



Faculté des Sciences Exacte et Informatique
Département de Mathématique

N° d'ordre :

N° de séries :

Mémoire de fin d'études

Présenté pour l'obtention du diplôme de

Master

Spécialité : Mathématiques.

Option : Probabilités et statistique .

Thème

Les copules dans \mathbb{R}^2

Présenté par :

- Tebboub Samira.
- Hamdellou Razika.

Devant le jury :

Président	: Chraitia Hassen	M.C.B Université de Jijel
Encadreur	: Gherda Mebrouk	M.A.A Université de Jijel
Examineur	: Yaakoubi Fatma	M.A.A Université de Jijel
Examineur	: Stihi Sara	M.C.B Université de Jijel

Promotion **2017/2018**

Dédicace

A nos chères parents,

A nos frères et A nos sœurs,

A nos amis,

Remerciements

En préambule à ce mémoire nous remerciant ALLAH qui nous aide et nous donne la patience et le courage durant ces longues années d'étude. Nous souhaitant adresser nos remerciements les plus sincères aux personnes qui nous ont apporté leur aide et qui ont contribué à l'élaboration de ce mémoire ainsi qu'à la réussite de cette formidable année universitaire. Ces remerciements vont tout d'abord au corps professoral et administratif de la Faculté des Sciences Exacte et Informatique, Département de Mathématiques, pour la richesse et la qualité de leur enseignement et qui déploient de grands efforts pour assurer à leurs étudiants une formation actualisée.

La première personne que nous tenons à remercier est notre encadreur **Mr.Mebrouk Gherda**, pour l'orientation, la confiance, la patience qui ont constitué un apport considérable sans lequel ce travail n'aurait pas pu être mené au bon port. Qu'il trouve dans ce travail un hommage vivant à sa haute personnalité.

nous exprimons par nos respectueux remerciement aux membres des jury d' avoir accepter de jury ce travail :

Mr.Chraitia Hassen pour avoir accepté de présider le jury de soutenance ainsi que *M^{me} Yaakoubi Fatima* et *M^{me} Setihi Sara* pour voir bien voulu juger ce travail en prenant la charge d'examinatrice.

Nous remercions toute le staff de le direction du Barage de Beni Haroune et en particulier son directeur **Mr.Dahwi Kamel** qui nous a accueilli avec une bienveillance qui n'a menager ni sur temps ni ses efforts pour nous aider.

Avant de continuer, nous offrons nos plus grands remerciements, gratitude, appréciation et amour à ceux qui portent le message le plus sacré de la vie...

A ceux qui nous ont ouvert la voie de la science et de la connaissance...

À tous nos professeurs distingués ...

"Soyez un érudit.. Si vous ne pouvez pas être éduqués, si vous ne pouvez pas aimer les scientifique, si vous ne pouvez pas ne pas les haïr."

On n'oublie pas nos parents pour leur contribution, leur soutien et leur patience.

En fin, nous adressons nos plus sincères remerciements à tout nos proches et amis, qui nous ont toujours encouragé au cours de la réalisation de ce mémoire.

Merci à tous et à toutes.

Table des matières

Dédicace	2
Remerciements	3
listoftables	6
listoffigures	7
introduction	8
Résumé	11
1 Notion de copule et ses propriétés	13
1.1 Les mesures de dépendance	13
1.1.1 Coefficient de corrélation linéaire	13
1.1.2 Le coefficient de corrélation de Kendall et le rho de Spearman . . .	14
1.2 la copule	15
1.2.1 Définition de copule	16
1.2.2 Propriétés des copules	17
1.2.3 Théorème de Sklar	18
1.2.4 Relation d'ordre sur les copules	21
1.2.5 Copule et variables aléatoires	21

1.2.6	Copules associées à une copule	24
2	Les copules paramétriques	27
2.1	La copule archimédienne	27
2.1.1	Propriétés d'une copule archimédienne	29
2.1.2	Exemples de copules archimédiennes	31
2.2	La famille de Marshall-Olkin	35
2.3	La copule Elliptique	37
2.3.1	Copule Gaussienne	39
2.3.2	Copule Student	40
2.4	La copule de valeurs extrêmes	42
2.4.1	Les copules archimax	44
2.5	Estimation de copule	45
2.5.1	Fonction de répartition empirique	46
2.5.2	Copule empirique	46
2.5.3	Méthode d'estimation	47
3	Utilisation des copules	51
3.1	Copules et modélisation	51
3.1.1	Prospection graphique de la dépendance	51
3.1.2	Mesures de dépendance	53
3.1.3	Sélection et adéquation de la copule	54
3.2	Application numérique	56
	Conclusion	60
	Annexes	61

Liste des tableaux

3.1	Les mesure de dépendance	57
-----	------------------------------------	----

Table des figures

2.1	Densité de la copule de Gumbel $\theta = 1.5$	31
2.2	Densité de la copule de Clayton ($\theta = 2.582$)	33
2.3	Densité de la copule de Frank	34
2.4	Densité de la copule Gaussienne avec $\rho = 0.5$	40
2.5	Densité de la copule de student avec $\rho = 0.5$ et $\nu = 1$	41
3.1	cote en fonction de température	57
3.2	Distribution de température	58
3.3	Distribution de cote	59

Introduction

Le terme copule vient du mot latin "copulae", qui signifie au sens figuré, lien, alliance, liaison ou union. Cette notion apparait sous d'autres appellations dans certains travaux de Fréchet [1951], Féron [1956] et Dall'Aglio [1956] portant sur l'étude des tables de contingence et lois multidimensionnelles à structures marginales fixées. Mais c'était en [1959] que le terme copule a été utilisé pour la première fois grâce à Sklar [1959] dans la théorie des lois multidimensionnelles.

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

L'introduction des copules à des fins statistiques est un phénomène relativement récent qui trouve sa source à la fin des années 50 dans des recherches portant initialement sur les tables de contingence. Au début, il était difficile de trouver des traces de la notion copule dans la littérature statistique. Mais à partir des années 70 et avec le développement de la théorie des processus empiriques, d'autres auteurs ont redécouvert le concept des fonctions copules sous d'autres appellations en établissant, de manière originale certaines de leurs propriétés. C'est ainsi qu'elles sont nommées "représentation uniforme" par Ki-meldorf et Sampson [1975], "fonction de dépendance" par Galambos [1978] et Deheuvels [1979], ou encore "la forme standard" par Cook et Johnson [1981]. D'une façon explicite, une copule met en évidence la relation de dépendance entre deux ou plusieurs variables aléatoires par la relation :

$$F(x_1, x_2) = C\left(F(x_1), F(x_2)\right), \quad x_1, x_2 \in \mathbb{R}$$

Dans cette expression, F est la fonction de répartition conjointe des variables aléatoires X_1 et X_2 . On peut changer les fonctions de répartition marginales F_1 et F_2 , à volonté tout en conservant la structure de dépendance (la copule C) intacte. Les fonctions de répartition marginales F_1 et F_2 pourraient être des lois différentes. Ceci est un premier atout pour les statisticiens, puisque les copules leur autorisent une sélection plus étendue des fonctions de répartitions conjointes et ce indépendamment des différentes lois marginales considérées. Grâce à cette flexibilité, la théorie des copules trouve des applications dans de nombreux domaines comme l'hydrologie (Salvadori, De Michele, Kottegoda et Rosso [2007]), les sciences actuarielles (Frees et Valdes [1998]) et en finance (Cherubini, Vecchiato et Lucciano [2004], Mc-Neil, Frey et Embrechts [2005]).

Pourquoi la copule et non le coefficient de corrélation ?

Modéliser la dépendance par des indicateurs statistiques est une chose, la modéliser à l'aide d'une fonction de dépendance en est une autre. Ce qui est le cas de la fonction copule qui répond à cet objectif. Cela explique le choix de la copule pour modéliser la dépendance et non le coefficient de corrélation, une mesure souvent utilisée par les praticiens car cette dernière possède plusieurs limites telles que :

- Le coefficient de corrélation n'est pas défini si les moments d'ordre deux des variables aléatoires ne sont pas finis. Ce n'est pas une mesure appropriée de la dépendance pour les distributions à queues lourdes où les variances peuvent être infinies.
- Le coefficient de corrélation peut être le même alors que la structure de dépendance est totalement différente notamment pour les valeurs extrêmes.
- Le coefficient de corrélation et sa portée de caractérisation ne fonctionnent que pour les variables gaussiennes, pour lesquelles corrélation et dépendance recouvrent la même réalité.
- La corrélation est une mesure de dépendance qui ne peut pas nous indiquer tout ce que nous voudrions savoir sur la structure de la dépendance.

Ainsi, à chaque fois qu'il est nécessaire de modéliser une structure de dépendance entre plusieurs variables aléatoires, nous pouvons faire appel à la fonction copule qui possède les avantages suivants :

- La copule est une fonction qui modélise les propriétés des structures de dépendance telles que la forme et l'intensité.
- La copule est capable de mesurer la dépendance pour les distributions à queues lourdes.
- Elles permettent de construire des modèles de distributions multidimensionnelles.
- Les copules permettent d'élaborer des modèles non gaussiens, c'est à dire qu'avec une

copule, on peut construire par exemple une distribution avec une marginale gaussienne et une autre uniforme.

Ce mémoire se divise en trois chapitres.

Dans le premier chapitre, nous présentons les outils de base liés à la théorie des copules. Nous fournissons un exposé de synthèse sur ce sujet, en présentant la définition et les propriétés fondamentales des copules.

L'objet principale du second chapitre est d'étudier quelques familles des copules paramétriques les plus utilisées et leurs propriétés fondamentales une attention particulière est accordée à la copule archimédienne, à la Marshall-Olkin, à la elliptique et à la copule des valeurs extrêmes, et le présenter quelque méthodes d'estimation des copules telles que :

- La méthodes du maximum de vraisemblance
- La méthode d'inférence sur les marginales
- La méthodes des moments

Au troisième chapitre utilisation des copules, nous donnons une application des copules en hydrologique, on calcule une mesure de dépendance et les copules très utilisées. L'implémentation se fera à l'aide du logiciel de statistique $\ll R \gg$

Résumé

Les copules représentent un outil innovant pour modéliser la structure de dépendance de plusieurs variables aléatoires. Introduit par Sklar [1959] pour résoudre un problème de probabilité énoncé par Maurice Fréchet, les copules deviennent essentielles à l'appréhension de nombreux domaines d'application tels que l'hydrologie (Salvadori, De Michele, Kotegoda, et Rosso [2007]), les sciences actuarielles (Frees et Valdez [1998]), ou la finance (Cherubini, Vecchiato, et Luciano [2004], Mc-Neil, Frey, et Embrechts [2005]). Le grand intérêt est qu'elles fournissent des expressions relativement simples des structures des dépendances liant les marges d'une loi multidimensionnelle. Plus précisément, dans le cas bidimensionnel, une copule C définie sur $[0, 1]^2$, associée à une distribution F de marges uniformes F_1 et F_2 , permet de représenter la fonction de répartition jointe $F(x_1, x_2)$ en fonction de ces marginales $F_1(x_1)$ et $F_2(x_2)$ par la relation :

$$F(x_1, x_2) = C\left(F_1(x_1), F_2(x_2)\right)$$

Cependant en pratique, la copule est inconnue, d'où l'utilité de l'estimer. Dans ce mémoire nous commençons par les définitions et les propriétés liées aux copules ainsi que les modèles paramétriques des copules et quelque méthode d'estimation paramétrique, puis expliquer les méthodes de la modélisation de copule.

Abstract

The copulas represent an innovation tool for modeling the dependence structure of several random variables. Introduced by Sklar [1959] to solve a problem of likelihood stated by Maurice Fréchet, copulas become essential to apprehension many fields of application such as hydrology (Salvadori, De Michele, Kottegoda, and Rosso [2007]), actuarial sciences (Frees and Valdez [1998]), or finance (Cherubini, Vecchiato, and Luciano [2004], Mc-Neil, Frey, and Embrechts [2005]). The great interest is that they provide relatively simple expressions of the structures dependencies linking the margins of a multidimensional law. More precisely, for the two-dimensional case, a copula C defined on $[0, 1]^2$, associated with a distribution F uniform margins F_1 and F_2 , to represent the distribution function attached $F(x_1, x_2)$ based on these marginal $F_1(x_1)$ and $F_2(x_2)$ by the relation :

$$F(x_1, x_2) = C\left(F_1(x_1), F_2(x_2)\right)$$

However, in practice, the copula is unknown, hence the utility of estimating it. In this memoir we start with the definitions and properties related to the copula as well as the parametric models of the copula and some parametric estimation method, then explain the methods of the copula modeling.

Chapitre 1

Notion de copule et ses propriétés

Introduction

Dans ce premier chapitre, nous introduisons quelques définitions de base qui nous permettront de présenter clairement le concept de copule, nous présentons les propriétés les plus importantes de la fonction copule et son rôle dans l'étude de la dépendance des variables aléatoires. Pour faciliter la tâche nous nous intéressons plus particulièrement à l'étude des copules bivariantes et on donnera une extension au cas multivarié.

1.1 Les mesures de dépendance

1.1.1 Coefficient de corrélation linéaire

Définition 1.1.1 *Soient X et Y deux variables aléatoires ayant des variances finies, le coefficient de corrélation linéaire des variables X et Y est donné par*

$$\rho(X, Y) = \frac{\text{cov}(X, Y)}{\sqrt{\text{var}(X)\text{var}(Y)}}$$

1.1.2 Le coefficient de corrélation de Kendall et le rho de Spearman

a- Notion de la concordance

Définition 1.1.2 Notons (x_1, y_1) et (x_2, y_2) deux observations d'un vecteur aléatoire continue (X, Y) , alors (x_1, y_1) et (x_2, y_2) sont dites

1 - Concordantes si

$$((x_1 - x_2))(y_1 - y_2) > 0 \Leftrightarrow (x_1 < x_2 \text{ et } y_1 < y_2) \text{ ou } (x_1 > x_2 \text{ et } y_1 > y_2)$$

. 2 - Discordantes si

$$(x_1 - x_2)(y_1 - y_2) < 0 \Leftrightarrow (x_1 < x_2 \text{ et } y_1 > y_2) \text{ ou } (x_1 > x_2 \text{ et } y_1 < y_2).$$

Généralement, soit $\{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$ un échantillon de n observations d'un couple (X, Y) . Il existe $C_n^2 = \frac{n!}{2!(n-2)!}$ paires de distributions de couples $(x_i, y_i), (x_j, y_j)$ qui sont, soit concordantes, soit discordantes.

