Ministère de l'enseignement supérieur et de la recherche scientifique

Université Yahia Farès de Médéa

Faculté des sciences -département de mathématiques & informatique



Projet de fin d'étude

MÉMOIRE

Présenté en vue de l'obtention du diplôme de
Master en Mathématiques

Option : Recherche Opérationnelle

Intitulé

Théorie des copules paramétriques et non paramétriques

Présenté par

Hadjer KORICHI

Composition du Jury :

M. A. LEGHMIZI	M.C.A	Université.de Médéa	Président.
K. SERIER ABDALLAH	M.A.A	Université de Médéa	Directrice du mémoire.
N. IKHLEF ESCHOUF	M.C.A	Université de Médéa	Examinateur.
A. BENDALI-BRAHAM	M.A.B	Université de Médéa	Examiatrice.

Juillet 2019

Dédicace

“A la mémoire de mon père "

“A ma trés chère mère“

“A mes frères, mes soeurs“

"A tous mes futures"

Remerciements

Je voudrais tout d'abord à remercier Mme K. SERIER ABDALLAH pour m'avoir encadrer d'une part, et d'autre part m'avoir proposer un sujet aussi riche et plus intéressant je le suit également reconnaissante pour ses qualités pédagogiques, ses précieux conseils, ses stimulants encouragements et sa disponibilité, j'ai pu mener à bien ce travail dans des conditions favorables. Elle m'a également permis d'approfondir mes connaissances et acquérir un savoir faire.

Mes remerciements vont à Mr M.A. LEGHMIZI qu'il m'a fait l'honneur en acceptant de présider le jury et de l'intérêt qu'il a porté à ce travail.

Mes remerciements vont aussi, aux membres de jury Mme A. BENDALI-BRAHAM *et Mr* N. IKHLEF ESCHOUF *qui ont accepté de le composé prenant du temps et du soin qu'ils ont apporté à ce travail.*

Je tiens à exprimer ma reconnaissance à toute ma famille, ,mes amies et tous ceux qui m'ont aidé de près ou de loin.

Résumé

Ce travail est consacré à l'étude de la théorie des copules paramétriques et non paramétriques. Les copules sont des outils mathématiques performants qui permettent de déceler la structure de dépendance d'un tel modèle. Ce sont des fonctions de répartition particulières, qui relie les fonctions de répartition multivariées de lois de probabilité dans $\mathbb{R}^d$ $(d \geq 2)$ et les fonctions de répartition marginales et leurs coordonnées. Sklar (1959) (voir [31]) a introduit le théorème fondamental de la théorie des copules.

Dans ce mémoire nous nous interessons à l'estimation des paramètres de copules par les méthodes paramétriques, les méthodes d'estimation non paramétriques sont aussi étudies.

Abstract

this work is devoted to the study of the theory of parametric and non-parametric copulas. Copulas are powerful mathematical tools that can detect the dependency structure of such a model. These are specific dispatch functions, which combine the multivariate distribution functions of probability laws in $\mathbb{R}^d$ ($d \geq 2$), and the marginal distribition functions and their coordinates. Sklar (1959) (see[30]) introduced the fundamental theorem of copula theory.

In this study we are interssed to the estimation of copula by the parameteric and no parametric methods.

Table des matières

Dédicace **ii**

Remerciements **iii**

Abréviations et notations **xi**

Introduction générale **1**

1 Copules et mesure de dépendence **4**

1.1 Fonction de répartition bivariée et ses proprietés 4

1.2 La copule 5

1.2.1 Copules bivariées 6

1.2.2 Propriétés des copules bivariées 7

1.2.3 Théorème de Sklar 10

1.3 Exemples 11

1.3.1 1.**Copule gaussienne** 11

1.3.2 2 D**istribution de Gumbel (cas bivarié)** 12

1.4 Fonction de répartition multivariée 13

1.4.1 Copules multivariées 13

1.4.2 Fonction de densité conjointe $d-$ dimensionnelle 14

1.4.3 Théorème de Sklar 15

1.4.4 Propriétés des copules multivariées 16

1.5 Relation entre les mesures de dépendance et les copules 19

1.5.1 Mesures de concordance 19
1.5.2 Dépendance de queue 23

2 Copules paramétriques et non paramétriques 26
2.1 Copules paramétriques 26
2.1.1 Copules elliptiques 26
2.1.2 Copule gaussienne 27
2.1.3 Copule de student 30
2.2 Copules archimédiennes 33
2.3 Propriétés 34
2.4 Quelques types des copules Archimédiennes 34
2.4.1 Copule de Franck 36
2.4.2 **Copule de Gumbel** 37
2.5 Copules des valeurs extrêmes 39
2.6 Copules Archimax 43
2.7 Copules non paramétriques 46
2.7.1 Copule empirique 46

3 Méthodes d'estimation des copules 48
3.1 Estimation paramétrique 48
3.1.1 Méthode des moments 49
3.1.2 Méthode du maximum de vraisemblance (MLE) 50
3.1.3 Méthode d'inférence sur les marginales (IFM) 51
3.2 Estimation non paramétriques 52
3.2.1 Copules empirique 53
3.2.2 Procédure de GENEST et RIVEST 53
3.2.3 Estimation à noyau 54
3.3 Conclusion générale 57

Annexe 59
3.3.1 La densité normale multivariée 59

3.3.2 La densité de student multivariée 59

Table des figures

2.4.1 Densité de la copule de clayton 36
2.4.2 Densité de copule franck 37
2.4.3 Densité de copule gumbel 38

Liste des tableaux

1.1 Tableau coefficients de quelques copules usuelles 25

2.1 Tableau de quelques types de copules archimédiennes 39

2.2 Quelquest types de distribution de copules 40

Abréviations et notations

$d.d.l$	Degré de liberté
EV	Extreme Value
i.e.,	C'est-à-dire
$i.i.d$	indépendante et identiquement distribuée
IFM	fonction d'inférence des marginales
GEV	Generalized Extreme Value Distribution
MDA	Max-domaine d'attraction
MLE	méthode de maximum de vraisemblance
$v.a$	Variable aléatoire
$(\Omega, \mathcal{A}, \mathbb{P})$	espace de probabilité
Γ	La fonction gamma
Θ	L'ensemble de tous les paramètres
$M_d(\mathbb{R})$	l'ensemble des matrices carrées réelles de taille d^2
I	L'intervalle $[0, 1]$
I^2	$I \times I = [0, 1] \times [0, 1]$

Introduction générale

Mesurer la dépendance entre deux ou plusieurs variables aléatoires est une pratique largement répendue par les statisticiens. Quelques mesures de dépendance entre les variables aléatoires en été proposé comme celui du coefficient de corrélation, le tau de Kendall, le rho de Spearman. Bien que, ces mesures soient simples à calculer et peuvent être facilement interprétées, elles ne sont pas en mesure de détecter toutes les formes de dépendances. Pour résoudre ce problème ; le terme "copule a été introduit en 1959. Ce terme c'est à-dire "copule" a été utilisé pour la première fois grâce à Sklar ([30]) dans la théorie des lois multidimensionnelles.

En effet, la fonction "copule" a l'avantage de modéliser complètement la dépendance entre les variables aléatoires. Les copules sont des fonctions de répartition particulières, qui relies les fonctions de répartition multivariées de lois de probabilité dans $\mathbb{R}^d$, pour $d \geq 2$, aux fonctions de répartition marginales. Par exemple si l'indépendance entre d variables aléatoires $X_1, X_2, ...X_d$ est assurée, alors la fonction de répartition conjointe de ces n variables s'exprime de la manière suivante :

$$\mathbb{F}(x) = \mathbb{P}[X \leq x] = \mathbb{P}(X_1 \leq x_1, X_2 \leq x_2, ..., X_d \leq x_d), \quad \forall (x_1, x_{2,...,}x_d)' \in \mathbb{R}^d$$

où $f_i(\cdot)$ est la i ^{ème} densité marginale univariée . Par contre, il n'en demeure pas moins qu'établir l'indépendance entre des variables n'est pas nécessairement simple, car, bien qu'elle implique un coefficient de corrélation nul, la réciproque se montre généralement fausse. Ainsi, les copules autorisent une sélection plus étendue des fonctions de répartition jointes et ce indépendamment des différentes lois marginales considérées. De plus, elles permettent de résumer la structure de dépendance d'un vecteur aléatoire. D'ailleurs, les

mesures de dépendance peuvent être exprimées explicitement en fonction de la copule. En particulier, le tau de Kendall et le rho de Spearman entre deux vecteurs aléatoires X_1 et X_2 s'expriment, respectivement, par

$$\tau = 4 \int_{I^2} \mathbb{C}(u,v) d\mathbb{C}(u,v) - 1$$
$$\rho_S = 12 \int_{I^2} \mathbb{C}(u,v) du dv - 3$$

La théorie des copules permet de construire des modèles paramétriques pour la loi jointe de d variables aléatoires réelles dépendantes $(X_1, ...X_d)'$. En pratique, d'après Fisher ([14]) elle joue un rôle trés important et contribuet de manière significative dans les différents domaine tel que l'économétrie,biologie, comme d'autres pour modéliser de manière plus réaliste la loi jointe de plusieurs variables aléatoires.

Ainsi,pour modéliser la structure de dépendance entre plusieurs variables aléatoires, on utilise la fonction copule qui modélise les stuctures et les mesure de dépendance qui permet de construire des modèles de distributions multidimensionnelles.

Il existe plusieurs approches pour l'estimation d'une copule. Ces méthodes sont subdivisées en trois approches , l'approche paramétrique, semi- paramétrique et non-paramétrique.

Dans l'approche paramétrique, on impose un modèle spécifier que pour la copule et les dis tributions marginales à la fois.

Finalement, l'approche non-paramétrique tient compte à la fois des modèles non-paramétriques pour la copule et pour les distributionsmarginales. Deheuvels (voir [8])a suggéré la distribution

mul-tidimensionnelle empirique pour estimer la fonction copule.

L'objectif principal de notre travail est de revoir les différentes méthodes d'estimation des copules à savoir l'approche paramètrique et non-paramètrique. La première approche c'est à dire l'approche paramètrique consiste à estimer le paramètre de la copule par l'une des méthodes (Maximum de vraisemblance exacte (EML), Fonctions d'inférence des marginales (IFM) et la méthode des moments).

Tandisque dans la deuxième approche on utilise les copules empiriques.- Procédure de GENEST et RIVEST.

Ce travail est composé de trois chapitres :

Chapitre 1 : Copules et mesure de dépendence

Dans ce chapitre, on présente la fonction de répartition bivariée ses proprietés et la théorie des copules et leurs propriétés fondamentales ainsi quelque exmeples des copules sont présentées. Enfin, on discute la relation existant entre la structure de dépendance et la théorie des copules.

Chapitre 2 : Les copules paramétriques et non paramétriques

Dans ce chapitre, on présente quelques familles de copules paramétriques c'est les familles les plus utilisées : Les copules Archimédiennes, les copules Elliptiques, des valeurs extrêmes et la copule Archimax. Les copules non paramétriques sont aussi présentées.

Chapitre 3 : Méthodes d'estimation des copules

Dans ce chapitre, on présente les méthodes d'estimation des copules paramétriques et non parametriques. Pour l'approche paramétrique on discute la méthode des moments, la méthode (MLE) et enfin la méthode (IFM) et pour l'estimation non paramétrique, nous présentons la procédure de GENEST et RIVEST et nous rappelons l'estimation à noyau.

Annexe A :

Dans cette annexe, on va rappeler quelques notions de base de probabilité concernant la densité normale multivariée et la densité de student multivariée ainsi la fonction gamma et leurs propriétés et quelque définition de tho de kandel et rho de Spearman.

Chapitre 1

Copules et mesure de dépendence

Introduction

La copule est un outil qui permet de modéliser la structure de dépendance de plusieurs variables aléatoires. Soit X un vecteur de dimension $d \geqslant 2$. La modélisation de copule est basée sur le théorème de sklar a été introduit en (1959)([30]),Sklar mathématicien américan dans son articleil parle de fonction de répartition à d dimension et leurs marginales, pour mieux comprendre on va commencer par le cas bivarié après on généralise .

Dans ce chapitre, on présente les propriétés les plus importantes de la fonction copule et son rôle dans l'étude de l'indépendance des variables aléatoires. La première section, est une introduction de la fonction de répartition bivariée. Dans la deuxième section, on donne la définition d'une copule bivariée, ainssi que ces différentes propriétés. Dand La troisième section, on va présenter quelques exemples de copule. Dans la quatrième section on généralise fonction de répartition. On termine le chapitre par la présentation du lien entre les mesures de dépendance et les copules.

1.1 Fonction de répartition bivariée et ses proprietés

Soit (Ω, A, P) un espace de probabilité et X un vecteur aléatoire de dimension 2 i,.e $X = (X_1, X_2)'$. La fonction de répartition bivariée du vecteur X est une fonction $F : \mathbb{R}^2 \longrightarrow I$ telle que :

$$\mathbb{F}(x) = \mathbb{P}[X \leq x] = \mathbb{P}(X_1 \leq x_1, X_2 \leq x_2), \quad \forall (x_1, x_2)' \in \mathbb{R}^2, \tag{1.1.1}$$

qui vérifie les conditions suivantes *:*

a) $\mathbb{F}$ est continue à droite, $\lim\limits_{x\longrightarrow x_0, y\longrightarrow y_0} \mathbb{F}(x,y) = \mathbb{F}(x_0,y_0)$

b) $\lim\limits_{x\longrightarrow -\infty} \mathbb{F}(x,y) = \lim\limits_{y\longrightarrow -\infty} \mathbb{F}(x,y) = 0$

c) $\lim\limits_{x\longrightarrow +\infty, y\longrightarrow +\infty} \mathbb{F}(x,y) = 1$

d) $\mathbb{F}$ est 2-croissante, pour tout $(a_1,a_2),(b_1,b_2)$ dans $\mathbb{R}^2$ avec $-\infty \leq a_1 \leq b_1 \leq +\infty$ et $-\infty \leq a_2 \leq b_2 \leq +\infty$, on a

$$\mathbb{F}(b_1,b_2) - \mathbb{F}(a_1,b_2) - \mathbb{F}(b_1,a_2) + \mathbb{F}(a_1,a_2) \geq 0$$

Les lois marginales $\mathbb{F}_1$et $\mathbb{F}_2$ peuvent s'exprimer en fonction de $\mathbb{F}$ de la manière suivante :

$$\begin{aligned} \mathbb{F}_1(x) &= P(X \leq x) = \lim_{x\longrightarrow +\infty} \mathbb{F}(x,y) = \mathbb{F}(x,+\infty) \text{ et} \\ \mathbb{F}_2(y) &= P(Y \leq y) = \lim_{y\longrightarrow +\infty} \mathbb{F}(x,y) = \mathbb{F}(+\infty,y). \end{aligned} \tag{1.1.2}$$

- Si $\mathbb{F}$ est croissante par rapport à x et par rapport à y n'implique pas que $\mathbb{F}$ est 2-croissante.
- Si $\mathbb{F}$ est 2-croissante n'implique pas que $\mathbb{F}$ est croissante par rapport à x et par rapport à y.
- Si $\mathbb{F}$ est 2-croissante et $\lim\limits_{x\longrightarrow -\infty} \mathbb{F}(x,y) = \lim\limits_{y\longrightarrow -\infty} \mathbb{F}(x,y) = 0$ implique que $\mathbb{F}$ est croissante par rapport à x et par rapport à y.
- Si $\mathbb{F}$ est de classe C^2, être 2-croissante équivalente à $\frac{\partial^2 \mathbb{F}(x,y)}{\partial x \partial y} \geq 0$, donc pour tout $(x,y) \in \mathbb{R}^2$,

$$\max(\mathbb{F}_1(x) + \mathbb{F}_2(y) - 1, 0) \leq \mathbb{F}(x,y) \leq \min(\mathbb{F}_1(x), \mathbb{F}_2(y)).$$

Proposition 1.1.1 *Deux variables aléatoires X et Y sont indépendantes si et seulement si*

$$\mathbb{F}_{(X,Y)}(x,y) = \mathbb{F}_1(x)\mathbb{F}_2(y).$$

1.2 La copule

Dans cette section, on va introduire la définition d'une copule, ainsi que ses propriétés, où on commence par le cas le plus simple i.e., le cas bivarié, ensuite on généralise.

1.2.1 Copules bivariées

Définition 1.2.1 *On appelle copule bivariée toute fonction* $\mathbb{C}$, $I^2 \longmapsto I$ *qui possède les propriétés suivantes :*

i) $\forall (u,v) \in I^2 : \mathbb{C}(u,0) = \mathbb{C}(0,v) = 0.$

ii) $\forall (u,v) \in I^2 : \mathbb{C}(u,1) = u$ *et* $\mathbb{C}(1,v) = v.$

iii) $\mathbb{C}$ *est 2-croissante i.e.,*$\forall (u_1,u_2),(v_1,v_2) \in I^2$ *telle que* $u_1 \leq v_1$ *et* $u_2 \leq v_2$ *on a*

$$\mathbb{C}(v_1,v_2) - \mathbb{C}(v_1,u_2) - \mathbb{C}(u_1,v_2) + \mathbb{C}(u_1,u_2) \geq 0. \tag{1.2.1}$$

1. La Définition 1.2.1 assure que la copule est une distribution de probabilité de marges uniformes.

Soient U_1 et U_2 deux $v.a$ uniformes. Considérons le vecteur aléatoire $U = (U_1,U_2)'$, on a

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

De la propriété $i)$ de la définition 1.2.1 on a

$$\mathbb{P}[U_1 \leq 0, U_2 \leq u_2] = \mathbb{P}[U_1 \leq u_1, U_2 \leq 0] = 0.$$

De la propriété $ii)$ de la définition 1.2.1 on a

$$\mathbb{P}[U_1 \leq u_1, U_2 \leq 1] = u_1 \text{ et } \mathbb{P}[U_1 \leq 1, U_2 \leq u_2] = u_2$$

Comme la copule $\mathbb{C}$ est une distribution de probabilité, alors

$$\mathbb{P}[u_1 \leq U_1 \leq v_1, u_2 \leq U_2 \leq v_2] = \mathbb{C}(v_1,v_2) - \mathbb{C}(v_1,u_2) - \mathbb{C}(u_1,v_2) + \mathbb{C}(u_1,u_2)$$
$$\forall (u_1,u_2),(v_1,v_2) \in I \times I$$

Exemple 1.2.1 $\forall u_1,u_2 \in I$, *la fonction* $\mathbb{M}(u_1,u_2) = \min(u_1,u_2)$ *est une copule.*

– $\forall u_1,u_2 \in I$, on a $\min(u_1,0) = 0 = \min(0,u_2)$, alors $\mathbb{M}$ vérifie la propriété $i)$.

$\forall u_1,u_2 \in I$, on a $\min(u_1,1) = u_1$ et $\min(1,u_2) = u_2$, alors $\mathbb{M}$ vérifie la propriété $ii)$.

$\forall u_1,u_2,v_1$ et $v_2 \in I$ avec $u_1 \leq v_1$ et $u_2 \leq v_2$, on a

$\min(u_2,v_2) \geq \min(u_2,v_1)$ et $\min(u_1,v_2) \geq \min(u_1,v_1)$, alors M vérifie la propriété $iii)$

Par conséquent $\mathbb{M}$ est une copule.

De la même manière on établit que les fonctions définies par $\mathbb{W}(u_1,u_2) = \max(u_1+u_2-1,0)$ et $\not> (u_1,u_2) = u_1u_2$ sont aussi des copules.

1.2.2 Propriétés des copules bivariées

Dans ce qui suit, on donne les propriétés et le théorème le plus important d'une copule bivariée.

Théorème 1.2.1 ***(Bornes de Fréchet (voir** [25]**)*** *Pour toute copule bivariée* $\mathbb{C}$ *on a*

$$\forall (u_1, u_2) \in I^2,\ \mathbb{W}(u_1, u_2) \leq \mathbb{C}(u_1, u_2) \leq \mathbb{M}(u_1, u_2).$$

où $\mathbb{M}(u, v)$ *est la copule minimum notée par* $\mathbb{M}(u, v) = \min(u, v)$ *et* $\mathbb{W}(u, v)$ *est la copule maximum notée par* $\mathbb{W}(u_1, u_2) = \max(0, u_1 + u_2 - 1)$.

Preuve. On a $\forall u_1, u_2 \in I; \mathbb{C}(u_1, u_2) \leq \mathbb{C}(u_1, 1) = u_1$ et $\mathbb{C}(u_1, u_2) \leq \mathbb{C}(1, u_2) = u_2$, alors, $\mathbb{C}(u_1, u_2) \leq \min(u_1, u_2))$,

Or, $\forall u_1, u_2 \in [0, 1]$; $\mathbb{C}(1, 1) - \mathbb{C}(1, u_2) - \mathbb{C}(u_1, 1) + \mathbb{C}(u_1, u_2) \geq 0$, d'où,

$$\mathbb{C}(u_1, u_2)) \geq u_1 + u_2 - 1.$$

Mais $\mathbb{C}(u_1, u_2) \geq 0, \forall u_1, u_2 \in I$, alors

$$\mathbb{C}(u_1, u_2) \geq \max(0, u_1 + u_2) - 1).$$

D'où

$$\max(0, u_1 + u_2) - 1) \leq \mathbb{C}(u_1, u_2) \leq \min(u_1, u_2).$$

■

Remarque 1.2.1 *Les copules* $\mathbb{W}$ *et* $\mathbb{M}$ *sont appelées borne inférieure (resp borne supérieure) de Fréchet-Hoeffding ou copule minimale (resp copule maximale). Elles sont notées resp par* $\mathbb{C}^+$ *et* $\mathbb{C}^-$.

