Universite Claude Bernard Lyon 1 **Institut de Science Financière et d'Assurances**

ESTIMATION DE COPULES

✧✧✧

PROJET

Estimation de copule

entre les actions BMW et VOLKSWAGEN

✧✧✧

Auteurs **: NGUYEN Thi Hanh**

LAI Phuong Hoa

Lyon, Avril 2023

Contents

Introduction

L'estimation de la copule entre les cours de bourse de deux entreprises est importante pour la gestion des risques financiers et l'optimisation de portefeuille. En estimant la copule, on peut comprendre la structure de dépendance entre les deux actions, qui n'est pas reflétée par les seules distributions marginales.

Prenons l'exemple de deux actions ayant des distributions marginales similaires, mais ayant des structures de dépendance différentes. Dans cette situation, il peut ne pas être optimal de diversifier simplement en investissant dans les deux actions car elles peuvent être fortement corrélées et ne pas offrir de réels avantages en termes de diversification.

En estimant la copule, on peut quantifier le degré de dépendance entre les deux actions et utiliser cette information pour optimiser son portefeuille en sélectionnant des actions moins corrélées et offrant de plus grands avantages de diversification.

Dans ce projet, nous analysons la dépendance entre les deux cours boursiers de deux grands constructeurs automobiles mondiaux, le Groupe BMW (BMW) et le Groupe VOLKSWAGEN (VOW3).

Pour cette analyse, nous avons recueilli les prix quotidiens de 2 actions sur la période allant du 1er janvier 2021 au 31 décembre 2022 à partir de `https://finance.yahoo.com`. Ceci afin de nous assurer que nous disposons d'une série temporelle suffisamment longue qui capture les informations historiques sur la structure de dépendance entre les deux actifs pour estimer plus précisément la copule, ainsi que d'obtenir une estimation fiable des paramètres de la copule.

Données de préparation

Dans un premier temps, nous calculons le taux de rendement des deux actions.

Le calcul des rendements des cours des actions de deux sociétés est important pour estimer la copule car il nous permet d'analyser la performance relative des deux actions et leur structure de dépendance. Les rendements représentent la variation en pourcentage de la valeur d'une action sur une période de temps et sont calculés comme la différence entre le cours de l'action à deux moments différents divisé par le cours initial de l'action. En utilisant les rendements au lieu des cours des actions, nous pouvons normaliser les données et faciliter la comparaison des performances des deux actions. Ces données standardisées sont ensuite utilisées pour estimer la copule, qui mesure la structure de dépendance entre les deux stocks.

La formule pour calculer le rendement d'une action est la suivante :

$$\text{Rendement} = \frac{\text{Prix(t)} - \text{Prix(t-1)}}{\text{Prix(t-1)}}$$

En pratique, les rendements sont souvent exprimés sous forme de rendements logarithmiques, ce qui présente certains avantages dans l'analyse statistique. La formule est:

$$\log \text{Rendement} = \log \frac{\text{Prix(t)}}{\text{Prix(t-1)}}$$

La deuxième étape consiste à vérifier la stationnarité : testez la stationnarité des rendements à l'aide de tests statistiques tels que le test Augmented Dickey-Fuller (ADF). La stationnarité est une hypothèse importante pour les modèles de copule, car elle garantit que les propriétés statistiques des données ne changent pas dans le temps.

