



---

## Statistiques des risques multiples

---

2023 - 2024

**Authors:**

GANIYU Isaac

JUMBONG Junior

**Enseignant:**

ZERRAD Ali

March 2024

# Contents

<b>List of figures</b>	<b>ii</b>
<b>List of tables</b>	<b>ii</b>
<b>Introduction</b>	<b>1</b>
<b>1 Présentation des données</b>	<b>2</b>
1.1 Probabilité de défaut instantanée des deux créances pour chaque maturité.	2
1.2 Taux de recouvrement de chaque créance . . . . .	2
1.3 Analyse exploratoire univariée et étude des facteurs de risques . . . . .	4
1.3.1 Analyse descriptive univariée . . . . .	4
1.3.2 Etude des lois marginales . . . . .	6
1.4 Etude de la dépendance entre les facteurs de risque sur la base de critères non paramétriques. . . . .	7
1.4.1 Etude graphique . . . . .	7
<b>2 Méthodologie : Calcul de la CreditVaR</b>	<b>9</b>
2.1 Modélisation de la perte . . . . .	9
2.2 Présentation de la copule . . . . .	10
2.2.1 Copules elliptiques . . . . .	10
2.2.2 Copules archimédiennes . . . . .	11
2.3 Méthodes d'inférence statistique . . . . .	12
2.4 Critères d'ajustement . . . . .	12
2.4.1 Kendall plot . . . . .	12
2.4.2 Tests d'adéquation et bilan comparatif des modélisations . . . . .	13
<b>3 Résultats</b>	<b>14</b>
3.1 Estimation des paramètres . . . . .	14
3.2 Kendall plot . . . . .	15
3.3 Comparaison du dépendogramme de l'échantillon à celui des copules théoriques	16
3.4 Tests d'adéquation et bilan comparatif des modélisations . . . . .	16
3.5 Résultats de la CreditVaR : simulation de la distribution par Monte Carlo	17
<b>Conclusion</b>	<b>18</b>
<b>References</b>	<b>a</b>

## List of Figures

1	Densité de probabilité des taux de recouvrement de BNP et de SG . . . . .	4
2	Rendement des deux entreprises . . . . .	5
3	QQ-plot de la loi de student à gauche et la loi normale inverse gaussienne à droite de BNP . . . . .	7
4	Ajustement linéaire et dépendogramme . . . . .	8
5	Kendall plot des copules elliptiques . . . . .	15
6	Kendall plot des copules archimédiennes . . . . .	15
7	Comparaison des dépendogrammes des copules elliptiques à celui de l'échantillon	16
8	Comparaison des dépendogrammes des copules archimédiennes à celui de l'échantillon . . . . .	16
9	Distribution de la creditVaR par Monte Carlo avec niveau de confiance de 99% . . . . .	17
10	Adéquation des rendements de BNP à une loi normale . . . . .	b
11	Adéquation des rendements de BNP à une loi de Student . . . . .	b
12	Adéquation des rendements de BNP à une loi normale inverse gaussienne .	c

## List of Tables

1	Probabilités de défaut implicites . . . . .	2
2	Paramètres de la loi Bêta de recouvrement estimés par la méthode des moments . . . . .	3
3	Statistiques descriptives des rendements pour BNP et SG . . . . .	5
4	P-valeurs du test de Kolmogorov-Smirnov pour différentes distributions sur les rendements de BNP et SG . . . . .	6
5	Paramètres estimés pour BNP et SG . . . . .	7
6	Mesures de dépendances entre BNP et SG . . . . .	8
7	Copules archimédiennes bivariées les plus courantes. . . . .	11
8	Estimation des paramètres des différentes copules par la méthode IFM. . .	14
9	P-value des copules . . . . .	17

# Introduction

La Value at Risk (VaR) est un indicateur important utilisé par les institutions financières pour évaluer le risque de crédit d'un portefeuille. La VaR représente la perte maximale potentielle d'un portefeuille sur un horizon donné, avec un niveau de confiance spécifié. En d'autres termes, elle résume la perte maximale sur un horizon cible qui ne sera pas dépassée avec un niveau de confiance donné. Son utilisation demande la caractérisation de la fonction de perte du portefeuille pour l'horizon cible. Si le portefeuille est constitué d'au moins deux créances, il est indispensable de prendre en compte les dépendances entre ces créances, en particulier si elles appartiennent au même secteur, pour mesurer le risque de contagion. En effet, les corrélations entre les différents actifs du portefeuille jouent un rôle essentiel dans l'évaluation du risque global.

On utilise souvent et par abus le coefficient de corrélation linéaire de Pearson comme indicateur de dépendance. Toutefois, son efficacité ne dépend que de la structure de dépendance linéaire ou de l'univers gaussien, un cadre d'analyse rare en finance. Lorsqu'on est confronté à cette situation, on utilise habituellement d'autres indicateurs de dépendance basés sur les discordances et les concordances observées dans un échantillon.

Cette première approche d'évaluation de dépendance à l'aide d'un indicateur est intéressante, mais l'analyse gagnerait en profondeur avec une méthode capable de considérer intégralement la structure de ce lien. La copule répond à ce besoin. L'avantage des copules est qu'elles permettent de relier des marginales suivant des lois différentes. Le but de ce document est d'évaluer le risque de crédit au moyen d'une Value at Risk de niveau de confiance 99% sur un portefeuille composé de deux actifs, issus du secteur bancaire. De ce fait, le risque de perte qui en résulte est conséquemment élevé. Il est donc crucial de considérer la dépendance entre ces deux actifs.

L'organisation de ce document est la suivante : Tout d'abord, nous abordons la description des données. Ensuite, nous détaillons la méthodologie de calcul de la CreditVar. Enfin, dans la dernière section, nous présentons nos résultats finaux. Nous utiliserons le logiciel python pour nos analyses.

# 1 Présentation des données

Le portefeuille sur lequel nous travaillons est constitué de deux actifs. Le premier est une obligation senior BNP avec un taux de recouvrement moyen ( $R_1$ ) de 60% et une volatilité ( $\sigma_1$ ) de 15%, tandis que le second est une obligation junior (ou subordonnée) de Société Générale, avec un taux de recouvrement moyen ( $R_2$ ) de 30% et une volatilité ( $\sigma_2$ ) de 25%. Il est important de noter que nous supposons l'absence de dépendance entre les taux de recouvrement de ces deux actifs, une hypothèse clé qui influence notre analyse du risque associé à ce portefeuille et qui simplifie l'analyse. Si on considère en plus que les spreads de CDS de BNP et SG, pour une maturité de 4 ans, sont respectivement de 100 bp (points de base) et de 120 bp, et pour une maturité de 5 ans, de 120 bp et de 150 bp ; les CDS étant évalués avec un taux de recouvrement égal à 40%, on peut, à travers le triangle de crédit, extraire les probabilités de défaut implicites pour différentes maturités.