Dans la suite nous présontons les propriétés fondamentales et le théorème le plus important des copules bivariées, introduit par Sklar (1959).

Théorème 1.2.2 *(**Continuité**([25]))* *Soit $\mathbb{C}$ une copule bivariée ; pour tout u_1, u_2, v_1 et $v_2 \in I$ avec $u_1 \leq u_2$ et $v_1 \leq v_2$, on a*

$$|\mathbb{C}(u_2, v_2) - \mathbb{C}(u_1, v_1)| \leq |u_2 - u_1| + |v_2 - v_1|. \tag{1.2.2}$$

Comme toute copule bivariée est 2-croissante donc $\forall u_1, u_2, v_1$ et $v_2 \in I$ où $u_1 \leq u_2$ et $v_1 \leq v_2$ soit γ_1 le graphe passant par les points (u_1, v_1) et (u_2, v_1) et γ_2 le graphe passant par les points (u_2, v_1) et (u_2, v_2). Il existe deux copules $\mathbb{C}_{\gamma_1}, \mathbb{C}_{\gamma_2}$ telle que $\mathbb{C}(u_1, v_1) = \mathbb{C}_{\gamma_1}(u_1, v_1)$ et $\mathbb{C}(u_2, v_2) = \mathbb{C}_{\gamma_2}(u_2, v_2)$, donc $\mathbb{C}(u_2, v_1) = \mathbb{C}_{\gamma_1}(u_2, v_1) = \mathbb{C}_{\gamma_2}(u_2, v_1)$

$$\begin{aligned}
|\mathbb{C}(u_2, v_2) - \mathbb{C}(u_1, v_1)| &= |\mathbb{C}(u_2, v_2) + \mathbb{C}(u_2, v_1) - \mathbb{C}(u_2, v_1) - \mathbb{C}(u_1, v_1)| \\
&\leq |\mathbb{C}(u_2, v_2) + \mathbb{C}(u_2, v_1)| + |\mathbb{C}(u_2, v_1) - \mathbb{C}(u_1, v_1)| \\
&= \left|\mathbb{C}_{\gamma_2}(u_2, v_2) + \mathbb{C}_{\gamma_2}(u_2, v_1)\right| + \left|\mathbb{C}_{\gamma_1}(u_2, v_1) - \mathbb{C}_{\gamma_1}(u_1, v_1)\right| \\
&\leq |u_2 - u_1| + |v_2 - v_1|.
\end{aligned}$$

Théorème 1.2.3 *(Différentiabilité([4]) Soit $\mathbb{C}$ une copule bivariée ;*

1 Les dérivées partielles $\frac{\partial \mathbb{C}(u_1,u_2)}{\partial u_j}$ existent $\forall u_1, u_2 \in I$ et $0 \leq \frac{\partial \mathbb{C}(u_1,u_2)}{\partial u_j} \leq 1$; $\forall j = 1, 2$

2 Les fonctions $u_1 \longmapsto \frac{\partial \mathbb{C}(u_1,u_2)}{\partial u_2}$ et $u_2 \longmapsto \frac{\partial \mathbb{C}(u_1,u_2)}{\partial u_1}$ sont bien définies et sont décroissantes sur I.

La dérivée partielle de la copule $\mathbb{W}(u_1, u_2) = \max(u_1 + u_2 - 1, 0)$ existe.

Si $u_1 + u_2 - 1 \neq 0$ i.e., $u_2 \neq 1 - u_1$ et on a $\max(u_1 + u_2 - 1, 0) = u_1 + u_2 - 1 \geq 0$, ce qui donne

$$\frac{\partial \mathbb{C}(u_1,u_2)}{\partial u_1} = \frac{\partial \mathbb{C}(u_1,u_2)}{\partial u_2} = \begin{cases} 0 \text{ si } u_2 < 1 - u_1 \\ 1 \text{ si } u_2 > 1 - u_1 \end{cases} \tag{1.2.3}$$

Théorème 1.2.4 (Invariance (voir [25])) *Soit $X = (X_1, X_2)'$ un vecteur de v.a continues de copule $\mathbb{C}$. Si α_1, α_2 sont deux fonctions strictement croissantes sur $\operatorname{Im} X_1$ et $\operatorname{Im} X_2$ respectivement, alors :*

$$\mathbb{C}(X_1, X_2) = \mathbb{C}(\alpha_1(X_1), \alpha_2(X_2))$$

Preuve. *On note repectivement $\mathbb{F}$ et $\mathbb{G}$ les fonctions de répartition conjointes de vecteurs aléatoires (X_1, X_2) et $(\alpha_1(X_1), \alpha_2(X_2))$, par $\mathbb{F}_1$, $\mathbb{F}_2$ les fonctions de répartition marginales*

de $\mathbb{F}$ *et par* $\mathbb{G}_1$, $\mathbb{G}_2$ *les fonctions de répartition marginales de* $\mathbb{G}$. *On remarque que :*

$$\begin{aligned}\mathbb{G}_1(X_1) &= \mathbb{P}[\alpha_1(X_1) \leq x_1] = \mathbb{P}\left[X_1 \leq \alpha_1^{-1}(x_1)\right] = \mathbb{F}_1\left(\alpha_1^{-1}(x_1)\right), \\ &\quad (\textit{car } \alpha_1 \textit{ est une fonction croissante})\end{aligned}$$

et on a donc $\mathbb{G}_1^{-1}(u_1) = \alpha_1\left(\mathbb{F}_1^{-1}(u_1)\right)$ *et* $\mathbb{G}_2^{-1}(u_2) = \alpha_2\left(\mathbb{F}_2^{-1}(u_2)\right)$, $\forall u_1, u_2 \in I$. *On en déduit que*

$$\begin{aligned}\mathbb{C}_{\alpha_1(X_1)\alpha_2(X_2)}(u_1,u_2) &= \mathbb{C}\left(\mathbb{G}_1^{-1}(u_1), \mathbb{G}_2^{-1}(u_2)\right) \\ &= \mathbb{P}\left[\alpha_1(X_1) \leq \mathbb{G}_1^{-1}(u_1), \alpha_2(X_2) \leq \mathbb{G}_2^{-1}(u_2)\right] \\ &= \mathbb{P}\left[X_1 \leq \alpha_1^{-1}\left(\mathbb{G}_1^{-1}(u_1)\right), X_2 \leq \alpha_2^{-1}\left(\mathbb{G}_2^{-1}(u_2)\right)\right] \\ &= \mathbb{P}\left[X_1 \leq \mathbb{F}_1^{-1}(u_1), X_2 \leq \mathbb{F}_2^{-1}(u_2)\right] \\ &= \mathbb{C}_{X_1X_2}(u_1,u_2)\end{aligned}$$

■

Définition 1.2.2 (L'inverse généralisé) *([23]) Soit* $\mathbb{F}$ *une fonction de distribution. L'inverse généralisé* $\mathbb{F}^{[-1]}$ *de* $\mathbb{F}$ *est une fonction définie par :*

$$\mathbb{F}^{[-1]}(t) = \inf\{x \in \mathbb{R}, \mathbb{F}(x) \geq t\} = \sup\{x \in \mathbb{R}, \mathbb{F}(x) \leq t\}, \tag{1.2.4}$$

Remarque 1.2.2 *Si la fonction* $\mathbb{F}$ *est strictement croissante alors la notion de (pseudo-inverse) coïncide avec la notion d'inverse ou de réciproque, i.e.,* $\mathbb{F}^{[-1]} = \mathbb{F}^{-1}$

Proposition 1.2.1 *Soit* $\mathbb{F}$ *une fonction réelle croissante et continue à droite, d'inverse généralisé* $\mathbb{F}^{-1}$, *alors on a les propriétés suivantes :*

a- La fonction $\mathbb{F}^{-1}$ est croissante, continue à gauche et on a l'équivalence suivante pour tout $x \in \mathbb{R}$ et tout $t \in [0,1]$

$$\mathbb{F}(x) \geq t \Leftrightarrow x \geq \mathbb{F}^{-1}(t),$$

b- Pour tout $t \in]0,1]$, on a $\mathbb{F}(\mathbb{F}^{-1}(t)) \geq t$ et avec égalité si $\mathbb{F}^{-1}(t) > -\infty$ et si $\mathbb{F}$ est continue en $\mathbb{F}^{-1}(t)$.

c- Soit U une variable aléatoire uniformement distribuée sur I, la fonction $\mathbb{F}$ est la fonction de répartition de la variable aléatoire $\mathbb{F}^{-1}(U)$.

d- Soit Y une variable aléatoire de fonction de répartition $\mathbb{F}$. Si $\mathbb{F}$ est continue, alors $\mathbb{F}(Y)$ est de loi uniforme sur $]0,1]$.

1.2.3 Théorème de Sklar

La théorie des copules est basée essentiellement sur le théorème introduit par sklar 1959 est donné par le théorème suivant :

Théorème 1.2.5 *(**Sklar** ([30]) Soit $\mathbb{F}$ une distribution conjointe bivariée de distributions marginales $\mathbb{F}_1$ et $\mathbb{F}_2$ respectivement. On pose $\mathbb{F}_i(x_i) = u_i$, $i = 1, 2$:*

$\mathbb{F}$ peut sécrire en terme de la fonction copule notée $\mathbb{C}(u, v)$, tel que :

$$\forall (x_1, x_2) \in \mathbb{R}^2, \qquad \mathbb{F}(x_1, x_2) = \mathbb{C}(\mathbb{F}_1(x_1), \mathbb{F}_2(x_2)) \tag{1.2.5}$$

Preuve. $\forall (x_1, x_2) \in \mathbb{R}$, les focntions de répartition marginales $\mathbb{F}_1(x_1)$ et $\mathbb{F}_2(x_2)$ sont uniformement continues sur I, étant

$$\begin{aligned} \mathbb{C}(u_1, u_2) &= \mathbb{P}[\mathbb{F}_1(x_1) \leq u_1, \mathbb{F}_2(x_2) \leq u_2] \\ &= \mathbb{P}\left[x_1 \leq \mathbb{F}_1^{(-1)}(u_1), x_2 \leq \mathbb{F}_2^{(-1)}(u_2)\right] \\ &= \mathbb{F}\left(\mathbb{F}_1^{(-1)}(u_1), \mathbb{F}_2^{(-1)}(u_2)\right). \end{aligned}$$

Et comme

$$\begin{aligned} \mathbb{F}(x_1, x_2) &= \mathbb{P}[X_1 \leq x_1, X_2 \leq x_2] \\ &= \mathbb{P}[\mathbb{F}_1(X_1) \leq \mathbb{F}_1(x_1), \mathbb{F}_2(X_2) \leq \mathbb{F}_2(x_2)] \\ &= \mathbb{C}(\mathbb{F}_1(x_1), \mathbb{F}_2(x_2)) \end{aligned}$$

Si les distributions marginales $\mathbb{F}_1$ et $\mathbb{F}_2$ sont continues, alors la copule $\mathbb{C}$ est unique ; sinon elle est déterminée de manière unique sur $(Ran\mathbb{F}_1) \times (Ran\mathbb{F}_2)$. ■

Réciproquement, si $\mathbb{C}$ est une copule et $\mathbb{F}_1$, $\mathbb{F}_2$ sont des distributions univariées, alors la fonction définie par (1.2.5) est la distribution conjointe dont les marges sont $\mathbb{F}_1$ et $\mathbb{F}_2$.

Définition 1.2.3 *Si c est la densité copule associée à la copule* $\mathbb{C}$ *existe alors elle est définie par*

$$c(u_1, u_2) = \frac{\partial^2 \mathbb{C}(u_1, u_2)}{\partial u_1 \partial u_2}$$

Définition 1.2.4 *Si la fonction de répartition conjointe* $\mathbb{F}$ *est absolument continue, d'après le théorème de Sklar, on peut exprimer la densité d'un vecteur aléatoitre* $(X_1, X_2)'$ *en fonction de la densité conjointe* f *et ces fonctions de répartitions marginales* $\mathbb{F}_1$ *et* $\mathbb{F}_2$ *par*

$$f(x_1, x_2) = c(\mathbb{F}_1(x_1), \mathbb{F}_2(x_2)) \prod_{i=1}^{2} f_i(x_i) \tag{1.2.6}$$

Preuve. Par définition $f(x_1, x_2) = \frac{\partial^2 \mathbb{F}(x_1, x_2)}{\partial x_1 \partial x_2}$ et $c(u_1, u_2) = \frac{\partial^2 \mathbb{C}(u_1, u_2)}{\partial u_1 \partial u_2}$, et en utilisant le théorème de Sklar on obtient

$$\begin{aligned} f(x_1, x_2) &= \frac{\partial^2 \mathbb{C}(u_1, u_2)}{\partial u_1 \partial u_2} = \frac{\partial^2 \mathbb{C}(\mathbb{F}_1(x_1), \mathbb{F}_2(x_2))}{\partial \mathbb{F}_1(x_1) \partial \mathbb{F}_2(x_2)} \\ &= c(\mathbb{F}_1(x_1), \mathbb{F}_2(x_2)) f_1(x_1) f_2(x_2) \\ &= c(\mathbb{F}_1(x_1), \mathbb{F}_2(x_2)) \prod_{i=1}^{2} f_i(x_i) \end{aligned}$$

■

1.3 Exemples

Dans cette section est réservé de quelque types copules connue, (gaussienne, gumbel).

1.3.1 1.Copule gaussienne

Définition 1.3.1 *Soit* $\mathbb{F}_\rho$ *la fonction de distribution normal bivariée de coefficient de correlation* ρ *telle que,* $\forall x, y \in \mathbb{R}$

$$\mathbb{F}_\rho(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \int_{-\infty}^{x} \int_{-\infty}^{y} \exp\left\{\left(\frac{-1}{2(1-\rho^2)}\left[\left(\frac{x-m_1}{\sigma_1}\right)^2 - \frac{2\rho(x_1-m_1)(x_2-m_2)}{\sigma_1\sigma_2} + \left(\frac{x_2-m_2}{\sigma_2}\right)^2\right]\right\} dx_1 dx_2$$

Soit $X \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, m \begin{pmatrix} m_1 \\ m_2 \end{pmatrix}, \sigma \begin{pmatrix} \sigma_1 \\ \sigma_2 \end{pmatrix}$

On a

$$f_{m,\Sigma} = \frac{1}{2\pi \det(\Sigma)^{\frac{1}{2}}} \exp(-\frac{1}{2}(X-m)'\Sigma^{-1}(X-m)) \tag{1.3.1}$$

telle que $\Sigma = \sigma_1\sigma_2 \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$ aprés les calculs 1.3.1 on obtient

$$f_{m,\Sigma} = -\frac{1}{2(1-\rho^2)} \left[\left(\frac{x_1-m_1}{\sigma_1} \right)^2 - 2\rho \left(\frac{x_1-m_1}{\sigma_1} \right) \left(\frac{\rho(x_2-m_2)}{\sigma_2} \right) + \left(\frac{x_2-m_2}{\sigma_2} \right)^2 \right]$$

D'où

$$\mathbb{F}_\rho(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \int_{-\infty}^{x}\int_{-\infty}^{y} \exp\left(-\frac{1}{2(1-\rho^2)} \left[\left(\frac{x_1-m_1}{\sigma_1} \right)^2 - 2\rho \left(\frac{x_1-m_1}{\sigma_1} \right) \left(\frac{\rho(x_2-m_2)}{\sigma_2} \right) + \left(\frac{x_2-m_2}{\sigma_2} \right)^2 \right] \right) \tag{1.3.2}$$

$\mathbb{F}_\rho$ est la fonction de répartition conjointe des v.a X et Y où $X, Y \sim \mathcal{N}(m_{_i}, \sigma_i)$ pour $i = 1, 2$, la copule associée à $\mathbb{F}_\rho$ est donnée par

$$\mathbb{F}_\rho(x,y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \int_{-\infty}^{x}\int_{-\infty}^{y} \exp\left\{ \left(\frac{-1}{2(1-\rho^2)} \left[\left(\frac{x-m_1}{\sigma_1} \right)^2 - \frac{2\rho(x_1-m_1)(x_2-m_2)}{\sigma_1\sigma_2} + \left(\frac{x_2-m_2}{\sigma_2} \right)^2 \right] \right\} dx_1 dx_2.$$

où $\phi_i^{-1}; i = 1, 2$ est la fonction inverse de la loi $\mathcal{N}(m_i, \sigma_i)$.

En particulier, si $X, Y \sim \mathcal{N}(0,1)$ où $\mathcal{N}(0,1)$ est la loi normale centrée réduit en substituant dans l'equation (1.3.2) avec $m_1 = m_2 = 0$ et $\sigma_1 = \sigma_2 = 1$ on obtient :

$$\mathbb{C}_\rho(u,v) = \frac{1}{(2\pi)(1-\rho^2)^{1/2}} \int_{-\infty}^{\Phi_1^{-1}(u)} \int_{-\infty}^{\Phi_1^{-1}(v)} \exp\left(-\frac{x^2-2\rho xy+y^2}{2(1-\rho^2)} \right) dxdy.$$

$\mathbb{C}_\rho$ est une copule de Gauss.

1.3.2 2 Distribution de Gumbel (cas bivarié)

Définition 1.3.2 *Soit $\mathbb{F}_\theta$ la fonction de disrtibution donnée par*

$$\mathbb{F}_\theta(x,y) = \begin{cases} 1 - e^{-x} - e^{-y} + e^{-(x+y+\theta xy)} & x \geq 0,\ y \geq 0 \\ 0 & \text{sinon} \end{cases}$$

où θ un paramètre dans $[0,1]$. Les fonctions de de distribution marginales $\mathbb{F}_1(x)$ et $\mathbb{F}_2(y)$ sont des exponentielles et leurs foncions inverses sont $\mathbb{F}_1(u) = -\ln(1-u)$ et $\mathrm{F}_2(v) =$

$-\ln(1-v)$ respectivement, pour u, v dans $[0,1]$, on en deduit que la copule correspondante es

$$\mathbb{C}_\theta(u,v) = u + v - 1 + (1-u)(1-v)\, e^{-\theta \ln(1-u)\ln(1-v)}$$

1.4 Fonction de répartition multivariée

Définition 1.4.1 *Soit (Ω, A, P) un espace de probabilité et X un vecteur aléatoire d-dimentionnel ie $X = (X_1, ..., X_d)'$. On définit la fonction de répartition d-dimentionnelle du vecteur X comme une fonction $\mathbb{F}_X : \mathbb{R}^d \longrightarrow I$ telle que :*

$$\mathbb{F}_X(x) = P[X \leq x] = P[(X_1, ..., X_d) \leq (x_1, \ldots, x_d)] = P(X_1 \leq x_1, ..., X_d \leq x_d),$$

qui vérifie les conditions suivantes :

(i) $\mathbb{F}$ est continue à droite relativement à chaque composante ;

(ii) pour tout $(a_1, ..., a_d), (b_1, ..., b_d)$ dans $\mathbb{R}^d$ tels que $a_i \leq b_i$, pour tout $i = 1, ..., d$

$$\Delta_{a_n}^{b_n} \Delta_{a_{n-1}}^{b_{n-1}} ... \Delta_{a_1}^{b_1} \mathbb{F}(x) \geq 0,$$

où $\Delta_{a_k}^{b_k} \mathbb{F}(x) = \mathbb{F}(x_1, ..., x_{k-1}, b_k, x_{k+1}, ..., x_d) - \mathbb{F}(x_1, ..., x_{k-1}, a_k, x_{k+1}, ..., x_d)$;

(iii) $\mathbb{F}(x_1, ..., x_{i-1}, -\infty, x_{i+1}, ..., x_d) = 0$ pour tout $i = 1, ..., d$, et $\mathbb{F}(\infty, ..., \infty) = 1$.

1.4.1 Copules multivariées

Définition 1.4.2 *Une copule est une fonction de répartition jointe de dimension d définie sur l'hypercube $[0,1]^d$ de sort que les marges univariées soient uniformes sur l'intervalle unitaire $[0,1]$.*

Une fonction $\mathbb{C} : I^d \longmapsto I$ est dite une copule à d dimensions si :

1 Pour tout $u = (u_1, \ldots, u_d)' \in I^d$ ayant au moins une composante nulle, on a $\mathbb{C}(u) = 0$;

2 Pour tout $u = (u_1, \ldots, u_d)' \in I^d$ ayant toutes ses composantes égales à 1 sauf la $k^{ème}$ composante ie $u = (1, \ldots, 1, u_k, 1, \ldots, 1)'$, on a $\mathbb{C}(u) = u_k$;

$\mathbb{C}$ est d-croissante i.e., pour tout $u = (u_1, \ldots, u_d)'$ et $v = (v_1, \ldots, v_d)'$ dans I^d tels que $u_i \leq v_i$, pour $i = 1, \ldots, d$, on a

$$\sum_{i_1=1}^{2} \ldots \sum_{i_d=1}^{2} (-1)^{i_1+\ldots+i_d} \times \mathbb{C}(x_{1i_1}, \ldots, x_{di_d}) \geq 0,$$

où $x_{1j} = u_j$ et $x_{2j} = v_j$ pour tout $j = \{1, \ldots, d\}$.

Remarque 1.4.1 *Une copule de dimension d est uniformément continue sur son domaine de définition.*

Et le théorème de continuité est donné par le théorème suivant :

Théorème 1.4.1 *Soit $\mathbb{C}$ une copule multivariée ; pour tout $u = (u_1, ..., u_d)' \in I^2$ et $v = (v_1, ..., v_d)' \in I$, on a*

$$|\mathbb{C}(u) - \mathbb{C}(v)| \leq \sum_{k=1}^{d} |v_k - u_k|.$$

où $\mathbb{C}$ est uniformement continue sur son domaine.

