République Algérienne Démocratique et Populaire

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

UNIVERSITÉ MOHAMED KHIDER, BISKRA

FACULTÉ des SCIENCES EXACTES et des SCIENCES de la NATURE et de la VIE

DÉPARTEMENT DE MATHÉMATIQUES

Mémoire présenté en vue de l'obtention du Diplôme :

MASTER en Mathématiques

Option : **Statistique**

Par

MREDEF Chaima

Titre :

Mesures de concordance des copules

Membres du Comité d'Examen :

Dr.	**ABDELLI Jihane**	UMKB	Président
Dr.	**BENELMIR Imane**	UMKB	Encadreur
Dr.	**BENAMEUR Sana**	UMKB	Examinatrice

Juin 2023

DEDICACE

Je dédie ce humble travail

À la memoire de mon défunt père.

Aucune dédicace ne saurait exprimer mon respect, mon amour éternel et ma considération pour les sacrifices que tu as consenti pour mon instruction et mon bien être.

À la lumière de mes jours, la source de mes efforts, la flamme de mon cœur, ma vie et mon bonheur, maman que j'adore.

À mes soeurs et mes frères.

À toute ma famille de proche ou de loin.

À mes amis avec qui j'ai passé des moments mémorables et agréables.

À tous mes professeurs et tous ceux qui ont contribué dans mon parcours d'étude.

À mes camarades de promotion 2022/2023.

À tous ceux que j'aime.

REMERCIEMENTS

*Mes remerciements vont tout premièrement à **ALLAH** le tout puissant pour la santé, la volonté, la perseverance et la patience qu'il m'a donné durant ces longues années d'étude et le courage pour terminer ce mémoire.*

Ensuite, j'offre mes remerciements à mon encadreur Mme **BENELMIR Imane** qui m'a été d'une aide précieuse avec son expérience ces connaissances et ces instructions pour réaliser ce modeste travail.

Je remercie aussi les membres du jury Dr. **ABDELLI Jihane** et Dr. **BENAMEUR Sana** pour avoir accepté d'évaluer mon mémoire, leurs commentaires qui m'ont permis d'améliorer la version finale de ce travail.

Enfin, une pensée toute spéciale ira à ma famille et à mes amis, pour leur support et leurs encouragements.

Merci à tous.

Table des matières

Table des figures

Liste des tableaux

Introduction

Le terme copule vient du mot latin copula, qui signifie au sens figuré lien, alliance, liaison ou union. Cette notion apparait sous d'autres appellations dans certains travaux de *Fréchet* [9], *Féron* et *Dall'Aglio* [7] portant sur l'étude des tables de contingence. Mais c'était en 1959 que le terme copule a été utilisé pour la première fois grâce à *Sklar* [19]. On peut dire que les copules sont des fonctions qui relient des distributions multivariées à leurs marges unidimensionnelles.

Les copules intéressent les statisticiens pour deux raisons principales premièrement, en tant que moyen d'étudier les mesures de la dépendance et deuxièmement, comme point de départ pour construire des familles de distributions bivariées.

La mesure de la dépendance entre des variables aléatoires est une pratique largement répandue par les statisticiens. Un riche ensemble de mesures de dépendance entre ces variables aléatoires a été proposé comme le coefficient de corrélation de Pearson, le tau de kendall, le rho de Spearman, etc. Bien que ces mesures sont simple à calculer et peuvent être facilement interprétées, elles ne sont pas en mesure de détecter toutes les formes de dépendances, donc il était indéniable de trouver un autre moyen pour résoudre ce problème. En effet, la fonction copule à l'avantage de modéliser complètement la dépendance entre les variables.

Ce mémoire se décompose de deux chapitres. Dans le premier chapitre on va présenter quelque concept de base à savoir des généralités (théorème de Sklar), définitions, propriétés, types des copules (usuelle, paramétrique, archimédienne, valeurs extrêmes, archimax et empririque), copules associées a une copule (survie, duale, co-copule, mixte).

Dans le deuxième chapitre on donnera les définitions sur les mesures de concordances (associations), types des mesures de concordance, la dépendance de queue et pour terminer ce chapitre une simulation sera étudier sur quelques copules à un paramètre sous logiciel de programmation R.

On cloturat ce modestre travail par une conclusion générale.

Chapitre 1

Généralité sur les copules

Le concept de copule a été introduit par *Abe Sklar* en 1959 [19] pour but de résoudre un problème de probabilité énoncé par *Maurice Fréchet* dans le cadre des espaces métriques aléatoires.

La copule permet d'introduire et de caractériser des formes très flexibles de dépendance entre différentes variables aléatoires (vas). Elle tient en considération, de façon meilleure de certains faits en en finance (asymétrie, dépendance des queues,...).

La copule s'exprime en terme des fonctions de distribution marginales (fds) des vas, permettant ainsi de faire le lien entre ces fds marginales et leur fd bivariée.

Dans ce chapitre, on s'intéresse aux cas des copules bivariées à un paramètre, on débute par les principaux éléments de la théorie des copules : définition, propriétés, types des copules (usuelles, paramétrique, archimédiennes, etc) et les copules associées à une copule.

1.1 Théorie des copules

En statistique, une copule est un objet mathématique venant de la théorie des probabilités. La copule joue un rôle important dans la modélisation de la dépendance entre les vas.

Soient X et Y deux vas de fds marginales F et G respectivement. Pour tout couple de vas réelles (X, Y), la fd jointe associée à ce couple notée H est définit par

$$\boxed{\forall (x, y) \in \mathbb{R}^2 : H(x, y) = \mathbb{P}(X \leq x, Y \leq y),}$$

avec $F(x) = \mathbb{P}(X \leq x)$ et $G(y) = \mathbb{P}(Y \leq y)$ qui suivent une loi uniforme dans l'intervalle $I = [0, 1]$.

1.1.1 Théorème de Sklar (1959)

Ce théorème a été découvert par *Abe Sklar*, il permet de modéliser la loi conjointe notée H du couple (X, Y).

On appelle copule à deux dimensions toute fd bivariée C ayant pour marginales F et G qui suivent la loi uniforme sur I [14].

Théorème 1.1.1 *Soit H est une fd bivariée de marges continues F et G. Il existe une unique fonction C telle que*

$$\boxed{\forall (x, y) \in \mathbb{R}^2 : H(x, y) = C(F(x), G(y)).} \tag{1.1}$$

On appel la fonction bivariée C copule.

On pose $F(x) = u$ et $G(y) = v$.

D'après 1.1, on déduit que la copule C est définit comme suit

$$\boxed{\forall (u,v) \in I^2 : C(u,v) = H\left(F^{-1}(u), G^{-1}(v)\right).} \tag{1.2}$$

1.1.2 Propriétés

Une copule C est une fd qui vérifie les conditions suivantes [8].

Pour tout $u, v \in I$, on a

1. $C(u,0) = C(0,v) = 0.$
2. $C(u,1) = u$ et $C(1,v) = v.$
3. $C(u,v) = C(v,u).$
4. $\forall u_1, v_1, u_2, v_2 \in I$ avec $u_1 \leq u_2$ et $v_1 \leq v_2$. La copule C est une fonction 2-croissante, on a

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

5. Soient deux copules C_1 et C_2, on dit que C_1 est plus petite que C_2 et inversement si

$$\forall u, v \in I : C_1(u,v) \leq C_2(u,v).$$

Preuve. Pour démontrer les propriétés précédentes, on prend comme exemple la copule minimum suivante [21]

$$\forall u, v \in I : M(u, v) = \min(u, v).$$

1. $\min(u, 0) = \min(0, v) = 0.$

2. $\min(u, 1) = u$ et $\min(1, v) = v.$

3. $\min(u, v) = \min(v, u).$

4. $\forall u_1, v_1, u_2, v_2 \in I$

   - Si $u_1 \leq u_2 < v_1 \leq v_2$, on a

     $\min(u_1, v_1) - \min(u_2, v_1) - \min(u_1, v_2) + \min(u_2, v_2) = u_1 - u_2 - u_1 + u_2 = 0.$

   - Si $v_1 \leq v_2 < u_1 \leq u_2$, on a

     $\min(u_2, v_2) - \min(u_2, v_1) - \min(u_1, v_2) + \min(u_1, v_1) = v_2 - v_1 - v_2 + v_1 = 0.$

   - Si $u_1 \leq v_1 < u_2 \leq v_2$, on a

     $\min(u_2, v_2) - \min(u_2, v_1) - \min(u_1, v_2) + \min(u_1, v_1) = u_2 - v_1 - u_1 + u_1 = u_2 - v_1 > 0.$

5. C'est évident.

■

Théorème 1.1.2 (Théorème d'invariance) *L'un des théorèmes essentiels à la théorie des copules est celui de l'invariance par transformation strictement croissantes [3]. Soient X et Y deux vas continues de marginales F et G et de copule C_{XY}.Si α et β sont deux fonctions strictement croissantes, alors*

$$C_{\alpha(X)\beta(Y)} = C_{XY}.$$

Preuve. On note respectivement F et G les fd jointes des vecteurs (X, Y) et $(\alpha(X), \beta(Y))$ respectivement.

Par la suite on note que F_1, F_2 sont les fds marginales de F et par G_1, G_2 les fds marginales de G. On remarque que les marges de G sont

$$\begin{aligned} G_1(x) &= \mathbb{P}[\alpha(X) \leq x] \\ &= \mathbb{P}[X \leq \alpha^{-1}(x)], \text{ car } \alpha \text{ est une fonction croissante} \\ &= F_1(\alpha^{-1}(x)). \end{aligned}$$

On fait la même chose pour la deuxième marge de G, on obtient alors

$G_2(y) = F_2(\beta^{-1}(y))$.