## 1.1 Probabilité de défaut instantanée des deux créances pour chaque maturité.

Le triangle de crédit est une notion très importante en finance qui lie la probabilité de défaut ( $\lambda$ ), le spread de crédit ( $S$ ) et le taux de recouvrement ( $R$ ). Il est donné par la formule  $S = \lambda(1 - R)$ . De cette formule, on obtient que  $\lambda = \frac{S}{1-R}$ . Ainsi, nous avons le tableau (table 1) avec les probabilités de défaut de chaque créance pour différentes maturités.

	4 ans	5 ans
BNP	0.017	0.02
SG	0.02	0.025

Table 1: Probabilités de défaut implicites

Nous pouvons tirer de ce tableau que la probabilité de défaut de BNP est plus faible que celle de Société Générale pour chaque maturité.

En considérant les caractéristiques des obligations de notre portefeuille, nous pouvons caractériser les distributions des taux de recouvrement de chaque créance.

## 1.2 Taux de recouvrement de chaque créance

Le taux de recouvrement est considéré ici comme aléatoire et distribué suivant une loi Bêta de densité :

$$f(x) = \frac{1}{B(a, b)} x^{a-1} (1-x)^{b-1} \mathbf{1}_{[0,1]}(x)$$

où  $B(a, b)$  est la fonction bêta, définie par :

$$B(a, b) = \int_0^1 t^{a-1}(1-t)^{b-1} dt.$$

La moyenne ( $\mu$ ) et la variance ( $\sigma^2$ ) sont données par :

$$\mu = \frac{a}{a+b}$$

$$\sigma^2 = \frac{ab}{(a+b)^2(a+b+1)}$$

Les paramètres  $a$  et  $b$  sont calibrés par la méthode des moments et donnés par :

$$\hat{a} = \frac{\mu^2(1-\mu)}{\sigma^2} - \mu$$

$$\hat{b} = \frac{\mu(1-\mu)^2}{\sigma^2} - (1-\mu)$$

.

Les paramètres estimés  $\hat{a}$  et  $\hat{b}$  de nos deux créances sont donnés dans le tableau ci-dessous (tableau 2) :

	$\alpha$	$\beta$
BNP	5.800	3.867
SG	0.708	1.652

Table 2: Paramètres de la loi Bêta de recouvrement estimés par la méthode des moments

Nous pouvons conclure à partir de ces résultats que la courbe de recouvrement de BNP est asymétrique à droite ( $\alpha > \beta$ ) et celle de société générale est asymétrique à gauche ( $\alpha < \beta$ ). Nous pouvons représenter les densités des distributions des taux de recouvrement de BNP et de SG sur un même graphique (figure 1) pour une comparaison. En effet, on constate que le taux de recouvrement de BNP plus élevé que celui de SG.

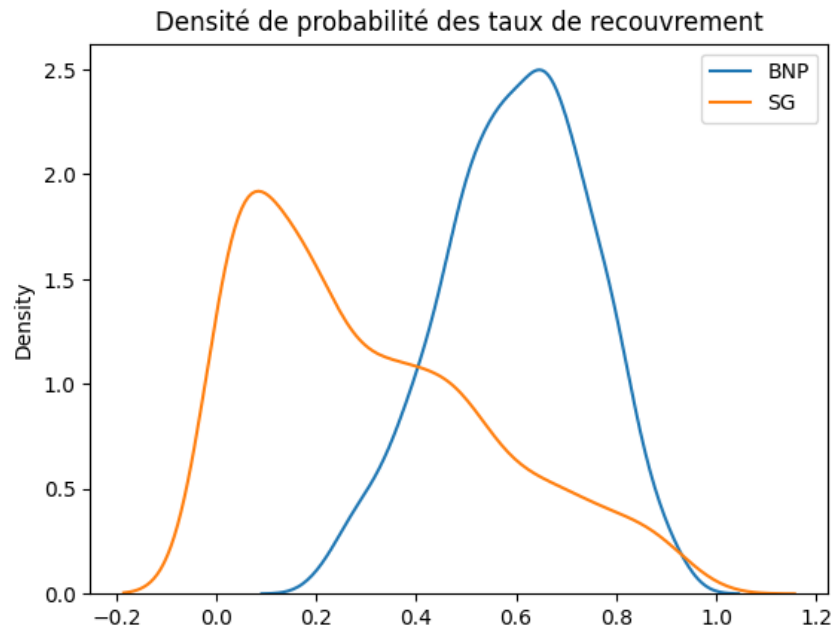


Figure 1: Densité de probabilité des taux de recouvrement de BNP et de SG

Nous allons maintenant nous intéresser à l'historique des données actions des deux entreprises.

## 1.3 Analyse exploratoire univariée et étude des facteurs de risques

### 1.3.1 Analyse descriptive univariée

Nous avons également analysé le prix des actions des deux entités dans ce portefeuille, transformant ces valeurs en rendements ( $r_t = \frac{s_t - s_{t-1}}{s_{t-1}}$ ) pour assurer la stationnarité des données et évaluer les performances des deux entreprises (voir Table 3). Le nombre total de mesures de rendement s'élève à 999. La moyenne des rendements est légèrement négative pour les deux entreprises, étant plus basse pour BNP (-0,059 %) que pour SG (-0,047 %). La volatilité des rendements est plus élevée chez BNP (2,43 %) que chez SG (2,06 %) (voir graphe 2).

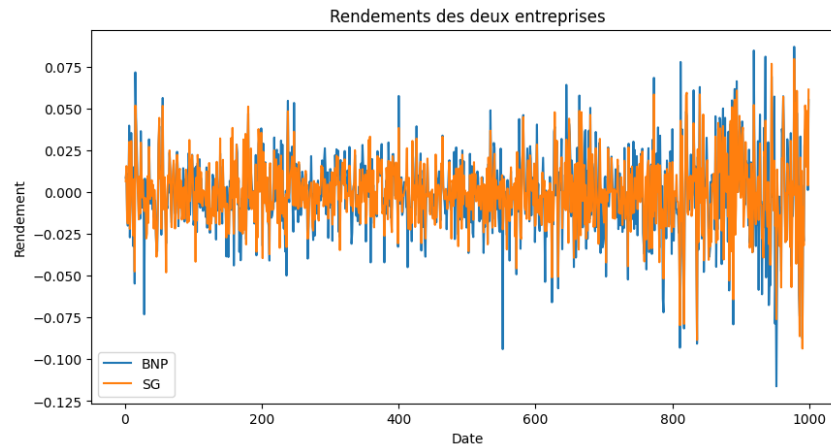


Figure 2: Rendement des deux entreprises

La valeur minimale des rendements pour la période étudiée est de -0,12% pour BNP et de -0,093% pour SG, ce qui démontre que BNP a enregistré au moins un jour une perte proportionnelle plus importante que SG. Concernant le premier quartile, les valeurs sont de -0,014 % pour BNP et de -0,012 % pour SG, soulignant que 25 % des rendements de chaque entreprise sont inférieurs à ces seuils, caractérisant ainsi les pertes subies.

