



Faculté des Sciences Exactes et Informatique
Département de Mathématiques

N° d'ordre :

N° de séries :

Mémoire de fin de cycle

Présenté pour l'obtention du diplôme de

Master

Spécialité : Mathématiques.

Option : Probabilités et statistique.

Thème

Inférence sur les copules: Approche Bayésienne

Présenté par :

Rayane Bouafia

Wissam Boumekrek

Devant le jury :

Président	: Hassen Cheraitia	M.C.B Université de Jijel
Encadreur	: Djowayda Ghouil	M.A.A Université de Jijel
Examineur	: Fatima Yakoubi	M.A.A Université de Jijel

Promotion **2019/2020**

Remerciements

*Nous tenons à remercier "**Allah**" le tout puissant pour toute la santé, la force, le courage et la volonté qu'il nous a donné pour l'achèvement de ce travail.*

*Nous tenons à remercier particulièrement et avec gratitude notre encadreur Mme : **Ghouil Djowayda** pour ses précieux conseils, ses apports appréciés et ses encouragements.*

*Nous adressons nos sincères remerciements au Mr : **Mabrouk Gherda** pour son assistance, sa patience, ses orientations et pour tous ses conseils très précieux qui ont étayé notre conduite dans la réalisation de ce mémoire.*

*Nous remercions Mr : **Cheraitia Hassen** pour ses conseils et pour toute l'aide qu'il nous apporté.*

*Nos précieux remerciements vont à le président et les membres de jury : Mr **Cheraitia Hassen** et Mme : **Yakoubi Fatima** pour l'honneur qu'ils nous ont fait en acceptant de juger nous travail et faire partie de ce jury.*

Un vif remerciement à nos famille pour nous avoir donné jour après jour autant d'amour, de soutien et d'encouragement.

Et enfin nous adressons nos remerciements à toute personne ayant contribué de près ou de loin à la concrétisation de ce travail.

Dédicace

Avec l'expression de ma reconnaissance, je dédie ce modeste travail à ceux qui, quels que soient les termes embrassés, je n'arriverais jamais à leur exprimer mon amour sincère :

*À l'homme, mon précieux offre du dieu, qui doit ma vie, ma réussite et tout mon respect : mon cher **Père**.*

*À la femme qui a souffert sans me laisser souffrir, qui n'a jamais dit non à mes exigences et qui n'a épargné aucun effort pour me rendre heureuse : mon adorable **Mère**.*

*À mes chères sœurs **Chahrazed** et **Meryem** et frères **Walid** et **Nabil** qui n'ont pas cessé de me conseiller, encourager et soutenir tout au long de mes études, que dieu les protège et leurs offre la chance et le bonheur.*

*À mon petite ange **Nadine** qui m'apporte toujours bonheur et joie.*

*À mes amis en particulier : **Dalal Djoumma** et **Sara Lobiod** pour leurs aides et supports dans les moments difficiles. merci pour leurs amours et leurs encouragements.*

*Sans oublier mon binôme **Rayane** pour son soutien moral, sa patience et sa compréhension tout au long de ce projet.*

♡Wissam♡

Dédicace

Bon Dieu , sous tes cieux , je te crois toujours mieux, j'avais de la foi, que je vais avoir le succès derrière cette table fait en bois, sachant que ce beau jour va arriver, et que tout le monde sera plain de joie, Que j'enlève le store de cette perle de toute ma vie scolaire jusqu'au universitaire, et que tous seront fières de moi...

*Le grand merci d'abord au **Dieu** et puis, Ceux que j'ai po plus cher qu'eux....*

*Celui qui a fait de moi une vrais femme, qui a supporté les peines pour protéger notre royaume, qui étais toujours supère, qui étais mon ami et mon **PÈRE** Je t'aime avant tout papa, et je t'offre ce modeste travaille.*

*Et puis A la lumière de mes jours, la source de mes efforts, la flamme de mon cœur, ma vie et mon bonheur ; **maman** que j'adore, tu mérites tout le bonheur autant de ce monde et sa largesse...*

*Aussi... Aux basilics de ma vie, Mon adorable frère, le généreux et le sage **Mohamed lamine**- la tendre **AMINA**- et dernièrement la Factieuse et la joie de la maison la petite **MALAK** ; que le dieu vous bénisse je vous aime trop, et je vous souhaite une vie pleine de succès et que votre destin soit mieux que le mien Et je vous offre cette carrière*