Soit (X_1, Y_1) et (X_2, Y_2) deux couples de variables aléatoires continues, de lois conjointes H_1 et H_2 mais de marges communes F et G (ie F pour X_1 et X_2) et (G pour Y_1 et Y_2).

Définition 1.1.3 On définit la fonction de concordance entre ces deux couples (X_1, Y_1) et (X_2, Y_2) par

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

b- Le taux kendall

Définition 1.1.4 Soit (X_1, Y_1) et (X_2, Y_2) deux vecteurs aléatoires continus (iid) de fonction de répartition conjointe H , de fonctions marginales F pour $(X_1 \text{ et } X_2)$ et G pour $(Y_1 \text{ et } Y_2)$, le taux de Kendall noté τ est défini par

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

Une estimation du taux de Kendall se construit à partir d'un échantillon $\{(x_1, y_1), \dots, (x_n, y_n)\}$ de (X, Y) de la façon suivante

$$\hat{\tau} = \frac{2}{n(n-1)} \sum_{j=2}^n \sum_{i=1}^{j-1} \text{signe}(x_j - x_i)(y_i - y_j)$$

où

$$\text{signe}(z) = \begin{cases} 1 & \text{si } z \geq 0 \\ -1 & \text{si } z < 0 \end{cases}$$

c- Rho de Spearman

Soient (\tilde{X}, \tilde{Y}) et (X', Y') deux vecteurs aléatoires copies du vecteur aléatoire (X, Y) alors, le rho de Spearman est donné par

$$\rho_s(X, Y) = 3 \left[\mathbb{P}\{(X - \tilde{X})(Y - Y') > 0\} - \mathbb{P}\{(X - \tilde{X})(Y - Y') < 0\} \right]$$

le rho de Spearman peut s'écrire en fonction du coefficient de corrélation de Pearson $\rho_s(X, Y) = \rho(F_X(x), F_Y(y))$ où F_X et F_Y sont les fonctions de répartition de X et Y .

Un estimateur de rho de Spearman se construit à partir d'un échantillon $\{(x_1, y_1), \dots, (x_n, y_n)\}$ de (X, Y) de la façon suivante :

$$\hat{\rho}(X, Y) = \frac{\sum_{i=1}^n (R_i - \bar{R})(S_i - \bar{S})}{\sqrt{\sum_{i=1}^n (R_i - \bar{R})^2} \sqrt{\sum_{i=1}^n (S_i - \bar{S})^2}}$$

avec

R_i : est un statistique de rang

S_i : est un statistique de signe

1.2 la copule

Introduction aux copules

Les copules permettent d'étudier la dépendance entre plusieurs variables aléatoires, avec l'idée que cette dépendance ne doit pas contenir d'information provenant des lois marginales des variables elles-mêmes.

1.2.1 Définition de copule

Définition 1.2.1 On appelle copule bivarée toute fonction C définie de $I^2 \rightarrow I$ qui possède les propriétés suivantes, avec $I = [0, 1]$

- 1- $\forall u \in [0, 1], C(u, 0) = C(0, u) = 0$
- 2- $\forall u, v \in [0, 1], C(u, 1) = u$ et $C(1, v) = v$
- 3- C est 2-croissante c'est-à-dire $\forall (u_1, u_2), (v_1, v_2) \in I * I$ avec $u_1 \leq v_1$ et $u_2 \leq v_2$

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

cette définition signifie que la copule est une distribution avec des marginales uniformes .

Soient U_1, U_2 deux variables aléatoires uniformes. Considérons le vecteur aléatoire

$U = (U_1, U_2)$, nous avons

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

la propriété 1) implique que

$$\mathbb{P}[U_1 \leq 0, U_2 \leq u] = \mathbb{P}[U_1 \leq u, U_2 \leq 0] = 0$$

la propriété 2) implique que

$$\mathbb{P}[U_1 \leq 1, U_2 \leq u] = \mathbb{P}[U_1 \leq u, U_2 \leq 1] = u$$

C est une distribution de probabilité, ce qui implique que

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

Exemple 1

$\forall u, v \in I$, la fonction $M(u, v) = \min(u, v)$ définit une copule.

En effet :

- $\forall (u_1, u_2) \in [0, 1]^2, \min(u_1, 0) = \min(0, u_2) = 0 \implies M$ vérifie (1)
- $\forall (u_1, u_2) \in [0, 1]^2, \min(u_1, 1) = u_1, \min(1, u_2) = u_2 \implies M$ vérifie (2)
- $\forall (u_1, u_2) \in [0, 1]^2, (v_1, v_2) \in [0, 1]^2$ avec $u_1 \leq u_2$ et $v_1 \leq v_2$, on a les cas suivantes :

1. Si $u_1 \leq u_2 < v_1 \leq v_2$ on a

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

2. Si $v_1 \leq v_2 < u_1 \leq u_2$ on a :

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

3. Si $u_1 < v_1 < u_2 < v_2$ on a :

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

$\Rightarrow C$ est 2-croissante.

Par conséquent, M est une copule. De même façon, on peut montrer que les fonction $W(u, v) = \max(u + v - 1, 0)$ et $\prod(u, v) = uv$ définissent aussi des copules. M , W et \prod sont appelées des copules usuelles.

1.2.2 Propriétés des copules

Les résultats suivants donnent les propriétés et les théorèmes les plus importants d'une copule bivariable

Théorème 1.2.1 (Continuité)

Soit C une copule bivariable pour tout $u_1, u_2, v_1, v_2 \in I$ avec $u_1 \leq u_2$ et $v_1 \leq v_2$, on a :

$$\left| C(u_2, v_2) - C(u_1, v_1) \right| \leq \left| u_2 - u_1 \right| + \left| v_2 - v_1 \right|$$

Théorème 1.2.2 (Différentiabilité)

Soit C une copule bivariable $\forall u_1, u_2 \in I$

1) les dérivés partielles $\frac{\partial C(u_1, u_2)}{\partial u_j}$ existent p.s et $0 \leq \frac{\partial C(u_1, u_2)}{\partial u_j} \leq 1$ avec $j = 1, 2$

2) les fonctions $u \rightarrow \frac{\partial C(u, v)}{\partial v}$ et $v \rightarrow \frac{\partial C(u, v)}{\partial u}$ sont bien définies et décroissantes sur I p.s

Proposition 1 Soit F une fonction réelle croissante et continue à droite d'inverse généralisé $F^{[-1]}$.

avec $F^{[-1]}(t) = \inf\{x/F(x) \geq t\} = \sup\{x/F(x) \leq t\}$

1. La fonction $F^{[-1]}$ est croissante et continue à gauche. De plus $\forall x \in \mathbb{R}, p \in]0, 1]$

$$F(x) \geq p \iff x \geq F^{[-1]}(p)$$

2. Soit X une variable aléatoire à valeurs dans \mathbb{R} de fonction de répartition F . Si F est continue, alors $F(X)$ est de loi uniforme sur I

1.2.3 Théorème de Sklar

Pour toute fonction de répartition bivariee H , avec F et G pour distribution marginales, il existe une copule C telle que

$$\forall (x, y) \in \mathbb{R} \times \mathbb{R}, H(x, y) = C(F(x), G(y)). \quad (1.1)$$

de plus, si F et G sont continues, alors C est unique, sinon elle est déterminée de manière unique sur $(ImF) \times (ImG)$.

Réciproquement, si C une copule et F, G sont des distributions univariées, alors la fonction H , d'après la définition du copule (1.1) est la distribution conjointe dont les marginales sont F et G .

Preuve :

Du faite F et G sont continues et en utilisant 2 du proposition (1.1), $F(X)$ et $G(Y)$ sont uniformes sur $[0, 1]$. Soit H la fonction de distribution du couple $F(x)$ et $G(Y)$ i.e,

$$\begin{aligned} C(u, v) &= \mathbb{P}[F(X) \leq u, G(Y) \leq v] \\ &= \mathbb{P}[X \leq F^{-1}(u), Y \leq G^{-1}(v)] \\ &= H(F^{-1}(u), G^{-1}(v)) \end{aligned}$$

ceci termine la démonstration première partie.

En utilisant 1 du théorème (1.2.2)

on a

$$\begin{aligned} H(x, y) &= \mathbb{P}[X \leq x, Y \leq y] \\ &= \mathbb{P}[F(X) \leq F(x), G(Y) \leq G(y)] \\ &= C(F(x), G(y)) \end{aligned}$$

cela termine la démonstration de la réciproque du théorème. □

Remarque 1

-Si C est une copule, si F_1 et F_2 sont deux fonctions de répartition alors

$$F(x, y) = C\left(F_1(x), F_2(y)\right)$$

est une fonction de répartition bivariée ayant F_1 et F_2 pour marginales.

-Les copules admettent des densités de probabilités :

Si la densité c associée à la copule C existe alors, elle est définie par

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

-Si la fonction de répartition conjointe H est absolument continue, en utilisant le théorème de Sklar, nous pouvons exprimer la densité d'un vecteur aléatoire (X, Y) en fonction de la densité de sa copule et de ses fonctions de répartition marginales F et G par

$$h(x, y) = C\left(F(x), G(y)\right)f(x)g(y)$$

Exemple 2

$$H(x, y) = \frac{1}{1 + e^{-x} + e^{-y}}$$

définie sur $\mathbb{R} \times \mathbb{R}$ les marginales F et G sont des exponentielles, leurs fonctions inverses sont $F^{-1}(u) = \ln\left(\frac{u}{1-u}\right)$, $G^{-1}(v) = \ln\left(\frac{v}{1-v}\right)$ respectivement, pour $u, v \in I$. On en déduit que la copule correspondante est

$$c(u, v) = \frac{uv}{u + v - uv} \text{ et est appelée la copule logistique de Gumbel}$$

Théorème 1.2.3 (Bornes de Fréchet)

Soit H une fonction de répartition conjointe d'un couple aléatoire (X, Y) de fonctions de répartition marginales F et G .

Pour toute copule bivariée C associée à H et $\forall (u, v) \in I^2$, on a

$$W(u, v) = \max(u + v - 1, 0) \leq C(u, v) \leq \min(u, v) = M(u, v)$$

Preuve :

On a

$$\forall u, v \in I \begin{cases} C(u, v) \leq C(u, 1) = u \\ C(u, v) \leq C(1, v) = v \end{cases} \Rightarrow C(u, v) \leq \min(u, v) = M$$

de la définition (1.2.1) dans 3) et $\forall u, v \in I$, on a

$$\begin{aligned} C(u, v) &\geq C(u, 1) + C(1, v) - C(1, 1) \\ &\Rightarrow C(u, v) \geq (u + v - 1) \end{aligned}$$

or $\forall u, v \in I$, $C(u, v) \geq 0$

$$\Rightarrow C(u, v) \geq \max(u + v - 1, 0) = W(u, v)$$

donc

$$W(u, v) \leq C(u, v) \leq M(u, v)$$

□

Proposition 2

Soit (X, Y) un couple de variables aléatoires continues de copule C , le taux de Kendall est la différence entre la probabilité de concordance et la probabilité de discordance son expression en terme de copule est la suivante

$$\begin{aligned} \tau(X, Y) &= 4 \int_0^1 \int_0^1 C(u, v) dC(u, v) \\ &= 4 \int_0^1 \int_0^1 \partial_u C(u, v) \partial_v C(u, v) dudv \end{aligned}$$

où τ peut s'écrire :

$$\tau(X, Y) = 4E(C(u, v)) - 1$$

avec $u, v \rightsquigarrow \mathbb{U}_{[0,1]}$.

Théorème 1.2.4

Soit (X, Y) un couple de variables aléatoires continues de copule C alors

$$\rho_s(X, Y) = 12 \int_0^1 \int_0^1 uv dC(u, v) - 3$$

$$\rho_s(X, Y) = 12 \int_0^1 \int_0^1 C(u, v) dudv - 3$$

1.2.4 Relation d'ordre sur les copules

Définition 1.2.2

Soient C_1 et C_2 deux copules, on dit que C_1 est plus petite que C_2 ou C_2 est plus grande que C_1 et on note $C_1 \prec C_2$ ou $C_2 \succ C_1$ si C_1 et C_2 sont telle que

$$C_1(u, v) \leq C_2(u, v), \forall u, v \in [0, 1]$$

cette relation est une relation d'ordre partiel

Définition 1.2.3

Une structure de dépendance positive est une fonction copule qui vérifie l'inégalité suivant $C^\perp \prec C \prec C^+$ où C^\perp est la copule produit

$$C^\perp(u, v) = uv \quad \text{et} \quad C^+(u, v) = \min(u, v)$$

cette relation est aussi d'ordre partiel, il existe plusieurs copules qui n'ont pas une structure de dépendance positive ni négative .

Définition 1.2.4

La copule paramétrique $C(u, v, \theta)$ note C_θ est dite positivement ordonnée si $C_{\theta_1} > C_{\theta_2}$ négativement ordonnée si $C_{\theta_1} < C_{\theta_2}$ pour $\theta_1 > \theta_2$

1.2.5 Copule et variables aléatoires

Soient X et Y deux variables de fonction de répartition conjointe H de fonctions répartition marginales F et G

Définition 1.2.5 (monotonie)

Deux variables aléatoires X et Y sont contre monotonie, s'il existe un variable Z telle que $X = f(Z)$, $Y = g(Z)$ avec f une fonction croissante et g une fonction décroissante

Théorème 1.2.5 (indépendance)

Deux variables aléatoires X et Y sont indépendante, si la structure de dépendance est la copule C^\perp noté aussi \prod .