1.4.2 Fonction de densité conjointe $d-$ dimensionnelle

Toute distribution multivariée, absolument continue, admet une densité conjointe et elle est de la forme

$$f(x_1, ..., x_d) = c(\mathbb{F}_1(x_1), \mathbb{F}_2(x_2), ..., \mathbb{F}_d(x_d)) \prod_{i=1}^{d} f(x_i), \tag{1.4.1}$$

où c la fonction densité copule de la copule $\mathbb{C}$ et est donnée par :

$$c(u_1, u_2, ...u_d) = \frac{\partial^d C(u_1, u_2, ...u_d)}{\partial u_1 ... \partial u_d}, \quad \forall (u_1, \ldots, u_d)' \in I^d$$

De La formule (1.4.1) la densité copule multivariée de dimension d nottée $\mathfrak{c}$ est donnée par

$$\mathfrak{c}(u_1, \ldots, u_d) = \frac{f\left(\mathbb{F}_1^{-1}(u_1), \ldots, \mathbb{F}_d^{-1}(u_d)\right)}{\prod_{i=1}^{d} f_i\left(\mathbb{F}_i^{-1}(u_i)\right)}, \text{avec} \quad f_i\left(\mathbb{F}_i^{-1}(u_i)\right) > 0, \forall i = 1, ..., d$$

où $\mathfrak{c}(u) = \frac{\partial^d \mathbb{C}(u)}{\partial u_1 ... \partial u_d} \geq 0$.

1.4.3 Théorème de Sklar

Le théorème suivant (Sklar (1959)) est fondamental dans la théorie des copules ; ce théorème précise le lien existant entre la copule $\mathbb{C}$ et les fonctions de répartition marginales univariées $\mathbb{F}_1, ..., \mathbb{F}_d$ et la distribution cojointe multivariée $\mathbb{F}$. Pour plus de détail voir Nelson(voir [25]), Sklar (voir [30]) ou ENPC (2008).

Théorème 1.4.2 *Soit $\mathbb{F}$ une distribution multivariée i.e., $\mathbb{F}$ la répartition conjointe de distributions marginales $\mathbb{F}_1, ..., \mathbb{F}_d$ respectivement, alors il existe une copule $\mathbb{C}$ de dimension d telle que pour tout $x = (x_1, \ldots, x_d)' \in \mathbb{R}^d$,*

$$\mathbb{F}(x) = \mathbb{F}(x_1, \ldots, x_d) = \mathbb{C}(\mathbb{F}_1(x_1), \ldots, \mathbb{F}_d(x_d)) \tag{1.4.2}$$

Preuve. Etant $\mathbb{F}_1, ..., \mathbb{F}_d$ sont continues de plus $\mathbb{F}_1(x_1), ..., \mathbb{F}_d(x_d)$ sont uniformes sur $[0, 1]$, $\forall x = (x_1, \ldots, x_d)' \in \mathbb{R}^d$. Soit $\mathbb{F}$ la fonction de distribution multnivariée

On a

$$\begin{aligned}
\mathbb{C}(u_1, \ldots, u_d) &= \mathbb{P}(\mathbb{F}_1(x_1) \leq u_1, ..., \mathbb{F}_d(x_d) \leq u_d) \\
&= \mathbb{P}\left(x_1 \leq \mathbb{F}_1^{-1}(u_1), ..., x_d \leq \mathbb{F}_d^{-1}(u_d)\right) \\
&= \mathbb{F}\left(\mathbb{F}_1^{-1}(u_1), ..., \mathbb{F}_d^{-1}(u_d)\right)
\end{aligned}$$

D'autre part

$$\begin{aligned}
\mathbb{F}(x_1, \ldots, x_d) &= \mathbb{P}[X_1 \leq x_1, ..., X_d \leq x_d] \\
&= \mathbb{P}[\mathbb{F}_1(X_1) \leq \mathbb{F}_1(x_1), ..., \mathbb{F}_d(X_d) \leq \mathbb{F}_d(x_d)] \\
&= \mathbb{C}(\mathbb{F}_1(x_1), ..., \mathbb{F}_d(x_d))
\end{aligned}$$

■

Remarque 1.4.2 *Si les distributions marginales $\mathbb{F}_1, ..., \mathbb{F}_d$ sont continues, alors $\mathbb{C}$ est unique. Dans le cas contraire $\mathbb{C}$ est uniquement déterminée sur le produit $(\operatorname{Im}\mathbb{F}_1) \times \ldots \times (\operatorname{Im}\mathbb{F}_d)$ et de ce fait perd beaucoup de son intérêt.*

Inversement, si $\mathbb{C}$ est une copule de dimension d et si $\mathbb{F}_1, \ldots, \mathbb{F}_d$ sont des fonctions de répartition univariées, alors la fonction $\mathbb{F}$ est défini par (1.4.2) une fonction de répartition d- dimensionnelle de marginales $\mathbb{F}_1, \ldots, \mathbb{F}_d$.

Corollaire 1.4.1 *Soit $\mathbb{F}$ une distribution multivariée de distributions marginales continues $\mathbb{F}_1, \ldots, \mathbb{F}_d$. Alors, la copule $\mathbb{C}$ associée à $\mathbb{F}$ est donnée par*

$$\mathbb{C}(u) = \mathbb{C}(u_1, \ldots, u_d) = \mathbb{F}\left(\mathbb{F}_1^{-1}(u_1), \ldots, \mathbb{F}_d^{-1}(u_d)\right), u = (u_1, \ldots, u_d)' \in I^d \tag{1.4.3}$$

Corollaire 1.4.2 *La relation entre la fonction de densité conjointe $f(x_1, \ldots, x_d)$ et la densité de la copule $\mathfrak{c}(u_1, \ldots, u_d)$ est donnée par*

$$f(x_1, \ldots, x_d) = \mathfrak{c}(\mathbb{F}_1(x_1), \ldots, \mathbb{F}_d(x_d)) \prod_{i=1}^{d} f_i(x_i) \tag{1.4.4}$$

Remarque 1.4.3 *Pour démontrer la relation* (1.4.4) *suivez les mêmes démarches du cas bivarié.*

1.4.4 Propriétés des copules multivariées

Théorème 1.4.3 (D'inveriance) ([25])*Soit $X = (X_1, \ldots, X_d)'$ un vecteur aléatoire continue, de fonction de répartition jointe $\mathbb{F}$ associée à une copule $\mathbb{C}$ et $(T_1, \ldots, T_d)$ est une suite de fonction strictement croissante, alors la fonction de répartition jointe du vecteur $T(X) = (T_1(X_1), \ldots, T_d(X_d))'$ est associée à la même copule $\mathbb{C}$.*

Ainsi, la copule est invariante par les transformations strictement croissantes de variables aléatoires.

Preuve. On a $\mathbb{F}_i$ et $\mathbb{G}_i$ les fonctions de répartition marginales des variables aléatoire X_i et $T_i(X_i)$ respectives, $\mathbb{C}$ et $\mathbb{D}$ les copules associées respectivement, aux vecteurs aléatoires X et $T(X)$.

Comme les transformations T_i sont croissantes, alors pour tout $x_i \in \overline{\mathbb{R}}$ avec $i = 1, \ldots, d$, on a

$$\mathbb{G}_i(x_i) = \mathbb{P}[T_i(X_i) \leq x_i] = \mathbb{P}\left[X_i \leq T_i^{-1}(x_i)\right] = \mathbb{F}_i \circ T_i^{-1}(x_i).$$

D'aprés l'equation (1.4.3) on a, pour tout $u = (u_1, \ldots, u_d)' \in I^d$,

$$\begin{aligned}
\mathbb{C}(u_1, \ldots, u_d) &= \mathbb{P}\left[X_1 \leq \mathbb{F}_1^{-1}(u_1), \ldots, X_d \leq \mathbb{F}_d^{-1}(u_d)\right] \\
&= \mathbb{P}\left[T_1(X_1) \leq T_1 \circ \mathbb{F}_1^{-1}(u_1), \ldots, T_d(X_d) \leq T_d \circ \mathbb{F}_d^{-1}(u_d)\right] \\
&= \mathbb{P}\left[T_1(X_1) \leq \mathbb{G}_i^{-1}(u_1), \ldots, T_d(X_d) \leq \mathbb{G}_i^{-1}(u_d)\right] \\
&= \mathbb{D}(u_1, \ldots, u_d).
\end{aligned}$$

Comme les variables $X_1, \ldots, X_d$ sont continues alors $Ran\mathbb{F}_1 = \ldots = Ran\mathbb{F}_d = I$ et donc $\mathbb{C} \equiv \mathbb{D}$ sur I^d. ■

Théorème 1.4.4 (Différentiabilité) *Soit $\mathbb{C} : I^d \longrightarrow I$ une copule. Les dérivées partielles de $\mathbb{C}(.)$ existent presque sûrement, pour tout $i = 1, ..., d$, on a*

$$0 \leq \frac{\partial \mathbb{C}(u)}{\partial u_i} \leq 1, \quad u = (u_1, ..., u_d)' \in I^d$$

De plus, les fonctions

$$u \longmapsto \frac{\partial \mathbb{C}(u)}{\partial u_i}$$

sont non décroissantes presque partout.

Théorème 1.4.5 (Convexité) *Soient $\lambda_i \geq 0$, pour $i = 1, ..., p$, avec $\sum_{i=1}^{p} \lambda_i = 1$, et $\mathbb{C}_1, ..., \mathbb{C}_p$ des fonctions copules, alors $\sum_{i=1}^{p} \lambda_i \mathbb{C}_i$ est aussi une copule.*

D'une manière générale, si $(\mathbb{C}_\theta)_{\theta \in \Theta}$ est une famille de copules et Λ est une mesure de probabilité sur l'ensemble Θ, alors $\mathbb{C}^*(.) = \int_\Theta \mathbb{C}_\theta(.)\Lambda(d\theta)$ est aussi une copule.

Remarque 1.4.4 *L'ensemble $\mathcal{C}_d$ des copules de dimension d est un sous-ensemble compact et convexe de l'espace des fonctions continues à valeurs réelles définies sur I^d, muni de la topologie de la convergence uniforme. Il suit que, dans $\mathcal{C}_d$, la convergence en tout point implique la convergence uniforme. La convexité provient directement de la définition de la copule. La compacité est due au fait que l'ensemble des fonctions continues à valeurs réelles définies sur I^d est un espace métrique compact et que le sous-ensemble des d-copules est un sous ensemble fermé.*

Théorème 1.4.6 *([25]) Soit $\mathbb{F}$ une fonction de répartition de dimension d d'un vecteur aléatoire $X = (X_1, ..., X_d)'$ de fonctions de répartition marginales $\mathbb{F}_1, ..., \mathbb{F}_d$: Pour toute copule $\mathbb{C}$ de dimension d associée à $\mathbb{F}$ et pour tout $u = (u_1, \ldots, u_d)' \in [0,1]^d$, on a*

$$\mathbb{W}(u) = \max\left(\sum_{i=1}^{d} u_i - d + 1, 0\right) \leq \mathbb{C}(u) \leq \mathbb{M}(u) = \min(u_1, ..., u_d) \tag{1.4.5}$$

où $\mathbb{M}(u)$ est la copule minimum multivariée et $\mathbb{W}(u)$ est la copule maximum multivariée.

Preuve. On pose $\mathbb{U}_i = \mathbb{F}_i(X_i)$ pour tout, comme on a $\mathbb{P}(\mathbb{U}_i \leq u_i) = u_i$, pour tout $i = 1, ..., d$.

La première inégalité de provient du fait que

$$\begin{aligned}\mathbb{P}(\bigcap_{1\leq i\leq d}(\mathbb{U}_i \leq u_i)) &= 1 - \mathbb{P}(\bigcup_{1\leq i\leq d}(\mathbb{U}_i \geq u_i)) \\ &\geq 1 - \sum_{i=1}^{d}\mathbb{P}(\mathbb{U}_i > u_i) \geq 1 - \sum_{i=1}^{d}(1-u_i) \geq 1 - d + \sum_{i=1}^{d}u_i\end{aligned}$$

Et comme $\mathbb{C}$ est définie sur $[0,1]^d$, alors

$$\mathbb{C}(u) \geq \max\left(0, 1 - d + \sum_{i=1}^{d}u_i\right)$$

La seconde inégalité de (1.4.4) provient du fait que

$$\bigcap_{1\leq i\leq d}(\mathbb{U}_i \leq u_i) \subset (\mathbb{U}_i \leq u_i), \text{ pour tout } i = 1, ..., d.$$

$$\text{ainsi, } \mathbb{C}(u) = \mathbb{P}(\bigcap_{1\leq i\leq d}(\mathbb{U}_i \leq u_i)) \leq \mathbb{P}(\mathbb{U}_i \leq u_i) = u_i, \text{ pour tout } i = 1, ..., d.$$

Par conséquent

$$\mathbb{C}(u) \leq \min(u_1, ..., u_d).$$

■

La fonction $\mathbb{M}^d$ est une copule de dimension d pour $d > 2$.

2 La fonction $\mathbb{W}^d$ n'est pas une copule de dimension d pour $d > 2$.

En effet,

on considère un d-volume $[1/2, 1]^d \subset [0,1]^d$, alors

$$\begin{aligned}V_{\mathbb{W}}\left([1/2,1]^d\right) &= \max(1 + ... + 1 - d + 1, 0) - \\ &\quad -d \times \max(1/2 + 1 + \ldots + 1 - d + 1, 0) + \\ &\quad +C_d^2 \max(1/2 + 1/2 + 1 + \ldots + 1 - d + 1, 0) + \\ &\quad \ldots \\ &\quad + \max(1/2 + ... + 1/2 - d + 1, 0) \\ &= 1 - d/2 + 0... + 0 \leq 0\end{aligned}$$

Donc $\mathbb{W}$ n'est pas une copule pour $d \geq 3$. La relation entre les mesures de dépendance et les copules.

1.5 Relation entre les mesures de dépendance et les copules

1.5.1 Mesures de concordance

L'inégalité des bornes de Fréchet-Hoeffding met en évidence un ordre partiel sur l'ensemble des copules comme suit :

Définition 1.5.1 *Soient $\mathbb{C}_1$ et $\mathbb{C}_2$ deux copules bivariées. On dit que $\mathbb{C}_1$ est plus petite que $\mathbb{C}_2$ et on note $\mathbb{C}_1 \prec \mathbb{C}_2$, si et seulement si*

$$\mathbb{C}_1(u_1, u_2) \leq \mathbb{C}_2(u_1, u_2), \text{ pour tout } (u_1, u_2) \in I^2$$

L'ordre $\prec$ est appelé l'ordre de concordance.

Remarque 1.5.1 *La relation d'ordre " $\prec$ " est partielle car on ne peut pas comparer toutes les copules entre elles. Considérons, par exemple, la copule cubique définie par*

$$\mathbb{C}_{cub}(u_1, u_2, \theta) = \prod_{i=1}^{2} u_i + \theta \prod_{i=1}^{2} \left[u_i \left(u_i - 1\right) \left(2u_i - 1\right)\right], \text{ avec } \theta \in [-1, 2].$$

On a

$$\mathbb{C}_{cub}(\tfrac{3}{4}, \tfrac{3}{4}, 1) = 0.5712 \geq \Pi(\tfrac{3}{4}, \tfrac{3}{4}) = \tfrac{3}{4} \times \tfrac{3}{4} = 0.5625$$

$$\mathbb{C}_{cub}(\tfrac{3}{4}, \tfrac{1}{4}, 1) = 0.1787 \leq \Pi(\tfrac{3}{4}, \tfrac{1}{4}) = \tfrac{3}{4} \times \tfrac{1}{4} = 0.1875$$

Pour toute copule $\mathbb{C}$ et d'aprés l'inégalité de Fréchet-Hoeffding, on a

$$\mathbb{W} < \mathbb{C} < \mathbb{M}.$$

Définition 1.5.2 ***(Mesure de concordance)****Une mesure numérique d'association entre deux v.a continues X et Y, dont la copule est $\mathbb{C}$, est une mesure de concordance si elle satisfait les propriétés suivantes :*

a) κ est définie pour toute paire (X, Y) de v.a continues

b) $-1 = \kappa_{<X,-X>} \leq \kappa_{<\mathbb{C}>} \leq \kappa_{<X,X>} = 1$

c) $\kappa_{<X,Y>} = \kappa_{<Y,X>}$

d) Si X et Y sont indépendantes, alors $\kappa_{<X,Y>} = 0$

e) $\kappa_{<-X,Y>} = \kappa_{<X,-Y>} = -\kappa_{<X,Y>}$

f) Si $\mathbb{C}_1 \prec \mathbb{C}_2$ alors $\kappa_{<\mathbb{C}_1>} \leq \kappa_{<\mathbb{C}_2>}$

g) Si (X_m, Y_m) est une suite de v.a continues dont la copule est $\mathbb{C}_m$ et si $\mathbb{C}_m$ converge vers $\mathbb{C}$ alors $\lim\limits_{m\to\infty} \kappa_{<\mathbb{C}_m>} = \kappa_{<\mathbb{C}>}$

A partir de la propriété (1.5.1) de la définition (1.5.2), on constate que l'ordre de concordance implique l'ordre sur κ. De plus, l'une des propriétés importantes de $\kappa_{<X,Y>}$ est que si α et β sont des fonctions strictement monotones, alors

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

Par exemple, si $Y = f(X)$ et f est décroissante, alors

$$\kappa_{<X,Y>} = \kappa_{<X,f(X)>} = \kappa_{<X,-X>} = -1$$

Parmi les mesures de concordance, trois mesures trés célèbres jouent un rôle important en statistique non-paramétrique, le **tau de Kendall**, le **rho de Spearman** [(1)] et l'**indice de Gini**. Ces mesures s'expriment en fonction de copules.

Théorème 1.5.1 (*voir [25]*)*Soient* (X_1, Y_1) *et* (X_2, Y_2) *deux vecteurs indépendants de v.a continues de fonctions de répartition jointes, respectives,* $\mathbb{F}_1$ *et* $\mathbb{F}_2$*. Soient* $\mathbb{G}_1$ *et* $\mathbb{G}_2$ *les marges associées, respectivement, à* X_1, X_2 *et à* Y_1, Y_2*. Soient* $\mathbb{C}_1$ *et* $\mathbb{C}_2$ *les copules associées, respectivement, à* $\mathbb{F}_1$ *et à* $\mathbb{F}_2$ *données par* $\mathbb{F}_1(x, y) = \mathbb{C}_1(\mathbb{G}_1(x), \mathbb{G}_2(y))$ *et* $\mathbb{F}_2(x, y) = \mathbb{C}_2(\mathbb{G}_1(x), \mathbb{G}_2(y))$*. Si* $\mathbb{Q}$ *est la mesure de concordance et de discordance de* (X_1, Y_1) *et* (X_2, Y_2)*, i.e.,*

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

alors

$$\mathbb{Q} = \mathbb{Q}(\mathbb{C}_1, \mathbb{C}_2) = 4 \int\limits_{[0,1]^2} \mathbb{C}_2(u, v)\, d\mathbb{C}_1(u, v) - 1$$

[(1)] voir Annex A

Preuve. Puisque toutes les variables aléatoires sont continues, alors

$$\mathbb{P}\left\{(X_1 - X_2)(Y_1 - Y_2) < 0\right\} = 1 - \mathbb{P}\left\{(X_1 - X_2)(Y_1 - Y_2) \geq 0\right\}$$

ainsi

$$\mathbb{Q} = 2\mathbb{P}\left\{(X_1 - X_2)(Y_1 - Y_2) \geq 0\right\} - 1$$

L'étude du premier terme

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

ces quantités peuvent être calculées par intégration

$$\begin{aligned}
\mathbb{P}\left\{X_1 > X_2, Y_1 > Y_2\right\} &= \mathbb{P}\left\{X_2 < X_1, Y_2 < Y_1\right\} \\
&= \int\int_{\mathbb{R}^2} \mathbb{P}\left\{X_2 < x, Y_2 < y\right\} d\mathbb{C}_1\left(\mathbb{G}_1(x), \mathbb{G}_2(y)\right) \\
&= \int\int_{\mathbb{R}^2} \mathbb{C}_2\left(\mathbb{G}_1(x), \mathbb{G}_2(y)\right) d\mathbb{C}_1\left(\mathbb{G}_1(x), \mathbb{G}_2(y)\right)
\end{aligned}$$

Par changement de variables $u = \mathbb{G}_1(x)$ et $v = \mathbb{G}_2(y)$, on obtient alors

$$\mathbb{P}\left\{X_1 > X_2, Y_1 > Y_2\right\} = \int\int_{I^2} \mathbb{C}_2(u, v)\, d\mathbb{C}_1(u, v)$$

D'une façon similaire

$$\begin{aligned}
\mathbb{P}\left\{X_1 < X_2, Y_1 < Y_2\right\} &= \int\int_{\mathbb{R}^2} \mathbb{P}\left\{X_2 > x, Y_2 > y\right\} d\mathbb{C}_1\left(\mathbb{G}_1(x), \mathbb{G}_2(y)\right) \\
&= \int\int_{\mathbb{R}^2} \left(1 - \mathbb{G}_1(x) - \mathbb{G}_2(y) + \mathbb{C}_2\left(\mathbb{G}_1(x), \mathbb{G}_2(y)\right)\right) d\mathbb{C}_1\left(\mathbb{G}_1(x), \mathbb{G}_2(y)\right) \\
&= \int\int_{I^2} \left\{1 - u - v + \mathbb{C}_2(u, v)\right\} d\mathbb{C}_1(u, v)
\end{aligned}$$