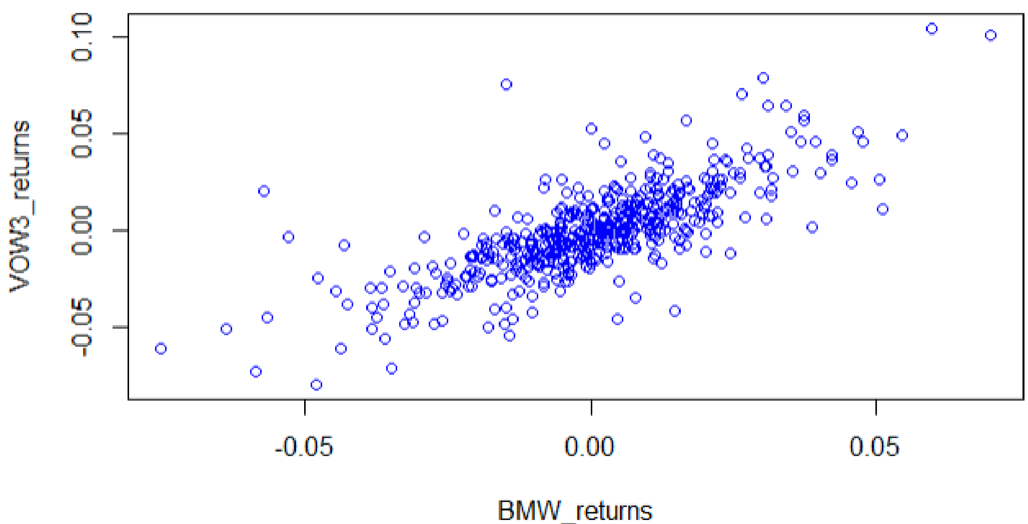
Nous avons les résultats du test statistique ADF des retours de BMW et VOW3 comme ci-dessous :

```
        Augmented Dickey-Fuller Test

data:  BMW_returns
Dickey-Fuller = -7.2949, Lag order = 7, p-value = 0.01
alternative hypothesis: stationary
```

```
	Augmented Dickey-Fuller Test

data:  VOW3_returns
Dickey-Fuller = -7.9858, Lag order = 7, p-value = 0.01
alternative hypothesis: stationary
```

Valeur $p - value = 0,01 < 0,05$, on peut donc conclure que 2 séries sont stationnaires.

L'étape suivante consiste à vérifier la dépendance. Le test de dépendance entre les rendements des deux actifs à l'aide de tests statistiques tels que le coefficient de corrélation de Pearson ou le coefficient de corrélation de rang de Spearman.



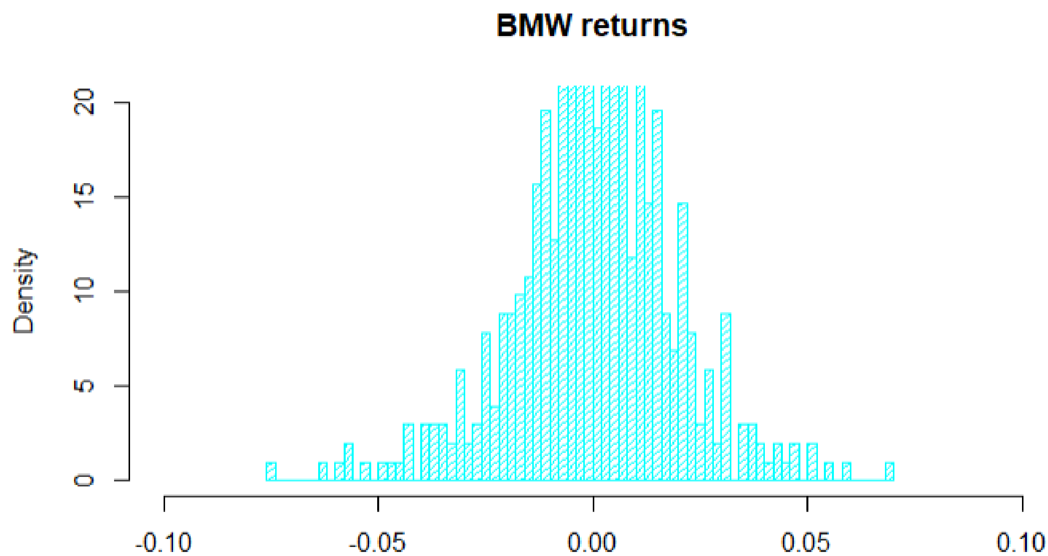
```
	Spearman's rank correlation rho

data:  BMW_returns and VOW3_returns
S = 5424530, p-value < 2.2e-16
alternative hypothesis: true rho is not equal to 0
sample estimates:
      rho 
0.7560772 


	Pearson's product-moment correlation

data:  BMW_returns and VOW3_returns
t = 26.379, df = 509, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.7207320 0.7943404
sample estimates:
      cor 
0.7599623 
```