La médiane des rendements est de -0,036 % pour BNP et de -0,055 % pour SG. Ces médianes, négatives et proches de zéro, révèlent que les rendements sont distribués autour de cette valeur, avec une tendance vers des pertes pour les deux entreprises.

Pour évaluer les gains, l'examen du troisième quartile révèle des valeurs de 0,013% pour BNP et de 0,011% pour SG, indiquant que 75 % des rendements sont inférieurs à ces niveaux, ce qui reflète les gains typiques sur les trois quarts des jours observés. Les performances maximales atteignent 0,088% pour BNP et 0,079% pour SG, illustrant le meilleur rendement enregistré pour chacune des entreprises durant la période étudiée.

	BNP	SG
count	999	999
mean	-0.059	-0.047
std	0.024	0.020
min	-0.17	-0.094
25%	-0.014	-0.012
50%	-0.036	-0.055
75%	0.013	0.011
max	0.088	0.079

Table 3: Statistiques descriptives des rendements pour BNP et SG



### 1.3.2 Etude des lois marginales

Nous avons étudié l'appartenance des lois marginales des deux créances à une famille de lois paramétriques : Loi normale, loi de Student et loi normale inverse gaussienne. Les paramètres ont été estimés par la méthode du maximum de vraisemblance, et la qualité d'ajustement a été validée graphiquement par les QQ-plots et aussi à travers les tests statistiques non paramétriques de Kolmogorov-Smirnov. Le tableau ci-dessous présente les résultats du test de Kolmogorov-Smirnov pour l'adéquation des différentes lois citées ci-dessus (Table 4) sur les rendements de BNP et de SG.

Distribution	P-valeur BNP	P-valeur SG
Loi gaussienne	0.0087	0.000
Loi de Student	0.44	0.519
Loi normale inverse gaussienne	0.48	0.686

Table 4: P-valeurs du test de Kolmogorov-Smirnov pour différentes distributions sur les rendements de BNP et SG

Nous allons utiliser un seuil de significativité de 1 %. Concernant les rendements de BNP par exemple, les p-value du test de Kolmogorov-Smirnov pour l'ajustement à une loi de Student et une distribution normale inverse gaussienne sont respectivement de 0,44 et 0,48 et sont supérieures à 1 %, ce qui indique que nous n'avons pas suffisamment de preuves pour rejeter l'hypothèse selon laquelle les données suivent ces deux lois. La p-value du test d'adéquation à une loi normale est de 0,0087, inférieure à 1 %, suggérant que les données ne semblent pas suivre une loi normale.

Nous avons finalement choisi la loi normale inverse gaussienne comme le meilleur ajustement aux données. En effet, le QQ-plot montre une meilleure adéquation aux extrémités inférieures comparé à ce que montre l'analyse basée sur la loi de Student (figure 3 pour les rendements de BNP). Ainsi, elle modélise mieux les queues de distributions des rendements, dont les pertes. Par la suite, nous supposons que les distributions marginales des rendements de BNP Paribas (BNP) et de Société Générale (SG) suivent une loi normale inverse gaussienne. Les paramètres de cette distribution sont estimés par la méthode du maximum de vraisemblance (Table 5). La loi normale inverse est caractérisée par quatre paramètres : le paramètre de localisation  $\mu$ , le paramètre  $\beta$  d'asymétrie, le paramètre d'échelle  $\delta$ , et le paramètre de queue  $\alpha$ .

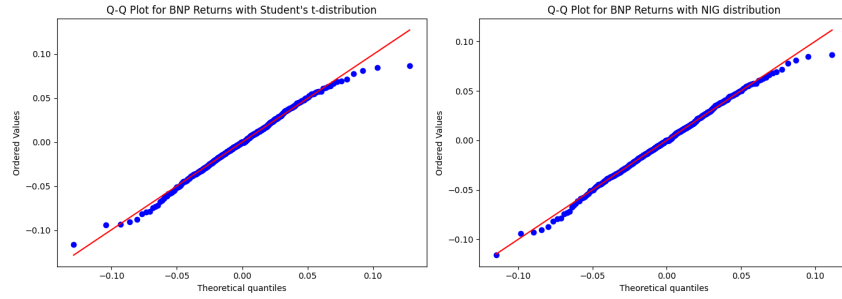


Figure 3: QQ-plot de la loi de student à gauche et la loi normale inverse gaussienne à droite de BNP

Paramètre	BNP	SG
$\alpha$	1.11	1.09
$\beta$	-0.016	0.053
$\mu$	-0.021	-0.15
$\delta$	0.025	0.021

Table 5: Paramètres estimés pour BNP et SG

Nous avons caractérisé les lois marginales de chacun de rendements par une loi normale inverse gaussienne. Nous allons maintenant étudier la structure de dépendance entre les facteurs de risque sur la base de critères non paramétriques.

## 1.4 Etude de la dépendance entre les facteurs de risque sur la base de critères non paramétriques.

### 1.4.1 Etude graphique

La modélisation des facteurs de risque, peut être efficacement réalisée en utilisant des outils graphiques basés sur des critères non paramétriques, tels que les nuages de points, les ajustements linéaires et le dépendogramme. Ces méthodes, choisies pour leur capacité à traiter des données sans présupposer une distribution spécifique, offrent une approche flexible et visuelle pour identifier et analyser les relations entre les variables de risque. Par exemple, les nuages de points permettent de visualiser la dispersion et la relation potentielle entre deux variables, tandis que les ajustements linéaires cherchent à modéliser la relation par une ligne droite, facilitant ainsi la compréhension des tendances générales. Le dépendogramme, représente la structure de dépendance sous la forme du nuage de points des marges uniformes extraites de l'échantillon  $n$  couples de données  $((x_{1,1}; x_{2,1}), \dots, (x_{1,n}; x_{2,n}))$ , soit, pour  $j \in [1, 2]$  :

$$u_{i,j} = \frac{1}{n} \sum_{k=1}^n 1_{\{x_{j,k} \leq x_{j,i}\}}$$

$i \in [1, n]$  le dépendogramme de l'échantillon est donc la représentation de  $n$  couples  $((u_{1,1}; u_{2,1}), \dots, (u_{1,n}; u_{2,n}))$ . Il permet d'observer le caractère plus ou moins simultané des réalisations issues de l'échantillon.

Le graphique ci-dessous (figure 4) présente, pour les créances BNP et SG, à gauche, l'ajustement linéaire de leurs rendements et, à droite, leur dépendogramme.