Or, $\mathbb{C}_1$ est la fonction de répartition jointe du vecteur $(U, V)'$ où U et V sont des variables aléatoires uniformes, donc $E(U) = E(V) = 1/2$ et par conséquent

$$\begin{aligned}
\mathbb{P}\left\{X_1 < X_2, Y_1 < Y_2\right\} &= 1 - \frac{1}{2} - \frac{1}{2} + \int\int_{I^2} \mathbb{C}_2(u, v)\, d\mathbb{C}_1(u, v) \\
&= \int\int_{I^2} \mathbb{C}_2(u, v)\, d\mathbb{C}_1(u, v)
\end{aligned}$$

Enfin, $\mathbb{P}\left\{(X_1 - X_2)(Y_1 - Y_2) > 0\right\} = 2\int\int_{I^2} \mathbb{C}_2(u, v)\, d\mathbb{C}_1(u, v)$. ■

Théorème 1.5.2 ***Preuve.*** *En regroupant ces résultats on déduit*

$$\mathbb{Q} = \mathbb{Q}\left(\mathbb{C}_1, \mathbb{C}_2\right) = 4\int\int_{I^2} \mathbb{C}_2\left(u, v\right) d\mathbb{C}_1\left(u, v\right) - 1$$

■

Corollaire 1.5.1 (*voir* [9])*Soient* $\mathbb{C}_1$, $\mathbb{C}_2$ *et* $\mathbb{Q}$ *définies comme dans le théorème précédent, alors*

1 $\mathbb{Q}$ est symétrique, i.e., $\mathbb{Q}\left(\mathbb{C}_1, \mathbb{C}_2\right) = \mathbb{Q}\left(\mathbb{C}_2, \mathbb{C}_1\right)$

2 $\mathbb{Q}$ conserve l'ordre, i.e., $\mathbb{C}_1 \leq \mathbb{C}_1'$ et $\mathbb{C}_2 \leq \mathbb{C}_2'$, alors $\mathbb{Q}\left(\mathbb{C}_1, \mathbb{C}_2\right) \leq \mathbb{Q}\left(\mathbb{C}_1', \mathbb{C}_2'\right)$

Remarque 1.5.2 *Pour les copules fréquemment utilisées* $\mathbb{M}$, $\mathbb{W}$ *et* $\mathbf{\Pi}$. *On peut facilement evaluer la mesure de concordance* $\mathbb{Q}$ *dans le cas bivarié. Rappelons que le support de* $\mathbb{Q}$ *est la diagonale* $u = v$, *alors que la masse de probabilité de* $\mathbb{W}$ *se située sur la diagonale secondaire* $u = 1 - v$. *Ainsi, on a*

1) $\mathbb{Q}(\mathbb{M}, \mathbb{M}) = 4\int\int_{I^2} \min(u, v) d\mathbb{M}(u, v) - 1 = \int_0^1 v dv - 1 = 1$

2) $\mathbb{Q}(\mathbb{M}, \Pi) = 4\int\int_{I^2} uv d\mathbb{M}(u, v) - 1 = \int_0^1 v^2 dv - 1 = 1/3$

3) $\mathbb{Q}(\mathbb{M}, \mathbb{W}) = 4\int\int_{I^2} \mathbb{W}(u, v) d\mathbb{M}(u, v) - 1 = 4\int_0^1 \mathbb{W}(u, v) dv - 1 = 0$

4) $\mathbb{Q}(\mathbb{W}, \Pi) = 4\int\int_{I^2} uv d\mathbb{W}(u, v) - 1 = 4\int_0^1 (1 - v) v dv - 1 = -1/3$

5) $\mathbb{Q}(\mathbb{W}, \mathbb{W}) = 4\int\int_{I^2} \max(u + v - 1, 0) d\mathbb{W}(u; v) - 1 = 4\int_0^1 0 dv - 1 = -1$

6) $\mathbb{Q}(\Pi, \Pi) = 4\int_0^1\int_0^1 uv du dv - 1 = 0$

Théorème 1.5.3 (Daniels (1950)**)** *Si* X *et* Y *sont deux variables aléatoires continues, alors*

$$-1 \leq 3\tau(X, Y) - 2\rho_S(X, Y) \leq 1$$

Théorème 1.5.4 ((Durbinet Stuart(1951) **(voir** [**22**]**))** *Si* X *et* Y *sont deux variables aléatoires continues, alors*

$$\frac{1 + \rho_S}{2} \geq \left(\frac{1 + \tau}{2}\right)^2 \text{ et } \frac{1 - \rho_S}{2} \geq \left(\frac{1 - \tau}{2}\right)^2$$

En combinant ces deux dernières inégalités, on obtient

$$\frac{3\tau-1}{2} \leq \rho_S \leq \frac{1+2\tau-\tau^2}{2} \qquad si\ \tau \geq 0$$

et

$$\frac{\tau^2+2\tau-1}{2} \leq \rho_S \leq \frac{3\tau+1}{2} \qquad si\ \tau \leq 0$$

Une autre mesure de concordance qui s'exprime a l'aide de la copule c'est l'indice de Gini. Elle est définie par :

$$\gamma = 2 \int\int_{[0,1]^2} (|u+v-1| - |u-v|)\, d\mathbb{C}(u,v).$$

Comme la différentielle d $\mathbb{C}$ intervient dans l'expression de la mesure γ, il est possible de la calculer via l'expression equivalente suivante.voir Nelsen ([24]),page 147

$$\gamma = 4 \int_{[0,1]} (\mathbb{C}(u,u) + \mathbb{C}(u,1-u) - u)\, du$$

Théorème 1.5.5 *Soit (X,Y) un couple elliptique et soit $r = cor(X,Y)$.Alors*

$$\tau(X,Y) = \frac{2}{\pi} \arcsin(r)$$
$$\rho(X,Y) = \frac{6}{\pi} \arcsin(r/2)$$

1.5.2 Dépendance de queue

Le concept de dépendance de queue fournit une description de la dépendance au niveau des queues de distribution, très intéressante pour étudier la survenance simultanée de valeurs extrêmes. C'est une mesure locale contrairement au tau de Kendall et au rho de Spearman qui mesurent la dépendance sur l'ensemble de la distribution.

Définition 1.5.3 (*voir [19])Soient $\mathbb{F}$ une fonction de répartition conjointe de dimension d, I et J deux sous-ensembles disjoints de $\{1,...,n\}$. Le coefficient de queue inférieure de la copule (respectivement le coefficient de queue supèrieure de la copule) d'un vecteur aléatoire $X = (X_1,...,X_d)'$, de marginales $\{\mathbb{F}_i : i = 1,...,d\}$ sont définis dans l'ordre par :*

$$\begin{aligned}\Lambda_L^{I,J}(\underline{X}) &= \lim_{t\to\infty} \mathbb{P}\left[X_i \leq \mathbb{F}_i^{-1}\left(\tfrac{x_i}{t}\right), \forall i \in I \quad / \quad X_j \leq \mathbb{F}_j^{-1}\left(\tfrac{x_j}{t}\right), \forall j \in J\right] \\ &= \lim_{t\to\infty} \mathbb{C}\left(\tfrac{x_i}{t}, \forall i \in I \quad / \quad \tfrac{x_j}{t}, \forall j \in J\right) \\ \Lambda_U^{I,J}(X) &= \lim_{t\to\infty} \mathbb{P}\left[X_i > \mathbb{F}_i^{-1}\left(1-\tfrac{x_i}{t}\right), \forall i \in I \quad / \quad X_j > \mathbb{F}_j^{-1}\left(1-\tfrac{x_j}{t}\right), \forall j \in J\right]\end{aligned}$$

si ces limites existent.

Dans le cas bivarié, le coefficient de queue inférieure de la copule (respectivement supérieure) défini précédemment subi une légère modification, pour tout $(x,y) \in \overline{\mathbb{R}}_+^2$

$$\Lambda_U(x,y) = y\Lambda_U^{\{1\},\{2\}}(x,y)$$
$$\Lambda_L(x,y) = y\Lambda_L^{\{1\},\{2\}}(x,y)$$

A partir de cela, on retrouve la notion de la dépendance de queue qui est une mesure simple et intuitive de la dépendance entre des événements extrêmes.Pour plus de détails voir Nelsen (voir [25]) et Joe(voir [22])

Définition 1.5.4 (Dépendance de queue) *Soient X et Y deux variables aléatoires de fonctions de répartitions $\mathbb{F}$ et $\mathbb{G}$ respectivement. Le coefficient de dépendance de queue supérieure de X et Y (resp.le coefficient de dépendance de queue inférieure de X et Y) est défini par :*

$$\lambda_U = \Lambda_U(1,1) = \lim_{t\to 1^-} \mathbb{P}\left[X > \mathbb{F}^{-1}(t) \,|Y > \mathbb{G}^{-1}(t)\right]$$

respectivement

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

lorsque ces limites existent, et $\lambda_U \in (0,1]$ (resp. $\lambda_L \in (0,1]$).

Remarque 1.5.3 *Si $\mathbb{F}$ et $\mathbb{G}$ sont continues, une définition équivalente de ces coefficients est donnée en fonction des copules par*

$$\begin{aligned}
\lambda_U &= \Lambda_U(1,1) = \lim_{t\to 1^-} \mathbb{P}\left[X > \mathbb{F}^{-1}(t) \,|Y > \mathbb{G}^{-1}(t)\right] > 0 \\
&= \lim_{t\to 1^-} \frac{\mathbb{P}\left[X > \mathbb{F}^{-1}(t), Y > \mathbb{G}^{-1}(t)\right]}{\mathbb{P}\left[Y > \mathbb{G}^{-1}(t)\right]} \\
&= \lim_{t\to 1^-} \frac{1-2t+\mathbb{C}(t,t)}{1-t} \\
&= \lim_{t\to 1^-} \frac{\overline{\mathbb{C}}(t,t)}{1-t}, \text{ où } \overline{\mathbb{C}}(t,t) \text{ est la copule de survie de } C^{(1)}
\end{aligned}$$

Si $\lambda_U \in]0,1]$, alors X et Y sont dites asympototiquement dépendantes au noveau supérieur de la queue de la distibution. Si $\lambda_U = 0$, alors X et Y sont dites asympototiquement

indépendantes au noveau supérieur de la queue de la distibution

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

Si $\lambda_L = 0$, alors elle n'a pas de dépendance de queue à gauche.

Si $\lambda_U \in]0,1]$, cela signifie qu'il y a une probabilité non nulle que l'une des variables dépassé t sachant que l'autre est déjà supérieure à t. Autrement dit, λ_U est la probabilité d'avoir un extrême sachant qu'on a déjà un extrême.

Ces expressions permettent de calculer les coefficients de dépendance pour une large famille de copules paramétriques, particulièrement les copules archimédiennes. Dans le tableau suivant on donne l'expression de ces coefficients pour quelques copules usuelles Pour plus de détails, voir Joe, H. (voir [21]) .

Copules	$\mathbb{C}(u,v)$	λ_L	λ_U
Indépendance	uv	0	0
Gumbel	$\exp\left\{-\left[(-\ln u)^\theta + (-\ln v)^\theta\right]^{1/\theta}\right\}$	0	$2-2^{1/\theta}$
Clayton	$\left(u^{-\theta}+v^{-\theta}-1\right)^{-\frac{1}{\theta}}$	$2^{-\frac{1}{\theta}}$	0
Borne supérieur de Fréchet	$\min(u,v)$	1	1
Borne inférieur de Fréchet	$\max(u+v-1,0)$	0	0

TAB. 1.1: Tableau coefficients de quelques copules usuelles

Chapitre 2

Copules paramétriques et non paramétriques

Introduction

Dans ce chapitre, on va introduire les copules paramétriques les plus importantes est leurs propriétés et on présente les copules non paramétriques. Ce chapitre est composé de cinq sections. Dans la première, section on présente les copules paramétriques (copules elliptiques, copules archemidiennes). Dans la deuxiemme section, on présente les copules des valeurs extrêmes. Dans la troisiémme section on va présenter les copules Archimax. Enfin, et dans la dernière section on parle de copules non paramétriques.

2.1 Copules paramétriques

Les copules paramamétriques sont des fonctions de répartition qui dépendent d'un ou plusieurs paramètres. Dans cette section, on s'interesse aux copules elliptiques et les copules archemidiennes.

2.1.1 Copules elliptiques

Les copules elliptiques sont des copules associées aux distributions elliptiques multivariées . Dans cette section on introduit la famille des copules elliptiques faciles à simuler sont les copules gaussiennes et les copules de Student .

Définition 2.1.1 *On appelle copule elliptique toute copule de la forme suivante :*

$$c_{\varrho} = \frac{1}{\sqrt{1-\rho^2}} \int_{-\infty}^{\phi_{(f_1)}(u)^{-1}} \int_{-\infty}^{\phi_{(f_2)}(v)^{-1}} f(\frac{x^2 - 2\rho xy + y^2}{\sqrt{\rho^2}}) d_x d_y = \mathbb{F}_{\varrho}(\phi^{-1}_{(f_1)}(u), \phi^{-1}_{(f_2)}) \quad (2.1.1)$$

où F_{ϱ} est la distribution jointe des variables X et Y, $\phi^{-1}_{(f_1)}(u)$ et $\phi^{-1}_{(f,2)}(v)$ sont leurs fonctions quantiles respectives et ϱ leurs coefficients de corrélation.

Remarque 2.1.1 *Cette famille donne la ralation entre la copule Gaussienne et la copule de Student .*

2.1.2 Copule gaussienne

La copule gaussienne ne présente pas de dépendance de queue au niveau supérieur de la distribution (respectivement au niveau inférieure) et donc n'est pas adaptée à des valeurs extrêmes. L'importance de cette copule réside dans le fait qu'elle est sous-jacente à la distribution normale multivariée.

En effet, pour modéliser la structure de dépendance d'un échantillon par une copule gaussienne est cohérent avec la mesure de cette dépendance par le coefficient de corrélation linéaire. La fonction de distribution de la copule gaussienne de dimension d, s'écrit :

$$\mathbb{C}(u_1, \ldots, u_d) = \Phi_{\Sigma}(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_d)), \ \forall u = (u_1, \ldots, u_d)' \in I^d. \quad (2.1.2)$$

Φ^{-1} est l'inverse de la distribution normale centrée réduite univariée.

La fonction $\Phi_{\Sigma} = \frac{1}{(2\pi)^{d/2}\det(\Sigma)^{1/2}} \exp\left(-\frac{1}{2} x \Sigma^{-1} x'\right)$ est la fonction de répartition de la loi normale centrée réduite et Σ sa matrice de variance covariance (égale à la matrice de corrélation dans ce cas).

Définition 2.1.2 *Une copule Gaussienne de dimension d et de matrice de covariance Σ est définie comme suit :*

$$\mathbb{C}(u_1, \ldots, u_d) = \int_{-\infty}^{\Phi^{-1}(u_1)} \cdots \int_{-\infty}^{\Phi^{-1}(u_d)} \frac{1}{(2\pi)^{d/2}\det(\Sigma)^{1/2}} \exp\left(-\frac{1}{2}\left(W^T \Sigma^{-1} W\right)\right) dw_1 \ldots dw_d.$$

où $W = (w_1, \ldots, w_d)' = (\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_d))'$, pour tout $u = (u_1, \ldots, u_d)' \in I^d$.

Théorème 2.1.1 *La densité d'une copule Gaussienne multivariée est donnée par*

$$\mathfrak{c}\left(u_1,\ldots,u_d\right)=\frac{1}{\det(\Sigma)^{1/2}}\exp\left(-\frac{1}{2}W^T\left(\Sigma^{-1}-Id_2\right)W\right) \tag{2.1.3}$$

où $W=\left(w_1,\ldots,w_d\right)'=\left(\Phi^{-1}\left(u_1\right),\ldots,\Phi^{-1}\left(u_d\right)\right)'$.

Preuve. En utilisant la relation et les fonctions de distribution Gaussiennes univariées et multivariées, on obtient :

$$\frac{1}{(2\pi)^{d/2}\det(\Sigma)^{1/2}}\exp\left(-\frac{1}{2}\left(X^T\Sigma^{-1}X\right)\right)=c\left(\Phi\left(x_1\right),\ldots,\Phi\left(x_d\right)\right)\prod_{i=1}^{d}\frac{1}{\sqrt{2\pi}}\exp\left(\frac{-1}{2}\left(x_i^2\right)\right).$$

où $x=\left(x_1,\ldots,x_d\right)'$.

En posant $u_i=\Phi\left(x_i\right),i=1,\ldots,d$, on trouve

$$\mathfrak{c}\left(u_1,\ldots,u_d\right)=\frac{\frac{1}{(2\pi)^{d/2}\det(\Sigma)^{1/2}}\exp\left(-\frac{1}{2}\left(X^T\Sigma^{-1}X\right)\right)}{\frac{1}{(2\pi)^{d/2}}\exp\left(-\frac{1}{2}W^TW\right)}=\frac{1}{\det(\Sigma)^{1/2}}\exp\left(-\frac{1}{2}\left(W^T\left(\Sigma^{-1}-Id\right)W\right)\right)$$

avec $W=\left(w_1,\ldots,w_d\right)'=\left(\Phi^{-1}\left(u_1\right),\ldots,\Phi^{-1}\left(u_d\right)\right)'$. ■

Définition 2.1.3 *Soit* $\left(X_1,\ldots,X_d\right)'$ *un vecteur aléatoire d'une copule Gaussienne* $\mathbb{C}$*, alors pour tout* i *et* j *telle que* $i,j=1,...,d$*, le tau de Kendall est donné par*

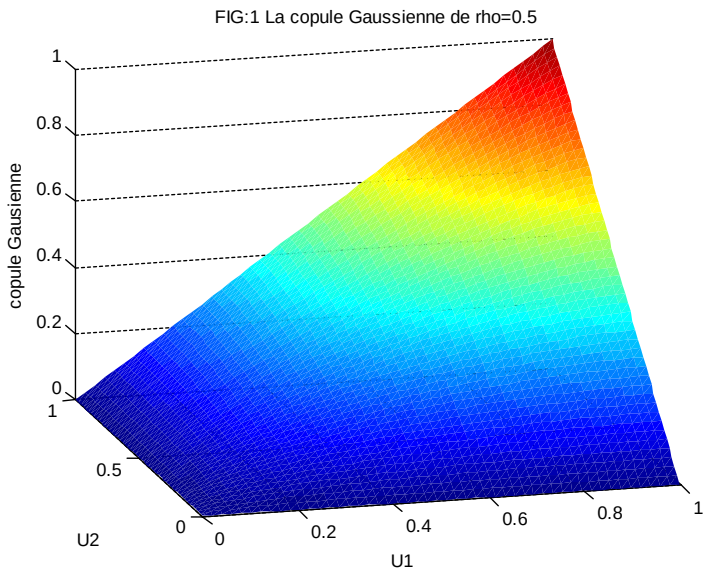
$$\tau_{X_i,X_j}=\frac{2}{\pi}\arcsin\rho_{i,j}$$

avec $\rho_{i,j}$ *est le coefficient de corrélation linéaire entre* X_i *et* X_j*.*

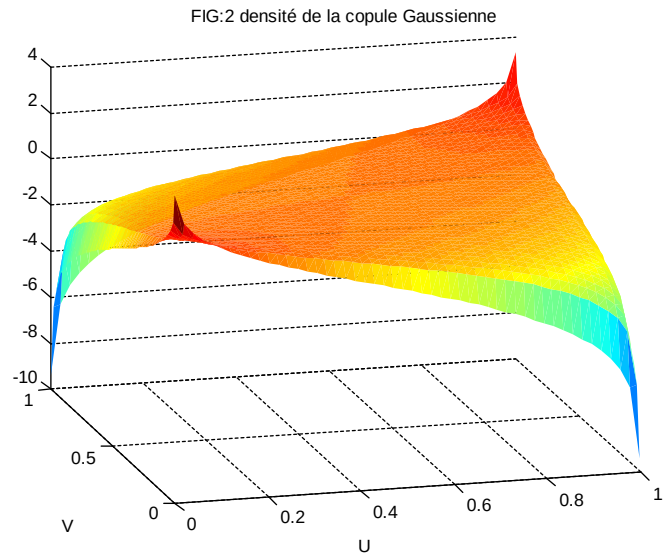
Soit $\left(X_1,\ldots,X_d\right)'$un vecteur aléatoire d'une copule Gaussienne $\mathbb{C}$, alors le rhô de Spearmen est donné par

$$\rho_{S_{X_i,X_j}}=\frac{6}{\pi}\arcsin\frac{\rho_{i,j}}{2},i,j=1,...,d$$

avec $\rho_{i,j}$ est le coefficient de corrélation linéaire entre X_i et X_j.



Copule gaussienne de rho=0.8



Densité des copules gaussienne

2.1.3 Copule de student

La copule de Student ($t-$copula) est la copule sous-jacente à une distribution multivariée de Student. Cette structure de dépendance capte les dépendances extrêmes positives et négatives. Elle est construite de la même manière que la copule gaussienne mais à partir de la distribution de Student centrée réduite.