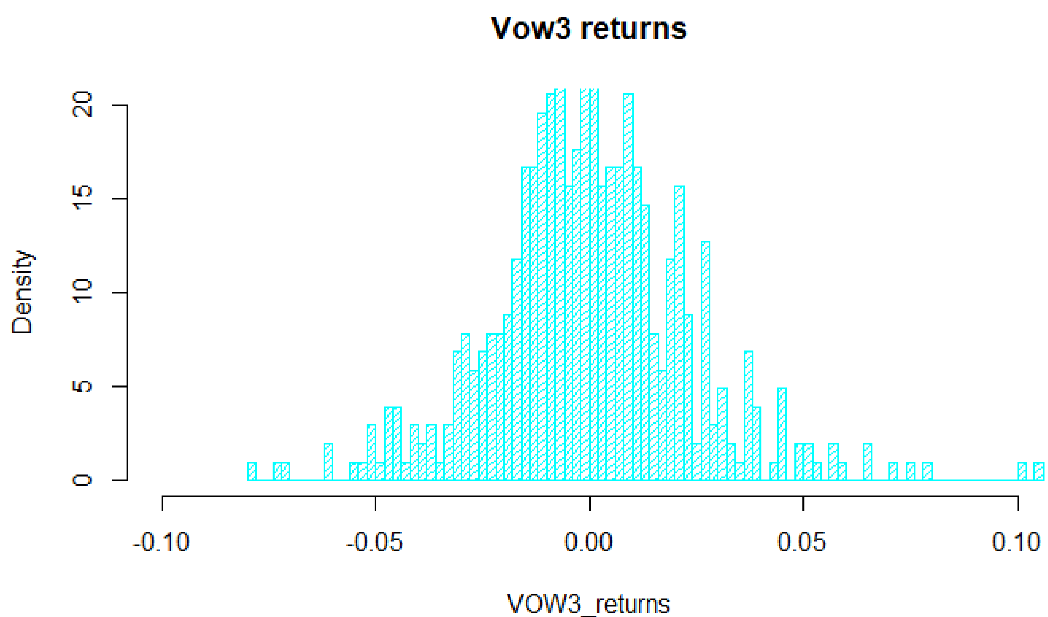
Ainsi, les rendements de BMW et de VOW3 dépendent l'un de l'autre.