*Je ne veux pas oublier tous mes oncles, mes tantes et leurs enfants du plus grand au plus petit Mes grandes mères que le Dieu les bénisse, et les garde toujours pour nous la source de '**baraka**'.*

*A mes amies **Abir, Kenza, Zineb et Khadidja**, Je ne peux trouver les mots justes et sincères pour vous exprimer mon affection et mes pensées, vous êtes pour moi des sœurs et des amies sur qui je peux compter. Je vous souhaite une vie pleine de santé et de bonheur.*

*Sans oublier la courageuse qui a supporté avec moi ce défi Qu'avec lui j'ai commencé et j'ai terminé **Wissam**.*

Dédicace et merci pour tout le monde Avec beaucoup d'amour

♡Rayane♡

Résumé

Les modèles de copule sont devenus l'un des outils les plus largement utilisés dans la modélisation appliquée de données multivariées. De même, les méthodes Bayésiennes sont de plus en plus utilisées pour mener une inférence efficace basée sur la vraisemblance. Cependant, à ce jour, il n'y a eu qu'une utilisation limitée des approches Bayésiennes dans la formulation et l'estimation des modèles de copule. Le but de ce travail est de guider le lecteur à se familiariser sur l'estimation Bayésienne en générale et particulièrement l'estimation Bayésienne des copules en donnant un exemple d'une copule archimédienne où les paramètres sont estimés par une méthode MCMC (M-H).

Abstract

Copula models have become one of the most widely used tools in the applied modelling of multivariate data. Similarly, Bayesian methods are increasingly used to obtain efficient likelihood-based inference. However, to date, there has been only limited use of Bayesian approaches in the formulation and estimation of copula models.

The purpose of this work is to guide the reader to familiarize himself with the Bayesian estimation in general and particularly the Bayesian estimation of copulas by giving an example of an Archimedean copula where the parameters are estimated by an MCMC (M-H) method.

Table des matières

Introduction	11
1 Les copules et les mesures de dépendance	14
1.1 La théorie des copules	14
1.1.1 Introduction	14
1.1.2 Définition d'une copule et le théorème d'existence	15
1.1.3 Bornes de Fréchet-Hoeffding	17
1.1.4 Propriétés des copules	18
1.1.5 Densité de la copule	20
1.1.6 Quelques types de copules paramétriques	21
1.2 Les mesures de dépendance	26
1.2.1 Le coefficient de Corrélacion linéaire de Pearson	27
1.2.2 Concepte de concordance	27
1.2.3 Les coefficients de corrélation de Kendall et rho de Spearman .	28
2 Estimation des copules par l'approche Bayésienne	31
2.1 Fondements de la statistique Bayésienne	31
2.1.1 Paradigme Bayésien	31
2.1.2 Théorème de Bayes	32
2.1.3 L'estimation Bayésienne	33
2.1.4 Choix de la distribution a priori	35

2.1.5	Méthodes de Monte-Carlo par Chaîne de Markov	40
2.1.6	Principe général des méthodes MCMC	43
2.2	Approche Bayésienne et copules	45
2.2.1	Exemple d'application	46
	Annexe	52

Table des figures

1.1	Nuages de points des copules Gaussiennes selon ρ	22
1.2	Densité c_ρ des copules Gaussiennes selon ρ	22
1.3	Nuages de points des copules de Student selon $\kappa = 5$	23
1.4	Densité $c_{\rho,\kappa}$ des copules de Student selon $\kappa = 5$	24
1.5	Nuages de points des copules Archimédiennes de paramètres θ	26
1.6	Densité $c_\theta(u, v)$ des copules Archimédiennes de paramètres θ	26
2.1	Nuage des points de la copule de Clayton selon θ	50
2.2	Les densités de la copule de Clayton selon θ	50

Liste des tableaux

1.1	Tau de Kendall et le rho de Spearman de quelques copules	30
2.1	Lois a priori conjuguées pour quelques familles exponentielles	39
2.2	Les valeurs de l'estimateur θ_B et Z_B	49