On pose $G_1(x) = u$ et $G_2(y) = v$, on obtient donc $G_1^{-1}(u) = \alpha F_1^{-1}(u)$ et

$G_2^{-1}(v) = \beta F_2^{-1}(v)$.

On en déduit alors que

$$\begin{aligned} C_{\alpha(X)\beta(Y)}(u, v) &= G(x, y) \\ &= G(G_1^{-1}(u), G_2^{-1}(v)) \\ &= \mathbb{P}[\alpha(X) \leq G_1^{-1}(u), \beta(Y) \leq G_2^{-1}(v)] \\ &= \mathbb{P}[X \leq \alpha^{-1}(G_1^{-1}(u)), Y \leq \beta^{-1}(G_2^{-1}(v))] \\ &= F(x, y) \\ &= C_{XY}(u, v). \end{aligned}$$

■

Théorème 1.1.3 *Si α est strictement croissante et β strictement décroissante alors*

$$C_{\alpha(X)\beta(Y)}(u, v) = u - C_{XY}(u, 1 - v).$$

Si α est strictement décroissante et β strictement croissante alors

$$C_{\alpha(X)\beta(Y)}(u, v) = v - C_{XY}(1 - u, v).$$

Si α et β est strictement décroissantes alors

$$C_{\alpha(X)\beta(Y)}(u,v) = u + v - 1 - C_{XY}(1-u, 1-v).$$

1.1.3 Densité

La copule C admet une densité de probabilité notée c définie comme suit [3]

Soit $c : I^2 \to \mathbb{R}^2$, alors

$$\boxed{c(u,v) = \frac{\partial^2 C(u,v)}{\partial u \partial v}.} \tag{1.3}$$

Théorème 1.1.4 *Soit la copule C, la dérivée partielle existe presque surement. On a*

$$\forall u \in I : 0 \leq \frac{\partial C(u,v)}{\partial u} \leq 1.$$

De plus les fonctions $\partial C(u,v)/\partial u$ et $\partial C(u,v)/\partial v$ sont finies et croissantes presque par tout sur I.

1.2 Copules usuelles

Il existe diverses copules usuelles on cite la copule produit, minimale et maximale [1].

1.2.1 Copule indépendante

La copule d'indépendance appelée aussi copule produit est définie par

$$\boxed{\Pi(u,v) = uv.} \tag{1.4}$$

1.2.2 Copule maximale

La copule maximale appelée aussi copule comonotone ou copule de dépendance positive est définie comme suit

$$\boxed{W(u,v) = \max(u+v-1,0).} \tag{1.5}$$

1.2.3 Copule minimale

La copule minimale appelée aussi copule anti monotone ou copule de dépendance négative est définie comme suit

$$\boxed{M(u,v) = \min(u,v).} \tag{1.6}$$

Remarque 1.2.1 (Borne de Fréchet- Hoeffeding) *On a*

$$\boxed{\forall u,v \in I : W(u,v) \leq C(u,v) \leq M(u,v).} \tag{1.7}$$

La fonction W 1.5 s'appelle la borne inférieure de Fréchet-Hoeffding , tandis que M 1.6 est la borne supérieure de Fréchet-Hoeffding.

À noter que M est une copule, alors que W ne l'est pas.

1.3 Copule Elliptique

Les copules Elliptiques sont des copules paramétriques associées aux distributions elliptiques. Parmi ces copules on a la copule Gaussienne et la copule de student [8].

Définition 1.3.1 *On appelle copule elliptique toute copule qui s'écrit de la forme suivante*

$$\boxed{C_\rho(u,v) = \frac{1}{\sqrt{1-\rho^2}} \int_{-\infty}^{\Phi_{g,1}^{-1}(u)} \int_{-\infty}^{\Phi_{g,2}^{-1}(v)} g\left(\frac{s^2+t^2-2\rho st}{\sqrt{1-\rho^2}}\right) dsdt = H_\rho\left(\Phi_{g,1}^{-1}(u), \Phi_{g,2}^{-1}(v)\right),}$$

où H_ρ est la distribution jointe des vas S et T, $\Phi_{g,1}^{-1}(u)$ et $\Phi_{g,2}^{-1}(v)$ sont leurs fonction quantiles respective et ρ leurs coefficients de corrélation. Dans cette famille se trouve entre autres la copule Gaussienne et la copule de student.

1.3.1 Copule Gaussienne

Une copule gaussienne est une mesure de dépendance entre deux vas. Cette copule est définie à partir de la famille des lois gaussiennes.

Définition 1.3.2 *Soient ρ le coefficient de corrélation, Φ la fd de la loi normale $N(0,1)$ et H_ρ la fd bivariée de la loi normale de matrice de covariance $\sum$ (matrice de corrélation dans ce cas) associée à ρ. La copule gaussienne C_ρ est définie par*

$$C_\rho(u,v) = H_\rho\left(\Phi^{-1}(u), \Phi^{-1}(v)\right),$$

où Φ^{-1} est l'inverse de la fd de Φ.

Remarque 1.3.1 *1. La fd bivariée de la loi normale centrée réduite est comme suit*

$$C_\rho(u,v) = \int_{-\infty}^{u} \int_{-\infty}^{v} \frac{1}{2\pi\sqrt{(1-\rho^2)}} \exp\left(-\frac{s^2+t^2-2\rho st}{2(1-\rho^2)}\right) dsdt.$$

2. Soit $\rho \in [-1, 1]$. Dans le cas ou $\rho = 0$, alors $C_0 = \Pi$.

1.3.2 Copule de Student

Cette copule est construite de la même manière que la copule gaussienne. La copule de student est la fonction de dépendance associée à la distribution de student.

Définition 1.3.3 *La copule de student $C_{\rho,\nu}$ possède deux paramètres, le coefficient de corrélation ρ et le degré de liberté v. Elle est définie par*

$$C_{\rho,v}(u, v) = H_{\rho,v}\left(t_v^{-1}(u), t_v^{-1}(v)\right),$$

où t_ν^{-1}est l'inverse de la fd univariée t_ν à v degré de liberté.

Remarque 1.3.2 *1. La fd bivariée de la loi de student est comme suit*

$$C_{\rho,v}(u, v) = \int_{-\infty}^{u} \int_{-\infty}^{v} \frac{\Gamma(\frac{v+2}{2})}{2\pi\Gamma(\frac{v}{2})\sqrt{1-\rho^2}} \times \left(1 + \frac{s^2 + t^2 - 2\rho st}{v(1-\rho^2)}\right)^{-\frac{v+2}{2}} dsdt.$$

2. Si le degré de liberté v tend vers l'infini, alors la copule de student converge vers la copule gaussienne.

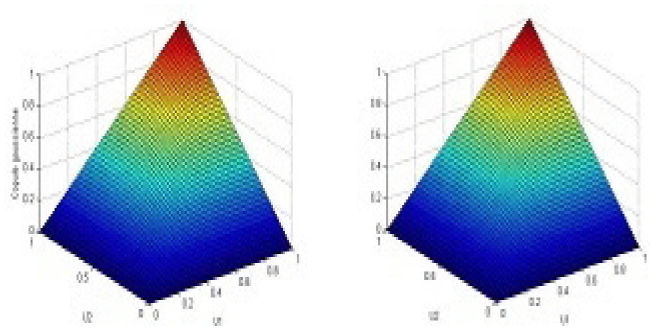

FIG. 1.1 – Copule Gaussienne (à gauche), copule Student (à droite).

1.4 Copules archimédiennes

Les copules archimédiennes ont l'avantage de décrire diverses structures de dépendance notamment les dépendances asymptotiques, les coefficients de queue inférieure et de queue supérieure. Les copules à un paramètre les plus utilisées dans cette famille sont la copule de Frank, Clayton et Gumbel mentionnées ci-dessous [14].

Définition 1.4.1 *Soit* Φ *une fonction décroissante convexe sur* $I \to [0, +\infty[$ *telle que* $\Phi(1) = 0$ *et* $\Phi(0) = +\infty$. *On appellera copule archimédienne stricte de générateur* Φ, *la copule définie par*

$$\boxed{\forall u, v \in I : C(u, v) = \Phi^{-1}(\Phi(u) + \Phi(v)).} \tag{1.8}$$

1.4.1 Copule de Frank

Soit le générateur $\Phi(x) = -\ln\left(\dfrac{e^{-ax} - 1}{e^{-a} - 1}\right)$ avec $a \neq 0$, la copule de Frank est donc définie comme suit

$$\boxed{C_a(u,v) = -\frac{1}{a}\ln\left(1+\frac{\left(e^{-au}-1\right)\left(e^{-av}-1\right)}{e^{-a}-1}\right).} \tag{1.9}$$

Remarque 1.4.1 *Les cas limites sont donc les suivants*

- $\lim\limits_{a\to+\infty} C_a = M.$
- $\lim\limits_{a\to 0} C_a = \prod.$
- $\lim\limits_{a\to-\infty} C_a = W.$

Théorème 1.4.1 *La fonction de densité d'une copule de Frank est définie par*

$$\forall a \in \mathbb{R}^* : c_a\left(u,v\right) = \frac{-a\left(e^{-a}-1\right)e^{-a(u+v)}}{((e^{-a}-1)+(e^{-au}-1)(e^{-av}-1))^2}.$$

1.4.2 Copule de Clayton

Soit le générateur $\Phi(t) = a(t^{-\frac{1}{a}} - 1)$ avec $a > 0$, la copule de Clayton est définie comme suit

$$\boxed{C_a\left(u,v\right) = \left(u^{-\frac{1}{a}} + v^{-\frac{1}{a}} - 1\right)^{-a}.} \tag{1.10}$$