Le choix de la famille de copules

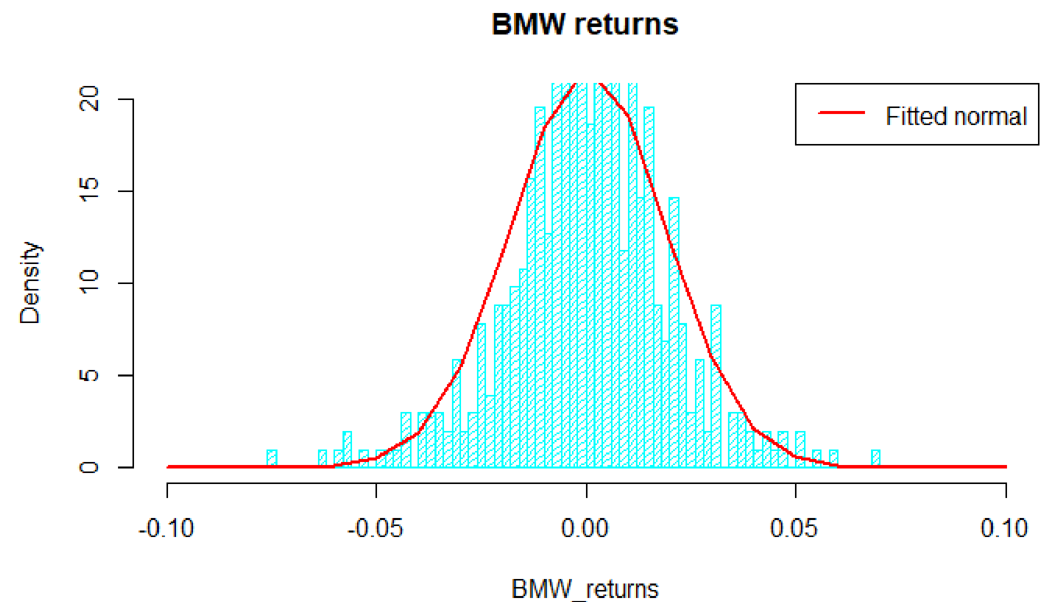
3.1 La distribution de BMW et VOW3

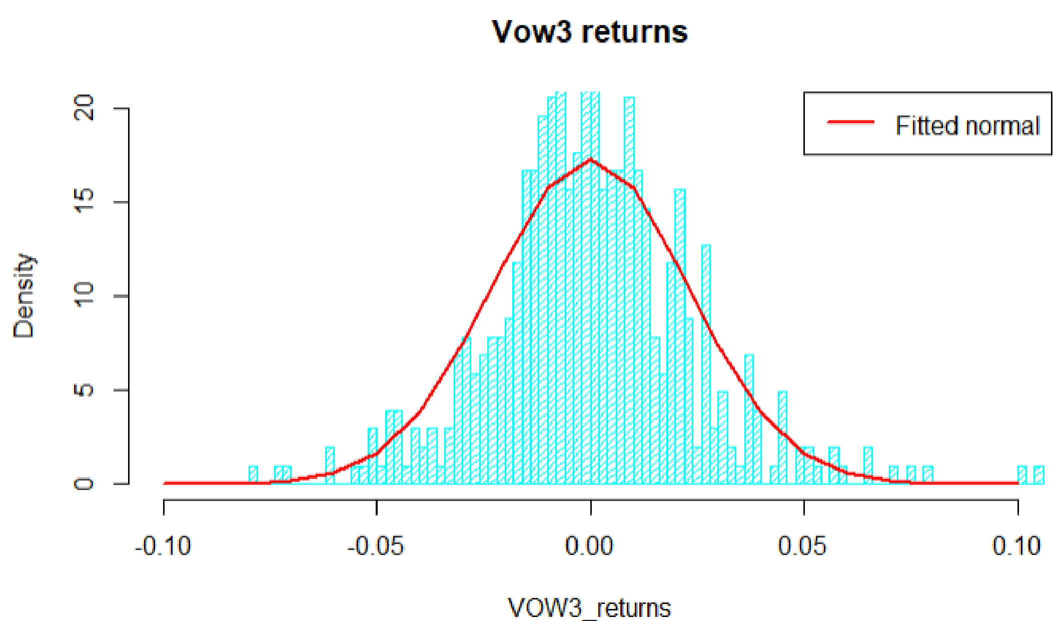
Tout d'abord, étudions les marginaux séparément. Pour ce faire, une vérification rapide de l'histogramme des deux variables pourrait fournir des informations.





Un bon aperçu du graphique de chaque variable est important afin de comprendre quelle distribution pourrait être la mieux adaptée. Il semble qu'une distribution normale puisse faire à la fois pour le retour BMW et le retour VOW3. En supposant que l'hypothèse de distribution est correcte, il ne reste plus qu'à estimer les paramètres. Estimons les paramètres, échantillonnons les observations à partir des distributions ajustées et effectuons une comparaison visuelle.



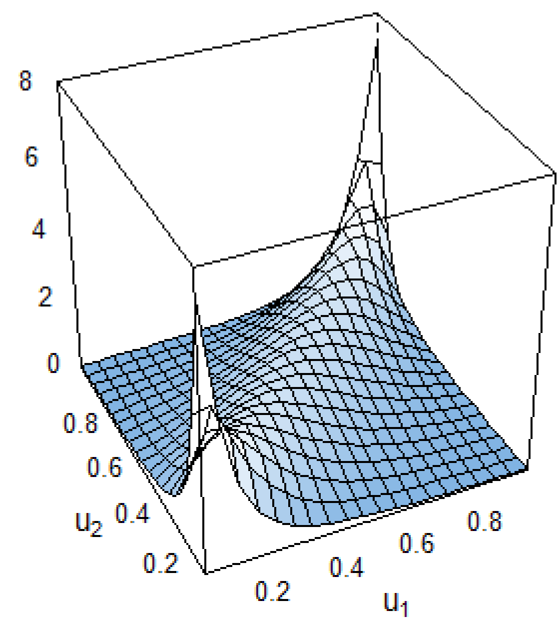


Nous pouvons voir en cyan les données observées et en rouge les données simulées pour BMW et VOW3. Cela ressemble à un bon match.