En effet :

– si X et Y sont indépendantes alors

$$\forall x, y \in \mathbb{R}, H(x, y) = F(x)G(y)$$

$$\begin{aligned} \forall u, v \in I, C_{XY}(u, v) &= H\left(F^{-1}(u), G^{-1}(v)\right) \\ &= F\left(F^{-1}(u)\right)G\left(G^{-1}(v)\right) = uv = \prod(u, v) \end{aligned}$$

– si $C_{XY} = \prod$ alors

$$\begin{aligned} \forall x, y \in I, C_{XY}(u, v) &= \prod(u, v) = uv \\ &\implies \forall x, y \in \mathbb{R}, H(x, y) = C_{XY}\left(F(x), G(y)\right) \\ &= \prod\left(F(x), G(y)\right) = F(x)G(y) \end{aligned}$$

Les variables aléatoires X et Y sont donc indépendantes.

Théorème 1.2.6

Soient X et Y deux variables aléatoires continues de copule $C_{X,Y}$. Si α_1 et α_2 sont deux fonctions strictement croissantes sur $Im(x)$ et $Im(y)$ respectivement alors

$$C_{X,Y} = C_{\alpha(X)\alpha(Y)}$$

Preuve :

Notons F et G les fonctions de répartition conjointes de (X, Y) et $(\alpha_1(X), \alpha_2(Y))$
 F_1 et F_2 les distributions marginales de F , G_1 et G_2 les distributions marginales de G

$$\begin{aligned} G_1(x) &= \mathbb{P}\left(\alpha_1(X) \leq x\right) = \mathbb{P}\left(X \leq \alpha_1^{-1}(x)\right) = F_1\left(\alpha_1^{-1}(x)\right) = u \\ G_2(y) &= \mathbb{P}\left(\alpha_2(Y) \leq y\right) = \mathbb{P}\left(Y \leq \alpha_2^{-1}(y)\right) = F_2\left(\alpha_2^{-1}(y)\right) = v \end{aligned}$$

$$G_1^{-1}(u) = \alpha_1\left(F_1^{-1}(u)\right) \quad \text{et} \quad G_2^{-1}(v) = \alpha_2\left(F_2^{-1}(v)\right)$$

$$\begin{aligned} C_{\alpha_1(u)\alpha_2(v)}(u, v) &= C\left(G_1^{-1}(u), G_2^{-1}(v)\right) \\ &= \mathbb{P}\left(\alpha_1(X) \leq G_1^{-1}(u), \alpha_2(Y) \leq G_2^{-1}(v)\right) \\ &= \mathbb{P}\left(X \leq \alpha_1^{-1}(G_1^{-1}(u)), Y \leq \alpha_2^{-1}(G_2^{-1}(v))\right) \\ &= \mathbb{P}\left(X \leq F_1^{-1}(u), Y \leq F_2^{-1}(v)\right) \end{aligned}$$

cela termine la démonstration □

Théorème 1.2.7

Soient X et Y deux variables de lois marginales F et G de copule C , si α_1 et α_2 deux fonctions strictement monotones sur $I(X)$ et $I(Y)$ respectivement

1) si α_1 est strictement croissante et α_2 est strictement décroissante alors

$$C_{\alpha_1(X)\alpha_2(Y)}(u, v) = u - C_{XY}(u, 1 - v)$$

2) si α_1 et α_2 sont strictement décroissantes alors

$$C_{\alpha_1(X)\alpha_2(Y)}(u, v) = u + v - 1 + C_{XY}(1 - u, 1 - v)$$

3) si α_1 est strictement décroissante et α_2 strictement croissante alors

$$C_{\alpha_1(X)\alpha_2(Y)}(u, v) = v - C_{XY}(1 - u, v)$$

Propriété 1.2.1

Soit une copule de la forme

$$H(x, y) = C\left(F(x), G(y)\right), \forall x, y \in \mathbb{R}$$

on a

$$\mathbb{P}\left(Y \leq y \mid X = x\right) = \frac{\partial C(u, v)}{\partial v} \frac{G(y)}{F(x)} \quad \text{et} \quad \mathbb{P}\left(X \leq x \mid Y = y\right) = \frac{\partial C(u, v)}{\partial u} \frac{F(x)}{G(y)}$$

Théorème 1.2.8

Soient X et Y deux variables aléatoires de fonctions de répartition F_1 et F_2 , un paramètre permettant de mesurer la dépendance de queue à droite est

$$\lambda_u = \lim_{t \rightarrow 1^-} \mathbb{P}\left(X_2 > F_2^{-1}(t) \mid X_1 > F_1^{-1}(t)\right).$$

à gauche est

$$\lambda_l = \lim_{t \rightarrow 0^+} \mathbb{P}\left(X_2 \leq F_2^{-1}(t) \setminus X_1 \leq F_1^{-1}(t)\right).$$

Si les limites existent alors

$$\begin{aligned} \lambda_u &= 2 - \lim_{u \rightarrow 1^-} \frac{1 - C(u, v)}{1 - u} \\ \lambda_l &= \lim_{u \rightarrow 0^+} \frac{C(u, v)}{u} \end{aligned}$$

Preuve :

on a

$$\begin{aligned} \lambda_u &= \lim_{u \rightarrow 1^-} \mathbb{P}\left(X_2 > F_2^{-1}(u) \setminus X_1 > F_1^{-1}(u)\right) \\ &= \lim_{u \rightarrow 1^-} \frac{1 - u + 1 - u - 1 + C(u, u)}{1 - u} \\ &= 2 - \lim_{u \rightarrow 1^-} \frac{1 - C(u, u)}{1 - u} \end{aligned}$$

et

$$\begin{aligned} \lambda_l &= \lim_{t \rightarrow 0^+} \mathbb{P}\left(X_2 \leq F_2^{-1}(t) \setminus X_1 \leq F_1^{-1}(t)\right) \\ &= \lim_{u \rightarrow 0^+} \frac{\mathbb{P}\left(F_2(X_2) \leq u, F_1(X_1) \leq u\right)}{\mathbb{P}\left(F_1(X_1) \leq u\right)} \\ &= \frac{C(u, u)}{u} \end{aligned}$$

cela termine la démonstration □

1.2.6 Copules associées à une copule

1. La copule de survie

Définition 1.2.6 Soit $\tilde{C}(u_1, \dots, u_n)$ la fonction définie par :

$$\tilde{C}(u_1, \dots, u_n) = \bar{C}(1 - u_1, \dots, 1 - u_n)$$

où

$$\bar{C}(u_1, \dots, u_n) = \mathbb{P}\left[U_1 > u_1, \dots, U_n > u_n\right]$$

alors $\tilde{C}(u_1, \dots, u_n)$ est appelée copule de survie (survival copula) de la copule C .

Remarque 2

Il ne faut pas confondre la copule de survie \tilde{C} et la distribution de survie \bar{C} associée à C (C est aussi une distribution), On a plutôt :

$$\forall u, v \in I, \quad \bar{C}(u, v) = 1 - u - v + C(u, v)$$

On peut vérifier que \tilde{C} et \bar{C} sont liées par la relation

$$\forall u, v \in I, \quad \tilde{C}(u, v) = \bar{C}(1 - u, 1 - v)$$

ou simplement

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

avec $\bar{u} = 1 - u$ et $\bar{v} = 1 - v$. Par un raisonnement similaire, on peut exprimer les probabilités :

$\mathbb{P}(X > x \text{ ou } Y > y)$, $\mathbb{P}(X \leq x, Y \leq y)$ et $\mathbb{P}(X > x, Y \leq y)$ en terme de copule.

2. La copule duale

elle est définie par :

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

3. La co-copule

elle est définie par :

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

comme $\bar{H}(x, y) = \bar{F}(x) + \bar{G}(y) - 1 + C(1 - \bar{F}(x), 1 - \bar{G}(y))$,

on a en posant

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

on peut vérifier que $C^*(u, v) = 1 - C(1 - u, 1 - v)$, pour tout u, v dans I

donc C^* est dite co-copule de C

4. La copule mixte

elle est définie par :

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

et $C''(u, v) = 1 - v - C(1 - u, 1 - v)$ est dite copule mixte de C , de même on peut établir que

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

Chapitre 2

Les copules paramétriques

Dans ce chapitre, on va présenter quelques copules paramétriques les plus utilisées ainsi que leurs propriétés.

2.1 La copule archimédienne

Les copules archimédiennes sont indiscutablement associées au statisticien canadien Christian Genest. Ce n'est pas lui qui les a inventées, mais il est le premier à avoir adopté une analyse statistique de ces fonctions copules. Et ses nombreuses publications ont largement contribué à les faire connaître. La classe de copules archimédiennes est une classe importante

Définition 2.1.1 Soit φ une fonction continue, strictement décroissante définie de $[0, 1]$ 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 valeurs dans $[0, 1]$ donnée par

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

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

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

si $\varphi(0) = \infty$, alors $\varphi^{[-1]} = \varphi^{-1}$

Lemme 2.1.1

Soit φ une fonction continue strictement décroissante définie de $[0, 1]$ dans $[0, \infty]$ telle que, $\varphi(1) = 0$, et soit $\varphi^{[-1]}$ son inverse généralisé. Soit C une fonction de $[0, 1]^2$ dans $[0, 1]$ définie par

$$C(u, v) = \varphi^{[-1]}(\varphi(u) + \varphi(v)) \quad (2.1)$$

C vérifier les propriétés fondamentales d'une copule

Preuve :

$$C(u, 0) = \varphi^{[-1]}(\varphi(u) + \varphi(0)) = 0, \quad \text{et} \quad \varphi^{[-1]}(\varphi(u) + \varphi(1)) = \varphi^{[-1]}(\varphi(u)) = u.$$

par symétrie ; $C(0, v) = 0$ et $C(1, v) = v$ □

Lemme 2.1.2

Soient φ , $\varphi^{[-1]}$ et C les fonctions satisfaisant les hypothèses du lemme (2.1.1), alors la fonction C est 2-croissante si et seulement si

$$C(u_2, v) - C(u_1, v) \leq u_2 - u_1 \quad \forall u_1 \leq u_2$$

Théorème 2.1.1

Soient φ une fonction continue strictement décroissante de $[0, 1]$ dans $[0, \infty]$ telle que, $\varphi(1) = 0$, et soit $\varphi^{[-1]}$ son inverse généralisée. La fonction C de $[0, 1]^2$ dans $[0, 1]$ définie par (2.1) est une copule si et seulement si φ est convexe

Les copules définie par (2.1) sont appelées les copules archémediennes et la fonction φ est plus connue sous le nom générateur de la copule.

Remarque 3

Si $\varphi(0) = \infty$, on dit que φ est un générateur stricte. Dans ce cas, $\varphi^{[-1]} = \varphi^{-1}$ et $C(u, v) = \varphi^{-1}(\varphi(u) + \varphi(v))$ serait une copule archimédienne stricte.

Une source importante de générateurs de copules archémediennes est l'inverse de la transformée de Laplace des fonctions de répartition. Il est donc simple de construire des copules archimédiennes de dimension n . Rappelons la définition de la transformée de Laplace :

Définition 2.1.2 La transformée de Laplace L_Θ d'une variable aléatoire Θ à valeur dans $[0, \infty]$, avec une fonction de distribution F_Θ est définie par

$$L_\Theta = \exp(-s\Theta) = \int_0^\infty \exp(-st) dF_\Theta(s)$$

2.1.1 Propriétés d'une copule archimédienne

Théorème 2.1.2

Soit C une copule archimédienne et φ son générateur alors

1. C est symétrique
2. C est associative, c'est à dire $C(C(u, v), w) = C(u, C(v, w))$ pour toute $u, v, w \in [0, 1]$

Preuve :

1. est un résultat directe de (2.1.1)
- 2.

$$\begin{aligned} C(C(u, v), w) &= \varphi^{[-1]}(\varphi(C(u, v)) + \varphi(w)) \\ &= \varphi^{[-1]}(\varphi(\varphi^{[-1]}(\varphi(u) + \varphi(v))) + \varphi(w)) \\ &= \varphi^{[-1]}(\varphi(u) + \varphi(v) + \varphi(w)) \\ &= \varphi^{[-1]}(\varphi(u) + \varphi(\varphi^{[-1]}(\varphi(v) + \varphi(w)))) \\ &= \varphi^{[-1]}(\varphi(u) + \varphi(C(v, w))) = C(u, C(v, w)) \end{aligned}$$

donc C est associative. □

Définition 2.1.3 Une copule C induit une mesure de probabilité μ_C dite C -mesure donnée par

$$\mu_C\{[u, 0][0, v]\} = C(u, v)$$

La C -mesure d'un sous ensemble mesurable A de I^2 est la probabilité que deux variables aléatoires uniformes $(U, V) \in I^2$ avec une fonction de distribution conjointe C à valeurs dans A .

Les ensembles de niveau pour une copule C sont donnée

$$A_t(C) = \{(u, v) \in [0, 1]^2 / C(u, v) = t\}$$

Théorème 2.1.3

Soit C une copule archimédienne générée par φ , soit $K_C(t)$ la C -mesure de l'ensemble $\{(u, v) \in [0, 1]^2 / C(u, v) \leq t\}$, alors $\forall t \in [0, 1]$

$$K_t(C) = t - \frac{\varphi(t)}{\varphi'(t^+)}$$

ou $\varphi'(t^+)$ est la dérivé a droite de φ

a- Le tau de Kendall

En général l'évaluation du tau de Kendall pour une certaine copule C nécessite le calcul du double intégrale. Pour les copules archimédiennes le tau de Kendall peut être évalué directement du générateur, grâce au théorème suivant.

Théorème 2.1.4

Soient X et Y deux variables aléatoires d'une copule archimédienne C générée par φ . Le tau de Kendall de X et Y est donnée par

$$\tau_C = 1 + 4 \int_0^1 \frac{\varphi(t)}{\varphi'(t)} dt$$

b- La dépendance de queue

Pour les copules archimédiennes la dépendance de queue peut être exprimée en terme de son générateur.