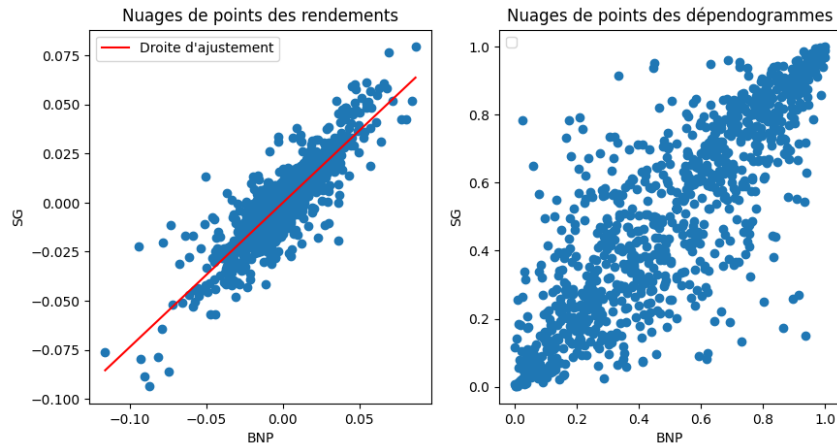


Figure 4: Ajustement linéaire et dépendogramme

Le graphique de gauche semble montrer une corrélation linéaire très forte entre les rendements de BNP et de SG, cependant, il ne permet pas de modéliser les valeurs extrêmes. Le graphique de droite, quant à lui, montre une relation de simultanéité très forte entre les deux actifs, ce qui est particulièrement marqué dans les valeurs extrêmes. Ainsi, si l'on veut recourir à une copule, nous privilégierons celles avec une forte dépendance de queues.

Nous pouvons compléter l'analyse par l'interprétation des mesures de dépendances usuelles (Table 6).

Table 6: Mesures de dépendances entre BNP et SG

	Spearman	Pearson	Tau de Kendall
Coefficient	0.84	0.86	0.67

Les résultats confirment une dépendance forte entre les facteurs de risque. Le coefficient de Spearman est de 0.84, assez élevé, indiquant une forte corrélation positive et monotone. Quant au coefficient de Pearson, s'élevant à 0.86, il suggère également une forte corrélation linéaire positive parmi les facteurs de risque. Enfin, le tau de Kendall est de 0.67, ce qui est significatif. Cela indique aussi une corrélation positive entre les facteurs de risque.

Rappelons que l'objectif de ce travail est de déterminer la Value at Risk de crédit (crédit VaR) à 99 % du portefeuille. Cette méthode se décompose en deux étapes

principales : premièrement, la caractérisation de la perte potentielle ; deuxièmement, l'estimation du quantile correspondant à un niveau de confiance de 99 %.

## 2 Méthodologie : Calcul de la CreditVaR

La Value at Risk est une mesure synthétique, incontournable qui permet de capturer le risque de perte d'un portefeuille. Tout d'abord, elle passe par la modélisation de perte pour une horizon donné, ensuite elle est définie comme un quantile.

### 2.1 Modélisation de la perte

Dans le cadre de la modélisation du risque de crédit, la perte est calculée en prenant en compte plusieurs paramètres, tels que la probabilité de défaut et la perte en cas de défaut, entre autres. Concernant nos deux créances, la perte à l'horizon  $t$  est définie par la formule :  $L_t = L_t^1 + L_t^2$ , où  $L_t^i = N_i(1 - R_i)\mathbf{1}_{\tau_i \leq t}$  représente la fonction de perte associée à la contrepartie  $i$ . Ici,  $N_i = 1000EUR$ , correspondant au montant notionnel de chaque créance  $i$ ,  $R_i$  désigne le taux de recouvrement, modélisé par une loi Bêta, et  $\tau_i$  le temps jusqu'au défaut de la firme  $i$ .

Une fois la fonction de perte caractérisée, la CreditVar(99%) est définie par :

$$\text{CreditVaR}(99\%) = \inf_{x \in \mathbb{R}} \{\mathbb{P}[L_t \leq x] \geq 99\%\}$$

On la détermine à partir des simulations Monte Carlo de la distribution de la perte du portefeuille de crédit et on sélectionne le quantile de niveau 99% sur le vecteur de ces pertes . Il faut tenir compte des dépendances , c'est-à-dire comment sont corrélées les temps de défaut  $(\tau_{BNP}, \tau_{SG})$ .

Face à l'absence de données spécifiques sur les temps de défaut des deux entreprises concernées, nous devons adopter une approche alternative pour modéliser le risque associé. Pour pallier cette lacune, deux méthodologies principales sont envisageables : soit estimer les dépendances à partir des données de rendement des entreprises, soit utiliser les spreads de crédit. Dans notre étude, nous avons opté pour l'utilisation des rendements des actions des entreprises comme méthode préférentielle pour simuler la structure de dépendances entre les deux créances. Cette décision repose sur la disponibilité et la fiabilité des données de rendement, qui permettent une simulation précise des interdépendances sans nécessiter d'informations directes sur les temps de défaut. Les copules sont un outil flexible qu'il est désormais possible d'utiliser pour représenter la dépendance entre risques de façon réaliste.

## 2.2 Présentation de la copule

Une copule est une fonction de répartition multivariée de marginales uniformes sur  $[0, 1]$ . Dans le cas bivarié, on a :

$$C(u_1, u_2) = P[U_1 \leq u_1, U_2 \leq u_2]$$

Dans la théorie de la copule, un théorème fondamental est celui de Sklar, qui stipule que, peu importe les distributions marginales, il est toujours possible de les relier entre elles au moyen d'une copule. Plus précisément, ce théorème permet de modéliser la dépendance entre des variables aléatoires présentant différentes distributions marginales. Le théorème s'énonce comme suit :

Soit  $F(u, v)$ , une distribution jointe de 2-dimensions avec  $F_1, F_2$ , ses marginales continues. Il existe alors une copule  $C$  unique telle que  $F(x_1, x_2) = C(F_1(x_1), F_2(x_2))$ , où  $(x_1, x_2) \in F_1 * F_2$

Nous allons étudier deux principales familles de copules : les copules elliptiques, qui comprennent la copule de Student et la copule Gaussienne, ainsi que les copules archimédiennes, composées des copules de Gumbel, Clayton, et Frank. Nous choisirons, via des tests, la copule la plus adéquate pour modéliser au mieux la dépendance entre les variables étudiées.

Nous pouvons aussi définir la densité bivariée en fonctions de la densité  $c$  de la copule et des densités marginales  $f_1$  et  $f_2$  :

$$f(x_1, y_2) = c(F_1(x_1), F_2(y_2)) \times f_1(x_1)f_2(y_2)$$

Cette équation est très importante en inférence statistique notamment pour l'estimation des paramètres.