3.2 Le choix de la famille de copules

Sur la base de la structure de dépendance entre les deux propriétés, nous choisissons la famille de copules appropriée en tant que t-copule avec le paramètre rho = 0,79, et le degré de liberté est de 3,89, approximativement égal à 4.

```
Bivariate copula: t (par = 0.79, par2 = 3.89, tau = 0.58)
```



Estimation pour des modèles de copules

4.1 Estimation par l'approche paramétrique

La fonction BiCopSelect évalue les paramètres de copule, mais une fois la copule sélectionnée, nous l'ajustons à nouveau à l'aide de la fonction fitCopula(). Ainsi, les paramètres de la copule ajustée sont estimés pour être identiques à ceux de la fonction BiCopSelect(). En outre, le tau de Kendall est le même que celui calculé à partir des valeurs observées de x et y.

```
    rho.1        df
0.7870951 3.8937693
[1] 0.579839
```

Il y a 2 méthodes existent pour estimer les paramètres inconnus: Maximum de vraisemblance et méthode IFM (Inference function for margins). Notre projet utilise la méthode du maximum de vraisemblance.

Il est supposé que nos deux ensembles de prix d'actions pour BMW et VOLKSWAGEN suivent la distribution normale, comme indiqué précédemment.

Les paramètres des données BMW selon la distribution normale sont:

$$\mu_1 = \text{mean(BMW returns)} = 0.58$$

$$\sigma_1 = \text{sd(BMW returns)} = 0.00052$$

Et les paramètres des données VOLKSWAGEN sont également normalement distribués comme:

$$\mu_2 = \text{mean(VOW3 returns)} = -0.000032$$

$$\sigma_2 = \text{sd(VOW3 returns)} = 0.0231$$

Les paramètres à estimer sont donc les paramètres des marges $(\mu_1, \sigma_1) = (0.58, 0.00052)$ et $(\mu_2, \sigma_2) = (-0.000032, 0.0231)$ et le paramètre $\theta = 0.79$ de la copule.

Afin d'obtenir une estimation des paramtrèes par MV, nous utilisons la fonction fitMvdc, qui demande des valeurs initiales. On obtient le résultat comme indiqué ci-dessous:

```
[1] 2759.354
Call: fitMvdc(data = mat_returns, mvdc = my_dist, start = para_list)
Maximum Likelihood estimation based on 511 2-dimensional observations.
Copula:  tCopula
Margin 1 :
  m1.mean     m1.sd
0.0006996 0.0190324
Margin 2 :
  m2.mean     m2.sd
-0.000106  0.024031
Copula:
 rho.1
0.8033
The maximized loglikelihood is 2760
Optimization converged
```

4.2 L'approche semiparamétrique

Nous utilisons la méthode CML (Canonical Maximum Likelihood) (Oakes, 1994) pour estimer. L'estimateur du maximum de pseudo-vraisemblance est fondée sur les rangs, puisque:

$$F_n(X_i) = \frac{R_i}{n}; \; G_n(Y_i) = \frac{S_i}{n}$$

pour tout $i = 1, ..., n$

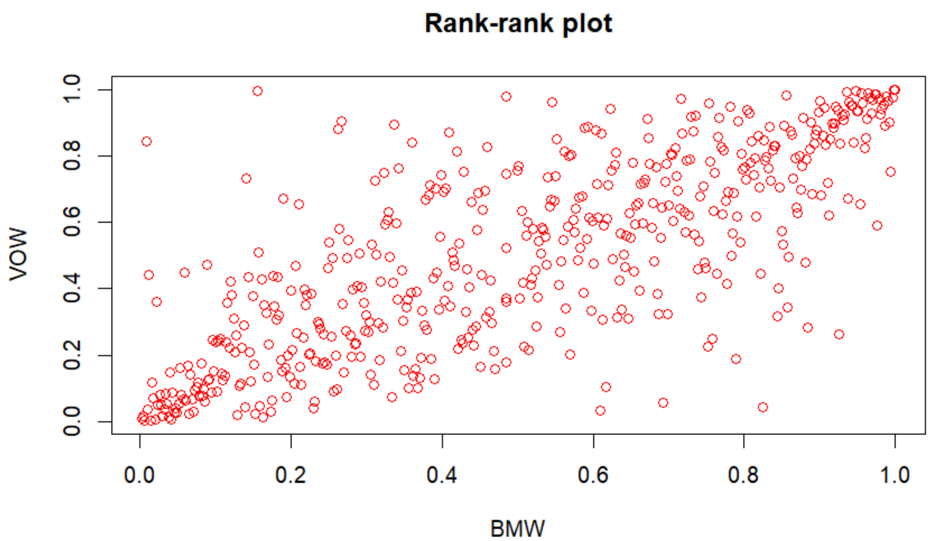
On obtient le résultat comme indiqué ci-dessous:

```
Fit based on "maximum likelihood" and 511 2-dimensional observations.
Copula: tCopula
 rho.1     df
0.7885 4.0000
The maximized loglikelihood is -2.247e+307
Optimization converged
```