Théorème 2.1.5

Soit C une copule archimédienne bivariée strict. Si $\varphi^{-1}(0)$ est fini alors

$$C(u, v) = \varphi^{[-1]}(\varphi(u) + \varphi(v))$$

n'a pas de dépendance de queue supérieure. Si C possède une dépendance de queue supérieure, $\varphi^{-1}(0) = -\infty$ et le coefficient de dépendance de queue au niveau supérieure de la distribution est donnée par

$$\lambda_U = 2 - 2 \lim_{s \rightarrow 0} \left[\frac{(\varphi^{-1})'(2s)}{(\varphi^{-1})'(s)} \right]$$

2.1.2 Exemples de copules archimédiennes

1- La famille de Gumbel

Soit la fonction $\varphi(t) = (-\ln(t))^\theta$, avec $\theta \geq 1$ et $t \in [0, 1]$. On remarque que $\varphi(t)$ est continue et $\varphi(1) = 0$, on a

$$\varphi'(t) = -\frac{\theta}{t} (-\ln t)^{\theta-1} \quad \text{et} \quad \varphi'(t) < 0$$

$$\varphi''(t) = \frac{\theta}{t^2} (-\ln t)^{\theta-2} [\theta - 1 - \ln t] \quad \text{et} \quad \varphi''(t) \geq 0$$

φ est une fonction continue strictement décroissante définie de $[0, 1]$ dans $[0, \infty]$, convexe et elle est un générateur strict, On trouve

$$C_\theta(u, v) = \varphi^{-1}(\varphi(u) + \varphi(v)) = \exp\left(-\left[(-\ln u)^\theta + (-\ln v)^\theta\right]^{\frac{1}{\theta}}\right)$$

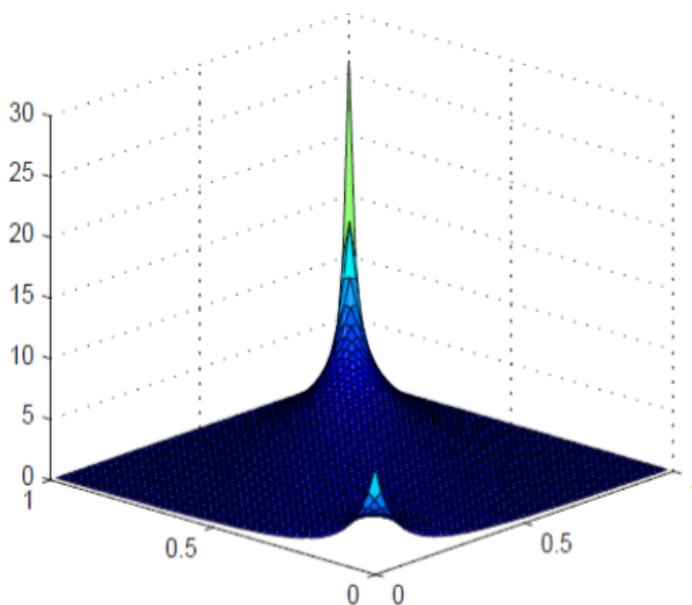


FIGURE 2.1 – Densité de la copule de Gumbel $\theta = 1.5$

- Le tau de Kendall

Le tau de Kendall de la copule de Gumbel s'écrit :

$$\begin{aligned} \tau_\theta &= 1 + 4 \int_0^1 \frac{\varphi(t)}{\varphi'(t)} dt \\ &= 1 - \frac{1}{\theta} \end{aligned}$$

- La dépendance de queue

La fonction génératrice inverse est donnée par

$$\varphi^{-1}(t) = \exp(-t^{1/\theta})$$

par conséquent

$$\varphi^{-1}(s) = \exp(-s^{1/\theta}) \quad \text{et} \quad \varphi^{-1}(s) = \frac{-s^{1/\theta-1} \exp(-s^{1/\theta})}{\theta}$$

en utilisant la théorème(2.1.5) on aura

$$\begin{aligned} \lambda_u &= 2 - 2 \lim_{s \rightarrow 0} \left[\frac{\varphi^{-1}(2s)}{\varphi^{-1}(s)} \right] \\ &= 2 - 2^{1/\theta} \lim_{s \rightarrow 0} \left[\frac{\exp(-(2s)^{1/\theta})}{\exp(-s^{1/\theta})} \right] \\ &= 2 - 2^{1/\theta} \end{aligned}$$

Donc cette famille de copule possède une dépendance de queue au niveau supérieur de la distribution. De la même manière on peut trouver le coefficient de dépendance de queue au niveau inférieur de la distribution.

$$\lambda_l = 0$$

2- La famille de Clayton

Soit la fonction génératrice $\varphi(t) = \frac{t^{-\theta} - 1}{\theta}$ telle que $\theta \in [-1, \infty] \setminus \{0\}$, la fonction de famille des copules de Clayton est donnée par

$$C_\theta(u, v) = \max([u^{-\theta} + v^{-\theta} - 1]^{-1/\theta}, 0)$$

Pour $\theta > 0$, les copules de Clayton sont strictes et l'expression de la copule sera donnée par

$$C_\theta(u, v) = (u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}$$

On a

$$\frac{\varphi(t)}{\varphi'(t)} = \frac{t^{\theta+1} - t}{\theta}$$

en utilisant la théorème (2.1.5), nous pouvons calculer le tau Kendall pour la famille de Clayton.

$$\begin{aligned} \tau_\theta &= 1 + 4 \int_0^1 \frac{t^{\theta+1} - t}{\theta} dt \\ &= 1 + \frac{4}{\theta} \left(\frac{1}{\theta - 2} - \frac{1}{2} \right) \\ &= 1 + \frac{4}{\theta} \left(\frac{-\theta}{2(\theta + 2)} \right) = \frac{\theta}{\theta + 2} \end{aligned}$$

On a

$$\varphi^{-1}(s) = (1 + \theta s)^{-1/\theta}$$

Le coefficient de dépendance de queue supérieur est donnée par

$$\begin{aligned} \lambda_L &= 2 \lim_{s \rightarrow \infty} \left[\frac{\varphi^{-1'}(2s)}{\varphi^{-1'}(s)} \right] \\ &= 2 \lim_{s \rightarrow \infty} \left[\frac{(1 + 2\theta s)^{-1/\theta-1}}{(1 + \theta s)^{-1/\theta-1}} \right] \\ &= 2 \cdot 2^{-1/\theta-1} \\ &= 2^{-1/\theta} \end{aligned}$$

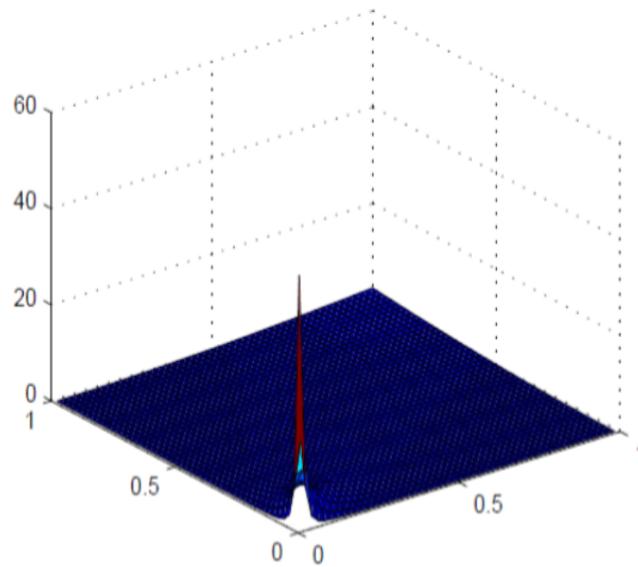


FIGURE 2.2 – Densité de la copule de Clayton ($\theta = 2.582$)

3- La famille de Frank

Soit la fonction génératrice $\varphi(t) = -\ln \left[\frac{\exp(-\theta t) - 1}{\exp(-\theta) - 1} \right]$ telle que $\theta \in \mathbb{R} \setminus \{0\}$, la copule de Frank est donnée par

$$C_{\theta}(u, v) = -\frac{1}{\theta} \ln \left(1 + \frac{(\exp(-\theta u) - 1)(\exp(-\theta v) - 1)}{\exp(-\theta) - 1} \right)$$

le tau de Kendall correspondre à cette famille est donné par la formule suivante

$$\tau_{\theta} = 1 - \frac{4}{\theta} (1 - D_1(\theta))$$

ou $D_k(x)$ est la fonction de Debye donnée par

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

de plus, le rho de Spearman est donnée par

$$\rho_{\theta} = 1 - \frac{12}{\theta} (D_1(\theta) - D_2(\theta))$$

Les copules de Frank sont des copules strictes.

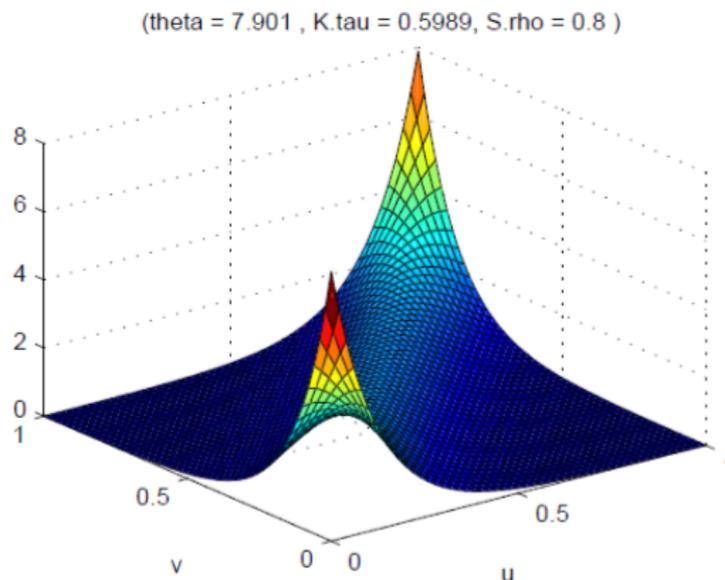


FIGURE 2.3 – Densité de la copule de Frank

Remarque 4

Les membres de la famille de Frank sont les copules archimédiennes qui satisfont l'équation $C(u, v) = \hat{C}(u, v)$.

-La dépendance de queue

On a

$$\varphi^{-1}(s) = -\frac{1}{\theta} \ln \left(1 - (1 - e^{-\theta})e^{-s} \right)$$

et

$$(\varphi^{-1})'(s) = \frac{\frac{1}{\theta} (1 - e^{-\theta}) e^{-s}}{1 - (1 - e^{-\theta}) e^{-s}}$$

puisque

$$(\varphi^{-1})'(0) = -\frac{\exp(-\theta) - 1}{\theta \exp(\theta)}$$

est finie, et selon le théorème (2.1.5) la famille de Frank n'a pas de dépendance de queue supérieure. En outre

$$\frac{(\varphi^{-1})'(2s)}{(\varphi^{-1})'(s)} = \exp(-s) \frac{1 - (1 - \exp(-\theta)) \exp(-s)}{1 - (1 - \exp(-\theta)) \exp(-2s)}$$

par conséquent,

$$\lim_{s \rightarrow \infty} \frac{(\varphi^{-1})'(2s)}{\varphi^{-1}(s)} = 0$$

ainsi la famille de Frank n'a pas de dépendance de queue inférieure.

2.2 La famille de Marshall-Olkin

Considérons un système de deux composants qui sont sujets à des chocs qui peuvent être fatals à au moins l'un des deux composants. Notons par X et Y les durées de vie des deux composants de répartition F et G respectivement. Soit H la fonction de répartition conjointe de (X, Y) . Supposons que les chocs sont des processus de Poisson indépendants de paramètres λ_1, λ_2 et λ_{12} respectivement. Les indices indiquent si les chocs effectuent seulement le premier composant, seulement le second composant ou tous les deux. Les temps d'occurrence des chocs Z_1, Z_2 et Z_{12} sont des lois exponentielles de paramètres λ_1, λ_2 et λ_{12} respectivement.

$$\begin{aligned} \bar{H}(x, y) &= \mathbb{P}(X > x, Y > y) \\ &= \mathbb{P}(Z_1 > x) \mathbb{P}(Z_2 > y) \mathbb{P}(Z_{12} > \max(x, y)). \end{aligned}$$

Les fonctions de survie univariées de X et Y sont respectivement :

$$\bar{F}(x) = \exp\left(-(\lambda_1 + \lambda_{12})x\right) \quad \text{et} \quad \bar{G}(y) = \exp\left(-(\lambda_2 + \lambda_{12})y\right)$$

Du fait que $\max(x, y) = x + y - \min(x, y)$

$$\begin{aligned}\bar{H}(x, y) &= \exp\left(-(\lambda_1 + \lambda_{12})x - (\lambda_2 + \lambda_{12})y + \lambda_{12} \min(x, y)\right) \\ &= \bar{F}(x)\bar{G}(y) \min\left(\exp(\lambda_{12}x), \exp(\lambda_{12}y)\right).\end{aligned}$$

Si on pose $u = \bar{F}(x)$, $v = \bar{G}(y)$, $\alpha = \lambda_{12}/(\lambda_1 + \lambda_{12})$ et $\beta = \lambda_{12}/(\lambda_2 + \lambda_{12})$, alors

$$\exp(\lambda_{12}x) = \bar{F}(x)^{-\alpha} \quad \text{et} \quad \exp(\lambda_{12}y) = \bar{G}(y)^{-\beta}$$

Par conséquent la copule de survie de $(X, Y)^T$ est donnée par

$$\bar{C}(u, v) = uv \min(u^{-\alpha}, v^{-\beta}) = \min(u^{1-\alpha}v, uv^{1-\beta})$$

Cette construction nous mène à la famille de copule donnée par

$$C_{\alpha\beta}(u, v) = \min(u^{1-\alpha}v, uv^{1-\beta}) = \begin{cases} u^{1-\alpha}v & \text{si } u^\alpha \geq v^\beta \\ uv^{1-\beta} & \text{si } u^\alpha \leq v^\beta. \end{cases}$$