Remarque 1.4.2 *Les cas limites sont donc les suivants*

- $\lim\limits_{a\to 0} C_a = \prod.$
- $\lim\limits_{a\to-\infty} C_a = W.$

Théorème 1.4.2 *La fonction de densité d'une copule de Clayton est définie par*

$$\forall a \in \mathbb{R}_+^* : c_a\left(u,v\right) = \left(a+1\right)\left(u.v\right)^{-a-1}\left(u^{-a}+v^{-a}-1\right)^{-\frac{1}{a}-2}.$$

1.4.3 Copule de Gumbel

Soit le générateur $\Phi(t) = (-\ln t)^a$ avec $a > 0$, la copule de Gumbel est définie comme suit

$$\boxed{C_a(u,v) = \exp\left\{-((-\ln u)^a + (-\ln v)^a)^{\frac{1}{a}}\right\}.} \tag{1.11}$$

Remarque 1.4.3 *Les cas limites sont donc les suivants :*

$\lim_{a\to+\infty} C_a = M.$

$\lim_{a\to 1} C_a = \prod.$

Théorème 1.4.3 *La fonction de densité c_a d'une copule de Gumbel est définie comme suit*

$$\frac{C_a(u,v)(-\ln u)^{a-1}(-\ln v)^{a-1}((-\ln u)^a + (-\ln v)^a)^{\frac{1}{a}-2}\left(a-1+((-\ln u)^a+(-\ln v)^a)^{\frac{1}{a}}\right)}{uv}.$$

Ci-dessous les graphes des trois copules mentionnées précédemment

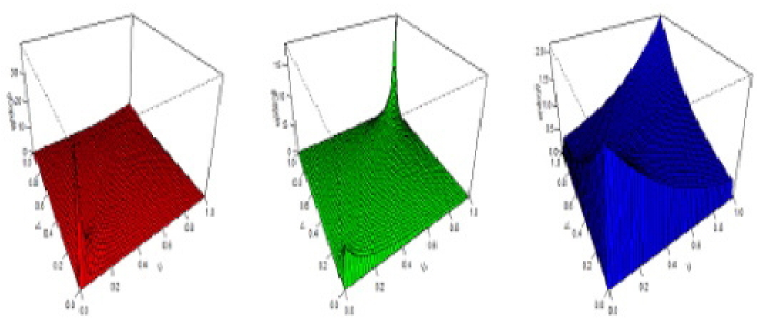

FIG. 1.2 – Copule de Frank, Clayton et Gumbel (de gauche vers la droite)

1.5 Copule des valeurs extrêmes

On appelle copule des valeurs extrêmes [1] toute copule C vérifiant

$$\forall t > 0 : C\left(u^t, v^t\right) = C^t\left(u, v\right). \tag{1.12}$$

Théorème 1.5.1 *Il existe une fonction convexe (fonction de dépendance) A définie de I dans $\left[\frac{1}{2}, 1\right]$ telle que*

$$C\left(u, v\right) = \exp\left\{\left(\ln u + \ln v\right) A\left(\frac{\ln u}{\ln u + \ln v}\right)\right\},$$

avec $\max\left\{(t, 1-t)\right\} \leq 1$. La copule C s'appelle alors copule des valeurs extrêmes bivariées.

Preuve. On prend comme exemple d'application, la copule de Galambos à un seul paramètre $\theta \geq 0$ comme suit

$$C_\theta(u, v) = uv \exp\left\{-\left((-\ln u)^\theta + (-\ln v)^\theta\right)^{-1/\theta}\right\}.$$

On a $\forall t > 0$:

$$\begin{aligned}
C_\theta(u^t, v^t) &= u^t v^t \exp\left\{-[(-\ln u^t)^\theta + (-\ln v^t)^\theta]^{-1/\theta}\right\} \\
&= u^t v^t \exp\left\{-[(-t \ln u)^\theta + t^\theta(-t \ln v)^\theta]^{-1/\theta}\right\} \\
&= u^t v^t \exp\left\{-[t^\theta(-\ln u)^\theta + t^\theta(-\ln v)^\theta]^{-1/\theta}\right\} \\
&= u^t v^t \exp\left\{-[t^\theta\left((-\ln u)^\theta + (-\ln v)^\theta\right)]^{-1/\theta}\right\} \\
&= u^t v^t \exp\left\{-t^{\theta/\theta}[(-\ln u)^\theta + (-\ln v)^\theta]^{-1/\theta}\right\} \\
&= u^t v^t \left(\exp\left\{-[(-\ln u)^\theta + (-\ln v)^\theta]^{-1/\theta}\right\}\right)^t \\
&= C_\theta^t(u, v).
\end{aligned}$$

Donc cette copule est une copule des valeurs extrêmes. ■

1.6 Copule Archimax

La copule Archimax est une copule à la fois à valeur extrême et Archimédienne [17]. La classe de ces copules est construite à partir du générateur Φ et de la fonction convexe A. Une copule Archimax est définie par

$$\boxed{C_{\Phi,A}\left(u,v\right)=\Phi^{-1}\left(\left(\Phi\left(u\right)+\Phi\left(v\right)\right)A\left(\tfrac{\Phi(u)}{\Phi(u)+\Phi(v)}\right)\right).} \tag{1.13}$$

- Si on pose $A\left(t\right)=1$, on obtient la copule Archimédienne ci-dessous

$$C_{\Phi,1}(u,v)=\Phi^{-1}\left[\left\{\Phi\left(u\right)+\Phi\left(v\right)\right\}\right].$$

- Si on pose $\Phi\left(t\right)=-\ln t$, on obtient la forme générale des copules à valeurs extrême ci-dessous

$$C_{\Phi,A}\left(u,v\right)=\exp\left\{\ln\left(uv\right)A\left(\frac{\ln u}{\ln uv}\right)\right\},$$

sachant que l'inverse de Φ est $\Phi^{-1}\left(s\right)=\exp\left(-s\right).$

Il est également à noter que si $A\left(t\right)=\max\left(t,1-t\right)$, on obtient

$$\begin{aligned}C_{\Phi,A,W}\left(u,v\right)&=\Phi^{-1}\left[\left(\Phi\left(u\right)+\Phi\left(v\right)\right)\min\left(\tfrac{\Phi(u)}{\Phi(u)+\Phi(v)},\tfrac{\Phi(v)}{\Phi(u)+\Phi(v)}\right)\right]\\&=\Phi^{-1}\left[\min\left(\Phi\left(u\right),\Phi\left(v\right)\right)\right]\\&=\max\left(u,v\right)\\&=W\left(u,v\right).\end{aligned}$$

Ceci est vrai peu importe le générateur Φ.

1.7 Copule empirique

Soit H une loi bivariée de marges continues F et G inconnues liées par une unique copule C. Étant donné un échantillon aléatoire $(x_1, y_1), ..., (x_n, y_n)$ de loi jointe H, comment peut-on estimer la couple C [19].

Pour répondre à cette question, on part du fait garanti par le théorème de Sklar que la copule C associée à H est la loi du couple (U, V). Si les marges F et G étaient connues, on pourrait estimer C par $\hat{C}$ après avoir posé $(u_i, v_i)_{i=1,\ldots,n} = (F_i, G_i)_{i=1,\ldots,n}$.

Comme ceci s'avère impossible, on fait plutôt l'estimation de C sur les pseudos observations $(\hat{u}_i, \hat{v}_i)_{i=1,\ldots,n} = \left(\hat{F}_i, \hat{G}_i\right)_{i=1,\ldots,n}$, où $\hat{F}$ et $\hat{G}$ sont les fds expérimentales respectives de F et de G, sachant que

$$\begin{cases} \hat{F}(x) = \dfrac{1}{n} \sum\limits_{i=1}^{n} \mathbf{1}(X_i \leq x). \\ \hat{G}(y) = \dfrac{1}{n} \sum\limits_{i=1}^{n} \mathbf{1}(Y_i \leq y). \end{cases}$$

L'estimateur de C est alors définit par

$$\boxed{\hat{C}(u, v) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}\left(\hat{U}_i \leq u, \hat{V}_i \leq v\right).} \tag{1.14}$$

La fonction $\hat{C}$ est appelée copule empirique, bien qu'il ne s'agit pas vraiment d'une copule. En effet, les marges de la fonction $\hat{C}$ ne sont pas uniformes sur l'intervalle I.

Il est bon de noter que $\hat{C}$ est une fonction des paires des rangs normés des observations. Soit R_i et S_i les rangs de X_i et de Y_i respectivement, on peut noter $\hat{F}(x) = R_i/n$ et $\hat{G}(y) = S_i/n$.

Soit un échantillon de taille n de données bivariées $(X_k, Y_k)_{k=1,\ldots,n}$ de couple de vas (X, Y) de fd jointes H et de marges F et G.

Définition 1.7.1 *La copule $\hat{C}$ définie sur le treillis*

$$L = \left\{ \left(\frac{i}{n}, \frac{j}{n} \right) : i, j = 1, ..., n \right\}.$$

Alors

$$\hat{C}\left(\frac{i}{n}, \frac{j}{n}\right) = \frac{1}{n}\sum_{k=1}^{n} \mathbf{1}\left(X_k \leq X_{(i)}, Y_k \leq Y_{(j)}\right).$$

$\hat{C}$ est une copule empirique, telle que $X_{(i)}$ et $Y_{(j)}$ sont les statistiques d'ordre de l'échantillon $(X_1, ..., X_n)$ et $(Y_1, ..., Y_n)$ respectivement.