### 2.2.1 Copules elliptiques

#### Copule gaussienne

La copule normale ou gaussienne bivariée est une copule elliptique de dimension 2 dont la fonction de distribution est donnée par :

$$\begin{aligned}
 C_\rho(u, v) &= \Phi_\Sigma(\Phi^{-1}(u), \Phi^{-1}(v)) \\
 &= \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi(1-\rho^2)^{\frac{1}{2}}} \exp\left(-\frac{x^2 - 2\rho xy + y^2}{2(1-\rho^2)}\right) dx dy
 \end{aligned}$$

Où  $\Phi_\Sigma$  est la fonction de répartition de la loi normale dont la matrice de variance covariance est la matrice  $\Sigma$ ,  $\Phi^{-1}$  est l'inverse d'une normale standard et  $\rho$  le coefficient de corrélation linéaire.

## Copule Student

La copule de Student bivariée est une copule elliptique de dimension 2 définie comme suit:

$$\begin{aligned}
 C_\rho(u, v) &= T_\Sigma(T_\nu^{-1}(u), T_\nu^{-1}(v)) \\
 &= \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi(1-\rho^2)^{\frac{1}{2}}} \left(1 + \frac{x^2 - 2\rho xy + y^2}{\nu(1-\rho^2)}\right)^{\frac{\nu+2}{2}} dx dy
 \end{aligned}$$

Où  $T_\nu^{-1}$  est la fonction inverse de la distribution t de Student,  $\nu$  est le nombre de degrés de liberté, et  $\rho$  est le coefficient de corrélation linéaire.

Les deux copules elliptiques possèdent un nombre différent de paramètres. De plus, une relation spécifique les unit : le coefficient de corrélation ( $\rho$ ) est lié au tau de Kendall ( $\tau$ ) par la relation suivante :  $\rho = \sin(\frac{\pi}{2}\tau)$ .

Une différence essentielle réside dans la gestion des dépendances extrêmes. Alors que la copule gaussienne est limitée pour étudier ces dépendances (sauf dans le cas où  $\rho = 1$ ), la copule de Student, quant à elle, capture les dépendances dans les queues à droite et à gauche, lorsque le coefficient de corrélation n'est pas égal à -1.

### 2.2.2 Copules archimédiennes

Les copules archimédiennes sont caractérisées par une fonction génératrice  $\phi$ , mais ne seront pas approfondies dans le cadre de ce document. Pour plus d'informations, voir (Genest and MacKay 1986). Nous examinerons trois exemples spécifiques de ces copules, chacune caractérisé par un paramètre,  $\theta$ , qui enrichit leur caractérisation (voir Table 7).

Nom	Générateur $\phi(u)$	Copule $C(u, v)$
Frank ( $\theta \neq 0$ )	$-\ln(\frac{e^{-\theta u}-1}{e^{-\theta}-1})$	$\frac{1}{\theta} \left( -\ln(1 + \frac{(e^{-\theta u}-1)(e^{-\theta v}-1)}{e^{-\theta}-1}) \right)$
Gumbel ( $\theta \geq 1$ )	$(-\ln(u))^\theta$	$\exp \left( -((-\ln(u))^\theta + (-\ln(v))^\theta)^{\frac{1}{\theta}} \right)$
Clayton ( $\theta > 0$ )	$\frac{u^{-\theta}-1}{\theta}$	$(u^{-\theta} + v^{-\theta} - 1)^{-\frac{1}{\theta}}$

Table 7: Copules archimédiennes bivariées les plus courantes.

Il est facile de déterminer les queues de dépendances et le taux de kendall de ces copules archimédiennes.

## 2.3 Méthodes d'inférence statistique

Pour l'estimation des paramètres des copules sélectionnées, plusieurs approches méthodologiques s'offrent à nous: la méthode des moments, la méthode du maximum de vraisemblance et l'approche IFM.

Nous privilégierons l'approche IFM (Inference Functions for Margins). Cette méthode procède en deux étapes principales : initialement, elle estime les paramètres des distributions marginales indépendamment de la structure de dépendance introduite par la copule. Dans un second temps, l'estimation des paramètres de la copule est effectuée en intégrant les estimations marginales obtenues précédemment:

- Etape 1 :  $\hat{\theta}_j = \arg \max_{\theta_j} \sum_{i=1}^n \ln f_j(x_{j,i}; \theta_j)$  pour  $j \in \{1, 2\}$
- Etape 2 :  $\hat{\theta}_c = \arg \max_{\theta_c} \sum_{i=1}^n \ln c \left( F_1(\mathbf{x}_{1,i}, \hat{\theta}_1); F_2(\mathbf{x}_{2,i}, \hat{\theta}_2), \theta_c \right)$

Cet algorithme a l'avantage d'être plus rapide que la méthode du maximum de vraisemblance.

## 2.4 Critères d'ajustement

Pour l'évaluation de l'ajustement des copules à la structure de dépendance d'un échantillon, nous utiliserons des outils graphiques tels que le dépendogramme et le Kendall plot. Nous utiliserons aussi les test d'adéquation. Nous avons précédemment présenté le dépendogramme.

### 2.4.1 Kendall plot

Le Kendall plot permet une comparaison directe entre la copule empirique et la copule théorique. Si l'on a observé  $n$  couples de données  $((x_{1,1}, x_{2,1}), \dots, (x_{1,n}, x_{2,n}))$  dont les marges uniformes associées sont notés :  $((u_{1,1}, u_{2,1}), \dots, (u_{1,n}, u_{2,n}))$ , l'algorithme de construction du Kendall plot est le suivant:

- 1) Partie empirique : on calcule, pour  $i \in [1, n]$ , la quantité

$$H_i = \frac{1}{n-1} \sum_{\substack{j=1 \\ j \neq i}}^n 1_{\{u_{1,j} \leq u_{1,i}; u_{2,j} \leq u_{2,i}\}}, \quad i \in [1, n]$$

On note alors  $H_{(i)}$  les statistiques d'ordre associées :  $H_{(1)} \leq \dots \leq H_{(n)}$ .

- 2) Partie théorique : on effectue  $S$  simulations Monte Carlo; pour  $s \in [1, S]$  : on génère une réalisation de la copule théorique  $((u_{1,1}^s, u_{2,1}^s), \dots, (u_{1,n}^s, u_{2,n}^s))$ ; on calcule,

pour  $i \in [1, n]$ :

$$H_i^{(s)} = \frac{1}{n-1} \sum_{\substack{j=1 \\ j \neq i}}^n 1_{\{u_{1,j}^s \leq u_{1,i}^s; u_{2,j}^s \leq u_{2,i}^s\}}, \quad i \in [1, n]$$

on détermine les statistiques d'ordre associées :  $H_{(1)}^{(s)} \leq \dots \leq H_{(n)}^{(s)}$ . On calcule alors, pour chaque  $i \in [1, n]$

$$Z_i = \frac{1}{S} \sum_{s=1}^S H_{(i)}^{(s)}$$

3) On représente les  $n$  couples  $(H_{(i)}, Z_i)$  pour obtenir le Kendall-plot.