Cette famille est dite la famille de Marshall-Olkin

Définition 2.2.1 Une copule C peut être décomposée on

$$C(u, v) = A_c(u, v) + S_c(u, v)$$

où A_c est dite composante absolument continue de C qui est donnée par

$$A_c(u, v) = \int_0^u \int_0^v C(\xi_1, \xi_2) d\xi_1 d\xi_2$$

S_c est la composante singulière de C , elle est donnée par

$$S_c(u, v) = C(u, v) - A_c(u, v)$$

-Le rho de Spearman

$$\begin{aligned}\rho_{\alpha,\beta} &= 12 \int_{I^2} C_{\alpha,\beta}(u, v) dudv - 3 \\ &= 12 \int_0^1 \left(\int_0^{u^{\alpha/\beta}} u^{1-\alpha}v dv + \int_{u^{\alpha/\beta}}^1 uv^{1-\beta} dv \right) du - 3 \\ &= \frac{3\alpha\beta}{2\alpha + 2\beta - \alpha\beta}\end{aligned}$$

-Le tau de Kendall

$$\begin{aligned}
\tau_{\alpha,\beta} &= 4 \int_{I^2} C_{\alpha,\beta}(u, v) dC_{\alpha,\beta}(u, v) - 1 \\
&= 4 \left(\frac{1}{2} - \int_{I^2} \frac{\partial C_{\alpha,\beta}(u, v)}{\partial u} \frac{\partial C_{\alpha,\beta}(u, v)}{\partial v} dudv \right) - 1 \\
&= \frac{\alpha\beta}{\alpha + \beta - \alpha\beta}
\end{aligned}$$

avec

$$\int_{I^2} C_{\alpha,\beta}(u, v) dC_{\alpha,\beta}(u, v) = \frac{1}{2} - \int_{I^2} \frac{\partial C_{\alpha,\beta}(u, v)}{\partial u} \frac{\partial C_{\alpha,\beta}(u, v)}{\partial v} dudv$$

Théorème 2.2.1

La copule de Marshall-Olkin possède une dépendance au niveau supérieur de la queue de distribution. Sans perte de généralité, supposons que $\alpha > \beta$ alors

$$\begin{aligned}
\lim_{u \rightarrow 1^+} \frac{\overline{C}(u, v)}{1 - u} &= \lim_{u \rightarrow 1^+} \frac{1 - 2u + u^2 \min(u^{-\alpha}, u^{-\beta})}{1 - u} \\
&= \lim_{u \rightarrow 1^+} \frac{1 - 2u + u^2 u^{-\beta}}{1 - u} \\
&= \lim_{u \rightarrow 1^+} (2 - 2u^{1-\beta} + \beta u^{1-\beta}) = \beta
\end{aligned}$$

Par conséquent $\lambda_u = \min(\alpha, \beta)$ est le coefficient de dépendance de queue au niveau supérieur de la distribution.

2.3 La copule Elliptique

Les copules elliptiques sont les copules associées aux distributions elliptiques multivariées, dans ce qui suit, nous donnons quelques définitions de la distribution elliptique ainsi que les deux exemples les plus connus de cette famille, la copule Gaussienne et la copule de Student.

Définition 2.3.1 *On dira que X suit une distribution elliptique si seulement si on peut écrire $X = \mu + \mathbf{R}AU$*

où

- $\mu = (u_1, \dots, u_n)$ est moyenne
- R une variable aléatoire positive indépendante de U
- A une matrice de dimension $n \times n$ telle que $\Sigma = \mathbf{A}\mathbf{A}^T$ est non singulière
- Σ une matrice de forme (matrice de variance-covariance)
- \mathbf{U} une variable aléatoire uniformément distribuée sur la sphère unité de \mathbb{R}^n .

Théorème 2.3.1

la fonction de densité d'une distribution elliptique est donnée par

$$f(x) = |\Sigma|^{-1/2} g((x - \mu)^T \Sigma^{-1} (x - \mu)), \forall x \in \mathbb{R}^n$$

où g est une fonction définie de $\mathbb{R}^+ \rightarrow \mathbb{R}$ dite générateur de densité, uniquement déterminée par la fonction de distribution de \mathbb{R}

Exemple 3

Soit la fonction suivante

$$g(t) = k \exp(-t/2)$$

génère la fonction de distribution normale multivariée et

$$g(t) = k(1 + t/m)^{(n+m)/2}$$

génère la fonction de distribution de student multivariée, avec k une constante de normalisation et $m \in \mathbb{N}$

Définition 2.3.2 Soit $X = (X_1, \dots, X_n)$ un vecteur aléatoire de distribution elliptique. La fonction de caractéristique $\phi_X(t)$, $t \in \mathbb{R}$ est donnée par

$$\begin{aligned} \phi_X(\mathbf{t}) &= E(\exp(i\mathbf{t}^T \mathbf{X})) \\ &= E(\exp(i\mathbf{t}^T \mu + \mathbf{R}\mathbf{A}U)) \\ &= E(\exp(i\mathbf{t}^T \mu) g(\mathbf{t}^T \Sigma \mathbf{t})) \end{aligned}$$

Notons par $X \sim E_n(\mu, \Sigma, g)$ la classe de la distribution elliptique de vecteur moyen μ , de matrice de covariance $\Sigma = (\sigma_{ij})$ et de générateur caractéristique g .

Définition 2.3.3 On appelle copule elliptique toute copule de la forme

$$\forall u, v \in I$$

$$C_\theta(u, v) = \int_{-\infty}^{\phi^{-1}(u)} \int_{-\infty}^{\phi^{-1}(v)} \frac{1}{2\pi\sqrt{1-\theta^2}} \left(-\frac{s^2 - 2\theta st + t^2}{2(1-\theta^2)} \right) ds dt \quad \text{où } \theta \in [-1, 1], \text{ et } \phi^{-1} \text{ est l'inverse de distribution Gaussienne univariée.}$$

2.3.1 Copule Gaussienne

La copule Gaussienne ne présente pas de dépendance de queue. L'importance de cette copule réside dans le fait qu'elle est sous-jacente à la distribution normal multivariée. En effet, modéliser la structure de dépendance d'un échantillon par copule Gaussienne est cohérent avec la mesure de cette dépendance par le coefficient de corrélation linéaire.

Définition 2.3.4 soit Σ une matrice diagonale définie positive avec $\text{diag}\Sigma = 1$ et ϕ_ρ la distribution normale multivariée standard de matrice de corrélation Σ . La copule normale est définie de la façon suivante

$$C(u_1, \dots, u_n; \Sigma) = \phi_\rho\left(\phi^{-1}(u_1), \dots, \phi^{-1}(u_n)\right)$$

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

Définition 2.3.5 La densité de la copule Gaussienne n -variée qui s'écrit :

$$c(u_1, \dots, u_n; \Sigma) = \frac{1}{\det(\Sigma)^{1/2}} \exp\left(-\frac{1}{2}\zeta^t(\Sigma^{-1} - I)\zeta\right)$$

avec I est la matrice identité de dimension $(n \times n)$ et $\zeta = \phi^{-1}(u_i)$

En deux dimension, la copule Gaussienne s'exprime de la façon suivante

$$C_\rho(u, v) = \int_{-\infty}^{\phi^{-1}(u)} \int_{-\infty}^{\phi^{-1}(v)} \frac{1}{2\pi(1-\rho^2)^{1/2}} \exp\left(-\frac{s^2 - 2\rho st + t^2}{2(1-\rho^2)}\right) ds dt$$

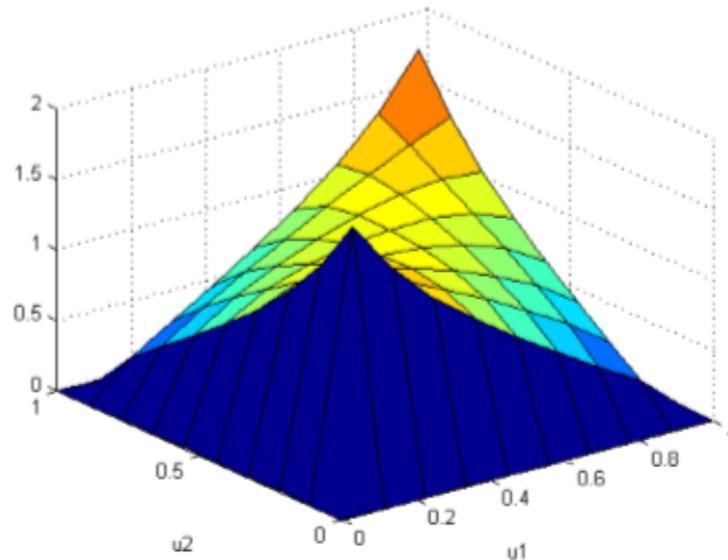
-le tau de Kendall et rho de Spearman

Théorème 2.3.2

Soit (X_1, \dots, X_n) un vecteur aléatoire d'une copule Gaussienne C , alors pour tout i et j dans $\{1, \dots, n\}$ le tau de Kendall est donné par

$$\tau_{X_i X_j} = \frac{2}{\pi} \arcsin \rho_{ij}$$

avec ρ_{ij} est le coefficient de corrélation linéaire entre X_i et X_j .

FIGURE 2.4 – Densité de la copule Gaussienne avec $\rho = 0.5$

Théorème 2.3.3

Soit (X_1, \dots, X_n) un vecteur aléatoire d'une copule Gaussienne C , alors pour tout i et j dans $\{1, \dots, n\}$ le rho de Spearman est donné par

$$\rho_{S_{ij}} = \frac{6}{\pi} \arcsin \frac{\rho_{ij}}{2}$$

avec ρ_{ij} est le coefficient de corrélation linéaire entre X_i et X_j

2.3.2 Copule Student

La copule de student est issue, de la distribution multivariée de student. Sa construction est dans le contenu de la copule gaussienne mais, contrairement à cette dernière, elle réussit à bien capter les dépendances aux extrêmes, tant positives que négatives, par ses queues lourdes.

Définition 2.3.6 Soient $\rho \in [-1, 1]$ le coefficient de corrélation de Pearson, T^{-1} l'inverse d'une fonction de répartition de Student centrée réduit univariée

$$T_{\rho,k}(u, v) = \int_{-\infty}^v \int_{-\infty}^u \frac{1}{2\pi\sqrt{1-\rho^2}} \left[1 + \frac{s^2 + t^2 - 2\rho st}{k(1-\rho^2)} \right]^{-(k+2)/2} ds dt$$

la distribution de la loi de Student de dimension 2 où $k \geq 0$ représente le nombre de degrés de liberté. Alors, pour $(u, v) \in [0, 1]^2$

$$C(u, v) = T_{\rho, k} \left(T^{-1}(u), T^{-1}(v) \right)$$

Pour obtenir la copule de Student, il suffit d'utiliser la définition de la densité de copule

$$\begin{aligned} c(u, v) &= \frac{\partial^2}{\partial u \partial v} T_{\rho, k} \left(T^{-1}(u), T^{-1}(v) \right) \\ &= \frac{f(u, v)}{f(u) \times f(v)} \\ &= \frac{k}{2\sqrt{1-\rho^2}} \frac{\Gamma(k/2)^2}{\Gamma((k+1)/2)^2} \frac{\left[1 + \frac{u^2 + v^2 - 2\rho uv}{k(1-\rho^2)} \right]^{-(k+2)/2}}{\left[\left(1 + \frac{u^2}{k} \right) \left(1 + \frac{v^2}{k} \right) \right]^{-(k+2)/2}} \end{aligned}$$

où Γ représente la fonction gamma, $f(u, v)$ la densité jointe d'une loi student et $f(u)$, $f(v)$ les densités marginales. Dans le cas bivarié, l'expression de la copule peut s'écrire comme suit

$$C_{\nu\rho}(u, v) = \int_{-\infty}^{t_v^{-1}(u)} \int_{-\infty}^{t_v^{-1}(v)} \frac{1}{2\pi(1-\rho^2)^{1/2}} \left(1 + \frac{s^2 - 2\rho st + t^2}{\nu(1-\rho^2)} \right)^{-(\nu+2/2)} ds dt$$

où ρ est le coefficient de corrélation linéaire.

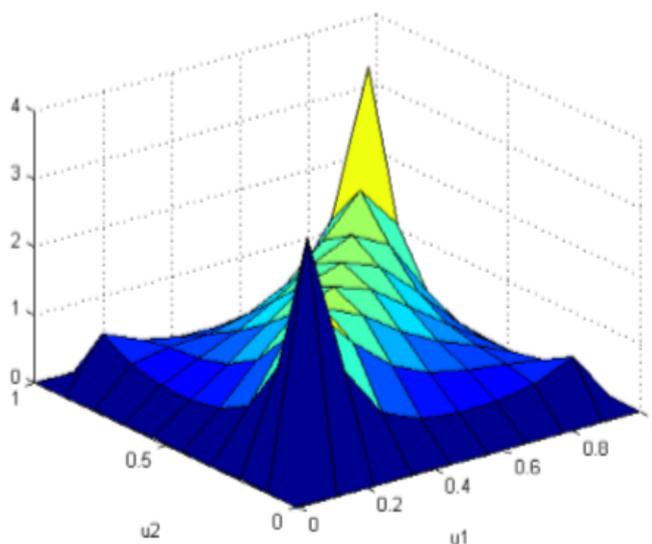


FIGURE 2.5 – Densité de la copule de student avec $\rho = 0.5$ et $\nu = 1$

Remarque 5

- $\Gamma(x) = \int_0^{\infty} \exp(-t)t^{x-1} dt$
- $\Gamma(n+1) = n!$
- $\Gamma(k + \frac{1}{2}) = (k - \frac{1}{2})\Gamma(k - \frac{1}{2})$
- $\Gamma(\frac{1}{2}) = \sqrt{\pi}$

Remarque 6

L'expression du tau de Kendall pour la copule de Student est identique à celle de la copule Gaussienne, pour le rho de Spearman il n'existe pas de forme explicite.