Définition 2.1.4 *Soit t_υ la fonction de disrtibution du loi de student univariée à υ d.d.l. On note par $t_{\Sigma,\upsilon}$ la fonction de disrtibution de student multivariée de matrice de corrélation Σ et à υ d.d.l. La copule de student multivariée associée à $t_{\Sigma,\upsilon}$ est définie par*

$$\mathbb{C}^t_{\Sigma,\upsilon}(u_1,\ldots,u_d) = t_{\Sigma,\upsilon}\left(t_\upsilon^{-1}(u_1),\ldots,t_\upsilon^{-1}(u_d)\right), u=(u_1,\ldots,u_d)' \in I^d \tag{2.1.4}$$

où t_υ^{-1} est l'inverse de la distribution de Student univariée à υ d.l.l

Définition 2.1.5 *La fonction de densité de la copule de Student multivariée est donnée par*

$$\mathfrak{c}^t_{\Sigma,\upsilon}(u_1,\ldots,u_d) = \frac{f_{\upsilon,\Sigma}\left(t_\upsilon^{-1}(u_1),\ldots,t_\upsilon^{-1}(u_d)\right)}{\prod_{i=0}^{n} f_\upsilon\left(t_\upsilon^{-1}(u_i)\right)}, u=(u_1,\ldots,u_d)' \in I^d$$

La fonction de distribution t_υ^{-1} est l'inverse de la distribution de Student centrée réduite univariée à υ d.d.l. La fonction $f_{\upsilon,\Sigma}$ est la densité de probabilité de la loi de Student centrée réduite, Σ sa matrice de corrélation et f_υ est la densité univariée de la loi de Student centrée réduite. A titre de rappel sa densité s'écrit pour tout $x=(x_1,\ldots,x_d)' \in \mathbb{R}^d$:

$$f_{\upsilon,\Sigma}(x) = \frac{\Gamma\left(\frac{\upsilon+d}{2}\right)}{\Gamma\left(\frac{\upsilon}{2}\right)\sqrt{(\pi\upsilon)^d \det(\Sigma)}}\left(1+\frac{x^t\Sigma^{-1}x}{\upsilon}\right)^{-(\upsilon+d)/2}$$

Définition 2.1.6 *La copule de student multivariée de dimension d de matrice de corrélation Σ et à υ d.d.l est donné par*

$$\mathbb{C}^t_{\Sigma,\upsilon}(u_1,\ldots,u_d) = \int_{-\infty}^{t_\upsilon^{-1}(u_1)} \cdots \int_{-\infty}^{t_\upsilon^{-1}(u_d)} \frac{\Gamma\left(\frac{\upsilon+d}{2}\right)\left(1+\frac{W^t\Sigma^{-1}W}{\upsilon}\right)^{-(\upsilon+d)/2}}{\det(\Sigma)^{1/2}\Gamma\left(\frac{\upsilon}{2}\right)(\upsilon\pi)^{d//2}} dw_1,\ldots dw_d$$

où $\mathbb{W}=(w_1,\ldots,w_d)'$.

Proposition 2.1.1 *Dans le cas bivarié, l'expression de la copule de student est donnée par*

$$\mathbb{C}^t_{\Sigma,\upsilon}(u,v) = \int_{-\infty}^{t_\upsilon^{-1}(u)} \int_{-\infty}^{t_\upsilon^{-1}(v)} \frac{1}{2\pi(1-\rho^2)^{1/2}}\left(1+\frac{s^2-2\rho st+t^2}{\upsilon(1-\rho^2)}\right)^{-(\upsilon+2)/2}$$

où ρ est le coéfficient de corrélation linéaire.

Preuve. Pour $d = 2, w = (x_1, x_2)^{'}$,

$$f_{\upsilon,\Sigma}(x_1, x_2) = \frac{\Gamma(\frac{\upsilon+2}{2})(1 + \frac{w^{'}\Sigma^{-1}w}{\upsilon})^{\frac{-\upsilon+2}{2}}}{\sqrt{\det(\Sigma)}\Gamma(\frac{\upsilon}{2})(\upsilon\pi)}$$

■

On pose $x_1 = t$ et $x_2 = s$, on obtient

$$f(s,t) = \frac{1}{2\pi\sqrt{1-\rho^2}}\left(1 + \frac{s^2 + t^2 - 2\rho st}{\upsilon(1-\rho^2)}\right)^{\frac{-(\upsilon+2)}{2}}$$

Théorème 2.1.2 *La densité copule de student multivariée est donnée par*

$$\mathfrak{c}^t_{\Sigma,\upsilon}(u_1, \ldots, u_d) = \frac{\Gamma\left(\frac{\upsilon+d}{2}\right)\left[\Gamma\left(\frac{\upsilon}{2}\right)\right]^{d-1}\left(1+\frac{W^t\Sigma^{-1}W}{\upsilon}\right)^{-(\upsilon+d)/2}}{\det(\Sigma)^{1/2}\left[\Gamma\left(\frac{\upsilon+1}{2}\right)\right]^d \prod_{i=1}^{d}\left(1+w_i^2/\upsilon\right)^{-(\upsilon+1)/2}}$$

avec $(w_1, \ldots, w_d) = (t_\upsilon^{-1}(u_1), ..., t_\upsilon^{-1}(u_d))$ *et* $\upsilon > 0$.

1 La distribution de student multivariée est formée par la combinaison de la copule de student avec des lois marginales de student.

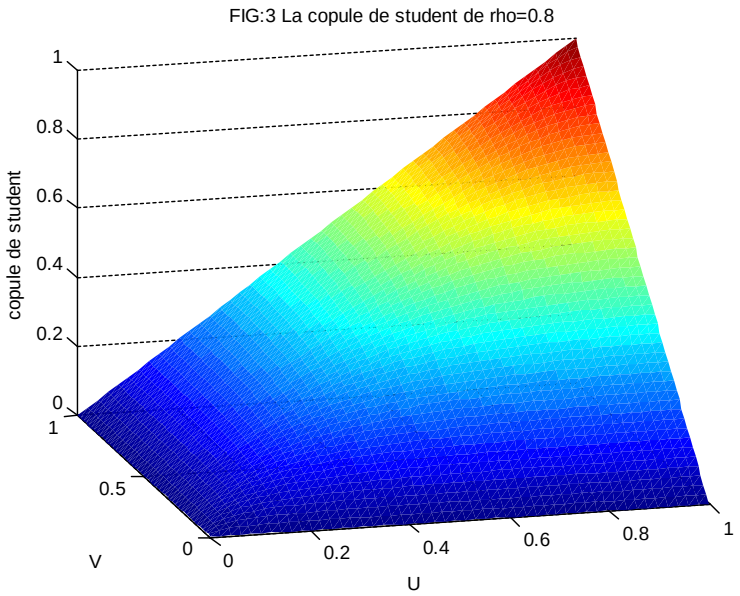
2 Le tau de kendall pour la copule de student est identique au cel de la copule Gaussienne, mais pour le rhô de Spearman il n'existe pas de forme explicite.

La densité copule de student bivariée est donnée par

$$\mathfrak{c}^t_{\Sigma,\upsilon}(u, v) = \frac{\upsilon}{2\sqrt{1-\rho^2}}\frac{\Gamma\left(\frac{\upsilon}{2}\right)^2}{\left(\Gamma\left(\frac{\upsilon+1}{2}\right)\right)^2}\frac{\left(1+x^2+y^2-2\rho xy\right)^{-(\upsilon+2)/2}}{((1+x^2/\upsilon)(1+x^2/\upsilon))^{-(\upsilon+1)/2}}$$

avec $x = t_\upsilon^{-1}(u), y = t_\upsilon^{-1}(v)$ et $\upsilon > 0$.

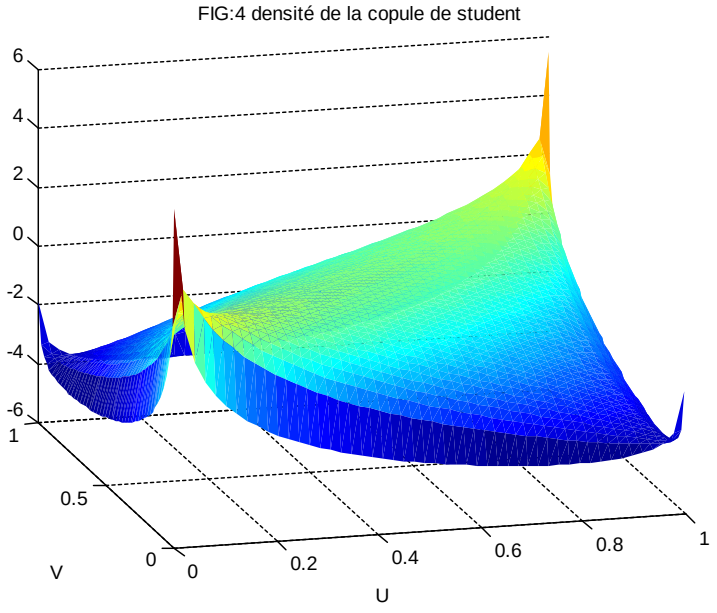
La représentation graphique de la copule de student et densité copule pour



Copule de student $rho = 0,8$

DZ

EEC



$_5$*.pdf*

Densité de copule student

2.2 Copules archimédiennes

La famille des copules archimédiennes permetrent de fqire une varietée de stricture de dépendance de la familles de copule, de plus leurs formes analytiques facilite

la simulation.

Définition 2.2.1 *Soit φ une fonction continue, strictement décroissante définie de I dans $[0,+\infty]$ telle que $\varphi(1)=0$. Le pseudo-inverse (inverse généralisée) de φ est la fonction $\varphi^{[-1]}$ de domaine de définition $[0,+\infty]$, à valeur dans $[0,1]$, donnée par*

$$\varphi^{[-1]}(t)=\begin{cases}\phi^{-1}(t) \;\; si\; 0\leq t\leq\varphi(0)\\ 0 \;\; si\; \varphi(0)\leq t\leq\infty\end{cases}$$

Remarquons que $\varphi^{[-1]}$ est continue, non croissante sur $[0,+\infty]$ est strictement décroissante sur $[0,\varphi(0)]$. En outre $\varphi^{[-1]}(\varphi(u))=u$ sur $[0,1]$ et

$$\begin{aligned}\varphi\left(\varphi^{[-1]}(t)\right) &= \begin{cases} t \;\; si\; 0\leq t\leq\varphi(0)\\ \varphi(0) \;\; si\; \varphi(0)\leq t\leq\infty\end{cases}\\ &= \min(t,\varphi(0)).\end{aligned}$$

Si $\varphi(0)=+\infty$, alors $\varphi^{[-1]}(t)=\varphi^{-1}(t)$.

Définition 2.2.2 *Soit φ une fonction strictement décroissante et continue sur $]0,1]$. La fonction définie pour tout $u=(u_1,...,u_d)'\in I^d$ par :*

$$\mathbb{C}(u_1,...,u_d)=\varphi^{-1}(\varphi(u_1)+...+\varphi(u_d))$$

est une copule si φ est d-fois dérivable sur $]0,1[$ et si pour tout $1\leq i\leq d$, $\varphi^i>0$ pour i paire et $\varphi^i<0$ sinon. Les fonctions s'écrivant de la sorte s'appelle les copules archimédiennes dont le générateur est φ.

De la définition précédente on déduit que la densité d'une copule archimédienne multivariée s'écrit :

$$c(u_1,...,u_d)=\left(\varphi^{-1}\right)^d(\varphi(u_1)+...+\varphi(u_d))\prod_{i=1}^{d}\varphi'(u_i).$$

Remarque 2.2.1 *Les copules archimédienne présentent un double intérêt. D'une part, elles permettent de construireune grande variété de familles de copules et donc représenter une grande variété de structures de dépendance. D'autre part, les copules ainsi générées ont des formes analytiques fermées et sont faciles à simuler. Pour plus d'information sur cette famille de copules le lecteur pourra se référer à Nelsen(voir [25])*

2.3 Propriétés

Les résultats du théorème suivant nous permettront l'extension a la copule Archimédienne multivariée

Théorème 2.3.1 *(voir [26]) Soit C une copule Archimédienne et φ son générateur, alors*

1. $\mathbb{C}$ est Symmétrique
$$\mathbb{C}(u_1, u_2) = \mathbb{C}(u_2, u_1), \forall (u_2, u_1) \in I^2.$$
2. $\mathbb{C}$ est Associative
$$\mathbb{C}(\mathbb{C}(u_1, u_2), u_3) = \mathbb{C}(u_1, \mathbb{C}(u_2, u_3)), \forall (u_2, u_1, u_3) \in I^3.$$

Preuve. On a d'apré le théorème (2.3.1)
$\mathbb{C}(\mathbb{C}(u_1, u_2), u_3) = \varphi^{-1}(\varphi(\mathbb{C}(u_{1,}u_2) + \varphi(u_3)$
$=\varphi^{-1}(\varphi(\varphi^{-1}(\varphi(u_1) + \varphi(u_2))) + \varphi(u_3)$
$= \varphi^{-1}(\varphi(u_1) + \varphi(u_2) + \varphi(u_3))$
$= \varphi^{-1}(\varphi(u_1) + \varphi(\varphi^{-1}(\varphi(u_2) + \varphi(u_3))))$
$= \varphi^{-1}(\varphi(u_1) + \varphi(C(u_{2,}u_3) = \mathbb{C}(u_{1,}C(u_{2,}u_3))$. ■
Donc $\mathbb{C}$ est associative

2.4 Quelques types des copules Archimédiennes

Les copules les plus utilisés dans cette famille sont les copule de Gumbel ; de Clayton et de Frank.

Copule de Cook – Johnson

La copule de Cook-Johnson, dite aussi : copule de Clayton ou copule de Kimeldorf-Sampson est une copule archimédienne dont le générateur est défini par

$$\varphi(u) = \alpha^{-1}\left(u^{-\alpha} - 1\right), \text{ et } \varphi^{-1}(s) = (1+\alpha s)^{-1/\alpha}, \ \alpha > 0, \ u \in]0,1]$$

La copule multivariée de Clayton s'écrit

$$\mathbb{C}_\alpha(u_1, ..., u_d) = \left(1 - d + \sum_{i=1}^{d} u_i^{-\alpha}\right)^{-1/\alpha}$$

Cette copule est différentiable et sa densité est donnée par

$$c_\alpha(u_1, ..., u_d) = \frac{\partial^d \mathbb{C}_\alpha(u_1,...,u_d)}{\partial u_1 ... \partial u_d} = \left(1 - d + \sum_{i=1}^{d} u_i^{-\alpha}\right)^{-d-\frac{1}{\alpha}} \prod_{j=1}^{d} \left(u_i^{-\alpha-1}(j\alpha - \alpha + 1)\right)$$

La copule de Clayton présente une dépendance asymptotique à gauche (sur les valeurs négatives) ce qui n'est pas le cas de la copule gaussienne.

Proposition 2.4.1 *Dans le cas bivarié, la copule de Clayton est donnée par*

$$\mathbb{C}_\alpha(u, v) = \left(u^{-\alpha} + v^{-\alpha} - 1\right)^{-1/\alpha}$$

Remarque 2.4.1 *Pour $\alpha \geq 0$ la copule de Clayton est complète ; i.e., $\mathbb{C}_{-1} = \mathbb{M}$, $\mathbb{C}_\infty = \Pi$ et $\lim_{\alpha \longrightarrow 0} \mathbb{C}_\alpha = \mathbb{W}$.*

Le tau de kendall pour la famille de Clayton est donné par :

$$\begin{aligned} \tau_\alpha &= 1 + 4\int_0^1 \frac{\varphi(u)}{\varphi'(u)} du = 1 + 4\int_0^1 \frac{u^{\alpha+1}+u}{\alpha} du \\ &= 1 + \frac{4}{\alpha}\left(\frac{1}{\alpha+2} - \frac{1}{2}\right) = \frac{\alpha}{\alpha+2} \end{aligned}$$

et le coéfficient de dépendance de queue inférieure (respectivement supérieur) est donné par $\lambda_U = 0$ (respectivement λ_L) telle que

$$\begin{aligned} \lambda_L &= 2\lim_{s\to\infty}\left(\frac{\varphi^{-1\prime}(2s)}{\varphi^{-1\prime}(s)}\right) \\ &= 2\lim_{s\to\infty}\frac{(1+2\alpha s)^{-1/\alpha-1}}{(1+\alpha s)^{-1/\alpha-1}} \\ &= 2.2^{-1/\alpha-1} \\ &= 2^{-1/\alpha} \end{aligned}$$

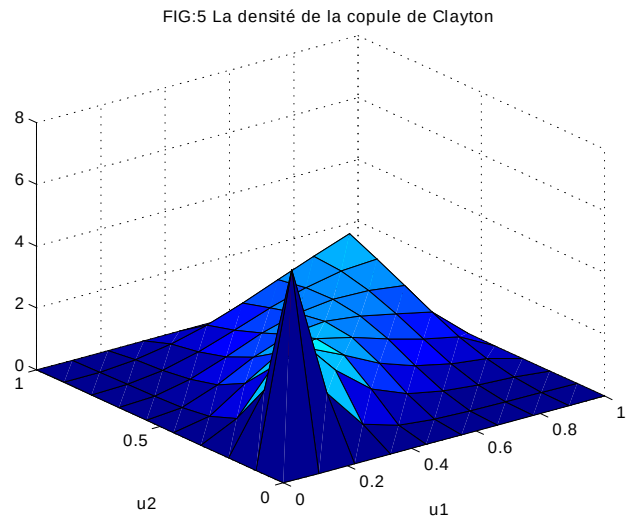


FIG. 2.4.1 : Densité de la copule de clayton

2.4.1 Copule de Franck

La copule de Franck permet de modéliser les dépendances aussi bien positives que négatives. Cependant, elle ne présente pas de dépendance de queue. Le générateur de cette copule archimédienne est :

$$\varphi(u) = -\ln\left(\frac{e^{-\alpha u}-1}{e^{-\alpha}-1}\right), \text{ et } \varphi^{-1}(s) = \frac{-1}{\alpha}\ln\left(1-\left(1-e^{-\theta}\right)e^{-s}\right)$$

où $\alpha \geq 0$ et $u \in]0,1]$.

La copule de Franck multivariée est donnée par

$$\mathbb{C}_\alpha(u_1,...,u_d) = -\frac{1}{\alpha}\ln\left(1+\frac{1}{(e^{-\alpha}-1)^{d-1}}\prod_{i=1}^{d}(e^{-\alpha u_i}-1)\right)$$

La densité copule de la copule de Franck bivariée

$$\mathfrak{c}_\alpha(u_1,u_2) = \frac{\alpha\left(1-e^{-\alpha}\right)e^{-\alpha(u_1+u_2)}}{\left((1-e^{-\alpha})-\left(1-e^{-\alpha u_1}\right)\left(1-e^{-\alpha u_2}\right)\right)^2}.$$

Le tau de Kendall de cette famille est donné par la formule suivante :

$$\tau_\alpha = 1-\frac{4}{\alpha}\left(1-D_1(\alpha)\right), \alpha \neq 0$$

où D_k est la fonction de Debye donnée par

$$D_k(x) = \frac{k}{x_k}\int_0^x \frac{t^k}{e^t-1}dt \quad \forall k \geq 0$$

et le rhô de Spearman est donné par

$$\rho_\alpha = 1 - \frac{12}{\alpha}\left(D_1(\alpha) - D_2(\alpha)\right)$$

Comme $\left[\varphi^{-1}\right]'(0) = -\frac{e^\alpha - 1}{\alpha}$, alors le coefficient de dépendance de queue supérieur de cette famille n'existe pas.

D'autre part $\frac{\varphi^{-1'}(2s)}{\varphi^{-1'}(s)} = e^{-s}\frac{1-\left(1-e^{-\alpha}\right)e^{-s}}{1-(1-e^{-\alpha})e^{-2s}}$, par conséquent, $\lim\limits_{s\longrightarrow\infty}\frac{\varphi^{-1'}(2s)}{\varphi^{-1'}(s)} = 0$, donc le coefficient de dépendance de queue inferieure de cette famille n'existe pas.

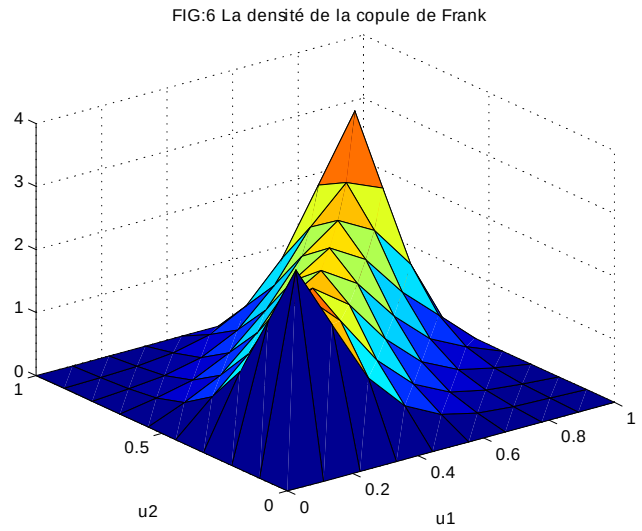


FIG. 2.4.2 : Densité de copule franck

2.4.2 Copule de Gumbel

La copule de Gumbel est appelée parfois, copule de Gumbel-Hougaard ; elle permet de modéliser les dépendances extrêmes. En effet, les copules constituent un outil statistique permettant de modéliser la d´ependance entre des variables aléatoires. La fonction copule relie en effet la densité jointe aux densités marginales et contient ainsi toute l'information sur la structure de d´ependance du modéle, nous traiterons ici le cas bivariée et nous finirons par une généralisation au cas multivariée. Elle appartient à la famille des copules archimédiennes où son générateur s'écrit :

$$\varphi(u) = (-\ln(u))^\alpha, \text{ avec } \alpha > 0,\ u \in]0,1] \text{ et } \varphi^{-1}(s) = \exp\left(-s^{1/\alpha}\right)$$

La copule de Gumbel s'écrit

$$\mathbb{C}_\alpha(u_1,\ldots,u_d) = \exp\left(-\left[\sum_{i=1}^{d}(-\ln(u_i))^\alpha\right]^{1/\alpha}\right)$$

Le tau de Kendall de la copule de Gumbel est

$$\tau_\alpha = 1 - \tfrac{1}{\alpha}$$

Cette famille de copule possède une dépendance de queue supérieur

$$\begin{aligned}\lambda_U &= 2 - 2\lim_{s\to\infty}\frac{\varphi^{-1\prime}(2s)}{\varphi^{-1\prime}(s)}\\ &= 2 - 2^{1/\alpha}\frac{\exp\left(-(2s)^{1/\alpha}\right)}{\exp\left(-s^{1/\alpha}\right)}\\ &= 2 - 2^{1/\alpha}\end{aligned}$$

et une dépendance de queue inférieur $\lambda_L = 0$.

