SP et DJ en fonction du temps

Application 2 : Rendements ETF SP/DJ

- Estimation du paramètre de la copule par méthode des moments et la formule du ρ de Pearson. (Ici $r = 0.96$)

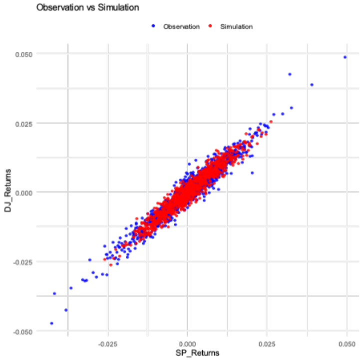


Figure: Scatter Plot Observations/Simulations

Conclusion

Dans cette étude, nous avons présenté les copules gaussiennes, leurs propriétes et leurs capacités à modéliser des rendements journaliers.

- En raison de leur simplicité et de leur adaptabilité, elles semblaient être un choix approprié pour cette tâche,
- Preuve empirique : copule gaussienne non adaptée à la modélisation des extrêmes,
- Pour les observations extrêmes : les copules archimédiennes ou modèles adaptés aux queues de distribution.

La modélisation des dépendances est cruciale, le choix judicieux de la copule appropriée peut avoir un impact significatif sur les résultats et les prévisions.