-La dépendance de queue

Si (X_1, X_2) suit une loi de student bivariée à ν degrés de liberté et de corrélation linéaire ρ et puisque $X_2 | X_1=x \sim t_{\nu+1}$, alors

$$\mathbb{E}(X_2 | X_1 = x) = \rho x \quad \text{Var}(X_2 | X_1 = x) = \left(\frac{\nu + x^2}{\nu + 1}\right)(1 - \rho^2)$$

ceci peut être employé pour montrer que la t-copule présente une dépendance de queue au niveau supérieur (inférieur) de la distribution.

$$\begin{aligned} \lambda_u &= 2 \lim_{x \rightarrow \infty} \mathbb{P}(X_2 > x | X_1 = x) \\ &= 2 \lim_{x \rightarrow \infty} \overline{t_{\nu+1}} \left(\left(\frac{\nu + 1}{\nu + x^2}\right)^{\frac{1}{2}} \frac{x + \rho x}{\sqrt{1 - \rho^2}} \right) \\ &= 2 \lim_{x \rightarrow \infty} \overline{t_{\nu+1}} \left(\left(\frac{\nu + 1}{\frac{\nu}{x^2} + 1}\right)^{1/2} \frac{\sqrt{1 - \rho}}{\sqrt{\rho + 1}} \right) \\ &= 2 \overline{t_{\nu+1}} \left(\frac{\sqrt{\nu + 1} \sqrt{1 - \rho}}{\sqrt{1 + \rho}} \right) \end{aligned}$$

On peut remarquer également que la dépendance de queue au niveau supérieur de la distribution est croissante pour ρ et décroissante pour ν .

2.4 La copule de valeurs extrêmes

Une autre classe particulière des copules est celle des valeurs extrêmes. Le nom «*extreme value copula*» suggère un lien entre la théorie des valeurs extrêmes et ses copules. Dans le cas bidimensionnelle, Geoffroy (1958), Tiago de oliveira(1958) et sibuya(1960) ont donné la forme générale des copules des valeurs extrêmes.

Définition 2.4.1 *Une copule de valeurs extrêmes est par définition une copule vérifiant, pour tout $\lambda > 0$*

$$C(u_1^\lambda, \dots, u_n^\lambda) = [C(u_1, \dots, u_n)]^\lambda$$

On peut vérifier par exemple qu'une copule de Gumble possède cette propriété.

Définition 2.4.2 Soit C une copule des valeurs extrêmes, la fonction $A : [0, 1] \rightarrow [1/2, 1]$ donnée par

$$A(t) = -\ln C\left(\exp-(1-t), \exp(-t)\right)$$

est appelée la fonction de dépendance de la copule des valeurs extrêmes C .

Théorème 2.4.1

Soit C une copule des valeurs extrêmes, alors

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

pour un choix approprié de la fonction A . En particulier, les contraintes suivantes doivent être vérifiées

1. $\max(t, 1-t) \leq A \leq 1$
2. A est convexe.

Exemple 4

Considérons la copule de Gumbel bivariée donnée par

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

pour $\theta \geq 0$ on a

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

La copule de Gumbel est donc une copule de valeurs extrêmes, se qui n'est pas le cas de la copule Farlie-Gumbel-Morgenstern qui est donnée par

$$C(u, v, \theta) = uv + \theta uv(1-u)(1-v)$$

on a

$$\begin{aligned} C(u^n, v^n, \theta) &= u^n v^n + \theta u^n v^n (1-u^n)(1-v^n) \\ &= u^n v^n (1 + \theta - \theta u^n - \theta v^n + \theta u^n v^n) \\ &\neq u^n v^n (1 + \theta - \theta u - \theta v + \theta uv)^n \\ &\neq C^n(u, v, \theta) \end{aligned}$$

b- Le tau de Kendall et le rho de Spearman

Le tau de Kendall et le rho de Spearman peuvent être exprimée en termes de la fonction de dépendance de queue A par

$$\tau = 4 \int \int_{[0,1]^2} C(u, v) dC(u, v) - 1 = \int_0^1 \frac{t(1-t)}{A(t)} dA'(t)$$

$$\rho_s = 12 \int \int_{[0,1]^2} uv dC(u, v) - 3 = 12 \int_0^1 \frac{1}{(1+A(t))^2} dt - 3$$

c- La fonction de distribution de Kendall

La fonction de distribution de Kendall associée à C peut s'écrire en fonction de τ , sa formule est donnée par

$$K(w) = \mathbb{P}(C(U, V) \leq w) = w - (1 - \tau)w \log w \quad w \in [0, 1]$$

Définition 2.4.3 (dépendance de queue)

On dit qu'une copule bivariée présente de la dépendance de queue si $\exists \lambda \in]0, 1]$ tel que

$$\lambda(u) = \frac{\tilde{C}(u, u)}{1 - u} \xrightarrow{u \rightarrow 1} \lambda$$

avec $\tilde{C}(u, v) = 1 - u - v + C(u, v)$ la fonction de survie conjointe. L'idée de ce quotient est de mesurer la probabilité que U dépasse le seuil u sachant que V la dépassé :

$$\lambda(u) = \mathbb{P}(U > u \mid V > u)$$

Cette notion est importante dans l'étude de la dépendance asymptotique entre deux variables aléatoires. Elle permet en effet de mesurer le niveau de dépendance dans les valeurs extrêmes, grandes ou petites.

2.4.1 Les copules archimax

Nous considérons une nouvelle famille de copules introduite par Capéra et al(2000) qui englobe la plupart des familles connues des copules, notamment les copules archimédiennes et toute les copules des valeurs extrêmes. Cette nouvelle famille offre plus de flexibilité pour la modélisations.

Définition 2.4.4 Une fonction bivariée est une copule archimax si et seulement si elle est de la forme

$$C_{\varphi,A}(x,y) = \varphi^{-1}\left[(\varphi(x) + \varphi(y))A\left(\frac{\varphi(x)}{\varphi(x) + \varphi(y)}\right)\right] \quad \forall x,y \in [0,1]$$

avec

1. $A : [0, 1] \rightarrow [1/2, 1]$ telle que $\max(t, 1 - t) \leq A(t) \leq 1$ pour toute $t \in [0, 1]$
2. $\varphi : (0, 1] \rightarrow [0, \infty)$ est une fonction convexe décroissante qui vérifie $\varphi(1) = 0$ avec la convention, $\varphi(0) = \lim_{t \rightarrow 0^+} \varphi(t)$ et $\varphi^{-1}(s) = 0$ pour toute $s \geq \varphi(0)$

Nous pouvons remarquer que l'ensemble des copules archimax $C_{\varphi,A}$ contient les copules de valeurs extrêmes ainsi que les copules archimédiennes. On effet, si l'on pose $\varphi(t) = \ln(1/t)$, la copule $C_{\varphi,A}$ est alors une copule de valeurs extrêmes, i.e.,

$$C_{\varphi,A}(x,y) = C_A(x,y) = \exp\left[\ln(xy) \frac{\ln y}{\ln(xy)}\right]$$

si on pose $A(t) = 1$, on retrouve la forme générale des copules archimédiennes

$$C_{\varphi,A}(x,y) = C_{\varphi}(x,y) = \varphi^{-1}(\varphi(x) + \varphi(y))$$

Exemple 5

Soit la fonction de dépendance de Tawn (1988) donnée par

$$A(t) = \theta t^2 - \theta t + 1, \quad 0 \leq \theta \leq 1 \quad t \in [0, 1]$$

considérons le générateur de Clayton

$$\varphi_{1,n}(t) = \frac{t^{-n} - 1}{n}, \quad n \geq 0$$

et celui de Frank

$$\varphi_{2,n}(t) = -\ln \frac{1 - \exp(-nt)}{1 - \exp(-t)}, \quad n \in \mathbb{R}$$

$\varphi_{1,n}(t)$ et $\varphi_{2,n}(t)$ sont des copules archimax.

2.5 Estimation de copule

L'estimation en statistique est une opération qui permet de calculer la valeur inconnue d'un paramètre θ du modèle statistique (Ω, P) à partir d'un échantillon observé (X_1, \dots, X_n) , on prend comme exemple la loi normale de paramètre μ et σ^2 ($N(\mu, \sigma^2)$), telle que l'estimation de μ est la moyenne $\overline{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$ et l'estimateur de σ^2 est la variance $S_n^2 = \frac{1}{n} \sum_{i=1}^n (X_i - \overline{X}_n)^2$. Il existe beaucoup de méthodes d'estimation telle que la méthode du maximum de vraisemblance, la méthode des moments...etc

2.5.1 Fonction de répartition empirique

Soit (X_1, \dots, X_n) un échantillon de taille n de loi F . La fonction de répartition empirique est définie par :

$$F_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}(X_i \leq x)$$

Cette estimation conduit à une estimation non paramétrique d'une copule.

Propriétés asymptotique

- F_n est un estimateur sans biais

$$E(F_n(x)) = F(x)$$

-Par la loi forte des grands nombres

$$F_n(x) \xrightarrow{p,s} F(x)$$

- F_n est asymptotiquement normale de variance $F(x)(1 - F(x))$

$$\sqrt{n}(F_n(x) - F(x)) \longrightarrow N(0, F(x)(1 - F(x)))$$

- F_n est estimateur uniformément consistant de F , c'est à dire

$$\| F_n - F \|_{\infty} \xrightarrow{p,s} 0$$

2.5.2 Copule empirique

La copule empirique a été induite et d'abord étudié par Deheuvels(1979), qui l'appelé fonction de dépendance empirique.

1-Cas des copules bivariées

Définition 2.5.1 Soit $\{(x_k, y_k)\}_{k=1}^n$ un échantillon de taille n d'une couple de variables aléatoires. La copule empirique est la fonction \hat{C} définie par

$$\hat{C}\left(\frac{i}{n}, \frac{j}{n}\right) = \frac{\text{Nombre des paires } (x, y) \text{ dans l'échantillon tels que } x \leq x_i \text{ et } y \leq y_j}{n}$$

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

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

$$\hat{c}\left(\frac{i}{n}, \frac{j}{n}\right) = \begin{cases} \frac{1}{n} & \text{si } (x_i, y_j) \text{ est un élément de l'échantillon} \\ 0 & \text{sinon} \end{cases}$$

2-Généralisation dans les cas des copules multivariées

En dimension T , $(x_t^1, \dots, x_t^n)_{t=1}^T$ est un échantillon de taille n , la copule empirique est définie de façon analogue au cas bivarié, \hat{C} est donnée comme suit

$$\hat{C}\left(\frac{t_1}{T}, \dots, \frac{t_n}{T}\right) = \begin{cases} \frac{1}{T} & \text{si } (x_t^1 \leq x_{t_1}^1, \dots, x_t^n \leq x_{t_n}^n) \\ 0 & \text{sinon} \end{cases}$$

Les copules empiriques peuvent être utilisées pour estimer les mesures de dépendance à savoir le ρ de Spearman et le τ de Kendall.

Théorème 2.5.1

Soient \hat{C} et \hat{c} la copule empirique et la fonction de densité respectivement de l'échantillon $(x_k, y_k)_{k=1}^n$. Le tau de Kendall et le rho de Spearman sont estimés par les formules suivantes

$$\hat{\tau} = \frac{2n}{n-1} \sum_{i=2}^n \sum_{j=2}^n \sum_{p=1}^{i-1} \sum_{q=1}^{j-1} \left(\hat{c}\left(\frac{i}{n}, \frac{j}{n}\right) \hat{c}\left(\frac{p}{n}, \frac{q}{n}\right) - \hat{c}\left(\frac{i}{n}, \frac{q}{n}\right) \hat{c}\left(\frac{p}{n}, \frac{j}{n}\right) \right)$$

$$\hat{\rho} = \left(\frac{12}{n^2 - 1} \right) \sum_{i=1}^n \sum_{j=1}^n \left(\hat{C}\left(\frac{i}{n}, \frac{j}{n}\right) - \frac{ij}{n^2} \right)$$

2.5.3 Méthode d'estimation

Méthode de Maximum de vraisemblance

De l'égalité $F(X_1, \dots, X_n) = C(F(X_1), \dots, F(X_n))$ on déduit par dérivation l'expression de la densité du vecteur (X_1, \dots, X_n) :

$$f(X_1, \dots, X_n) = c\left(F_1(X_1), \dots, F_n(X_n)\right) \prod_{i=1}^n f_i(X_i)$$

où $c(u_1, \dots, u_n) = \frac{\partial^n C(u_1, \dots, u_n)}{\partial u_1 \dots \partial u_n}$ désigne la densité de la copule. L'expression de la log vraisemblance de l'échantillon $(X_1^k, \dots, X_n^k)_{1 \leq k \leq K}$ s'en déduit immédiatement :

$$L(\theta) = \sum_k \ln \left(c\left(F_1(X_1^k; \theta), \dots, F_n(X_n^k; \theta)\right) \right) + \sum_{k=1}^K \sum_{i=1}^n f_i(X_i^k; \theta)$$

il reste maximiser cette expression en θ , ce qui peut s'avérer en pratique fastidieux.

Exemple 6 (la copule Gaussienne multivariée)

Soit la fonction de distribution de la copule Gaussienne multivariée C_{Ga} définie comme suit

$$C_{Ga}(u_1, \dots, u_n; R) = \Phi_R\left(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_n)\right)$$

où R est sa matrice de corrélation, sa fonction de densité est donnée par

$$\frac{1}{(2\pi)^{n/2} |R|^{1/2}} \exp\left(-\frac{1}{2}x'R^{-1}x\right) = c(\Phi(x_1), \dots, \Phi(x_n)) \prod_{i=1}^n \left(\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x_i^2\right)\right)$$

on a donc

$$c(\Phi(x_1), \dots, \Phi(x_n)) = \frac{1}{(2\pi)^{n/2} |R|^{1/2}} \exp(-1/2x'R^{-1}x) \prod_{i=1}^n \left(\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x_i^2\right)\right)$$

soit $u_i = \Phi(x_i) \Leftrightarrow x_i = \Phi^{-1}(u_i)$, on peut alors récrire la densité de la copule comme suite :

$$c(u_1, \dots, u_n) = \frac{1}{|R|^{1/2}} \exp\left(-\frac{1}{2}S'(R^{-1} - I)S\right)$$

avec $S = (\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_n))'$ Soit l'échantillon $\{x_1^t, \dots, x_n^t\}_{t=1}^T$, la fonction log-vraisemblance est donnée par

$$L_\theta = -\frac{T}{2} \ln |R| - \frac{1}{2} \sum_{t=1}^T S_t'(R^{-1} - I)S_t$$

où θ est l'ensemble de tous les paramètres et $S_t = \Phi(\Phi^{-1}(u_1^t), \dots, \Phi^{-1}(u_n^t))$.