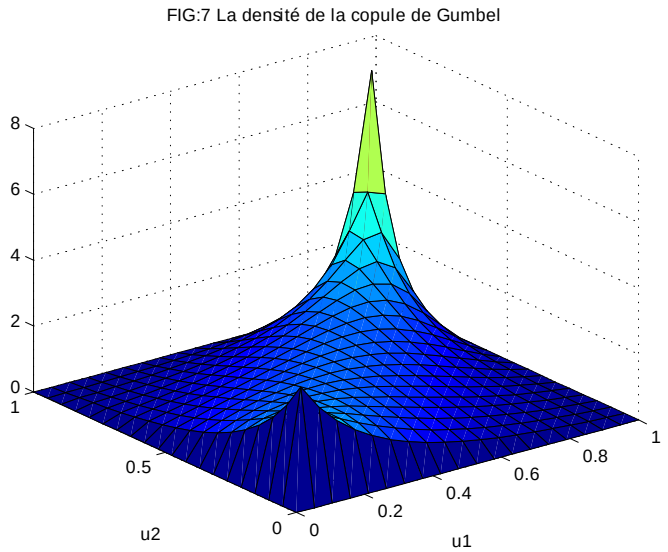


FIG. 2.4.3 : Densité de copule gumbel

Le tableau suivant regroupe quelques types de copules archimédiennes

Copules	$\varphi(u)$	$\mathbb{C}(u)$
Gumbel	$(-\ln u)^{\theta}, \theta \geq 1$	$\exp\left\{-\left[\sum_{i=1}^{d}(-\ln u_i)^{\theta}\right]^{\frac{1}{\theta}}\right\}$
Frank	$-\ln\frac{\exp(-\theta u)-1}{\exp(-\theta)-1}, \theta \in \mathbb{R}\backslash\{0\}$	$-\frac{1}{\theta}\ln\left(1+\frac{\prod_{i=1}^{d}(\exp(-\theta u_i)-1)}{(\exp(-\theta)-1)^{d-1}}\right)$
Clayton	$u^{-\theta}-1, \theta > 0$	$\left(\sum_{i=1}^{d} u_i^{-\theta}-d+1\right)^{-\frac{1}{\theta}}$

TAB. 2.1: Tableau de quelques types de copules archimédiennes

2.5 Copules des valeurs extrêmes

Les copules extrêmes jouent un rôle trés important dans la théorie des extrêmes multivariés. Deheuvels (1979a)(voir [7]), à fait usage de cette famille, pour obtenir une solution au problème de convergence de type $\left\{\underset{1\leq i\leq d}{sup}\, X_i(1), \ldots, \underset{1\leq i\leq d}{sup}\, X_i(n)\right\}$, où $\{X_i(1), \ldots, X_i(n), i \geq 1\}$ est une suite de vecteurs aléatoires indépendantes et de même loi à valeurs dans $\mathbb{R}^n$.Deheuvels (1978)(voir [6]),Deheuvels(1978a)(voir [7]) a déterminé, de manière exhaustive, les lois limites possibles par une représentation explicite et a fourni des conditions simples de convergence vers cette limite, il a obtenu des résultats permettant de caractériser le comportement asymptotique des extrêmes multivariés, même lorsqu'il n'y a pas de loi limite.

Définition 2.5.1 *Une copule de valeurs extrêmes (extreme value copula ou copule EV) vérifie la relation suivante*

$$\mathbb{C}(u_1^t, \ldots, u_d^t) = \mathbb{C}^t(u_1, \ldots, u_n), \ \forall t > 0 \tag{2.5.1}$$

Soit $(X_{i,1}, \ldots, X_{i,d}), i = 1, \ldots, n$ un échantillon de n vecteurs aléatoires $i.i.d$ de fonction de répartition multivariée $\mathbb{F}$, de marges $\mathbb{F}_1, ..., \mathbb{F}_d$ et de copule $\mathbb{C}$. On définit le maximum d'ordre $n \geq 1$, noté $\mathbb{M}_n = (\mathbb{M}_{n,1}, \ldots, \mathbb{M}_{n,d})'$, dont les coordonnées sont définies par

$$\mathbb{M}_{n,j} = max(X_{1,j}, \ldots, X_{n,j}), j = 1, \ldots, d \tag{2.5.2}$$

Si il existe des suites réelles $a_{n,j} \in \mathbb{R}$ et $b_{n,j} \in \mathbb{R}_+^*$,$j = 1, \ldots, d$, la théorie multidimensionnelle des valeurs extrêmes s'intéresse à la loi limite

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{\mathbb{M}_{n,1}-b_{n,1}}{a_{n,1}} \leq x_1, \ldots, \frac{\mathbb{M}_{n,d}-b_{n,d}}{a_{n,d}} \leq x_d\right) = \mathbb{F}^*(x_1, \ldots, x_d) \tag{2.5.3}$$

avec $\mathbb{F}^*$ est la fonction de répartition de $\mathbb{M}_n$ dont les lois marginales sont $\mathbb{F}_1^*, \ldots, \mathbb{F}_d^*$ associées à $\mathbb{M}_{n,1}, \ldots, \mathbb{M}_{n,d}$ ne sont pas dégénérées, alors $\mathbb{F}^*$ est appelée loi multivariée de valeurs extrêmes de maximum.

On dit alors que la fonction de répartition $\mathbb{F} \in MDA\left(\mathbb{F}^*\right)$.

Le théorème suivant permet de caractériser la loi des distributions des valeurs extrêmes dans le cas univarié.

Théorème 2.5.1 (*voir [11])Supposons d v.a* $X_1, ..., X_d$ *indépendantes et de même loi de distribution* $\mathbb{F}$*. S'il existe des constantes* a_n *et* b_n *et une distribution limite non dégénérée* $\mathbb{F}^*$ *telles que*

$$\lim_{n\to\infty}\left(\frac{\mathbb{M}_n-a_n}{b_n} \leq x\right) = \mathbb{F}^*\left(x\right) \qquad \forall x \in \mathbb{R}$$

alors $\mathbb{F}^*$ appartient à l'un des trois types suivants de distribution

Type 1 (Gumbel)	$\mathbb{F}^*\left(x\right) = \exp\left(-e^{-x}\right)$	$x \in \mathbb{R}$	$f^*\left(x\right) = \exp\left(-x - e^{-x}\right)$
Type 2 (Fréchet)	$\mathbb{F}^*\left(x\right) = \begin{cases} 0 \\ \exp\left(-x^{-\alpha}\right) \end{cases}$	$\begin{matrix} x \leq 0 \\ x > 0 \end{matrix}$	$f^*\left(x\right) = \begin{cases} 0 \\ \alpha x^{-(\alpha+1)} \exp\left(-x^{-\alpha}\right) \end{cases}$
Type 3 (Weibell)	$\mathbb{F}^*\left(x\right) = \begin{cases} \exp\left(-\left(-x\right)^{\alpha}\right) \\ 1 \end{cases}$	$\begin{matrix} x \leq 0 \\ x > 0 \end{matrix}$	$f^*\left(x\right) = \begin{cases} \alpha\left(-x\right)^{\alpha-1} \exp\left(-\left(-x\right)^{\alpha}\right) \\ 0 \end{cases}$

TAB. 2.2: Quelquest types de distribution de copules

Remarque 2.5.1 *Pour distinguer les trois distributions, on utilise généralement les notations suivantes :* Λ *pour la distribution Gumbel,* Φ_α *pour la distribution Fréchet et* Ψ_α *pour la distribution* ***Weibull (Resnick*** (1987)*).*

Théorème 2.5.2 *Toute loi limite* $\mathbb{F}^*$ *vérifiant* (2.5.3) *est continue, de plus ces lois marginales sont de type GEV. Du fait que* $\mathbb{F}^*$ *est continue, d'aprés le théorème de Sklar elle peut s'écrire*

$$\mathbb{F}^*(x_1, \ldots, x_d) = \mathbb{C}^* \left(\mathbb{F}_1^* (x_1), .., \mathbb{F}_d^* (_d)\right) \tag{2.5.4}$$

avec $\mathbb{C}^*$ *est la copule associée à* $\mathbb{F}^*$

Théorème 2.5.3 (*voir* [20])$\mathbb{C} \in MDA(\mathbb{C}^*)$ *si* $\mathbb{C}$ *satisfait la relation suivante*

$$\lim_{t\to\infty} \mathbb{C}^t \left(u_1^{1/t}, ..., u_d^{1/t}\right) = \mathbb{C}^* (u_1, .., u_d), \quad (u_1, ..., u_d) \in I^d \tag{2.5.5}$$

Remarque 2.5.2 *La famille des copules des valeurs extrêmes peut être interprétée comme limite des copules* $\mathbb{C}^*$ *quand* n *tend vers l'infini.*

Parfois, il est difficile de vérifier la relation (2.5.5). La condition suivante est plus facile à vérifier (2.5.3)

Théorème 2.5.4 $\mathbb{C} \in MDA(\mathbb{C}^*)$ *si et seulement si*

$$\lim_{u\to 0} \frac{1-\mathbb{C}((1-u)^{w_1}, ..., (1-u)^{w_d})}{u} = B(w_1, \ldots, w_d) \tag{2.5.6}$$

Dans le cas bidimensionnel, cette condition devient

$$\lim_{u\to 0} \frac{1-\mathbb{C}\left((1-u)^{1-t}, ..., (1-u)^{t}\right)}{u} = A(t) \quad \textit{pour tout } t \in I \tag{2.5.7}$$

Dans ce qui suit on utilise le concept de max-stabilité afin de simplifier la représentation des copules EV.

Définition 2.5.2 *Une copule multivariée est dite max-stable si elle vérifiée*

$$\mathbb{C}(u_1, ..., u_d) = \mathbb{C}\left(u_1^{1/m}, .., u_d^{1/m}\right)^m, m \in \mathbb{N}^+ \textit{ et } (u_1, ..., u_d) \in I^d \tag{2.5.8}$$

Remarque 2.5.3 *Toute copule max-stable est dans son propoe domaine d'attaction.*

Théorème 2.5.5 *Une copule est dite de valeurs extrêmes si et seulement si elle est max-stable.*

Preuve. Il est clair que toute copule max-stable est dite de valeurs extrêmes. Inversement si $\mathbb{C}^*$ une copule valeurs extrêmes, avec elle est vérifiée la relation (2.5.6) pour une certainne copule $\mathbb{C}$. Soit un réel positif r

Pour un entier fixé $m \geq 1$ et pour $n = mk$, on a

$$\mathbb{C}^{*r}\left(u_1^{1/r}, ..., u_d^{1/r}\right) = \lim_{t\to\infty} \mathbb{C}^{rt}\left(u_1^{1/rt}, .., u_d^{1/rt}\right)^t = \mathbb{C}^*\left(u_1, .., u_d\right)$$

donc $\mathbb{C}^*$ est max-stable. ■

Exemple 2.5.1 *copule de Gambel bivariée*

$$\mathbb{C}_\alpha\left(u_1, u_2\right) = \exp\left(-\left[\left(-\ln u_1\right)^\alpha + \left(-\ln u_2\right)^\alpha\right]^{1/\alpha}\right)$$

pour $\alpha \geq 0$, *on a*

$$\begin{aligned}
\mathbb{C}_\alpha\left(u_1^n, u_2^n, \alpha\right) &= \exp\left(-\left[\left(-\ln u_1^n\right)^\alpha + \left(-\ln u_2^n\right)^\alpha\right]^{1/\alpha}\right) \\
&= \exp\left(-\left[n^\alpha\left\{\left(-\ln u_1\right)^\alpha + \left(-\ln u_2\right)^\alpha\right\}\right]^{1/\alpha}\right) \\
&= \left[\exp\left(-\left[\left(-\ln u_1\right)^\alpha + \left(-\ln u_2\right)^\alpha\right]\right)\right]^n \\
&= \mathbb{C}_\alpha\left(u_1, u_2, \alpha\right)^n
\end{aligned}$$

La copule de Gambel est donc une copule de valeurs extrêmes. Ce qui n'est pas le cas de la copule Farlie-Gambel-morgenstern qui est donnée par

$$\mathbb{C}_\alpha\left(u_1, u_2, \alpha\right) = u_1u_2 + \alpha u_1u_2\left(1 - u_1\right)\left(1 - u_2\right)$$

on a

$$\begin{aligned}
\mathbb{C}_\alpha\left(u_1^n, u_2^n, \alpha\right) &= u_1^n u_2^n + \alpha u_1^n u_2^n\left(1 - u_1^n\right)\left(1 - u_2^n\right) \\
&= u_1^n u_2^n\left(1 + \alpha - \alpha u_1^n - \alpha u_2^n + \alpha u_1^n u_2^n\right) \\
&\neq u_1^n u_2^n\left(1 + \alpha - \alpha u_1 - \alpha u_2 + \alpha u_1 u_2\right)^n \\
&\neq \mathbb{C}_\alpha^n\left(u_1, u_2, \alpha\right)
\end{aligned}$$

La famille des copules de valeurs extrêmes coïncide avec l'ensemble des copules des distributions de valeurs extrêmes, i.e., la classe des distributions limites avec des marginales non dégénérés de

$$\left(\frac{\mathbb{M}_{n,1} - b_{n,1}}{a_{n,1}}, \ldots, \frac{\mathbb{M}_{n,d} - b_{n,d}}{a_{n,d}}\right) \tag{2.5.9}$$

2.6 Copules Archimax

Capéraà, Fougéres et Genest (2000) (voir [3]) ont introduit les copules archimax qui englobe la plupart des familles connues des copules, notamment les copules Archimédiennes et toutes les copules de valeurs extrêmes. Cette famille offre une grande flexibilité pour la modélisation. En effet, on peut connaître a priori les différents max-domaines d'attraction (MDA).

Définition 2.6.1 *Une fonction bivariée est une copule Archimax si et seulement si elle est de la forme :*

$$\mathbb{C}_{\varphi,A}(u_1,u_2)=\varphi^{-1}\left[\left(\varphi\left(u_1\right)+\varphi\left(u_2\right)\right)A\left(\frac{\varphi(u_1)}{\varphi(u_1)+\varphi(u_2)}\right)\right],\ \textit{pour tout}\ 0\leq u_1,u_2\leq 1 \quad (2.6.1)$$

avec

1. $A:[0,1]\rightarrow[1/2,1]$ tel que $max(t,1-t)\leq A(t)\leq 1$ pour tout $0\leq t\leq 1$.
2. $\varphi:(0,1]\rightarrow[0,\infty)$ est une fonction convexe, décroissante qui vérifie $\varphi(1)=0$ avec la convention suivante : $\varphi(0)=\lim\limits_{t\rightarrow 0^{+}}\varphi(t)$ et $\varphi^{-1}\left(s\right)=0$ pour tout $s\geq\varphi\left(0\right)$

Remarque 2.6.1 *Les copules Archimax contient toutes les copules Archimédiennes et toutes les copules de valeurs extrêmes. Si $\varphi(t)=ln(1/t)$, on obtient*

$$\mathbb{C}_{\varphi,A}(u_1,u_2)=\mathbb{C}_{A}(u_1,u_2)=\exp\left(\ln\left(u_1u_2\right)A\left(\frac{\ln(u_2)}{\ln(u_1u_2)}\right)\right),0\leq u_1,u_2\leq 1. \quad (2.6.2)$$

On reconnaît la forme générale des copules bivariées de valeurs extrêmes ; pour $A(t)=1$, on obtient la forme générale des copules Archimédiennes et elle est donnée par

$$\mathbb{C}_{\varphi,A}(u_1,u_2)=\mathbb{C}_{\varphi}(u_1,u_2)=\varphi^{-1}\left(\varphi\left(u_1\right)+\varphi\left(u_2\right)\right). \quad (2.6.3)$$

L'intérêt de la construction (2.6.2) réside dans le fait qu'avec un choix judicieux de φ et A, on obtient une copule qui sera dans le MDA de n'importe quel "attracteur" $\mathbb{C}_{A^*}$ donné.

Proposition 2.6.1 *Soit $\mathbb{C}_{\varphi,A}$ une copule Archimax telle que $\varphi(1-1/t)\in RV_{-m}$ pour un certain $m\geq 1$. Alors $\mathbb{C}_{\varphi,A}$ appartient au max-domaine d'attraction de la copule $\mathbb{C}_{A^*}$ où la fonction A^* est donnée par*

$$A^{*}(t)=\left[t^{m}+(1-t)^{m}\right]^{1/m}A^{1/m}\left(\frac{t^{m}}{t^{m}+(1-t)^{m}}\right),t\in I. \quad (2.6.4)$$

Cet attracteur peut être considéré de deux points de vue : $\mathbb{C}_{A^*}$ *est une copule Archimax de générateur* A^* *et* $\varphi(t) = ln(1/t)$ *ou* $\mathbb{C}_{A^*}$ *est une copule Archimax de générateurs* A *et* $\varphi^*(t) = ln^m(1/t)$*. De plus,* A^* *et* A *coincident si et seulement si* $m = 1$*.*

Donc on peut remarquer que la seule copule Archimédienne et de valeurs extrêmes est la copule de Gumbel.

En faisant un raisonnement en sens inverse, alors $\mathbb{C}_{\varphi,A} \in MDA(\mathbb{C}_{A^*})$ si

$$\begin{cases} \varphi(1-1/t) \in RV_{-m} & \text{et} \\ A(t) = \left[t^{1/m} + (1-t)^{1/m}\right]^m \left\{A^*\left(\frac{t^{1/m}}{t^{1/m}+(1-t)^{1/m}}\right)\right\}^m \end{cases} \tag{2.6.5}$$

La relation (2.6.5) définit-elle une fonction de dépendance ?

Supposons que : $D(t) = d\ln(A^*(t))/dt$

$$H^*(t) = t + (1-t)D(t) \text{ et } h^*(t) = dH^*(t)/dt$$

Capéraà, Fougères et Genest (2000) ont montré que la relation (2.6.4) est vraie si et seulement si

$$\frac{h^*(t)}{H^*(t)(1-H^*(t))} \geq \frac{m}{t(1-t)}, \text{ pour } t \in I \tag{2.6.6}$$

Cette relation reste valable pour tout $1 \leq m' \leq m$. Ainsi, si $A(t)$ est définie par (2.6.5), $C_{\varphi,A} \in MDA(C_{A^*})$ pour vu que $\varphi(1-1/t) \in RV_{-m'}$ pour $1 \leq m' \leq m$, on peut aussi remarquer que pour des générateurs A et φ quelconques tels que $\varphi(1-1/t) \in RV_{-m}$, on a

$$\lambda = \lim_{u \to \to 1} \frac{1-2u+C_{\varphi,A}(u,u)}{1-u} = 2 - (2A(1/2))^{1/m} \leq 1 \tag{2.6.7}$$

Donc les copules Archimax présentent une dépendance au niveau supérieur de la queue de distribution sauf si $m = 1$ et $A \equiv 1$.

Exemple 2.6.1 *considérons le générateur de Tawn* (1988) *défini par*

$$A_\alpha(t) = \alpha t^2 - \alpha t + 1, \text{ pour } 0 \leq \alpha \leq 1 et\, t \in [0,1] \tag{2.6.8}$$

et le combiner au générateur de la copule de Clayton

$$\varphi_{1,\eta}(t) = \frac{t^{-\eta}-1}{\eta} \tag{2.6.9}$$

ou celles de Genest et Ghoudi (1994) (voir [17])

$$\varphi_{2,\eta}(t) = (1 - t^{\eta})^{1/\eta} \tag{2.6.10}$$

On remarque que $\varphi_{2,\eta}(t) \in RV_{1/\eta}$. Ainsi, pour $\eta \neq 1$, $\mathbb{C}_{\varphi_{2,\eta},A}(x, y)$ n'est pas dans le domaine d'attraction de $\mathbb{C}_A$. Enfin, supposons que l'on veut construire une copule dans le max-domaine d'attraction de

$$\begin{gathered} A^{*}_{\alpha,\beta}(t) = \left[\alpha^k t^k + \beta^k (1-t)^k\right]^{1/k} + \left[(1-\alpha)^k t^k + (1-\beta)^k (1-t)^k\right]^{1/k}, \\ 0 \leq t \leq 1, 0 \leq \alpha, \beta \leq 1 \text{et } k \geq 1 \end{gathered} \tag{2.6.11}$$

Pour les valeurs $\alpha = 1/10$, $\beta = 3/10$, $m = 3/2$ et $k = 1.6$, la condition (2.6.5) est vérifiée. La copule originale A est donc celle définie par l'expression (2.6.6).