Remarque 1.7.1 *On peut construire la copule empirique $\hat{C}$ par les marges empiriques $\hat{F}, \hat{G}$ et $\hat{H}$ comme suit*

$$\hat{C}(u, v) = \hat{H}\left(\hat{F}^{-1}(u), \hat{G}^{-1}(v)\right).$$

Définition 1.7.2 *La densité empirique de $\hat{C}$ notée $\hat{c}$ est donnée par*

$$\hat{c}\left(\frac{i}{n}, \frac{j}{n}\right) = \begin{cases} 1/n & si \ \left(X_{(i)}, Y_{(j)}\right) \in (X_k, Y_k) : 1 \leq k \leq n \\ 0 & si \ non. \end{cases}$$

Il existe une relation entre les fonction $\hat{C}$ et $\hat{c}$ définit comme suit

$$\hat{C}\left(\frac{i}{n}, \frac{j}{n}\right) = \sum_{p=1}^{i}\sum_{q=1}^{j} \hat{c}\left(\frac{p}{n}, \frac{q}{n}\right); 1 \leq p, q \leq n.$$

1.8 Copules associées à une copule

Il existe diverses copules associées a d'autres, on va citer : la copule survie, duale, co-copule et la copule mixte [21].

1.8.1 Copule survie

La copule de survie notée $\tilde{C}$ est définie comme suit

$$\boxed{\tilde{C}(u,v) = \mathbb{P}(X > x, Y > y) = u + v - 1 + C(1-u, 1-v).} \tag{1.15}$$

Remarque 1.8.1 *Soit $\bar{C}$ la distribution de survie définie comme suit*

$$\bar{C}(u,v) = 1 - u - v + C(u,v) \tag{1.16}$$

On peut définir la fonction de survie $\tilde{C}$ en terme de la fonction de distribution $\bar{C}$.

$$\tilde{C}(u,v) = \bar{C}(1-u, 1-v) = \bar{C}(\bar{u}, \bar{v}),$$

avec $\bar{u} = 1 - u$ et $\bar{v} = 1 - v$.

Par un raisonnement similaire, on peut exprimer les probabilités $\mathbb{P}(X \leq x \text{ ou } Y \leq y)$, $\mathbb{P}(X > x \text{ ou } Y > y)$, $\mathbb{P}(X > x \text{ ou } Y \leq y)$ en terme de copule.

1.8.2 Copule duale

La copule duale notée $\overset{*}{C}$ est définie comme suit

$$\boxed{\overset{*}{C}(u,v) = u + v - C(u,v).} \tag{1.17}$$

On a

$$\begin{aligned}
\overset{*}{C}(u,v) &= \mathbb{P}(X \leq x \text{ ou } Y \leq y) \\
&= \mathbb{P}(X \leq x) + \mathbb{P}(Y \leq y) - \mathbb{P}(X \leq x, Y \leq y) \\
&= F(x) + G(y) - H(x,y) \\
&= F(x) + G(y) - C(F(x), G(y)) \\
&= u + v - C(u,v).
\end{aligned}$$

1.8.3 Co-copule

La fonction co-copule notée $\dot{C}$ est définie comme suit

$$\boxed{\dot{C}(u,v) = 1 - C(u,v).} \tag{1.18}$$

On a

$$\begin{aligned}
\dot{C}(u,v) &= \mathbb{P}(X > x \text{ ou } Y > y) \\
&= \mathbb{P}(X > x) + \mathbb{P}(Y > y) - \mathbb{P}(X > x, Y > y) \\
&= 1 - \mathbb{P}(X \leq x) + 1 - \mathbb{P}(Y \leq y) - (1 - \mathbb{P}(X \leq x, Y \leq y)) \\
&= 1 - F(x) + 1 - G(y) - (1 - H(x,y)) \\
&= \bar{F}(x) + \bar{G}(y) - \bar{H}(x,y) \\
&= \bar{F}(x) + \bar{G}(y) - (1 - u - v + H(x,y)) \\
&= 1 - H(x,y) \\
&= 1 - C(u,v).
\end{aligned}$$

,

Remarque 1.8.2 *On note F la fonction de survie d'une va X associée à F tel que*

$$\forall x \in \mathbb{R}, \bar{F}(x) = 1 - F(x).$$

La fonction de survie H d'un couple (X,Y) associé à la fd H est donnée par

$$\bar{H}(x,y) = \mathbb{P}(X > x, Y > y) = 1 - F(x) - G(y) + H(x,y).$$

Cette dernière peut également s'écrire comme ci dessous

$$\begin{aligned}\bar{H}(x,y) &= \mathbb{P}(X>x, Y>y)\\ &= 1-\mathbb{P}(X\leq x \text{ ou } Y\leq y)\\ &= 1-(F(x)+G(y)-\mathbb{P}(X\leq x, Y\leq y))\\ &= 1-F(x)-G(y)+H(x,y)\end{aligned}$$

1.8.4 Copule mixte

La copule mixte notée C" est définit comme suit

$$\boxed{C\text{"}(u,v) = 1-u+C(u,v).} \tag{1.19}$$

On a

$$\begin{aligned}C\text{"}(u,v) &= \mathbb{P}(X>x \text{ ou } Y\leq y)\\ &= \mathbb{P}(X>x)+\mathbb{P}(Y\leq y)-\mathbb{P}(X>x, Y\leq y)\\ &= 1-F(x)+G(y)-(G(y)-C(u,v))\\ &= 1-F(x)+G(y)-G(y)+C(u,v)\\ &= 1-u+C(u,v).\end{aligned}$$

Remarque 1.8.3 *On a* $\mathbb{P}(X>x, Y\leq y) = G(y)-C(F(x), G(y))$ *car* $\mathbb{P}\left(\bar{A}\cap B\right) = \mathbb{P}(B)-\mathbb{P}(A\cap B)$.

La copule mixte peut aussi s'écrire comme suit

$$C\text{"}(u,v) = \mathbb{P}(X\leq x \text{ ou } Y>y) = 1-v+C(u,v).$$

car $\mathbb{P}(X\leq x, Y>y) = F(x)-C(F(x), G(y))$.

Chapitre 2

Mesures d'associations

Lorsque deux vas X et Y ne sont pas indépendantes on s'intéresse alors à la liaison (dépendance ou corrélation) entre elles. Il s'agit donc du phénomène d'association.

Dans ce chapitre on va étudier les mesures d'associations appelée aussi mesures de concordances les plus connues et utilisées à savoir le tau de Kendall, le rho de Spearman, le beta de Blomkvist et le gamma de Gini. Puis on donnera leurs définitions en terme d'échantillon et de population. On termine ce chapitre par la définition et les types de dépendances de queues.

2.1 Mesure de concordance

Il existe diverses mesures de concordance qui ont été proposées dans la littérature. Lorsqu'on parle de dépendance, le premier mot qui vient à l'esprit est celui du coefficient de corrélation linéaire de Pearson. Il est tout à fait approprié lorsqu'on étudie des distributions normales ou de Student, mais celui-ci perd son efficacité si le modèle est différent.

2.1.1 Notion de concordane et de discordance

Soit $(x_1, y_1), \ldots, (x_n, y_n)$ un échantillon du couple (X, Y). Il existe $C_n^2 = n!/2!\,(n-2)!$ pairs de distribution des couples (x_i, y_i) et (x_j, y_j) qui sont soit concordantes ou discordantes [14].

On dit que (x_i, y_i) et (x_j, y_j) sont concordantes si

$$\boxed{(x_i - x_j)(y_i - y_j) > 0 \Longleftrightarrow (x_i > x_j \text{ et } y_i > y_j) \text{ ou } (x_i < x_j \text{ et } y_i < y_j).}$$

On dit que (x_i, y_i) et (x_j, y_j) sont discordantes si

$$\boxed{(x_i - x_j)(y_i - y_j) < 0 \Longleftrightarrow (x_i > x_j \text{ et } y_i < y_j) \text{ ou } (x_i < x_j \text{ et } y_i > y_j).}$$

Définition 2.1.1 *Une mesure numérique d'association notée $\kappa_{X,Y}$ entre deux vas X et Y dont la copule associée est C est une mesure de concordance si et seulement si elle satisfait les propriétés suivantes [6]*

- *$\kappa_{X,Y}$ est définie pour chaque couple (X, Y) de vas continues.*
- *$-1 \leq \kappa_{X,Y} \leq 1$ sachant que $\kappa_{X,X} = 1$ et $\kappa_{X,-X} = -1$.*
- *$\kappa_{X,Y} = \kappa_{Y,X}$.*
- *Si les vas X et Y sont indépendantes, alors $\kappa_{X,Y} = 0$.*
- *Si les copules respectives C_1 et C_2 de (X_1, Y_1) et (X_2, Y_2) sont telle que $C_1 \leq C_2$, alors $\kappa_{X_1,Y_1} \leq \kappa_{X_2,Y_2}$.*
- *$\kappa_{-X,Y} = \kappa_{X,-Y} = -\kappa_{X,Y}$.*
- *Si (X_n, Y_n) est une suite de vas continues de copule associée C_n et que C_n tend vers C, alors*

$$\lim_{n\to\infty} \kappa_{X_n,Y_n} = \kappa_{X,Y}.$$

– *Si* $\alpha(X)$ *et* $\beta(Y)$ *sont des fonctions strictement croissantes, alors*

$$\kappa_{\alpha(X),\beta(Y)} = \kappa_{X,Y}.$$

2.1.2 Fonction de concordance

La fonction de concordance [11] entre deux couples indépendant (X_1, Y_1) et (X_2, Y_2) est définit par

$$\boxed{Q = \mathbb{P}\left[(X_1 - X_2)(Y_1 - Y_2) > 0\right] - \mathbb{P}\left[(X_1 - X_2)(Y_1 - Y_2) < 0\right].} \tag{2.1}$$