L'estimateur du maximum de vraisemblance de R est donnée par

$$R^{MLE} = \frac{1}{T} \sum_{t=1}^T S_t' S_t$$

Méthode des fonctions d'inférence des marginales (IFM)

Cette méthode consiste à maximiser la fonction log-vraisemblance $L(\theta)$ mais en séparant entre les paramètres des distributions marginales γ_j et les paramètres de dépendance.

La procédure d'estimation effectuée en deux étapes, introduites par Joe et Xu (1996) consiste à :

-Estimer les paramètres de distribution marginale univariées γ_j , qui sont définis par :

$$\hat{\gamma} = \arg \max L_j(\gamma_j) \quad (2.2)$$

telle que

$$L_j(\gamma_j) = \sum_{i=1}^n \ln f_i(x_{ij})$$

où f_i est la densité de F_j .

-Puis, on estime θ , en utilise les estimateurs $\hat{\gamma}_j$ obtenues par (2.2) nous avons alors

$$\hat{\theta}_n^{IFM} = \arg \max L(\theta)$$

où

$$L(\theta) = \sum_{i=1}^n \ln c_{\theta}\{F_{\hat{\gamma}_1}(x_{i1}), \dots, F_{\hat{\gamma}_n}(x_{in})\}$$

Méthode de moment

Cette méthode consiste à estimer les paramètres $\beta_i, i = 1, \dots, n$ des lois marginales et le paramètre α de la copule par la méthode des moments, i.e.

1. Résoudre le système des n équation à n inconnues

$$\left\{ \begin{array}{l} \bar{X}_t = f(\beta_1, \dots, \beta_n) \\ S_t^2 = g(\beta_1, \dots, \beta_n) \\ \mu_{3,t} = h(\beta_1, \dots, \beta_n) \\ \cdot \\ \cdot \\ \cdot \end{array} \right.$$

où n désigne le dimension de α ; f, g et h sont les expressions des moments (ordinaires) d'ordre 1,2 et 3 en fonction des paramètres β_i . Répéter cette étape pour toutes les marginales.

2. Inverser le tau de kendall ou le rho de Spearman pour obtenir le paramètre α de la copule.

Exemple 7

Pour la copule de Gumbel de paramètre θ , on a

$$\tau = 1 - \frac{1}{\theta}$$

Nous en déduisons que :

$$\theta = \frac{1}{1 - \tau}$$

Si nous avons une estimation $\hat{\tau}$ du tau de Kendall, nous pouvons obtenir une estimation du paramètre de la copule en posons

$$\hat{\theta} = \frac{1}{1 - \hat{\tau}}$$

Dans le cas générale, on a l'estimateur non paramétrique du tau de Kendall est donnée par :

$$\hat{\tau} = \frac{c - d}{c + d}$$

ou c et d respectivement le nombre de paires disjointes concordantes et discordantes.

Chapitre 3

Utilisation des copules

L'objet de cet application est la mise en évidence de la structure de dépendance de deux variables Hydrologiques :Cote et Température. La dépendance est prospectée par Chi plot et Kendall(k)plot et mesurée grâce à deux coefficients de corrélation de rang le τ de Kendall et le ρ de Spearman, la distribution conjointe du copule Cote et Température.est construite à l'aide de la copule Gumbel. L'identification de la copule parmi trois modèles courants (Gumbel, Frank et Clyton) est réalisée par :

- (a) La comparaison entre les versions empiriques des fonctions K, J, M, L et R
- (b) Letest de χ^2 bidimensionnel. Cette approche par les copules a permis la reconstitution par simulation de l'échantillon cote et température

3.1 Copules et modélisation

3.1.1 Prospection graphique de la dépendance

Deux méthodes graphiques récentes basées sur les statistiques des rangs sont utilisées pour la détection de la dépendance entre deux variables à savoir chi plots introduite par Fisher et Switzer (1985, 2001) et Kendall plots (k plot) développée par Genest et Boies (2003) montrent que ces méthodes peuvent détecter non seulement la corrélation moyenne ou global entre les variables mais aussi la dépendance asymptotique

Chi plots : Il s'agit de porter sur un graphique les paires (λ_i, χ_i) et voir si le nuage obtenu des points est à l'intérieur ou à l'extérieur de la bande de contrôle. S'il est à l'intérieur cela indique l'indépendance entre les variables aléatoires X et Y , dans le cas contraire, on

peut conclure à la dépendance.

Ils définissent pour tout couple (X_i, Y_i) tel que $1 \leq i \leq n$

$$H_i = \frac{1}{n-1} \{j \neq i : X_j \leq X_i, Y_j \leq Y_i\}$$

$$F_i = \frac{1}{n-1} \{j \neq i : X_j \leq X_i\}$$

$$G_i = \frac{1}{n-1} \{j \neq i : Y_j \leq Y_i\}$$

$$\chi_i = \frac{H_i - F_i * G_i}{\sqrt{F_i(1-F_i)G_i(1-G_i)}} \quad \text{et} \quad \lambda_i = 4 \text{sign}(\tilde{F}, \tilde{G}) \max(\tilde{F}^2, \tilde{G}^2), \chi_i \text{ et } \lambda_i \in [-1, 1]$$

avec $\tilde{F} = F_i - 0.5$ et $\tilde{G} = G_i - 0.5$, λ_i représente la distance du couple (X, Y) par rapport au centre du nuage de points, et $\sqrt{n}\chi_i$ est la racine carrée (positive ou négative) de la statistique du χ^2 qui est utilisée pour tester l'indépendance du tableau de contingence (X, Y) . La largeur de la bande de contrôle est définie par $\chi = c_p/\sqrt{n}$ qui correspond à un pourcentage des (λ_i, χ_i) se trouvant à l'intérieur. Fisher et Switzer (2001) ont trouvé par simulations de Monte Carlo, des valeurs de c_p égales à 1.54, 1.78 et 2.18, respectivement, pour $p = 0.90, 0.95$ et 0.99 . En outre, ils recommandent d'exclure les observations douteuses et de ne considérer dans l'analyse que les couples tels que $\lambda_i < 4(1/(n-1) - 0.5)^2$ (Fisher et Switzer, 1985).

Abberger (2005) détaille la méthode de prospection de la dépendance asymptotique en examinant la forme des chi plot correspondant aux quadrants supérieur droit et inférieur gauche du nuage de points (X, Y) , si la dépendance est positive et à l'inverse supérieur gauche et inférieur droit si elle est négative. Lorsque la forme montre une courbure qui tend vers $\chi_i = 0$, alors il conclut qu'il y a une indépendance asymptotique.

K plot : Genest et Boies (2003) ont élaboré une méthode graphique inspirée de la notion de QQ plot (quantile-quantile plot). La procédure utilisée est la suivante :

- Calculer les quantiles H_i définies en précédant
- Ordonner $H(i)$ de telle sorte que $H(1) < H(2) < \dots < H(n)$
- Porter sur un graphique les paires $(W_{i:n}, H(i))$, pour tout $i \in [1, n]$.

La quantité $W_{i:n}$ représente l'espérance (moyenne) de la i ème statistique d'ordre associée à l'échantillon aléatoire H de taille n . Par définition, la i ème statistique d'ordre s'exprime par

$$W_{i:n} = n \binom{n-1}{i-1} \int_0^1 w k_0(w) \{k_0(w)\}^{i-1} \{1 - k_0(w)\}^{n-1} dw$$

où k_0 est la densité correspondante de k_0 , définie par

$$k_0(w) = \mathbb{P}(UV \leq w) = \int_0^1 \mathbb{P}(U \leq \frac{w}{v})dv = \int_0^w 1dv + \int_w^1 \frac{w}{v}dv = w - w \log w, W \in [0, 1]$$

où U et V sont des variables aléatoires uniformes indépendance dans l'intervalle $[0, 1]$.

3.1.2 Mesures de dépendance

Plus classiquement, les mesures d'association couramment utilisées (D'Agostino et Stephens, 1986, Joe, 1997, Genest et Favre, 2007) pour la caractérisation de la dépendance sont :

- Le coefficient de corrélation linéaire ρ de Pearson.
- Le coefficient de corrélation des rangs ρ_s de Spearman .
- Le coefficient de corrélation des rangs τ de Kendall.

Le coefficient de corrélation de Pearson mesure la dépendance linéaire entre deux variables aléatoires. Pour l'étude d'une dépendance non linéaire, les autres coefficients sont plus adaptés puisqu'ils ne dépendent que du rang de chaque observation par opposition au coefficient de corrélation de Pearson qui utilise les réalisations des variables. En outre, ρ_s et τ sont invariants par transformation strictement croissante des variables aléatoires, X et Y .

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

avec n la taille de l'échantillon, R_i et S_i les rangs respectifs des réalisations des variables aléatoires X et Y .

Le τ de Kendall est défini par la probabilité de concordance moins celle discordance :

$$\tau(X, Y) = \mathbb{P}[(X - X')(Y - Y') \geq 0] - \mathbb{P}[(X - X')(Y - Y') \leq 0] \quad -1 \leq \tau \leq 1$$

Genest et Boies (2003) montrent qu'une estimation de τ peut être calculée par :

$$\hat{\tau} = -1 + 4 \frac{1}{n} \sum_{i=1}^n H(i)$$

En outre, les distributions de ρ_s et τ sont proches de lois normales de moyennes nulles et de variances respectives $\frac{1}{n-1}$ et $\frac{2(2n+5)}{9n(n-1)}$. X et Y peuvent être considérés comme

dépendants au seuil α , si $:\sqrt{n-1}|\rho_s| > z_{\alpha/2}$ et respectivement $\sqrt{\frac{9n(n-1)}{2(2n+5)}}|\tau| > z_{\alpha/2}$

3.1.3 Sélection et adéquation de la copule

Comme pour les fonctions de distribution à une variable, une famille de copules est choisie pour ses particularités à représenter un ou plusieurs aspects des données. Ensuite, l'ajustement doit être réalisé pour retenir la copule qui modélise au mieux les observations.

A cet effet, Venter (2002, 2003) et Belguise (2001) proposent d'utiliser des fonctions descriptives K, J, M, L et R pour tester l'adéquation d'une copule ou pour sélectionner une copule ayant des caractéristiques spécifiques par exemple au niveau des queues de distribution. Il s'agit de comparer les valeurs empiriques et les valeurs estimées pour les différentes copules. Cette comparaison peut être faite par un test de χ^2 au seuil $\alpha\%$. Pour un ajustement complémentaire des copules, Hürlimann (2004) propose le test du χ^2 bidimensionnel.

La fonction $K(z)$

Il s'agit de la fonction de répartition de la variable aléatoire $C(U, V)$. Pour une copule archimédienne, $K(z)$ (Genest et Revest, 1993) est définie par

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

La version empirique de la fonction $K(z)$ est exprimée par

$$K_{emp}(z) = \{\text{nombre de } z_i \leq z\}/n$$

La fonction $J(z)$ ou de τ cumulatif

la fonction $J(z)$ s'écrit

$$J(z) = -1 + 4 \int_0^z \int_0^z C(u, v) c(u, v) du dv / C(z, z)^2$$

La double intégrale représente la moyenne pondérée de $C(u, v)$, le poids étant $C(z, z)$ le quotient représente la moyenne de $C(u, v)$ qui croît en fonction de z , le second $C(z, z)$ compare la moyenne de C à sa valeur maximale.

On a $J(1) = \tau$ La version empirique de la fonction $J(z)$ est exprimée par

$$J(z) = -1 + 4I(z)/C(z, z)^2$$

où $I(z)$ est définie par

$$I(z) = \frac{10}{n} \sum_{i=1}^n z_i \times \mathbb{I}\{u_i < z \text{ et } v_i < z\}$$

avec \mathbb{I} la fonction indicatrice et $u_i = \text{rang}(x_i)/(n+1)$ et $v_i = \text{rang}(y_i)/(n+1)$

La fonction $M(z)$

Est la moyenne conditionnelle cumulative, elle est définie par

$$M(z) = E(V \mid U < z) = \frac{\left(\int_{u=0}^z \int_{v=0}^1 v C(u, v) du dv \right)}{z}$$

et vérifiant $M(1) = \frac{1}{2}$

En définissant $D(z)$ et $N(z)$ comme suit

$D(z) = \sum_{i=1}^n \mathbb{I}\{u_i < z\}$ et $N(z) = \sum_{i=1}^n v_i \mathbb{I}\{u_i < z\}$, la version empirique de $M(z)$ est donnée par

$$M(z) = \frac{N(z)}{D(z)}$$

avec $D(1) = n$ et $N(1) = \frac{n}{2}$

Les fonctions $L(z)$ et $R(z)$

Ces deux fonctions montrent les concentrations des queues gauche (Left) et droite (Right) de la distribution. Les événements extrêmes sont mis en évidence particulièrement aux limites qui, si elles existent, représentent la dépendance asymptotique définie par Joe (1997) :

$$\lambda_{min} = \lim_{u \rightarrow 0} \mathbb{P}\left(Y \leq F_Y^{-1}(u) \mid X \leq F_X^{-1}(u)\right)$$

pour les petites valeurs extrêmes (queue gauche)

et

$$\lambda_{max} = \lim_{u \rightarrow 1} \mathbb{P}\left(Y > F_Y^{-1}(u) \mid X > F_X^{-1}(u)\right)$$

pour les grandes valeurs extrêmes (queue droite).

Les deux fonctions $L(z)$ et $R(z)$ sont alors

$$L(z) = \frac{\mathbb{P}(U < z, V < z)}{z^2}$$

$$R(z) = \frac{\mathbb{P}(U > z \text{ ou } V > z)}{(1-z)^2} = \frac{(1-2z + C(z, z))}{(1-z)^2}$$

On s'intéresse à la fonction L pour tout $z \in [0, \frac{1}{2}]$ et à la fonction R pour tout $z \in [\frac{1}{2}, 1]$, ces deux fonctions ont leurs versions empiriques.