4.3 L'approche nonparamétrique

4.3.1 Rank-rank plots

Nous optons pour la copule de Student pour nos deux jeux de données. Notre méthode pour ce diagramme rang-rang consiste à utiliser des observations fictives $(\frac{R_i}{n}, \frac{S_i}{n})$, qui sont des valeurs de substitution proches des paires non observées $(U_i, V_i) = (F(X_i), G(Y_i))$, pour former un échantillon de la copule de Student.

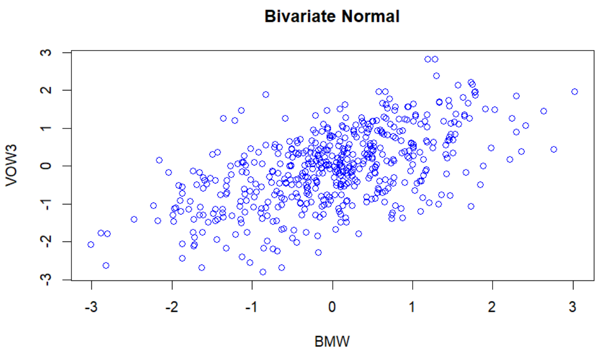


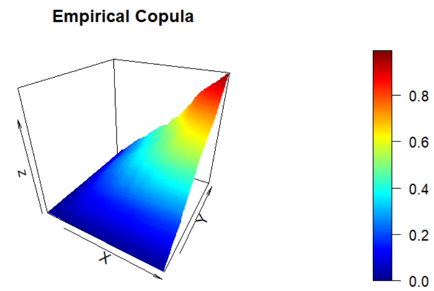
C'est rank-rank plots pour un échantillon de taille $n = 511$ d'une copule de Student $(\theta = 0.79)$ avec marges $F = \mathcal{N}(\mu_1, \sigma_1)$, $G = \mathcal{N}(\mu_2, \sigma_2)$.

4.3.2 Copule empirique

La copule empirique peut être comprise comme une mesure de la force et de la direction de la dépendance entre les variables aléatoires. Cette loi de probabilité place une masse de 1/n en chaque paire de rangs normalisés.

C'est la copule empirique pour un échantillon de taille $n = 511$ d'une copule de Student $(\theta = 0.79)$. Dans ce cas, les marges sont $F = \mathcal{N}(\mu_1, \sigma_1)$, $G = \mathcal{N}(\mu_2, \sigma_2)$.



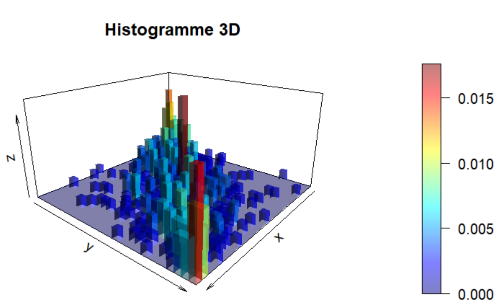


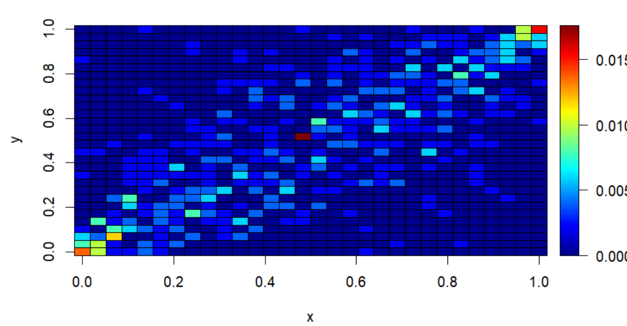
D'après la figure ci-dessus, nous pouvons observer qu'il s'agit d'une copule positive, indiquant une corrélation positive entre les deux variables. Autrement dit, lorsque le prix d'une action augmente, le prix de l'autre action a tendance à augmenter également.