Plus le Kendall plot se rapproche d'une droite, plus l'ajustement entre la structure de dépendance de l'échantillon et la copule estimée sur ce même échantillon est bon.

#### 2.4.2 Tests d'adéquation et bilan comparatif des modélisations

Les tests d'adéquation pour les copules consistent à tester l'appartenance à une famille paramétrique de copules notée  $\{C_\theta, \theta \in \Theta\}$ . Le test d'adéquation est donc un test d'hypothèses s'écrivant de la manière suivante :

$$\mathcal{H}_0 : C \in \{C_\theta, \theta \in \Theta\} \quad \text{vs} \quad \mathcal{H}_1 : C \notin \{C_\theta, \theta \in \Theta\}$$

La plupart des tests d'adéquation consacrés aux copules se fondent sur les pseudoobservations déduites des statistiques de rang, pour  $j \in [1, 2]$ .

$$\hat{U}_{j,i} = \frac{1}{n} \sum_{k=1}^n 1_{\{x_{j,k} \leq x_{j,i}\}}, \quad i \in [1, n]$$

et on note  $\hat{U}_i = (\hat{U}_{1,i}, \hat{U}_{2,i})$ . Les pseudo-observations  $\hat{U}_i, i \in [1, n]$  sont assimilables à un échantillon de la copule sous-jacente  $C$ .

On désigne par «copule empirique» la quantité suivante, définie pour  $(u_1, u_2) \in [0, 1]^2$ , par

$$\hat{C}_n(u_1, u_2) = \frac{1}{n} \sum_{i=1}^n 1_{\{\hat{O}_{1,i} \leq u_1; \hat{O}_{2,i} \leq u_2\}}$$

À partir d'une estimation  $C_{\theta_n}$  de  $C$  obtenue sous  $\mathcal{H}_0$ , on établit la statistique de test suivante :

$$T_n = \sum_{i=1}^n \left( \hat{C}_n(\hat{U}_i) - C_{\theta_n}(\hat{U}_i) \right)^2$$

Sur la base de ce test, on choisira donc la copule pour laquelle la p-value est la plus grande, c'est-à-dire, pour laquelle  $\mathcal{H}_0$  est la plus plausible. La p-value pour  $T_n$  peut être



rapprochée par une procédure de bootstrap paramétrique:

- 1) Évaluer la copule empirique  $\hat{C}_n$  et estimer  $\theta_n$  à partir de  $\hat{U}_1, \dots, \hat{U}_n$ .
- 2) Calculer la statistique  $T_n$ .
- 3) Pour un entier  $S$  grand, répéter les étapes suivantes pour  $s \in [1, S]$  :
  - Générer un échantillon aléatoire  $\left( \left( X_{1,1}^{(s)}, X_{2,1}^{(s)} \right), \dots, \left( X_{1,n}^{(s)}, X_{2,n}^{(s)} \right) \right)$  à partir de la copule  $C_{\theta_n}$  et calculer les pseudo-observations associées  $U_1^{(s)}, \dots, U_n^{(s)}$ .
  - En définissant  $\hat{C}_n^{(s)}(u_1, u_2) = \frac{1}{n} \sum_{i=1}^n 1_{\{\hat{U}_{1,i}^{(s)} \leq u_1, \hat{U}_{2,i}^{(s)} \leq u_2\}}$ , calculer un estimateur  $\theta_n^{(s)}$  de  $\theta$  à partir de  $U_1^{(s)}, \dots, U_n^{(s)}$ .
  - Calculer une réalisation  $T_n^{(s)}$  de  $T_n$  de la manière suivante :

$$T_n^{(s)} = \sum_{i=1}^n \left( \hat{C}_n^{(s)} \left( \hat{U}_i^{(s)} \right) - C_{\theta_n^{(s)}} \left( \hat{U}_i^{(s)} \right) \right)^2$$

- 4) L'estimation de la  $p$ -value est alors donnée par:

$$\frac{1}{S} \sum_{s=1}^S 1_{\{T_n^{(s)} \geq T_n\}}$$

## 3 Résultats

### 3.1 Estimation des paramètres

Comme annoncé, nous modélisons la structure de dépendance de nos données par plusieurs copules:

- Copules elliptiques: Gaussienne et Student
- Copules archimédiennes: Clayton, Gumbel et Frank

Nous implémentons ces différentes copules et nous estimons les paramètres par la méthode IFM.

Copule	Paramètres estimés
Gaussienne	$\rho = 0.860$
Student	$\rho = 0.866, \nu = 3.044$
Clayton	$\theta = 2.474$
Gumbel	$\theta = 2.991$
Frank	$\theta = 10.140$

Table 8: Estimation des paramètres des différentes copules par la méthode IFM.

Il nous faut maintenant mesurer l'adéquation de ces différentes copules à la structure de dépendance de nos données.

## 3.2 Kendall plot

Nous représentons le Kendall plot pour chacune des copules.