Théorème 2.1.1 *Soit* (X_1, Y_1) *et* (X_2, Y_2) *deux vas indépendantes de fd* H_1 *et* H_2 *avec des marge communes* F *et* G *respectivement. Soient* C_1 *et* C_2 *les copules associées aux fds* H_1 *et* H_2 *respectivement. On peut montrer que la fonction de concordance peut aussi être exprimée en fonction de la copule*

$$\boxed{Q_{C_1,C_2} = 4 \iint_{I^2} C_2(u,v)\, dC_1(u,v) - 1.} \tag{2.2}$$

Preuve. On pose $u = F(x)$ et $v = G(y)$

$$\begin{aligned} Q &= \mathbb{P}\left[(X_1 - X_2)(Y_1 - Y_2) > 0\right] - \mathbb{P}\left[(X_1 - X_2)(Y_1 - Y_2) < 0\right] \\ &= \mathbb{P}\left[(X_1 - X_2)(Y_1 - Y_2) > 0\right] - \left(1 - \mathbb{P}\left[(X_1 - X_2)(Y_1 - Y_2) > 0\right]\right) \\ &= 2\mathbb{P}\left[(X_1 - X_2)(Y_1 - Y_2) > 0\right] - 1. \end{aligned}$$

Avec

$$\mathbb{P}\left[(X_1 - X_2)(Y_1 - Y_2) > 0\right] = \mathbb{P}(X_1 < X_2, Y_1 < Y_2) + \mathbb{P}(X_1 > X_2, Y_1 > Y_2).$$

On a

$$
\begin{aligned}
\mathbb{P}\left(X_1 > X_2, Y_1 > Y_2\right) &= \mathbb{P}\left(X_2 < X_1, Y_2 < Y_1\right) \\
&= \iint\limits_{\mathbb{R}^2} \mathbb{P}\left(X_2 < X_1, Y_2 < Y_1\right) dC_1\left(F\left(x\right), G\left(y\right)\right) \\
&= \iint\limits_{\mathbb{R}^2} C_2\left(F\left(x\right), G\left(y\right)\right) dC_1\left(F\left(x\right), G\left(y\right)\right) \\
&= \iint\limits_{I^2} C_2\left(u, v\right) dC_1\left(u, v\right).
\end{aligned}
$$

De façon similaire

$$
\begin{aligned}
\mathbb{P}\left(X_1 < X_2, Y_1 < Y_2\right) &= \mathbb{P}\left(X_2 > X_1, Y_2 > Y_1\right) \\
&= \iint\limits_{\mathbb{R}^2} \mathbb{P}\left(X_2 > X_1, Y_2 > Y_1\right) dC_1\left(F\left(x\right), G\left(y\right)\right) \\
&= \iint\limits_{\mathbb{R}^2} \bar{H}\left(x, y\right) dC_1\left(F\left(x\right), G\left(y\right)\right) \\
&= \iint\limits_{\mathbb{R}^2} \left[1 - F\left(x\right) - G\left(y\right) + C_2\left(F\left(x\right), G\left(y\right)\right)\right] dC_1\left(F\left(x\right), G\left(y\right)\right) \\
&= \iint\limits_{I^2} \left[1 - u - v + C_2\left(u, v\right)\right] dC_1\left(u, v\right).
\end{aligned}
$$

Sachant que C_1 est la fd du couple (U, V) d'une loi uniforme sur l'intervalle I, alors $E\left(U\right) = E\left(V\right) = \frac{1}{2}$. Donc

$$
\begin{aligned}
\mathbb{P}\left(X_1 < X_2, Y_1 < Y_2\right) &= 1 - \frac{1}{2} - \frac{1}{2} + \iint\limits_{I^2} C_2\left(u, v\right) dC_1\left(u, v\right) \\
&= \iint\limits_{I^2} C_2\left(u, v\right) dC_1\left(u, v\right).
\end{aligned}
$$

Alors

$$
\boxed{Q_{C_1, C_2} = 4 \iint\limits_{I^2} C_2\left(u, v\right) dC_1\left(u, v\right) - 1.}
$$

■

2.2 Types des mesures de concordances

Il existe diverses mesures d'associations parmis eux le rho de Spearman, le tau de Kendall, le Gamma de Gini, le bêta de Blomkvist, ect.

2.2.1 Rho de Spearman

La mesure d'association la plus connue est le rho de Spearman, elle est basée sur la concordance et la discordance. Cette mesure a été découverte par *Spearman* [20] en 1904.

Définition 2.2.1 *Soient* $(x_1, y_1), (x_2, y_2)$ *et* (x_3, y_3) *trois paires indépendantes de vas de même fd jointe H. Le coefficient de corrélation de Spearman est définit comme suit*

$$\boxed{\rho = 3\left[\mathbb{P}\left(X_1 - X_2\right)\left(Y_1 - Y_3\right) > 0\right] - \mathbb{P}\left[\left(X_1 - X_2\right)\left(Y_1 - Y_3\right) < 0\right].} \tag{2.3}$$

Autrement dit, il s'agit de la différence entre la probabilité de concordance et celle de la discordance.

Théorème 2.2.1 *Soit* (X, Y) *deux vas continues de copule* C *.Le rho de Spearman de ce couple est défini par*

$$\boxed{\rho_C = 3Q_{C,\Pi} = 12 \iint_{I^2} uvdC\left(u, v\right) - 3.} \tag{2.4}$$

2.2.2 Tau de Kendall

Une autre mesure d'association très connue est celle du tau de Kendall, cette dernière est définie en termes de concordance et de discordance. Elle est nommée ainsi en hommage à *Maurice Kendall* [**13**] qui a développé cette mesure en 1938.

Définition 2.2.2 *Soient* (x_1, y_1) *et* (x_2, y_2) *deux paire indépendantes de vas, ayant pour fd marginales communes* F *(pour* X_1 *et* X_2*) et* G *(pour* Y_1 *et* Y_2*). Le tau de Kendall se définit comme suit*

$$\boxed{\tau = \mathbb{P}\left[\left(X_1 - X_2\right)\left(Y_1 - Y_2\right) > 0\right] - \mathbb{P}\left[\left(X_1 - X_2\right)\left(Y_1 - Y_2\right) < 0\right].} \tag{2.5}$$

Théorème 2.2.2 *Soit (X,Y) deux vas de copule C ayant pour fd jointe H. Le tau de Kendall de ce couple est défini par*

$$\boxed{\tau_C = Q_{C,C} = 4\iint_{I^2} C\left(u,v\right) dC\left(u,v\right) - 1.} \tag{2.6}$$

Le tableau ci dessous présent le tau de kendall et le rho de Spearman de quelques copules archimediennes

Copule	τ_C	ρ_C
Clayton	$a/\left(a+2\right)$	/
Frank	$1-\left[4\left(1-D_1\left(a\right)\right)/a\right]$	$1-\left[12\left(D_1\left(a\right)-D_2\left(a\right)\right)/a\right].$
Gumbel	$\left(a-1\right)/a$	/

TAB. 2.1 – Tau de Kendall et rho de Spearman de quelques copules archimédiennes.

[1]

2.2.3 Gamma de Gini

Une autre mesure de concordance est la mesure gamma de Gini qui a été développée en 1910 par *Corrado Gini* [10].

Définition 2.2.3 *Soient R_i et S_i les rangs des observation dans un échantillon de taille n de deux vas continues X et Y respectivement, alors*

$$\boxed{\gamma = \frac{1}{\lfloor n^2/2\rfloor}\left[\sum_{i=1}^{n}\left|R_i+S_i-n-1\right| - \sum_{i=1}^{n}\left|R_i-S_i\right|\right],} \tag{2.7}$$

où $\lfloor t\rfloor$ est la partie entière de t.

[1] Fonction de Debye : $D_k\left(a\right) = \dfrac{k}{a^k}\displaystyle\int_0^a \frac{t^k}{e^t-1}dt; k\geq 1.$

Théorème 2.2.3 *Soit (X, Y) deux vas de fd F et G respectivement. On pose $u = F(x)$ et $v = G(y)$, la mesure gamma de Gini est définie comme suit*

$$\boxed{\gamma_C = 2 \iint\limits_{I^2} (|u+v-1| - |u-v|)\, dC(u,v)\,.} \tag{2.8}$$

Remarque 2.2.1 *Sachant qu'il y'a une distance entre C et Π, il y'a aussi une autre distance entre C et W ou C et M ou les deux. Soit (X, Y) deux vas de copule associée C. Le gamma de Gini peut aussi s'écrire comme suit*

$$\gamma_G = Q_{C,M} + Q_{C,W} = 4\left[\int_0^1 C(u, 1-u)\, du - \int_0^1 (u - C(u,u))\, du\right]. \tag{2.9}$$

Preuve. Soit

$$\begin{aligned} Q_{C,M} &= 4 \iint\limits_{I^2} C(u,v)\, dM(u,v) - 1 &&= 4 \int\limits_0^1 C(u,u)\, du - 1. \\ Q_{C,W} &= 4 \iint\limits_{I^2} C(u,v)\, dW(u,v) - 1 &&= 4 \int\limits_0^1 C(u,1-u)\, du - 1. \end{aligned}$$

On trouve

$$\gamma_G = 4\left[\int_0^1 C(u, 1-u)\, du - \int_0^1 (u - C(u,u))\, du\right].$$

■

2.2.4 Beta de Blomkvist

Il existe une autre mesure de concordance est la mesure Beta de Blomkvist découverte en 1950 par *Blomkvist* [4].

Définition 2.2.4 *Soit (X, Y) deux vas et $(\tilde{x}, \tilde{y})$ est la médianes de X et Y respectivement. Le beta de Blomkvist est défini comme suit*

$$\boxed{\beta = \mathbb{P}\left[(X - \tilde{x})(Y - \tilde{y}) > 0\right] - \mathbb{P}\left[(X - \tilde{x})(Y - \tilde{y}) < 0\right].} \tag{2.10}$$

Théorème 2.2.4 *Soit* (X, Y) *deux vas continues de copule associée* C *ayant pour fd jointe* H *avec* $F(\tilde{x}) = G(\tilde{y}) = 1/2$. *Le beta de Blomkvist de ce couple est défini comme suit*

$$\boxed{\beta_C = 4C\left(\frac{1}{2}, \frac{1}{2}\right) - 1.} \tag{2.11}$$

Il est a noter qu'il existe d'autres mesures d'associations moins connues que celles citées précédemment comme celle de *Schweizer et Wolf*.