4.3.3 Histogramme 3D

Le résultat de l'histogramme 3D nous donnerait une représentation visuelle de la distribution conjointe des rendements de deux prix des actions. Nous pouvons l'utiliser pour identifier les zones de forte densité, où les deux prix des actions ont tendance à coexister fréquemment, ainsi que les zones de faible densité, où les deux prix des actions ont tendance à coexister moins fréquemment.

Ci-dessous, un histogramme 3D pour un échantillon de taille $n = 511$ d'une copule de Student ($\theta = 0.79$). Dans ce cas, les marges sont $F = \mathcal{N}(\mu_1, \sigma_1)$, $G = \mathcal{N}(\mu_2, \sigma_2)$.





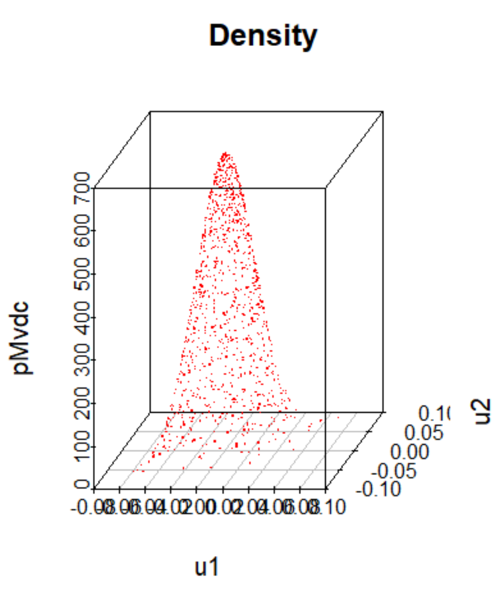
La validation de le modèle de copule

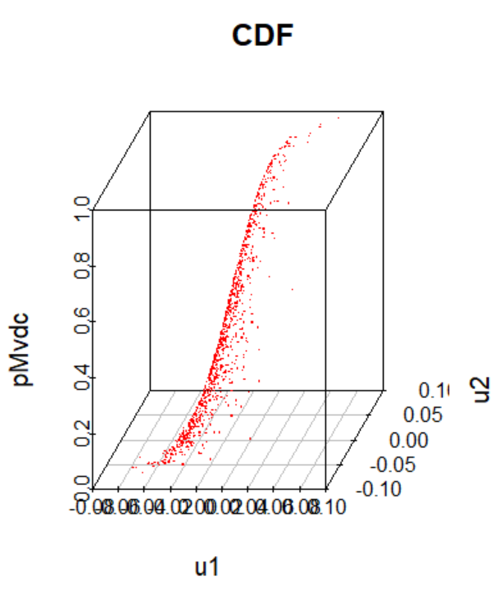
Une fois que la copule est ajustée, nous pouvons procéder à un test pour déterminer si l'ajustement est satisfaisant ou non. Nous utilisons la fonction gofCopula() pour réaliser ce test GOF (Goodness of fit test). On obtient les résultats ci-dessous:

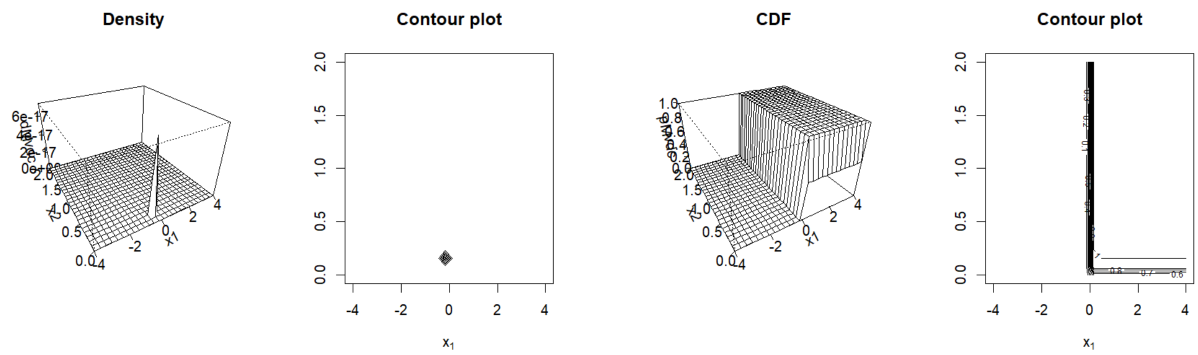
```
        Parametric bootstrap-based goodness-of-fit test of t-copula, dim. d = 2, with
        'method'="Sn", 'estim.method'="ml":

data:  x
statistic = 0.0098619, parameter = 0.78789, p-value = 0.9158
```