Test du χ^2 bidimensionnel $\chi^2(x, y)$

Il s'agit ici de mettre en place le test du χ^2 bidimensionnel pour l'adéquation de la distribution bidimensionnelle $F(x, y)$ obtenue grâce à la copule. La méthodologie adoptée reprend celle proposée par Hurliman (2004). l'espace est subdivisé en pavés $[x_{i-1}, x_i[\times [y_{j-1}, y_j[$. Les fréquences théoriques $f_{i,j}$ de chacun des pavés sont calculées par

$$f_{i,j} = C\left(F_X(x_i), F_Y(y_j)\right) - C\left(F_X(x_{i-1}), F_Y(y_j)\right) - C\left(F_X(x_i), F_Y(y_{j-1})\right) + C\left(F_X(x_{i-1}), F_Y(y_{j-1})\right)$$

Ces pavés sont regroupés en k classes, telles que chacune d'elles contient au moins 5 des observations afin de pouvoir utiliser la statistique suivante du χ^2 :

$$\chi^2 = \sum_{k=1}^m (O_k - E_k)^2 / E_k$$

où m est le nombre de classes, O_k et E_k sont les effectifs respectivement observé et théorique de la class. k

3.2 Application numérique

Dans cette section, les techniques de la copule décrit seront appliquées sur un ensemble de données se composant de deux séries de température et cote de Barage Beni Haroune. L'ensemble des données est étalé du 1 janvier 2004 au 31 décembre 2017

Le traitement des données sera effectué informatiquement. La programmation du logiciel de statistique R.

"R" est un logiciel de statistique que l' on peut télécharger et utiliser gratuitement. Réalisé par des professionnels en statistique, il est révisé et complété très régulièrement

Exemple

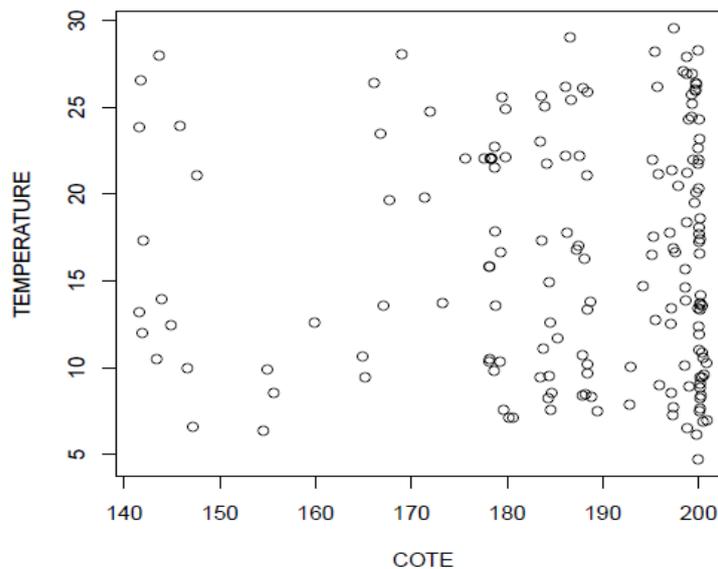


FIGURE 3.1 – cote en fonction de température

la figure (3.1) présente la variable aléatoire cote en fonction de variable aléatoire température.

Coefficient	Valeur
ρ pearson	-0,05807074
τ kandell	-0,731624
ρ_s spearman	-0,923008

TABLE 3.1 – Les mesure de dépendance

-Le coefficient de corrélation de Kendall varie dans l'intervalle $[-1, 1]$ dans ce résultats la corrélation de Kendall entre la température et cote est de $-0,731624$ on a le corrélation est négative comprendre le vecteur des rangs associés à l'une des variables est fonction décroissante de l'autre, plus l'intervalle $[-1, -0,7]$ plus converge vers $|\tau|$ et plus la corrélation entre les deux variables est de plus en plus forte

-Utilisez le coefficient de corrélation de Spearman pour examiner l'importance et la direction de la relation monotone entre deux variables continues. Dans une relation monotone les variables ont tendance à se déplacer dans la même direction relative.

Le coefficient de corrélation peut avoir une valeur entre -1 et 1 , plus la valeur absolue du coefficient est importante, plus la relation linéaire entre les variables est forte.

Dans ces résultats la corrélation de Spearman entre cote et température est de $-0,923008$, la relation entre ces variables est négative.

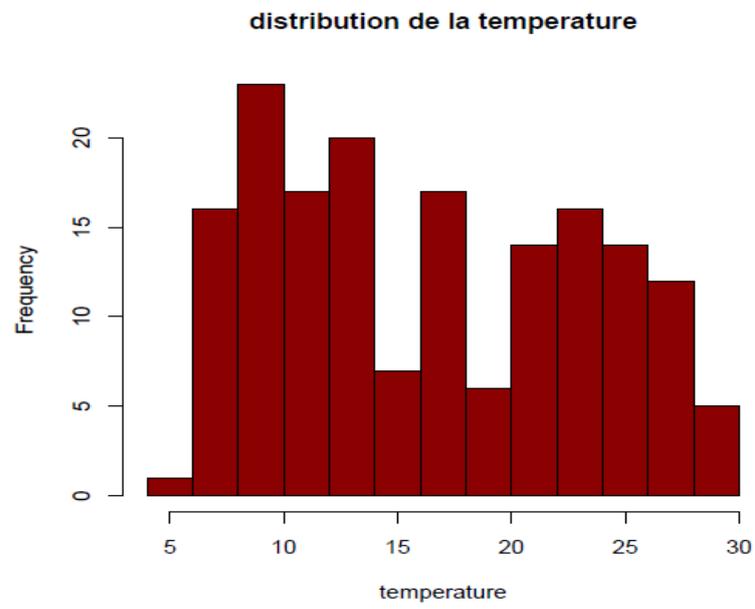


FIGURE 3.2 – Distribution de température

La figure (3.2) présente la distribution marginale de température.

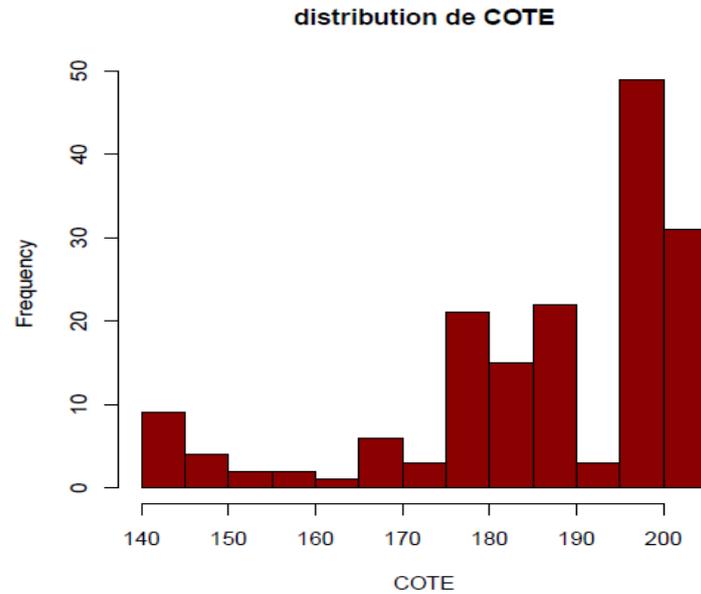


FIGURE 3.3 – Distribution de cote

La figure (3.3) présente la distribution marginale de la variable aléatoire cote.

Estimation du paramètre de dépendance

Pour mesurer le degré de dépendance entre cote et température les coefficients de corrélation ρ_s Spearman et de Kendall τ ont été estimés. On trouve

$$\hat{\rho}_s = -0,923008 \text{ avec } \sqrt{(n-1)}|\rho_s| = 1,69 > 1,64 \text{ valeur correspondant à } z_{\alpha/2} \text{ au seuil}$$

$$\alpha = 10\% \text{ et } \hat{\tau} = -0.731624 \sqrt{\frac{9n(n-1)}{2(2n+5)}}|\tau| = 1,73 > 1,64$$

Ces valeurs indiquent qu'on rejette l'hypothèse d'indépendance entre ces variables.

Conclusion

La donnée d'une copule comme modèle de dépendance d'un vecteur aléatoire est très riche.

La notion de mesure d'association sert à résumer cette structure de dépendance dans une collection de scalaires. Tout modèle probabiliste multivarié possède (au moins) une copule. La théorie des copules permet de modéliser la structure de dépendance entre les variables indépendamment des lois marginales (Sklar, 1959)

La notion de dépendance stochastique est exactement couverte par le concept de copule. Traiter cette notion uniquement à l'aide de corrélations linéaires est (en général) une très mauvaise idée.

En ce qui concerne l'approche statistique, divers auteurs l'ont adoptée :

Kurothe et al. (1997) et Goel et al. (2000) ont modélisé la relation entre l'intensité moyenne et la durée de pluie par une loi jointe bi-dimensionnelle (modèle bivarié). De même, De Michele & Salvadori (2003) analysent la dépendance entre les variables en modélisant chacune des lois marginales par la loi Pareto généralisée et leur loi jointe par la copule de Frank.

Annexes

Calcul de l'inverse de la transformée de Laplace de l'inverse du générateur d'une copule de Clayton

On cherche à inverser la transformation de Laplace g , de paramètre θ , définie par

$$g(u) = (1 + u)^{-1/\theta}.$$

On montre facilement :

$$g^n(u) = (-1)^n \left(\prod_{i=1}^n \left(\frac{1}{\theta} + n - i \right) \right) (1 + u)^{\frac{-1}{\theta} - n}$$

D'où, par la formule de Post-Widder, en notant f la densité à calculer :

$$\forall x, f(x) = \lim_{n \rightarrow \infty} \frac{(-1)^n}{n!} \left(\frac{n}{x} \right)^{n+1} (-1)^n \left(1 + \frac{n}{x} \right)^{\frac{-1}{\theta} - n} \left(\prod_{i=1}^n \left(\frac{-1}{\theta} + n - i \right) \right)$$

Et, en simplifiant :

$$f(x) = \frac{x^{\frac{-1}{\theta} - 1}}{n! n^{1/\theta - 1}} \left(\frac{x}{n} + 1 \right)^{\frac{-1}{\theta} - n} \prod_{i=1}^n \left(\frac{1}{\theta} + n - i \right)$$

Or, on montre facilement, en passant au logarithme et en utilisant la règle de l'Hôpital, que :

$$\lim_{n \rightarrow +\infty} \left(\frac{x}{n} + 1 \right)^{\frac{-1}{\theta} - n} = e^{-x}$$

D'où

$$f(x) = x^{\frac{1}{\theta} - 1} e^{-x} \lim_{n \rightarrow +\infty} \frac{n^{\frac{-1}{\theta}}}{(n-1)!} \prod_{i=0}^{n-1} \left(\frac{1}{\theta} + i \right)$$

$$\text{or, on a } \Gamma\left(n + \frac{1}{\theta}\right) = \prod_{i=0}^{n-1} \left(i + \frac{1}{\theta} \right) \Gamma\left(\frac{1}{\theta}\right)$$

Ainsi

$$f(x) = \frac{x^{\frac{1}{\theta} - 1} e^{-x}}{\Gamma\left(\frac{1}{\theta}\right)} \lim_{n \rightarrow +\infty} \frac{n^{\frac{-1}{\theta}} \Gamma\left(\frac{1}{\theta}\right)}{\beta\left(n; \frac{1}{\theta}\right)}$$

or, pour b fixé, $\beta(a, b) \sim a^{-b} \Gamma(b)$.

D'où

$$\forall f(x) = \frac{x^{\frac{1}{\theta} - 1} e^{-x}}{\Gamma\left(\frac{1}{\theta}\right)}$$

Et l'on peut conclure :

$$LS^{-1}(\phi^{-1}) \rightsquigarrow \Gamma\left(\frac{1}{\theta}; 1\right)$$

Calcul de l'inverse de la transformée de Laplace de l'inverse du générateur d'une copule de Frank

Soit θ le paramètre d'une copule de Frank de générateur ϕ . On a alors

$$\phi^{-1}(s) = -\frac{1}{\theta} \ln \left(1 - e^{-s} (1 - e^{-\theta}) \right)$$

On se restreint à $\theta > 0$ et on pose $\alpha = 1 - e^{-\theta}$, on a donc

$$\phi^{-1}(s) = \frac{-1}{\theta} \ln(1 - e^{-s}\alpha)$$

Où $e^{-s}\alpha \in [0, 1[$.

Or on a le développement en série entière suivant : $\ln(1-x) = -\sum_{t=1}^{\infty} \frac{x^t}{t}$ pour $x \in [0, 1[$

Doù, pour $\theta > 0$

$$\phi^{-1} = \frac{1}{\theta} \sum_{t=1}^{\infty} e^{-st} \frac{\alpha^t}{t} = \sum_{t=1}^{\infty} e^{-st} \frac{\alpha^t}{t\theta}$$

Et l'on reconnaît la transformée de Laplace de la distribution X telle que

$$\forall n > 0 \quad \mathbb{P}[X = n] = \frac{\alpha^n}{n\alpha} = \frac{\alpha^n}{\ln(1-\alpha)n}$$

Bibliographie

- [1] **Roger B.Nelsen** :*An Introduction to Copulas*,
- [2] **Frédéric Planchet** :*Introduction à la théorie des copules*, Modèles financiers en assurance et analyses dynamiques, support de cours 2017-2018
- [3] **Gildas Mazo** :*Construction et estimation de copules en grande dimension*, Docteur de l'université de Grenoble
- [4] **Lounis Fadhila** :*Modélisation de la dépendance par les copules et applications*, mémoire de magister ,université de Mouloude Mammeri, Tizi-Ouzou
- [5] **Alexandre Popier** :*Copules*, université de Maine, Le Mans, septembre 2010
- [6] **Myriam Chabot** :*Concepts de dépendance et copules*, université de Sherbrooke
- [7] **Emna Gargouri-Ellouze & Assia Chabchoub** : Modélisation de la structure de dépendance hauteur-durée d'événements pluvieux par la copule de Gumble, Hydrological sciences journal
- [8] **Kadi Nabil** :*Estimation non-paramétrique de la distribution et densité de copule* , université de Sherbrooke, Québec, Canada, avril 2014
- [9] **David Cadoux, Jean-Marc Loizeau** :*Copules et dépendance*, Institut des Actuaire, France