2.3 Dépendance de queue

Parmis les concepts de dépendance on a le concept de dépendance de queue. Il existe deux coefficients de dépendance de queue : le coefficient de dépendance inférieure noté λ_L et le coefficient de dépendance supérieure noté λ_U [14].

Définition 2.3.1 *Soit* X *et* Y *deux vas continues de fds respectives* F *et* G.

- *Le coefficient de dépendance supérieure* λ_U *est définit par*

$$\boxed{\lambda_U = \lim_{u \to 1^-} \mathbb{P}\left(X > F^{-1}(u) / Y > G^{-1}(u)\right).} \tag{2.12}$$

- *Le coefficient de dépendance inférieure* λ_L *est définit par*

$$\boxed{\lambda_L = \lim_{u \to 0^+} \mathbb{P}\left(X \leq F^{-1}(u) / Y \leq G^{-1}(u)\right).} \tag{2.13}$$

Théorème 2.3.1 *Soit (X,Y) deux vas de copule associée C ayant pour fd jointe H. Les coefficients de dépendance inférieure et supérieure peuvent aussi s'écrire en terme de copule comme suit*

$$\boxed{\lambda_U = 2 - \lim_{u\to 1^-} \frac{1 - C(u,u)}{1-u}.} \tag{2.14}$$

Et

$$\boxed{\lambda_L = \lim_{t\to 0^+} \frac{C(u,u)}{u}.} \tag{2.15}$$

Preuve. On a

$$\begin{aligned}
\lambda_U &= \lim_{u\to 1^-} \mathbb{P}\left(X > F^{-1}(u) / Y > G^{-1}(u)\right) \\
&= \lim_{u\to 1^-} \mathbb{P}\left(F(X) > u / G(Y) > u\right) \\
&= \lim_{u\to 1^-} \left[\mathbb{P}\left(F(X) > u \cap G(Y) > u\right)\right] / \mathbb{P}\left(G(Y) > u\right) \\
&= \lim_{u\to 1^-} \left[\mathbb{P}\left(F(X) > u \cap G(Y) > u\right)\right] / 1 - \mathbb{P}\left(G(Y) \leq u\right) \\
&= \lim_{u\to 1^-} \left[1 - \mathbb{P}\left(F(X) \leq u, G(Y) \leq u\right)\right] / 1 - \mathbb{P}\left(G(Y) \leq u\right) \\
&= \lim_{u\to 1^-} \left[1 - C(u,u)\right] / 1 - u \\
&= \lim_{u\to 1^-} \bar{C}(u,u) / 1 - u \\
&= \lim_{u\to 1^-} \left[1 - 2u + C(u,u)\right] / 1 - u \\
&= 2 - \lim\nolimits_{t\to 1^-} \left[1 - C(u,u)\right] / 1 - u.
\end{aligned}$$

Puis

$$\begin{aligned}
\lambda_L &= \lim_{u\to 0^+} \mathbb{P}\left(X \leq F^{-1}(u) / Y \leq G^{-1}(u)\right) \\
&= \lim_{u\to 0^+} \mathbb{P}\left(F(X) \leq u / G(Y) \leq u\right) \\
&= \lim_{u\to 0^+} \left[\mathbb{P}\left(F(X) \leq u \cap G(Y) \leq u\right)\right] / \mathbb{P}\left(G(Y) \leq u\right) \\
&= \lim_{u\to 0^+} \left[C(u,u)\right] / u.
\end{aligned}$$

■

Remarque 2.3.1 - *Si* $\lambda_L \in \left]0,1\right]$ *alors la copule* C *a une dépendance de queue inférieure.*

- Si $\lambda_L = 0$ alors la copule C n'a pas de dépendance de queue inférieure.
- Si $\lambda_U \in \left]0,1\right]$ alors la copule C a une dépendance de queue supérieure.
- Si $\lambda_U = 0$, alors la copule C n'a pas de dépendance de queue supérieure.

Exemple 2.3.1 *Le tableau ci-dessous contient les mesures de dépendances de queue de quelques copules.*

Copule	λ_L	λ_U
Frank	0	0
Clayton	$2^{\frac{-1}{a}}$	0
Gumbel	0	$2 - 2^{\frac{1}{a}}$

TAB. 2.2 – Mesures de dépendances de queue

2.4 Simulation des copules bivariées

Les copules archimédiennes sont très faciles à construire, de nombreuses familles paramétriques appartiennent à cette classe et ont une grande variété de structure de dépendance. C'est pourquoi de nombreux chercheurs travaillent avec ces copules, la plupart d'entre eux étudient le cas bivarié [15].

Pour simuler des réalisations d'un vecteur bivarié $X = (X_1, X_2)$ à partir de la décomposition de H, il faut être capable de simuler des réalisations d'un vecteur $U = (U_1, U_2)$ ayant des marginales uniformes et dont la distribution est la copule C, puisqu'on a l'égalité en loi

$$X = \left(F_1^{-1}(U_1), F_2^{-1}(U_2)\right).$$

Tous se ramène donc à savoir simuler la copule C.

2.4.1 Méthode des distributions conditionnelles

On se donne un vecteur U dont la copule associée est C, on simule (v_1, v_2) uniformes et indépendantes.

On pose alors $u_1 = v_1$, puis $u_2 = C_{u_1}^{-1}(v_2)$ avec $C_{u_1}(u_2) = \mathbb{P}(U_2 \leq u_2 \diagdown U_1 = u_1)$.

Cette expression se calcule à partir de la copule C comme suit

$$\mathbb{P}(U_2 \leq u_2 \diagdown U_1 = u_1) = \frac{\partial C(u_1, u_2)}{\partial u_1} / \frac{\partial C(u_1, 1)}{\partial u_1}.$$

[2]

Toutefois en pratique il n'est pas aisé d'inverser la fonction conditionnelle.

Exemple 2.4.1 (Copule de Frank) *Dans le cas de cette copule, la méthode des distributions conditionnelles peut être mise en œuvre simplement. D'après 1.9, on déduit que*

$$\begin{aligned} C_{U_2 \backslash U_1}(u_1, u_2) &= \partial C_a(u_1, u_2) / \partial u_1 \\ &= e^{-au_1}(e^{-au_2} - 1) / \left[e^{-a} - 1 + (e^{-au_1} - 1)(e^{-au_2} - 1)\right]. \end{aligned}$$

On peut alors inverser cette équation (en résolvant l'équation $C_{U_2/U_1}(u_1, u_2) = u$), on obtient

$$C_{U_2/U_1}^{-1}(u_1, u) = -\frac{1}{a} \ln\left(1 + \frac{u(e^{-a} - 1)}{u + (1 - u) e^{-au_1}}\right).$$

Donc

$$C_{U_2/U_1}^{-1}(u_1, u_2) = -\frac{1}{a} \ln\left(1 + \frac{u_2(e^{-a} - 1)}{u_2 + (1 - u_2) e^{-au_1}}\right).$$

La simulation peut donc être mise en œuvre simplement.

[2] On note
$C_{U_2/U_1}(u_1, u_2) = \partial C(u_1, u_2) / \partial u_1$.
$C_{U_2/U_1}(u_1, 1) = \partial C(u_1, 1) / \partial u_1$.

Exemple 2.4.2 (Copule de Gumbel) *D'après 1.11 on en déduit que*

$$\begin{aligned} C_{U_2/U_1}(u_1, u_2) &= \partial C_a(u_1, u_2)/\partial u_1 \\ &= \frac{1}{u_1}\left[1 + \left(\frac{\ln u_2}{\ln u_1}\right)^a\right]^{-1+1/a} C_a(u_1, u_2). \end{aligned}$$

L'inversion de cette expression n'est pas aisée, et on aura ici recours à des méthodes numériques.

2.4.2 Autre méthode

Il existe une autre méthode spécifique à d'autres copules particulières, par exemple pour la copule de Clayton on a l'algorithme suivant

Exemple 2.4.3 (Copule de Clayton)

1. *On génère des vas uniformes s et t à partir desquelles on détermine $x = -\ln(s)$ et $y = -\ln(t)$.*
2. *Simuler une va de loi Gamma de paramètres 1 et θ : $z \backsim \Gamma(1, \theta)$.*
3. *Déterminer les réalisations de la copule Clayton à partir des expressions suivantes : $u = \left(1 + \frac{x}{z}\right)^{-\theta}$ et $v = \left(1 + \frac{y}{z}\right)^{-\theta}$.*

2.4.3 Exemples d'application

Dans cette partie on va essayer d'appliquer les méthodes citées précédemment en utilisant le logiciel R sur quelques copules archimédiennes comme la copule de Clayton, Gumbel et Frank.