2.7 Copules non paramétriques

Les méthodes non paramétriques sont essentilement basé sur la copule empirique donnée par l'exprission suivante

$$\widehat{\mathbb{C}}(u_{1,...,}u_2) = \widehat{\mathbb{F}}(\widehat{\mathbb{F}}_1^{-1}(u_1), ..., \widehat{\mathbb{F}}_d^{-d}(u_d)),$$

ou $\widehat{\mathbb{F}}$ est un estimateur de la fonction de la fonction de répartition

2.7.1 Copule empirique

La notion de copule empirique à été introduite par Deheuvels en (1979) (voir [8]), qui est connue sous le nom « Fonction de dépendance empirique ».

Définition 2.7.1 *Soit* $\{x_1^t, ..., x_d^t\}_{t=1}^T$ *un échantillon d'un vecteur aléatoire* X *de dimension* d*. La fonction de répartition empirique est donnée par*

$$\mathbb{F}_d(x_1, \ldots, x_d) = \frac{1}{T}\sum_{t=1}^{T}\mathbb{I}_{\left\{X_1^t \leq x_1, ..., X_d^t \leq x_d\right\}}$$

avec $\mathbb{I}$ *est la fonction indicatrice.*

a Cas des copules bivariées

Définition 2.7.2 *Soit* $\{(x_t, y_t)\}_{t=1}^T$ *un échantillon de taille* T *d'un couple de v.a. La copule empirique est la fonction* $\widehat{\mathbb{C}}$ *définie par*

$$\widehat{\mathbb{C}}\left(\tfrac{i}{T}, \tfrac{j}{T}\right) = \frac{\textit{Nombre des paires } (x,y) \textit{ dans l'échantillon tels que } x \leq x_{(i)} \textit{ et } y \leq y_{(j)}}{T}$$

où $x_{(i)}$ *et* $y_{(j)}$ *représentent les statistiques d'ordre associées à l'échantillon.*

La fonction densité empirique de la copule $\hat{C}$ notée $\hat{c}$ est donnée par

$$\hat{c}\left(\tfrac{i}{T}, \tfrac{j}{T}\right) = \begin{cases} \frac{1}{T} & \text{si } \left(x_{(i)}, y_{(j)}\right) \text{ est un élément de l'échantillon} \\ 0 & \text{si non} \end{cases} \qquad i, j = 1, 2$$

IL existe une relation entre $\widehat{\mathbb{C}}$ et $\hat{c}$ donnée par

$$\widehat{\mathbb{C}}\left(\tfrac{i}{T}, \tfrac{j}{T}\right) = \sum_{p=1}^{i}\sum_{q=1}^{j}\hat{c}\left(\tfrac{p}{T}, \tfrac{q}{T}\right)$$

et

$$\hat{c}\left(\tfrac{i}{T}, \tfrac{j}{T}\right) = \widehat{\mathbb{C}}\left(\tfrac{i}{T}, \tfrac{j}{T}\right) - \widehat{\mathbb{C}}\left(\tfrac{i-1}{T}, \tfrac{j}{T}\right) - \widehat{\mathbb{C}}\left(\tfrac{i}{T}, \tfrac{j-1}{T}\right) + \widehat{\mathbb{C}}\left(\tfrac{i-1}{T}, \tfrac{j-1}{T}\right)$$

b Généralisation dans les cas des copules multivariées

Définition 2.7.3 *Soit* $\{x_1^t, ..., x_d^t\}_{t=1}^T$ *un échantillon de taille* T *d'un vecteur aléatoire* X *de dimension* d *, soit* $(r_1^t, \ldots, r_d^t)$ *la statistique de rang associée à cet échantillon multivarié. La statistique de rang représente tout simplement le rang d'une observation, et donc sa place dans l'échantillon une fois ordonné. Cela signifie que pour tout* $1 \leq i \leq d$*,* r_i^t *est le rang de* x_i^t *dans* $(x_i^t)_{1\leq t\leq T}$*. La copule empirique introduite par Deheuvels est définie sur l'ensemble* $\left\{\left(\frac{t_1}{T}, \ldots, \frac{t_d}{T}\right), 1 \leq i \leq d, t_i = 0, 1, \ldots, T\right\}$ *par*

$$\widehat{\mathbb{C}}_T\left(\tfrac{t_1}{T}, \ldots, \tfrac{t_d}{T}\right) = \tfrac{1}{T}\sum_{t=1}^{T}\prod_{i=1}^{d}\mathbb{I}_{r_i^t \leq t_i}$$

Il est toutefois possible d'estimer la fonction de densité de la copule empirique de Deheuvels à partir de la copule $\hat{C}_T$, avec la relation suivante (voir [25])

$$\hat{c}_T\left(\tfrac{t_1}{T}, \ldots, \tfrac{t_d}{T}\right) = \sum_{j_1=1}^{2} \cdots \sum_{j_d=1}^{2} (-1)^{j_1+\cdots+j_d}\, \widehat{\mathbb{C}}_T\left(\tfrac{t_1-j_1+1}{T}, \ldots, \tfrac{t_d-j_d+1}{T}\right)$$

Chapitre 3

Méthodes d'estimation des copules

L'estimation en statistique est une opération qui permet de calculer la valeur inconnue d'un paramètre θ du modéle statistique à partir d'un échantillion observé de taill n,par exemple pour la loi normale de paramètres μ et σ^2, telle que l'estimateur de moyenne $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ et l'estimateur de σ^2 est la variance $\mathrm{S}_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X_n})^2$. Parmet les méthodes d'estimation paramètrique, on cite la méthode du maximum de vraisemblance (méthode classique), la méthode des moments, la méthode d'inférence sur les marginales (IFM),...etc, et comme les copules sont des fonctions de répartition paramètriques de copule et comme on a déja vu (voir chapitre 02), donc on a besoin d'estimer ces paramètres. Dans troisième chapitre, on va présenter les méthodes d'estimation paramètriques et non paramètriques d'une copule.

3.1 Estimation paramétrique

Lorsqu' il s'agit d'une estimation paramètrique, il est clair que la distribution conjointe dépend d'un paramétre ou plusieurs ensembles de paramètres que l'on cherche à estimer. Le but de cette section est de présenter la méthode d'estimation paramètrique (on commence par la méthode des moments puis en discute la méthode du maximum de vraisemblance (MLE) enfin, on présente la méthode ces l'inférence sur les marginales (IFM).

3.1.1 Méthode des moments

Cette méthode est utilisée pour les mesures de dépendance, l'estimateur des moments de la mesure de dépendance considérée est obtenu en égalant l'expression paramétrique (analytique) de la mesure avec un estimateur non paramétrique de cette même mesure. Cette méthode consiste à estimer les paramètre θ_i, $i = 1, ..., d$ de lois marginales et le paramètre θ_C de la copule par la méthode des moments, i.e.,

1. Résoudre le système de d équations à d inconnues $\begin{cases} \overline{X_t} = f(\theta_1, ..., \theta_d) \\ S_t^2 = g(\theta_1, ..., \theta_d) \\ \mu_{3,t} = h(\theta_1, ..., \theta_d) \\ \vdots \end{cases}$
2. où d désigne la dimension de f, g et h sont les expressions des moments (ordinaires) d'ordre $1, 2$ et 3 en fonction des paramètres marginales θ_i et θ_C le paramètre de la copule. Répéter cette étape pour toutes les marginales.
3. Inverser le tau de Kendall ou le rhô de Spearman pour obtenir le paramètre θ_C de la copule.

Exemple 3.1.1 *(Copule de Clayton)*

le tau de Kendall de la copule de Clayton est $\theta = (\theta + 2)$ *avec* $\theta \in [-1, 0[\cup]0, \infty[$.

L'estimateur par la méthode des moments est :

$$\hat{\theta}_M = \frac{2\widehat{\tau}}{1-\widehat{\tau}}.$$

Exemple 3.1.2 *(Copule de Gumbel)*

Soit la copule de Gambel de pamamètre θ*,tel que son tau de Kendall est égale à* $(\theta - 1)$ *avec* $\theta \geq 1$, *ayant des marginales exponentielles de paramétre* β, *alors*

$$\widehat{\beta} = \frac{1}{\overline{X_n}} = \hat{\theta}_M = \frac{1}{1-\widehat{\tau}}$$

$\widehat{\beta}$ est l'estimateur de β.

Pour plus de détail sur l'estimation des paramètres de la copule par la méthode des moments voir Brahimi et Necir (voir [1]).

3.1.2 Méthode du maximum de vraisemblance (MLE)

Soit un échantillon $(x_1^t, ..., x_d^t)_{1\leq t\leq T}$ de fonction de répartition multivariée conjointe $\mathbb{F}(x, \theta)$ qui dépend du vecteur des paramètres $\theta \in \Theta \subset \mathbb{R}^p$; telle que $\theta = \left(\alpha_1, \alpha_2^{'}\right)^{'}$ où $\alpha_1 = \left(\theta_1^{'}, ..., \theta_d^{'}\right)^{'}$ paramètre marginales et α_2 est le paramètre de la copule, d'aprés le théorème de Sklar on a

$$\mathbb{F}(x_1, ..., x_d, \theta) = \mathbb{C}\left(\mathbb{F}_1(x_1, \theta_1), \ldots, \mathbb{F}_d(x_d, \theta_d); \alpha_2\right)$$

on a

$$f(x_1, ..., x_d, \theta) = c\left(\mathbb{F}_1(x_1, \theta_1), \ldots, \mathbb{F}_d(x_d, \theta_d); \alpha_2\right) \prod_{i=1}^{d} f_i(x_i, \theta_i)$$

désigne la densité copule de dimension d $\mathbb{C}(u_1, ..., u_d; \alpha_2)$ et f_i sont les densités des fonctions de répartition marginales $\mathbb{F}_i$, pour $i = 1, ..., d$, alors l'expréssion du log-vraisemblance de l'échantillon $(x_1^t, ..., x_d^t)_{1\leq t\leq T}$ est

$$\ell(x; \theta) = \ln \prod_{t=1}^{T} f(x_1^t, ..., x_d^t, \theta) = \sum_{t=1}^{T} \ln f(x_1^t, ..., x_d^t, \theta) \tag{3.1.1}$$

On peut récrire l'équation (3.1.1) comme suit

$$\begin{aligned}
\ell(x; \theta) &= \sum_{t=1}^{T} \ln\left(c\left(\mathbb{F}_1\left(x_1^t; \theta_1\right), ..., \mathbb{F}_d\left(x_d^t; \theta_d\right); \alpha_2\right)\right) \\
&\quad + \sum_{t=1}^{T} \ln f_1\left(x_1^t; \theta_1\right) + \ldots + \sum_{t=1}^{T} \ln f_d\left(x_d^t; \theta_d\right) \\
&= \sum_{t=1}^{T} \ln\left(c\left(\mathbb{F}_1\left(x_1^t; \theta_1\right), ..., \mathbb{F}_d\left(x_d^t; \theta_d\right); \alpha_2\right)\right) + \sum_{t=1}^{T}\sum_{i=1}^{d} \ln f_i\left(x_i^t; \theta_i\right)
\end{aligned}$$

où $\theta = \left(\alpha_1, \alpha_2^{'}\right)^{'}$ est le vecteur des paramètres (des marges et de la copule).

L'estimateur du maximum de vraisemblance de θ est donné par

$$\hat{\theta}_{MLE} = \arg\max_{\theta} L(x; \theta) \tag{3.1.2}$$

On peut montrer sous les conditions de régularité que l'éstimateur $\hat{\theta}_{MLE}$ existe ; il est consistant et asymptotiquement efficace de plus il vérifié la propriété de la normalité asymptotique. i.e.,

$$\sqrt{T}\left(\hat{\theta}_{MLE} - \theta_0\right) \rightarrow \aleph\left(0, \mathfrak{F}^{-1}(\theta_0)\right) \tag{3.1.3}$$

où $\mathfrak{F}^{-1}(\theta_0)$ est l'information usuelle de Ficher et θ_0 est la vraie valeur du θ.

Exemple 3.1.3 (La copule Gaussienne multivariée) *Soit la fonction de distribution de la copule Gaussienne multivariée* $\mathbb{C}$ *donnée par* (2.5.6) *et sa fonction de densité donnée par* (2.1.3).

Soit l'échantillon $(x_1^t, ..., x_d^t)_{1 \leq t \leq T}$. La fonction log-vraisemblance est

$$\ell_\theta = -\frac{T}{2} \ln (\det \Sigma) - \frac{1}{2} \sum_{t=1}^{T} W_t' \left(\Sigma^{-1} - I\right) W_t \tag{3.1.4}$$

où $W_t = \left(\Phi^{-1}(u_1^t), \ldots, \Phi^{-1}(u_d^t)\right)'$.

L'estimateur de maximum de vraisemblance de Σ est donné par [(1)]

$$\hat{\Sigma}_{MLE} = \frac{1}{2} \sum_{t=1}^{T} W_t' W_t$$

L'estimation par la méthode du maximum de vraisemblance est difficile à l'utiliser lorsque d est assez grand ; ceci introduit un grand nombre de paramètres à estimer ce qui rend l'estimation difficile. Shih et Louis (1995) (voir [28]), Joe et Xue (1996) (voir [22]) ont proposé la méthode suivante

3.1.3 Méthode d'inférence sur les marginales (IFM)

Shih et Louis (1995), Joe et Xue (1996) ont proposé de faire une estimation en deux étapes suivant l'algorithme suivant :

1. Estimer les paramètres $\theta_1, ..., \theta_d$ des marges. Obtenir $\widehat{\theta}_1, ..., \widehat{\theta}_d$ par la méthode MLE.
2. Etant $\widehat{\theta}_1, ..., \widehat{\theta}_d$ sont obtenus dans la première étape, estimer donc le paramètre de la copule θ_C. On injecte $\widehat{\theta}_1, ..., \widehat{\theta}_d$ dans l'expression de la partie des marginales

Cette méthode d'estimation est appelée par Shih et Louis « two-stage parametric ML method ». Joe et Xu (1996) utilisent le terme « inference functions for margins » (IFM). L'idée sous-jacente est d'écrire la vraisemblance sous la forme suivante

$$\ell(x; \theta) = \sum_{t=1}^{T} \ln \left(c\left(\mathbb{F}_1(x_1^t; \theta_1), ..., \mathbb{F}_d(x_d^t; \theta_d); \theta_C\right)\right) + \sum_{t=1}^{T} \sum_{i=1}^{d} \ln f_i(x_i^t; \theta_i)$$

avec

$$\theta = \left(\theta_1', ..., \theta_d', \theta_C'\right)'$$

[(1)] pour plus detail voir Magnus et Neudecker

chaque paramètre θ_i (peut être multidimentionnel) pour $i = 1, ..., d$ est estimé par la méthode du maximum de vraisemblance sans tenir compte de la copule, i.e.,

$$\hat{\theta}_i = \arg\max \sum_{t=1}^{T} \ln f_i\left(x_i^t; \theta_i\right), i = 1, ..., d$$

Le paramètre θ_C de la copule est ensuite estimé en tenant compte des estimations précédentes, i.e., on va utiliser les $\hat{\theta}_i$ obtenus dans la première étape pour obtenir $\hat{\theta}_C$

$$\hat{\theta}_C = \arg\max_{\theta} \sum_{t=1}^{T} \ln\left(c\left(\mathbb{F}_1\left(x_1^t; \hat{\theta}_1\right), ..., \mathbb{F}_d\left(x_d^t; \hat{\theta}_d\right); \theta_C\right)\right)$$

L'estimateur par la méthode *IFM* noté $\hat{\theta}_{IFM}$ est donné par

$$\hat{\theta}_{IFM} = \left(\hat{\theta}_1, ..., \hat{\theta}_d, \hat{\theta}_C\right)'$$

Joe (1997) a montré que l'estimateur *IFM* vérifié dans les conditions de régularité la propriété de la normalité asymptotique, i.e.,

$$\sqrt{T}\left(\hat{\theta}_{IFM} - \theta_0\right) \rightarrow \aleph\left(0, \mathcal{G}\left(\theta_0\right)\right)$$

où $\mathcal{G}\left(\theta_0\right)$ est la matrice d'information de Godambe et elle est donée par

$$\mathcal{G}\left(\theta_0\right) = D^{-1} M \left(D^{-1}\right)'$$

avec

$$D = E\left[\frac{\partial \mathfrak{g}(\theta)}{\partial \theta^T}\right]$$

et

$$M = E\left[\mathfrak{g}\left(\theta\right)^T \mathfrak{g}\left(\theta\right)\right]$$

et$\mathfrak{g}\left(\theta\right)$ n'est rien d'autre que la fonction score définite par

$$\frac{\partial \ell_1}{\partial \theta_1}, \ldots, \frac{\partial \ell_d}{\partial \theta_1}, \frac{\partial \ell_C}{\partial \theta_C}$$

3.2 Estimation non paramétriques

Dans certaine études il est possible que l'approche paramétrique conduit un trés mauvais ajoustement au données ; dans cette situation il est préferable d'utilisé l'approche d'estimation non paramétrique.

3.2.1 Copules empirique

La notion des copules empiriques (voir la subsection 2.7.1), sera utilisé pour construire des estimateurs non paramétrique de la mesure associer comme par exemple le tau de kandal et rho de spermane.

Théorème 3.2.1 *soient $\widehat{\mathbb{C}}_T$ et $\hat{c}_T$ la copule empirique et sa fonction de densité respectivement de l'échantillon $\{(x_t, y_t)\}_{t=1}^T$. Le tau de Kendall et le Rho de spearman sont estimés par les formules suivantes*

$$\hat{\tau} = \frac{2T}{T-1} \sum_{i=1}^{T} \sum_{j=1}^{T} \sum_{p=1}^{i-1} \sum_{q=1}^{j-1} \left[\hat{c}_T \left(\tfrac{i}{T}, \tfrac{j}{T} \right) \hat{c}_T \left(\tfrac{p}{T}, \tfrac{q}{T} \right) - \hat{c}_T \left(\tfrac{i}{T}, \tfrac{q}{T} \right) \hat{c}_T \left(\tfrac{p}{T}, \tfrac{j}{T} \right) \right]$$

$$\hat{\rho} = \left(\frac{12}{T^2 - 1} \right) \sum_{i=1}^{T} \sum_{j=1}^{T} \left[\widehat{\mathbb{C}}_T \left(\tfrac{i}{T}, \tfrac{j}{T} \right) - \tfrac{i}{T} \tfrac{j}{T} \right]$$

Pour la démonstration de ce théorème, voir Nelsen ([25]) et Deheuvels ([7]).

3.2.2 Procédure de GENEST et RIVEST

Christian Genest et Louis-Paul Rivest (1993) (voir [18]) ont proposé une procédure d'estimation pour sélectionner une copule appartenant à la famille Archimédienne. Leurs méthodologie est basée sur le générateur de la copule.

Soit X un vecteur de d v.a, $\mathbb{C}$ est la copule associée au générateur φ et K la fonction définie par

$$K(u) = \mathbb{P}(\mathbb{C}(u, v) \leq u)$$

La fonction K de la copule Archimédienne est

$$K(u) = u - \frac{\varphi(u)}{\varphi'(u)}$$

Un estimateur non paramétrique de K est alors donné par

$$\hat{K}(u) = \frac{1}{T} \sum_{t=1}^{T} \mathbb{I}_{\vartheta_t \leq u}$$

où

$$\vartheta_t = \frac{1}{T-1} \sum_{t=1}^{T} \mathbb{I}_{\left\{x_1^j \leq x_1^t, x_2^j \leq x_2^t\right\}}$$

Remarque 3.2.1 *La fonction K est liée au tau de Kendall du fait que*

$$\tau = 4\int_0^1 (1 - K(u))\, du - 1$$

Pour plus détails sur cette méthode voir (Denuit (2005) *(voir [4]) et Embrechts* (1997) *(voir [10]).*

3.2.3 Estimation à noyau

Des estimateurs de la copule peuvent être obtenus en employant une approche basée sur les noyaux. On considère un vecteur aléatoire X de dimention d de fonction de répartition $\mathbb{F}$, d'une copule $\mathbb{C}$ et de fonctions marginales $\mathbb{F}_i$, $i = 1, \ldots, d$ telle que le théorème de Sklar est vérifié. Soit l'échantillon $\{x_1^t, ..., x_d^t\}_{t=1}^T$. Avant de construire notre estimateur on va introduire les noyaux, i.e., des fonctions intégrables $k_i : \mathbb{R} \longrightarrow \mathbb{R}$ tels que

$$\int_{\mathbb{R}} k_i(x)\, dx = 1, \qquad i = 1, \ldots, d$$

Soit le noyau de dimention d

$$k(x) = \prod_{i=1}^{d} k_i(x_i)$$

et sa fonction de répartition

$$K(x) = \prod_{i=1}^{d} \int_{-\infty}^{x} k_i(x_i) = \prod_{i=1}^{d} K_i(x_i)$$

Pour simplifier, on va travailler avec les produits des noyaux univariés et par prolongement on obtient un résultat aux noyaux plus généraux, i.e.,

$$k(x, h) = \prod_{i=1}^{d} k_i\left(\tfrac{x_i}{h_i}\right) \text{ et } K(x, h) = \prod_{i=1}^{d} K_i\left(\tfrac{x_i}{h_i}\right)$$

où h est une matrice diagonale, avec des éléments h_i, $i = 1, \ldots, d$ et un déterminant $|h|$ dit la fenêtre de lissage.