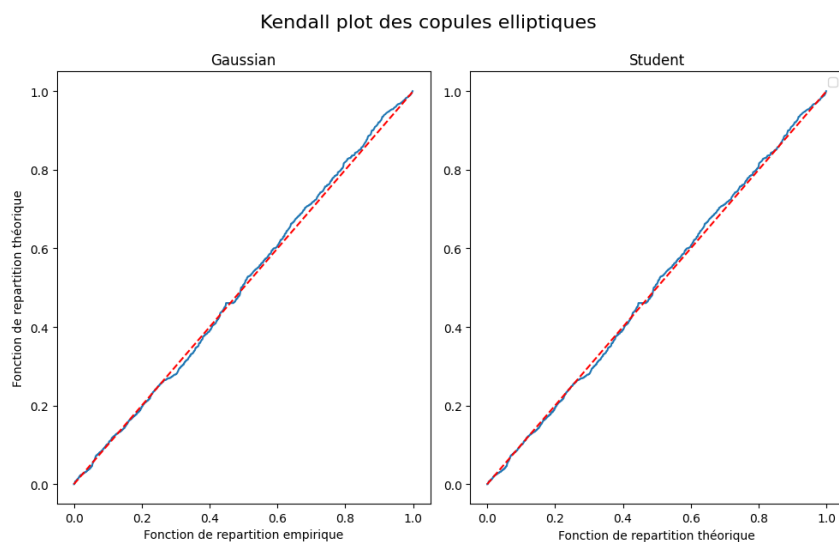


Figure 5: Kendall plot des copules elliptiques

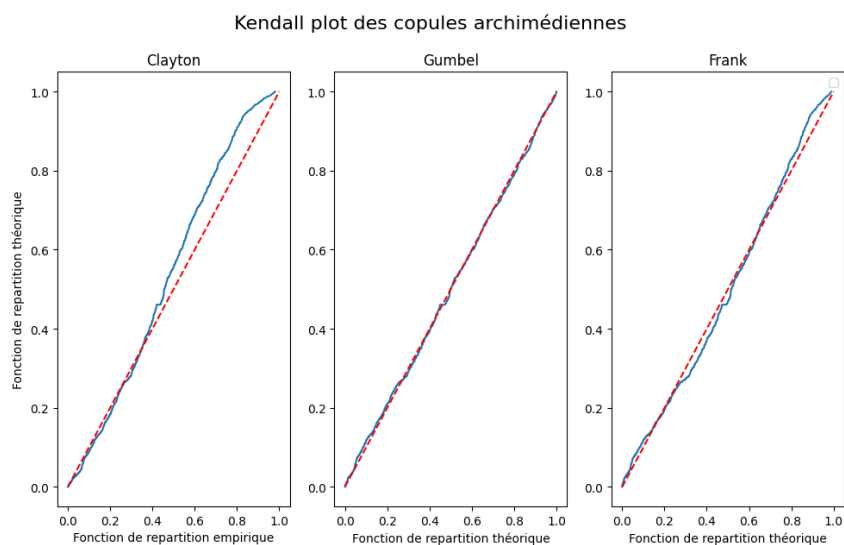


Figure 6: Kendall plot des copules archimédiennes

Graphiquement, les copules de Student et de Gumbel semblent être les plus alignées sur la première bissectrice. Ces graphiques nous donnent une première idée sur l'adéquation des copules choisies à la structure de dépendance de nos données.

En plus de ces graphiques, nous jugeons de l'adéquation grâce au dépendogramme.

### 3.3 Comparaison du dépendogramme de l'échantillon à celui des copules théoriques

Nous avons représenté ci-dessous les dépendogrammes de l'échantillon en rouge et celui des différentes copules théoriques. Ainsi nous pourrions nous forger une première idée sur la pertinence de la modélisation envisagée.

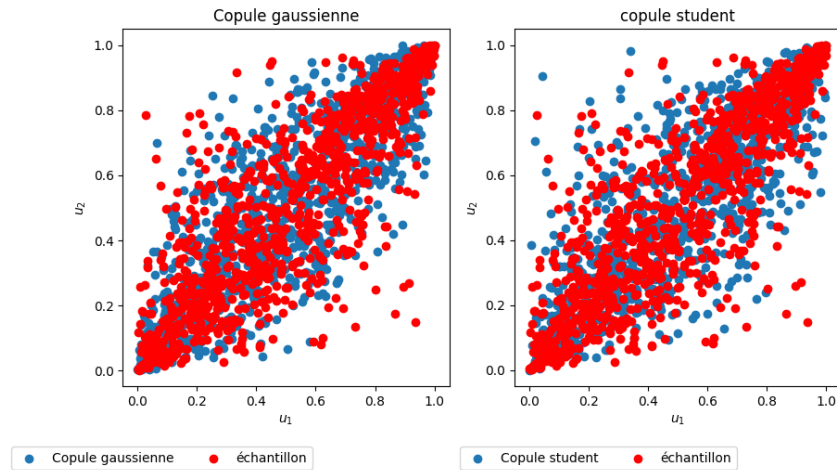


Figure 7: Comparaison des dépendogrammes des copules elliptiques à celui de l'échantillon

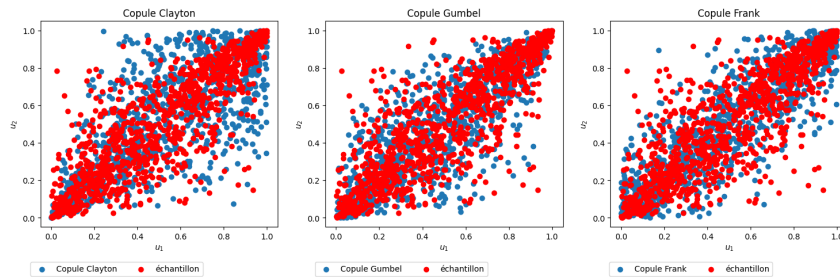


Figure 8: Comparaison des dépendogrammes des copules archimédiennes à celui de l'échantillon

Graphiquement, les copules gaussiennes, les copules de Student et de Gumbel semblent être les plus adaptées pour modéliser la structure de dépendance de notre portefeuille. En effet, les points rouges et bleus semblent se superposer.

### 3.4 Tests d'adéquation et bilan comparatif des modélisations

Nous implémentons maintenant la procédure du test d'adéquation, nous obtenons les p-values suivantes:

Copule	p-value
Gaussienne	0.002
Student	0.684
Clayton	0.976
Gumbel	0.992
Frank	0.000

Table 9: P-value des copules

Selon ces p-value, nous sélectionnons la copule de Gumbel pour la modélisation de la structure de dépendance de nos données. En effet, ayant la p-value la plus grande, cette copule est celle pour laquelle l'hypothèse  $\mathcal{H}_0$  est la plus pausable.

### 3.5 Résultats de la CreditVaR : simulation de la distribution par Monte Carlo

Nous déterminons, à partir de simulations de Monte Carlo, la distribution de la perte du portefeuille. En calculant le quantile de niveau 99% sur les vecteurs des pertes simulées et en répétant l'opération plusieurs fois, nous obtenons un estimateur ponctuel par la moyenne des résultats. En variant le nombre de simulations de 100 à 100 000, par incréments de 100, nous estimons la CreditVaR de niveau 99% à 1285,04. Le graphique ci-dessous (voir 9) représente les différentes valeurs de la créditVar.