Les packages et fonctions couramment utilisés dans ce logiciel sont : copula, Copula, iTau, claytonCopula, plot, gumbelCopula, frankCopula, rCopula, qnorm.

a. Méthode des distributions conditionnelles :

1. Copule de Clayton :

– Pour $u_1 = 0.05$:

```
library("copula")
tau=0.5                                  # Tho de Kendall.
a=iTau(claytonCopula(), tau = tau)       # Paramètre de la copule (a > 0).
a=2
dim=2                                    # Cas bivarié.
clayton=claytonCopula(a, dim=2)          # Générer la copule de Clayton.
n=1000                                   # Taille de l'échantillon.
set.seed(271)
u1=0.05                                  # Vecteur U1.
U=cCopula(cbind(u1, runif(n)),           # Définir la copule de Clayton.
copula=clayton, inverse=TRUE)
par(mfrow=c(1,2))
plot(U[,2], ylab = quote(U[2]),          # Nuage de points de la copule.
xlab=quote(U[1]),col="blue",
main = "u1=0.05")
```

– Pour $u_1 = 0.95$:

```
u1=0.95
U=cCopula(cbind(u1, runif(n)),copula=clayton, inverse=TRUE)
plot(U[,2], ylab = quote(U[2]),xlab=quote(U[1]),col="red",main="u1=0.95")
```

– Graphes :

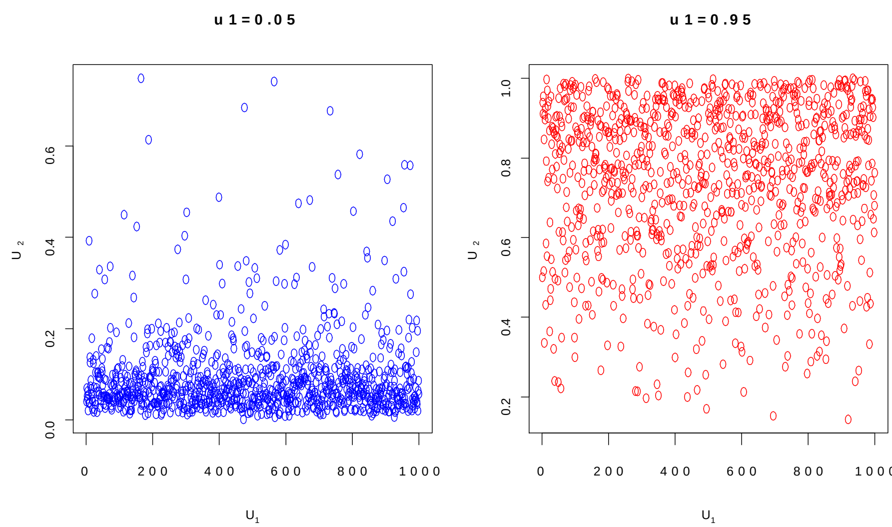

FIG. 2.1 – Nuages de points de la copule de Clayton pour n=1000 et a=2.

Commentaire :

La copule de Clayton possède une dépendance de queue inferieure, avec une moyenne mesure d'association.

2. Copule de Gumbel

– Pour $u_1 = 0.1$:

Rho=0.75

a=iRho(gumbelCopula(), Rho) # Rho de Spearman

a=2.29

dim=2

gumbel=gumbelCopula(a, dim =2) # Générer la copule de Gumbel.

n=1000

set.seed(271)

u1 = 0.1

U=cCopula(cbind(u1, runif(n)), # Définir la copule de Gumbel.

copula=gumbel, inverse=TRUE)

par(mfrow=c(1,2))

plot(U[,2], ylab = quote(U[2]),

xlab=quote(U[1]),col="blue",

main = "u1=0.1")

– Pour $u_1 = 0.90$:

u1=0.90

U=cCopula(cbind(u1, runif(n)), copula=gumbel, inverse=TRUE)

plot(U[,2], ylab = quote(U[2]),xlab=quote(U[1]),col="red",main="u1=0.90")

Graphes :

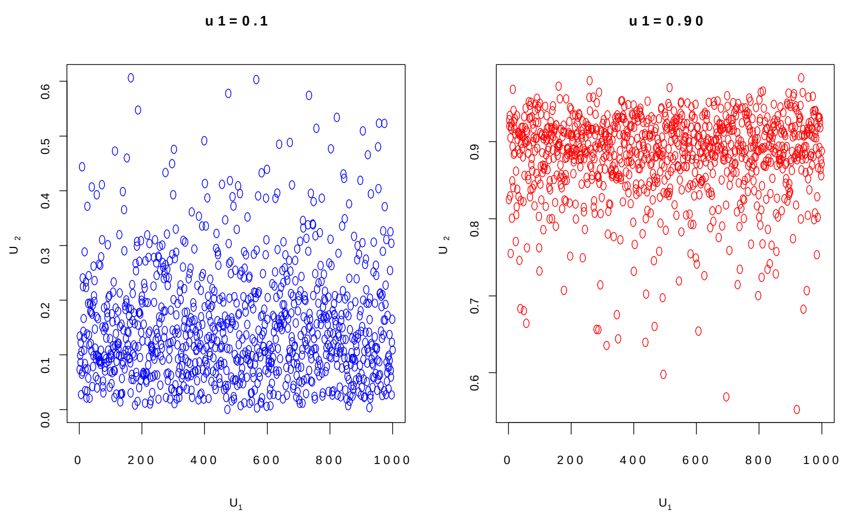

FIG. 2.2 – Nuages de points de la copule de Gumbel pour n=1000 et a=2.29.

Commentaire :

La copule de Gumbel possède une dépendance de queue supérieure, avec une forte mesure d'association.

3. Copule de Frank

– Pour $u_1 = 0.2$:

```
tau=0.1
a=iTau(frankCopula(), tau = tau)
a=0.91
dim=2
frank=frankCopula(a, dim = 2)        # Générer la copule de Frank.
n=1000
set.seed(271)
u1 = 0.2
U = cCopula(cbind(u1, runif(n)),     # Définir la copule de Frank.
copula = frank, inverse = TRUE)
par(mfrow=c(1,2))
plot(U[,2], ylab = quote(U[2]),
xlab=quote(U[1]),col="blue",
main = "u1=0.2")
```

– Pour $u_1 = 0.8$:

```
u1=0.8
U=cCopula(cbind(u1, runif(n)), copula=frank, inverse=TRUE)
plot(U[,2], ylab = quote(U[2]),xlab=quote(U[1]),col="red",main="u1=0.8")
```

Graphes :

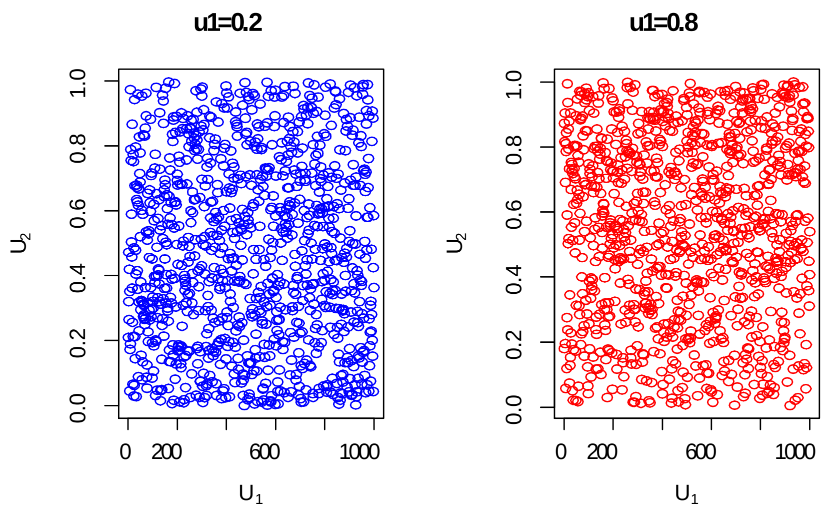

FIG. 2.3 – Nuages de points de la copule de Frank pour n=1000 et a=0.91.

Commentaire :

La copule de Frank ne possède ni une dépendance de queue supérieure ni une dépendance de queue inferieure, avec une très faible mesure d'association.

a. Autre méthode :

1. Copule de Clayton :

```
clayton_copula =claytonCopula(2, dim = 2)
n=1000
u1=rCopula(n, clayton_copula)                    # Générer le vecteur U.
v1=rCopula(n, clayton_copula)                    # Générer le vecteur V.
x1=qnorm(u1)                                     # Fonction quantile.
y1=qnorm(v1)
par(mfrow=c(1,3))
plot(x1,y1,col="red",main="Copule de Clayton pour a=2")
```

2. Copule de Gumbel :

```
gumbel_copula =gumbelCopula(2.29, dim = 2)
n=1000
u2=rCopula(n, gumbel_copula)
v2=rCopula(n, gumbel_copula)
x2=qnorm(u2)
y2=qnorm(v2)
plot(x2,y2,col="blue",main="Copule de Gumbel pour a=2.29")
```

3. Copule de Frank :

```
frank_copula =frankCopula(0.91, dim = 2)
n=1000
u3=rCopula(n, frank_copula)
v3=rCopula(n, frank_copula)
x3=qnorm(u3)
y3=qnorm(v3)
plot(x3,y3,col="black",main="Copule de Frank pour a=0.91")
```

Graphes :

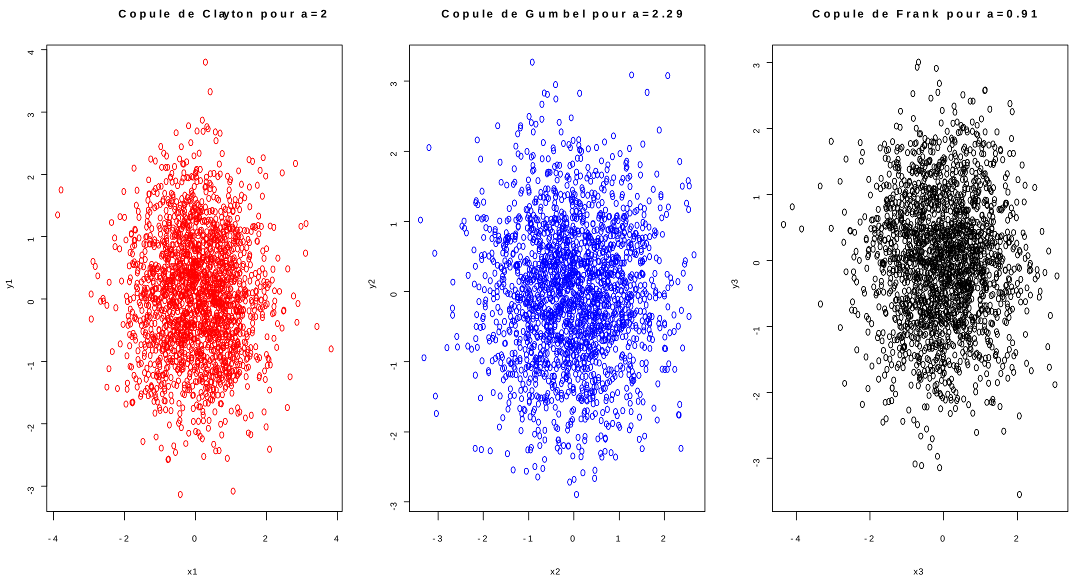

FIG. 2.4 – Nuages de points des copules archimediennnes pour n=1000.