Les fenêtres de lissage individuelles h_i sont des fonctions de T telle que

$$h_i \to 0 \text{ quand } T \to \infty$$

Les éstimateurs à noyau des densités marginales des X_i^t au point x_i sont données par

$$\hat{f}_i(x_i) = \frac{1}{T h_i} \sum_{t=1}^{T} k_i\left(\tfrac{x_i - X_i^t}{h_i}\right)$$

et la fonction de densité conjointe de X^t au point $x = (x1, ..., x_d)'$ est donnée par

$$\hat{f}(x) = \frac{1}{T|h|}\sum_{t=1}^{T}\prod_{i=1}^{d} k(x - X^t, h_i)$$

Par conséquent un éstimateur de la fonction de répartition conjointe de X_i^t est donnée par

$$\hat{\mathbb{F}}_i(x_i) = \int_{-\infty}^{x} \hat{f}_i(y)\, dy$$

et un estimateur de fonction de rérartition de X^t au point $x = (x_1, ..., x_d)'$ est donnée par

$$\hat{\mathbb{F}}(x) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_d} \hat{f}(y)\, dy$$

Si on définit le noyau Gaussien

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$$

Les estimateurs à noyau de $\mathbb{F}_i(x)$ et $\mathbb{F}(x)$ est donnés par

$$\hat{\mathbb{F}}_i(x_i) = \frac{1}{T}\sum_{t=1}^{T} \Phi\left(\frac{x_i - X_i^t}{h_i}\right)$$

et

$$\hat{\mathbb{F}}(x) = \frac{1}{T}\sum_{t=1}^{T}\prod_{i=1}^{d} \Phi\left(\frac{x_i - X_i^t}{h_i}\right)$$

où

$$\Phi(x) = \int_{-\infty}^{x} \varphi(t)\, dt$$

Dans la pratique, on choisit d'habitude $h_i = \hat{\sigma}_i \left(\frac{4}{3T}\right)^{1/5}$ où σ_i est l'écart-type de l'échantillon $\{x_1^t, ..., x_d^t\}_{t=1}^T$. Définissons $\hat{q}$ le vecteur où le $i^{ème}$ composante est le u_i-quantile de $\hat{\mathbb{F}}_i$,

$$\hat{q}_i = \inf_{x \in \mathbb{R}} \{x : \mathbb{F}_i \geq u_i\} \qquad u_i \in i$$

L'estimateur à noyau de la copule $\mathbb{C}$ est donné par

$$\widehat{\mathbb{C}}(u_1, \dots, u_d) = \hat{\mathbb{F}}(\hat{q})$$

Remarque 3.2.2 *Sous les conditions de régularité, l'estimateur à noyau est asymptotiquement Gaussien, et on a*

$$\left(T\prod_{i=1}^{d} h_i\right)^{1/2} \left(\widehat{\mathbb{C}}(u) - \mathbb{C}(u)\right) \xrightarrow{d} \mathcal{N}(0, \mathbb{C}(u))$$

Aprés la construction l'estimateur à noyau, on peut le différencier par rapport à u_i. Il est facile d'obtenir un estimateur de la dérivée partielle de la copule par rapport à une (ou plus) variable. Par exemple l'estimateur à noyau de la dérivée partielle du premier ordre de la copule $\mathbb{C}$ par rapport à u_i est donné par

$$\widehat{\frac{\partial \mathbb{C}(u)}{\partial u_i}} = \frac{\partial \widehat{\mathbb{C}}(u)}{\partial u_i} = \frac{1}{\hat{f}_i(\hat{q}_i(u_i))} \frac{\partial \hat{\mathbb{F}}(\hat{q}_i(u_i))}{\partial \hat{q}_i(u_i)} \quad \text{avec } u = (u_1, \ldots, u_d)'$$

où $\hat{f}_i$ est l'estimateur à noyau de la densité marginale de X_i

$$\hat{f}_i(x_i) = \frac{1}{Th_i} \sum_{t=1}^{T} \varphi\left(\frac{x_i - x_i^t}{h_i}\right)$$

et $\partial_i \hat{\mathbb{F}}$ est la dérivée partielle de $\hat{\mathbb{F}}$ par rapport à la $i^{ème}$ variable.

De la même mannière, sous les conditions de régularité, on peut montrer que

$$\left(T \prod_{i=1}^{d} h_i\right)^{1/2} \left(\frac{\partial \widehat{\mathbb{C}}(u)}{\partial u_i} - \frac{\partial \mathbb{C}(u)}{\partial u_i}\right) \xrightarrow{d} \mathcal{N}\left(0, \frac{1}{f_i(q_i(u_i))} \frac{\partial \mathbb{C}(u)}{\partial u_i}\right)$$

En appliquant le même type d'argument, on peut estimer les dérivées partielle d'ordre supérieur de la copule C, on obtient

$$\widehat{\frac{\partial^{ik} \mathbb{C}(u)}{\partial u_{i_1}, \ldots, \partial u_{i_k}}} = \frac{\partial^{ik} \widehat{\mathbb{C}}(u)}{\partial u_{i_1}, \ldots, \partial u_{i_k}} = \frac{\partial_{i_1, \ldots, i_k} \hat{\mathbb{F}}(\hat{q}(u))}{\hat{f}_{i_1}\left(\hat{q}_{i_1}\left(u_{i_1}\right)\right), \ldots, \hat{f}_{i_k}\left(\hat{q}_{i_k}\left(u_{i_k}\right)\right)} \quad \text{avec } u = (u_1, \ldots, u_d)$$

où tout les i_j sont différents et $k \leq d$, si $k = d$, on obtient l'estimateur à noyau de la densité de la copule $\mathbb{C}$ définit par

$$\hat{c}(u) = \frac{\hat{f}(\hat{q}(u))}{\hat{f}_1(\hat{q}_1(u_1)), \ldots, \hat{f}_d(\hat{q}_d(u_d))}$$

où $\hat{f}$ est l'estimateur à noyau de la densité conjointe de X

$$\hat{f}(x) = \frac{1}{T \prod_i h} \sum_{t=1}^{T} \prod_{i=1}^{d} \varphi\left(\frac{x_i - X_i^t}{h_i}\right)$$

3.3 Conclusion générale

Dans ce mémoire, un nouveau concept de modélisation multivarié de la structure de dépendance est utilisé c'est bien la fonction copule. Quelques chercheurs ont intéressés plus particulièrement aux quelques familles de copules pour répondre à la question du choix de la copule adéquate qui permet de modéliser la structure de dépendance entre des variables aléatoires. A la fin, on a présenté les approches les plus utilisées pour l'inférence statistique.

Perspectives

- Choisir une meilleur copule sappuyant, sur des tests statistiques. Par exemple le test du chi-deux utilisé dans le cadre de l'ajustement d'une distribution paramétrique à une distribution empirique constitue à cet égard un outil intéressant pour le choix d'une copule.
- Utiliser le test de Cramer-Von-Mises, test de Kolomogorov-Smirnov pour choisir la meilleur copule.
- Comparer les résultats des deux tests.

Annexe

Dans cette annexe on va rappeler quelques notions de base de probabilité concernant la densité normale multivariée et la densité de student multivariée ainsi la fonction gamma et leurs propriétés

3.3.1 La densité normale multivariée

Définition 3.3.1 *On appelle densité de probabilité suit la loi normale,la fonctiot* $f_X(x)$ *définie sur* $\mathbb{R}$ *par*

$$f_X(x) = \frac{1}{(2\pi)^{d/2}\det(\Sigma)^{1/2}} \exp\left(-\frac{1}{2}\left(X^T\Sigma^{-1}X\right)\right)$$

où $x = (x_1, \ldots, x_d)'$.

3.3.2 La densité de student multivariée

La densité copule de student multivariée est donnée par

$$\mathfrak{c}^t_{\Sigma,\upsilon}(u_1, \ldots, u_d) = \frac{\Gamma\left(\frac{\upsilon+d}{2}\right)\left[\Gamma\left(\frac{\upsilon}{2}\right)\right]^{d-1}\left(1+\frac{W^t\Sigma^{-1}W}{\upsilon}\right)^{-(\upsilon+d)/2}}{\det(\Sigma)^{1/2}\left[\Gamma\left(\frac{\upsilon+1}{2}\right)\right]^d \prod_{i=1}^{d}\left(1+w_i^2/\upsilon\right)^{-(\upsilon+1)/2}}$$

avec $(w_1, \ldots, w_d) = (t_\upsilon^{-1}(u_1), ..., t_\upsilon^{-1}(u_d))$ et $\upsilon > 0$.

Définition 3.3.2 (Loi Gamma) *On dit qu'un v.a suit une loi gamma si sa densité est donnée par*

$$f_X(x) = \begin{cases} \frac{x^{\alpha-1}}{\beta\Gamma(\alpha)}\exp(-x/\beta) & si\ x \geq 0 \\ 0 & sinon \end{cases}$$

$\Gamma(\alpha)$ *est la loi gamma définie par* $\Gamma(\alpha) = \int x^{\alpha-1}e^{-x}dx$.

– **proprietés de la fonction gamma**
– $\Gamma(\alpha+1)=\alpha\Gamma(\alpha)$
– $\Gamma(\alpha+1)=n$
– $\Gamma(1\setminus 2)=(\pi)^{1/2}$

Définition 3.3.3 (Tau (τ) de Kendall) *Soient (X_1,Y_1) et (X_2,Y_2) deux vecteurs aléatoires indépendants et identiquement distribués. Le tau de kendall du vecteur aléatoire (X_1,Y_1) est défini par*

$$\tau(X_1,Y_1)=\mathbb{P}\left\{(X_1-X_2)(Y_1-Y_2)>0\right\}-\mathbb{P}\left\{(X_1-X_2)(Y_1-Y_2)<0\right\} \tag{3.3.1}$$

L'expression de τ donnée dans la définition précédente, représente la valeur théorique du tau de Kendall que l'on peut définir, aussi, d'une manière empirique. Soit un échantillon de taille n de données bivariées $\{(x_k,y_k)\}_{k=1}^{n}$. On définit les nombres c et d comme étant respectivement le nombre de paires concordantes et discordantes dans cet échantillon. On définit la version empirique du tau de Kendall par

$$\tau_n=\frac{c-d}{c+d}=\frac{c-d}{\mathbf{C}_2^n}\ \text{avec}\ \mathbf{C}_2^n=\frac{n!}{2!\,(n-2)!}=\frac{n\,(n-1)}{2} \tag{3.3.2}$$

Cette version empirique peut être exprimée en terme de la densité copule associée

$$\tau_n=\frac{2n}{n-1}\sum_{i=1}^{n}\sum_{j=1}^{n}\sum_{p=1}^{i-1}\sum_{q=1}^{j-1}\left[c_n\left(\tfrac{i}{n},\tfrac{j}{n}\right)c_n\left(\tfrac{p}{n},\tfrac{q}{n}\right)-c_n\left(\tfrac{i}{n},\tfrac{q}{n}\right)c_n\left(\tfrac{p}{n},\tfrac{j}{n}\right)\right]$$

Théorème 3.3.1 *Soit (X,Y) un couple de variables aléatoires continues de copule $\mathbb{C}$. Compte tenu de (3.3.2) et de (3.3.1) , le tau de Kendall pour X et Y est donnée par*

$$\tau(X,Y)=\mathbb{Q}(\mathbb{C},\mathbb{C})=4\int\int_{I^2}\mathbb{C}(u,v)d\mathbb{C}(u,v)-1=4E\left\{\mathbb{C}(U,V)\right\}-1$$

où la paire (U,V) est de loi $\mathbb{C}$.

D'un point de vue numérique, il est difficile de manier cette expression car elle fait intervenir la différentielle $d\mathbb{C}(u,v)$. C'est pourquoi, il est facile de la calculer avec l'expression équivalente suivante :

$$\tau=1-4\int\int_{I^2}\partial_{u_1}\mathbb{C}\left(u_1,u_2\right)\partial_{u_2}\mathbb{C}\left(u_1,u_2\right)du_1du_2$$

Corollaire 3.3.1 *Dans le cas bivarié, si $\mathbb{C}$ est une copule archimédienne de génératrice ϕ alors*

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

Une extension de l'expression du tau de Kendall dans le cas multivarié est possible et elle est donnée par :

$$\tau(\mathbb{C}) = \frac{1}{2^{d-1}-1}\left(2^d \int \mathbb{C}(u_1, u_2, ..., u_d)\, d\mathbb{C}(u_1, u_2, ..., u_d) - 1\right)$$

Tout comme le tau de Kendall, le rho de Spearman est basé sur la concordance et la discordance de couples de variables aléatoires.

Définition 3.3.4 (rho (ρ_S) de Spearman) *Soient $(X_1, Y_1), (X_2, Y_2)$ et (X_3, Y_3) des copies indépendantes de vecteurs aléatoires. Le rho de Spearman, noté ρ_S, est défini par*

$$\rho_S(X, Y) = 3\left(\mathbb{P}\{(X_1 - X_2)(Y_1 - Y_3) > 0\} - \mathbb{P}\{(X_1 - X_2)(Y_1 - Y_3) < 0\}\right)$$

Théorème 3.3.2 *([25])Soit $(X, Y)'$ un vecteur de variables aléatoires continues associé à la copule $\mathbb{C}$. Alors le rho de Spearman de $(X, Y)'$ est donné par :*

$$\rho_S(X, Y) = 3\mathbb{Q}(\mathbb{C}, \Pi) = 12uvd\mathbb{C}(u, v) - 3 = 12\int_{[0,1]^2} \mathbb{C}(u, v)\, dudv - 3$$

1 On sait que pour toute copule $\mathbb{C}$, on a $\mathbb{W} \leq \mathbb{C} \leq \mathbb{M}$, le résultat précédent assure que $-1/3 = \mathbb{Q}(\mathbb{W}, \Pi) \leq \mathbb{Q}(\mathbb{C}, \Pi) \leq \mathbb{Q}(\mathbb{M}, \Pi) = 1/3$, ce qui justifie la multiplication de $\mathbb{Q}$ par le facteur 3 permettant que $\rho_S \in [-1, 1]$.

2 Si $X \sim \mathbb{F}$ et $Y \sim \mathbb{G}$, soit $U = \mathbb{F}(X)$ et $V = \mathbb{G}(Y)$, on a

$$\begin{aligned}\rho_S(X, Y) &= 12\int_{[0,1]^2} uvd\mathbb{C}(u, v) - 3 = 12E(UV) - 3 \\ &= \frac{E(UV)-1/4}{1/12} = \frac{Cov(U,V)}{\sqrt{Var(U)}\sqrt{Var(V)}} = \rho(\mathbb{F}(X), \mathbb{G}(Y))\end{aligned}$$

Donc le rho de Spearman pour une paire (X, Y) est equivalent au coefficient de corrélation usuel de Pearson de la paire (U, V).

3 On peut considérer que ρ_S est une mesure de distance moyenne entre la copule $\mathbb{C}$ et la copule d'indépendance car

$$\rho_S(\mathbb{C}) = 12\int_{[0,1]^2} \{\mathbb{C}(u, v) - uv\}\, dudv$$

4 Le tau de Kendall et le rho de Spearman ne dépendent pas du comportement des lois marginales, mais uniquement de la structure de dépendance.

5 Pour toutes fonctions croissantes f et g, on a $\tau(f(X); g(Y)) = \tau(X, Y)$ et $\rho(f(X), g(Y)) = \rho(X, Y)$.

6 Soit un échantillon de taille n de données bivariées $\{(x_k, y_k)\}_{k=1}^{n}$. On définit la version empirique du rho de Spearman en terme de densité de la copule associée par

$$\rho_n = \frac{12}{n^2 - 1} \sum_{i=1}^{n} \sum_{j=1}^{n} \left[c_n \left(\tfrac{i}{n}, \tfrac{j}{n} \right) - \tfrac{i}{n} \tfrac{j}{n} \right]$$

Les résultats obtenus concernant le rho de Spearman dans le cas bidimensionnel peuvent être étendus dans le cas multidimensionnel de la manière suivante :

$$\rho(\mathbb{C}) = \frac{1}{(d+1)^{-1} - 2^{-d}} \int_{[0,1]^d} u_1 \ldots . u_d d\mathbb{C}(u_1, ..., u_d) - 2^{-d}$$

Généralement, les valeurs τ et ρ pour une copule $\mathbb{C}$ sont différentes . En revanche, certaines relations existent entre ces deux mesures de concordance.

Bibliographie

[1] Brahimi, B., Necir, A., 2012. A semiparametric estimation of copula models based on the method of moments. Statistical Methodology 9 (4), 467-477.

[2] Cadoux D, J-M Loiseau, Copules et dépendance : application pratique à la détermination de besoin en fonds propres d'un assureur non vie.

[3] Capéraà, P., A-L. Foudères et C. Genest (2000), Bivariate distributions with given extreme value attractor, Journal of Multivariate Analysis, 72, $30-49$.

[4] Denuit, M. Dhaene, J. Goovaerts, M. & Kaas, R. Actuarial Theory for Dependent Risks Measures, Wiley , England,(2005).

[5] Durante F. & Sempi C., Copula Theory : An Introduction. , Dans : Jaworski, P., Durante, F., Haerdle, W., Rychlik, T. :(eds). Copula Theory and Its Applications, Proceedings of theWorkshop,Warsaw 25-26 sept 2009, Springer, Dordrecht (2010).

[6] Deheuvels, P. (1978). Caractérisation complète des lois extrêmes multivariées et de la convergence des types extrêmes. Publ.Inst.Statist. Univ.Paris, $1-2(2), 1-37$.

[7] Deheuvels, P. (1979a). Détermination complète du comportement asymptotique en loi des valeurs extrêmes multivariées d'un échantillon de vecteurs aléatoires indépendants. C. R. Acad. Sci. Paris S er. A-B, 288(3),A217−A220.

[8] Deheuvels, P. (1979b). La fonction de dépendance empirique et ses propriétés. Un test non paramétrique d'indépendance. Acad. Roy. Belg. Bull. Cl. Sci. $(5), 65(6), 274-292$.

[9] Embrechts, P., Hoing, A., et Juri, A. (2003). Using copulae to bound the value-atrisk for functions of dependent risks. Finance and Stochastics, $7(2), 145-167$.

[10] Embrechts, P., Kluppelberg, C. et T. Mikosch (1997), Modelling Extremal Events for Insurance and Finance, Springer-Verlag, Berlin.

[11] FANG K., KOTZ S. & NG W. (1990) Symmetric multivariate and related distributions, Chapman & Hall.

[12] Frees, J.-D. et Valdez, E. (1998). Understanding relationschips using copulas. North. Amer. Actu. Journ., $(2), 1 - 25$.

[13] Fermanian, J.D. and O. Scaillet, Nonparametric estimation of copulas for time series, Journal of Risk 5, 25-54, (2002).

[14] Fisher, N. I., 1997. Copulas. Encyclopedia of Statistical Sciences 1, 159-163.

[15] Genest, C. (1987), frank's family of bivariate distributions, Biometrika, 74.

[16] Genest, C. et MacKay, R. J. (1986). Copules archimédiennes et familles de lois bidimensionnelles dont les marges sont données. Canad. J. Statist., $14, 145 - 159$.

[17] Genest, C. et K. Ghoudi (1994), Une famille de lois bidimentionnelles insolite, Comptes Rendus de l'Académie des sciences de Paris, 318, série I, $351 - 354$.

[18] Genest, C. et L-P. Rivest (1993), Statistical inference procedures for bivariate Archimedean copulas, *Journal of the American Statistical Association*, $88, 1034 - 1043$.

[19] Gudendorf, G. and J. Segers Extreme-Value Copulas. Dans : Jaworski, P., Durante, F., Haerdle, W., Rychlik, T. :(eds).Copula Theory and Its Applications, Proceedings of theWorkshop,Warsaw 25-26 sept 2009, Springer, Dordrecht (2010)

[20] Ghoudi, Abdous et Khoudraji. (1999)Non-Parametric Estimation of the Limit Dependence Function of Multivariate Extremes

[21] Joe, H. (1997), Multivariate Models and Dependence Concepts, *Monographs on Statistics and Applied Probability*, 73, Chapmann & Hall, London.

[22] Joe, H. et Xu, J. (1996). The estimation method of inference functions for margins for multivariate models, departement of statistics,university of britich columbia. Technical Report, (166).

[23] L. Fadhila, Modélisation de la dépendance par les copules et applications, (2011) Université de Mouloude Mammeri,Tizi-Ouzou.12.

[24] Nelsen, R. B. An Introduction to Copulas, 2nd ed. New York : Springer (2006).

[25] Nelsen, R. B. (1999). An introduction to copulas, volume 139 of Lecture Notes in Statistics. Springer-Verlag, New York.

[26] Resnick, S.I. (1987), Extreme Values, Point Processes and Regular Variation, Springer-Verlag, New York.

[27] Ruymgaart, F. H. (1973). Asymptotic theory of rank tests for independence. Mathematisch Centrum, Amsterdam. Mathematical Centre Tracts, 43.

[28] Shih, J. et Louis, T. (1995). Inferences on the association parameter in copula models for biariate survival data. Biometrics, $51, 1384 - 1399$.

[29] Tawn, J.A. (1988), Bivariate extreme value theory : models and estimation, Biometrika, $75, 397 - 415$.

[30] Sklar, A. (1959). Fonctions de répartition a n dimensions et leurs marges. Publ. Inst. Statist. Univ. Paris, $8, 229 - 231$.

[31] Schweizer, B. et Wolff, E. F. (1981). On nonparametric measures of dependence for random variables. *Ann. Statist.*, $9(4), 879 - 885$.

[32] Yaari, M. (1987), The dual theory of choice under risk. Econometrica, 55, $95 - 115$.