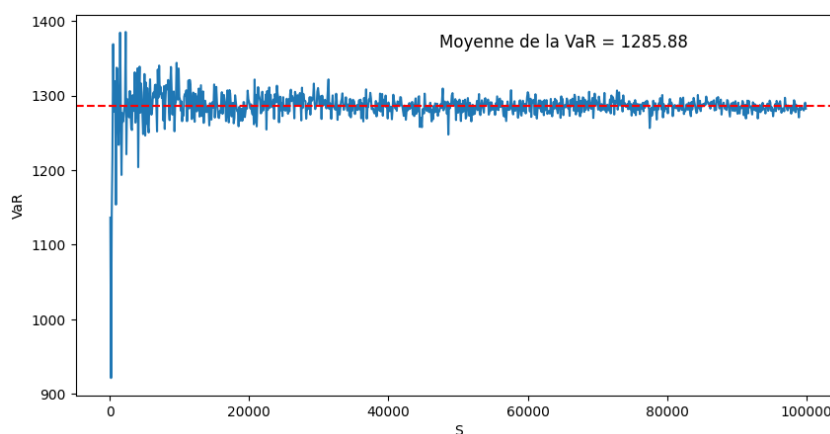


Figure 9: Distribution de la creditVaR par Monte Carlo avec niveau de confiance de 99%

## Conclusion

Ce document consiste en une analyse du risque de crédit à l'aide de copules. Nous avons couvert divers aspects, de la présentation des données utilisées à la présentation de nos résultats en passant par l'explication théorique de la méthodologie. Nous avons pu calculer la CreditVar associé à notre portefeuille hypothétique. Pour ce faire il a fallu simuler les pertes de chacune des créances de notre portefeuille. Cette simulation passe par le calcul des probabilités de défaut implicites, la caractérisation des taux de recouvrement et enfin par la caractérisation de la structure de dépendance selon la méthode IFM. Notre estimation de la CreditVar fluctue de manière plutôt robuste autour de la valeur **1285.04** lorsque nous faisons varier la taille de l'échantillon

## References

- Genest, Christian and R Jock MacKay (1986). “Copules archimédiennes et familles de lois bidimensionnelles dont les marges sont données”. In: *Canadian journal of statistics* 14.2, pp. 145–159.
- Genest, Christian, Bruno Rémillard, and David Beaudoin (2009). “Goodness-of-fit tests for copulas: A review and a power study”. In: *Insurance: Mathematics and economics* 44.2, pp. 199–213.

## Appendix

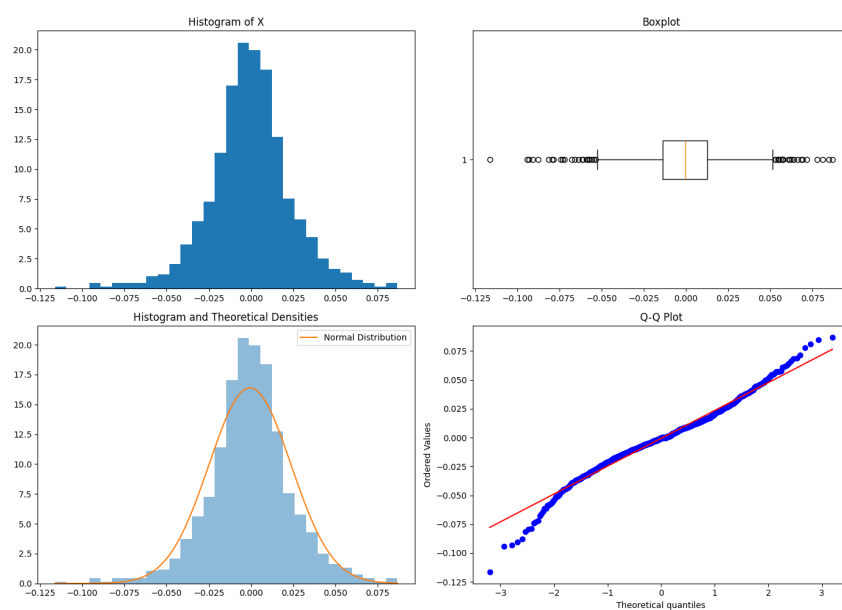


Figure 10: Adéquation des rendements de BNP à une loi normale

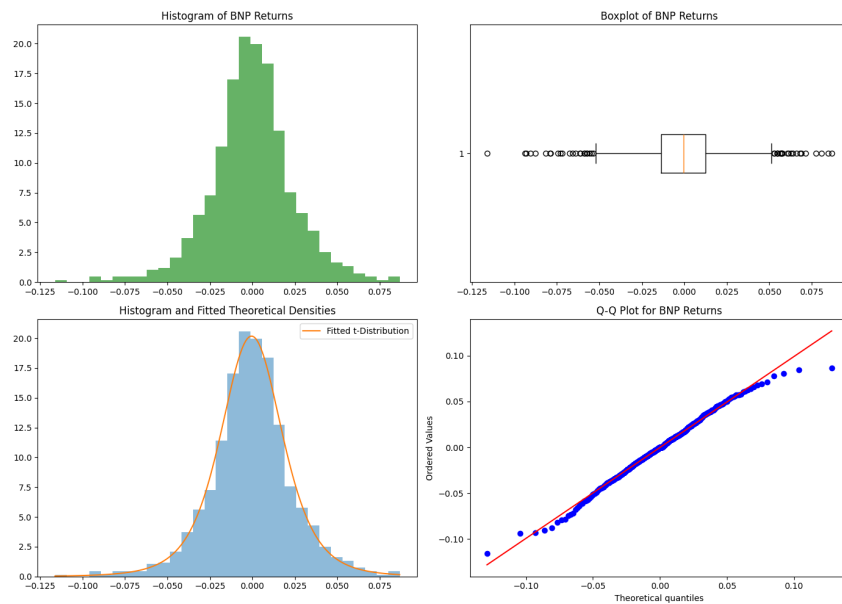


Figure 11: Adéquation des rendements de BNP à une loi de Student

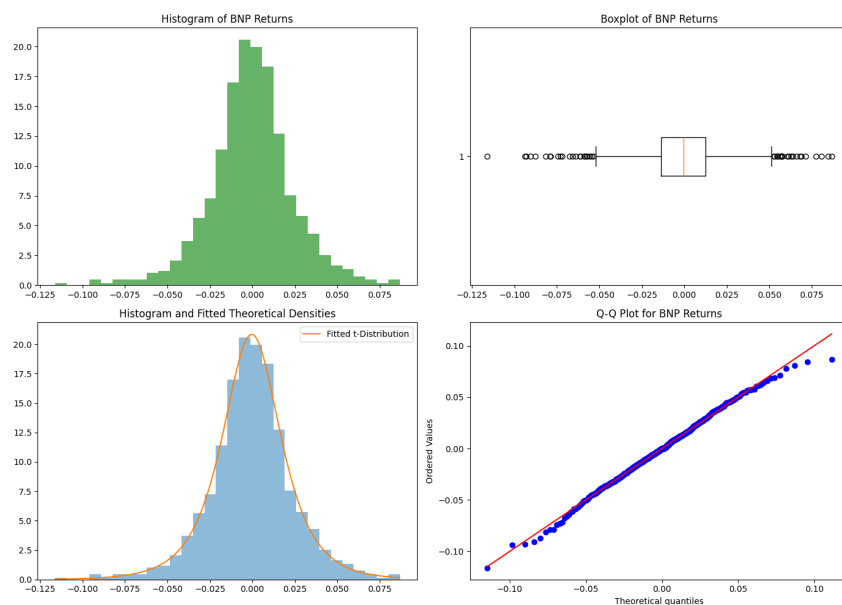


Figure 12: Adéquation des rendements de BNP à une loi normale inverse gaussienne