Commentaire :

- Dans cet exemple, on a d'abord défini les trois copules à un paramètre (Clayton, Gumbel, Frank) en utilisant leurs fonctions du package "copula". Ensuite, on a généré un échantillon de deux vas de taille 1000 suivant ces copules en utilisant la fonction "rCopula()". Enfin, on a transformé ces vas en variables normales en utilisant la fonction "qnorm()". Ces vas transformées suivront alors une distribution normale bivariée avec les copules associées.

- Ses exemples illustre l'utilisation des copules pour modéliser la dépendance entre deux vas et la transformation de ces vas en une distribution normale bivariée. Cette méthode montre comment les copules peuvent être utilisées pour modéliser la dépendance entre les vas et comment cette dépendance peut être transformée en une distribution normale bivariée pour faciliter l'analyse statistique.

Conclusion

Les copules sont des outils mathématiques utiles pour modéliser les dépendances entre les vas, en particulier dans des situations où les distributions marginales des vas ne sont pas connues ou ne peuvent pas être facilement modélisées. Les copules permettent de séparer la structure de dépendance des marges, ce qui permet de mieux comprendre et modéliser les relations entre les vas.

Les mesures d'association sont des outils statistiques utilisées pour quantifier la force et la direction de la relation entre deux vas. Les mesures d'association couramment utilisées comprennent le coefficient de corrélation de Pearson, le coefficient de corrélation de Spearman et le coefficient de corrélation de Kendall.

En combinant les copules avec des mesures d'association, il est possible de mieux comprendre et quantifier la dépendance entre deux vas, en particulier dans des situations où les relations ne sont pas linéaires ou où les marges ne sont pas normalement distribuées.

En conclusion, les copules et les mesures d'association sont des outils utiles pour modéliser et quantifier les relations entre les vas. Leur utilisation peut être particulièrement précieuse dans des situations où les distributions marginales des vas ne sont pas connues ou ne peuvent pas être facilement modélisées.

Bibliographie

[1] Aleiyouka, M. (2018). *Sur la dépendance des queues de distributions.* Doctoral dissertation, Normandie Université.

[2] Ben Ghachem, M. (2016). *Méthodologies statistiques pour la vérification de relations d'ordre stochastique entre deux variables aléatoires.* Doctoral dissertation, Université du Québec à Trois-Rivières.

[3] Benelmir, I. (2018). *Modélisation de la Dépendance par les Copules.* Doctoral dissertation, Université Mohamed Kheider-Biskra.

[4] Blomkvist, N., (1950) . *On a mesure of dependance between two vas.* The Annals of mathematical statistic, 21 (4) .

[5] Charpentier, A. (2010). *Copules et risques corrélés.* Journées d'Etudes Statistique.

[6] Chine, A., (2018) . *About copula statistic*, Université Mohammed kheider-Biskra.

[7] Dall'Aglio, G. (1956). *Sugli estremi dei momenti delle funzioni di ripartizione doppia.* Annali della Scuola Normale Superiore di Pisa-Classe di Scienze, 10(1 − 2).

[8] Fontaine, C. (2016). *Utilisation de copules paramétriques en présence de données observationnelles : cadre théorique et modélisations.* Doctoral dissertation, Université Montpellier.

[9] Fréchet, M. (1951). *Sur les tableaux de corrélation dont les marges sont données.* Ann. Univ. Lyon, $3^{\text{ème}}$ serie, Sciences, Sect. A, 14, $53 - 77$.

[10] Gini, C. (1910). *Indice di concentrazione e di dipendenza.* Bibl.

[11] Guessouri, A., (2019) . *Sur les copules et applications*, Université du Mohamed kheider-Biskra.

[12] Ihaka, R., & Gentleman, R. (1996). *R : a language for data analysis and graphics. Journal of computational and graphical statistics*, 5(3), $299 - 314$.

[13] Kendall, M. G. (1938). *A new measure of rank correlation.* Biometrika, 30(1/2), $81 - 93$.

[14] Nelsen, R. B. (2006). *An Introduction to Copulas.* Springer, New York. MR2197664.

[15] Planchet, F. (2006), *Simulation Techniques avancées. Winter and Associés.*

[16] Popier, A. (2010). *Copules.* Université du Maine, Le Mans.

[17] Reggas, K. (2018). *Copules multidimensionnelles et stratégies pour estimer leurs paramètres.* Doctoral dissertation, Université du Québec à Trois-Rivières.

[18] Salem, K. (2015). *À propos de la covariance limite du processus de copule empirique.* Doctoral dissertation, Université du Québec à Trois-Rivières.

[19] Sklar, M. (1959). *Fonctions de répartition à n dimensions et leurs marges.* In Annales de l'ISUP (Vol. 8, No. 3, pp. $229 - 231$).

[20] Spearman, C. (1961). *" General Intelligence" Objectively Determined and Measured.*

[21] Tebboub, S., Hamdellou, R. (2018). *Les copule dans* $\mathbb{R}^2$, Université Mohammed Seddik Ben Yehai-Jijel.

Annexe A : Logiciel R

2.5 Qu'est-ce-que le langage R ?

Le langage R est un langage de programmation et un environnement mathématique utilisé pour le traitement des données. Il permet de faire des analyses statistiques aussi bien simples que complexes comme des modèles linéaires ou non-linéaires, des tests d'hypothèse, de la modélisation des séries chronologiques, de la classification, etc. Il dispose également de nombreuses fonctions graphiques très utiles et de qualité professionnelle.

Le logiciel R a été créé par *Ross Ihaka et Robert Gentleman* [**12**] à l'Université d'Auckland, Nouvelle Zélande et a été développé par le R Developement Core Team.

L'origine du nom du langage provient d'une part, des initiales des prénoms des deux auteurs (*Ross Ihaka et Robert Gentleman*) et d'autre part, d'un jeu de mots sur le nom du langage S auquel il est apparenté.

Annexe B : Abréviations et Notations

Les différentes abréviations et notations utilisées tout au long de ce mémoire sont expliquées ci-dessous :

vas	: Variables aléatoires.
X, Y	: Vas de la loi jointe.
U, V	: Vas de la copule.
fd	: Fonction de distribution.
n	: Taille de l'échantillon.
I	: Intervalle $[0, 1]$.
F, G	: Fds marginales de X et Y.
H	: Fd jointe.
F^{-1}, G^{-1}	: Fonction inverse.
C	: Fd du couple (U, V).
c	: Densité de la copule.
α, β	: Fonctions de x.
a, θ	: Paramètres des copules.
$\prod$	: Copule produit.
M	: Copule minium.

W : Copule maximum.

C_ρ : Copule Gaussienne.

$C_{\rho,v}$: Copule de Student.

Φ_ρ : Fonction de la loi normale bivariée.

Φ : Fonction de la loi normale standard.

C_a : Copule archimédienne.

$\Phi(x)$: Générateur de la copule.

$C_{\theta,A}$: Copule archimax.

$A(t)$: Fonction de dépendance.

$\hat{F}, \hat{G}$: Distributions empiriques de X et Y respectivement.

$\hat{C}$: Copule empirique.

$\tilde{C}$: Copule survie.

$\bar{C}$: Distribution de la copule de survie.

$\overset{*}{C}$: Copule duale.

$\dot{C}$: Co-copule.

C" : Copule mixte.

$\kappa_{X,Y}$: Mesure numérique d'association.

Q_{C_1,C_2} : Fonction de concordance.

ρ_C : Rho de Spearman.

τ_C : Tho de Kendall.

D_k : Fonction de Debye.

β_C : Beta de Blomkvist.

γ_C : Gamma de Gini.

$\lfloor t \rfloor$: Partie entière.

λ_L : Dépendance de queue inférieure.

λ_U : Dépendance de queue supérieure.

الملخــــص:

تعتبر نظرية الكوبيل، التي نشأت من أبحاث سكلار في عــام 1959 لتمثيل الاعتمادية بين المتغيرات العشوائية. يهدف هذا البحث إلى دراسة مختلف مقاييس الارتباط التي يمكن تعريفها بالاعتماد الى الكوبيل.

الكلمات الرئيسية: الكوبول، نظرية سكلار، ارخميدس كوبيل، مقياس الارتباط، طو الكندال، رو سبيرمان، غاما جيني، بيتا بلومكفيست، اعتمادية الذيل.

Résumé:

La théorie des copules, issue des travaux de Sklar en 1959, permet une modélisation flexible de la dépendance entre les variables aléatoires. L'objectif de ce mémoire est d'étudier les différentes mesures d'associations qui peuvent être définies en termes de copules.

Mot clés : Copule, Théorème de Sklar, Copule archimédienne, Mesure d'association, Tau de Kendall, Rho de Spearman, Gamma de Gini, Beta de Blomkvist, Dépendance de queue.

Abstract:

Copula theory, stemming from Sklar's work in 1959, flexible modeling of dependence between random variables. The objective of this thesis is to study different association measures which can be defined by copula terms.

Keywords: Copula, Sklar's theorem, Archimedean copula, Association measure, Kendall's tau, Spearman's rho, Gini's gamma, Blomkvist's beta, tail dependence.