Selon les résultats obtenus, la p-valeur est de 0,91, ce qui est considérablement plus grande que le niveau de signification de 0,05. Par conséquent, nous ne pouvons pas rejeter l'hypothèse nulle et conclure que la copule convient aux données.

Ensuite, nous procédons à la construction d'une distribution à deux variables en utilisant la copule. Après avoir sélectionné et ajusté la copule aux données ainsi que sélectionné les paramètres appropriés, nous pouvons tenter de modéliser la relation générale entre les variables.

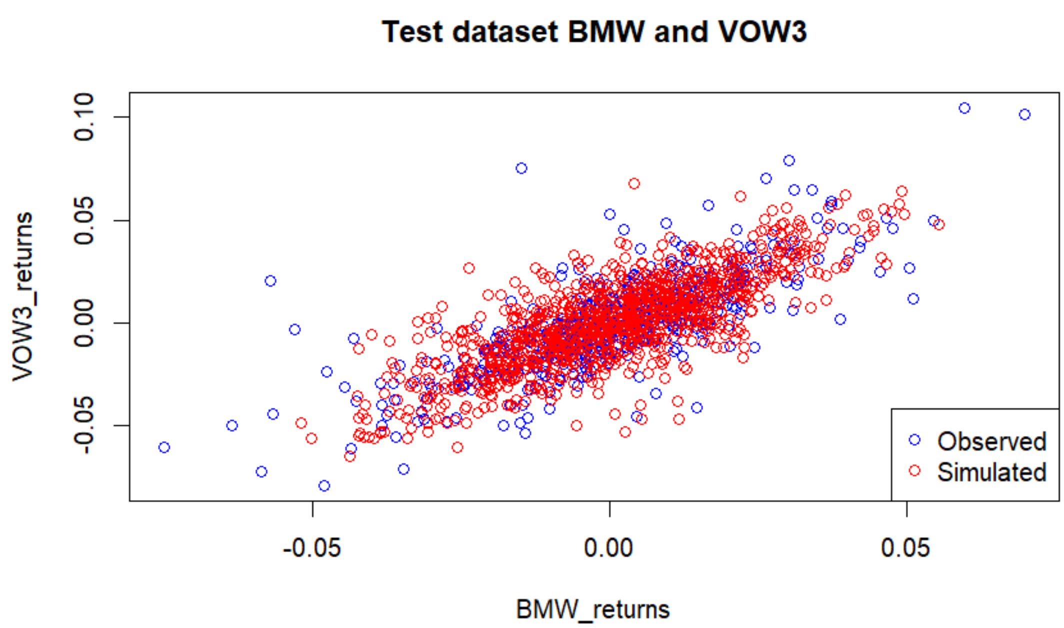






On peut échantillonner des observations à partir de la distribution commune estimée et vérifier comment elles se comparent à la distribution observée.

Enfin, nous pouvons observer l'échantillon provenant d'une distribution complètement nouvelle à deux variables. Il convient de noter que le tau de Kendall reste le même pour les données simulées ($\tau = 0.58$). Les informations sur la structure de corrélation ont été capturées par la copule indépendamment des marginales.



Conclusion

En utilisant une variété de méthodes pour estimer le modèle de copule, nous pouvons conclure que le modèle de copule qui correspond aux actions BMW et VOLKSWAGEN est la t-copule. De plus, il existe une forte dépendance positive entre les deux actions, et cette dépendance est lourde. Cela signifie que des événements extrêmes sont plus susceptibles de se produire que dans une distribution normale et que la corrélation entre les deux actions peut être plus prononcée en période de tension sur le marché. Ces résultats ont des implications pour la diversification du portefeuille et les stratégies de gestion des risques.