Introduction

La copule est un outil relativement innovant pour modéliser la structure de dépendance de plusieurs variables aléatoires. La connaissance de cet outil probabiliste est essentielle à l'appréhension de nombreux domaines d'application. Le concept de copule a été introduit par Sklar (1959) pour résoudre un problème de probabilité énoncé par Maurice Fréchet. À l'époque, Sklar et Schweizer travaillent sur les travaux de Menger concernant les espaces métriques aléatoires (Probabilistic Metric Space ou PMS), qui sont une généralisation de l'espace métrique usuel introduit par Fréchet (1906). Même si les copules occupent une place importante dans l'oeuvre de Sklar et Schweizer, elle ne sont pas l'objet central de leurs recherches. L'utilisation des copules par Sklar et Schweizer est assez originale, elles interviennent pour résoudre certains problèmes et ne font pas l'objet véritablement d'études appropriées. Pendant de nombreuses années, les copules sont peu (ou pas) utilisées en statistiques. Il y a les travaux sur la dépendance de Kimeldorf et Sampson dans les années (1975) ou encore les recherches de Paul Deheuvels à la fin des années (1970), mais l'intérêt pour la recherche dans ce domaine s'est plutôt développé durant les années 80 et 90. Par exemple, les copules Archimédiennes apparaissent dans Genest & MacKay (1986), Marshall & Olkin(1988) et ensuite dans Joe (1997). Elles sont également devenues des outils employés couramment dans plusieurs domaines. On cite par exemple les travaux de Li (2000), Cherubini & Luciano (2002), et Bennet & Kennedy (2004). Aussi les copules sont de plus en plus utilisées pour modéliser les distributions multidimensionnelles avec des marges continues en hydrologie (Salvadori, De Michele, Kottegoda, et Rosso 2007), en sciences actuarielles (Frees et Valdez 1998), en finance (Cherubini, Vecchiato, et Luciano 2004; Mc-Neil, Frey, et Embrechts 2005), en marketing (Danaher and Smith 2011), en études de transport (Bhat and Eluru 2009; Smith and Kauermann 2011), en statistiques médicales (Lambert and Vandenhende 2002; Nikoloulopoulos and Karlis 2008) et en économétrie (Smith 2003; Cameron et al. 2004; Patton 2006).

La démarche de la statistique Bayésienne se distingue de celle de la statistique classique par le fait qu'elle représente l'incertitude portant sur le paramètre θ du modèle par une distribution de probabilité nommée "loi a priori" de θ , elle s'interprète

comme la représentation formelle sous forme probabiliste de la connaissance sur le paramètre détenue par l'expert. Cette loi a priori doit être claire et indépendantes des données (Berry, 1996), sinon la même source d'information interviendrait deux fois.

Le fondement de la statistique Bayésienne s'est basé sur le théorème d'inversion des probabilités connu sous le nom du théorème de Bayes (Bayes, 1763). Les résultats de Bayes ont été redécouverts et étendus par Laplace (1774), lequel n'était pas apparemment au fait du travail de Bayes.

Bien que les premiers travaux d'inspiration Bayésienne datent du *XVII^{ème}* siècle, cette méthode connaît un regain de popularité depuis quelques décennies. Ce renouveau est sensible dans des domaines très variés, en grâce à la disponibilité de algorithmes MCMC calculateurs, mais aussi à une évolution de la pensée statistique et des problèmes abordés.

La statistique Bayésienne est très utilisée en sciences sociales et politiques, car les données y sont rares et coûteuses à collecter (Gelman et al., 2004). Elle sert aussi en physique des particules (Cousins, 1995 ; Demortier, 2006), en thermodynamique (Chatterjee et al., 1998), en mécanique statistique (Jaynes, 1957), en chimie (Vines et al., 1993 ; Pohorille and Darve, 2006), en génétique (Smyth, 2004 ; Chan et al., 2006), en bioinformatique (Wilkinson, 2007), en sciences cognitives (Kording, 2004), et en intelligence artificielle (Bessière et al. 1998a,b).

Pour un modèle paramétré bien défini de densité $f(x | \theta)$ où θ est le paramètre indiquant la densité, à valeurs dans l'espace Θ , la loi a priori de θ notée $\pi(\theta)$ représente pour un statisticien Bayésien, l'ensemble des informations a priori disponibles sur le paramètre ainsi que les imprécisions qui s'y rattachent, et dans un contexte pratique, elle regroupe aussi l'ensemble des opinions d'experts.

Le choix de cette loi a priori constitue le point le plus critiqué de l'analyse Bayésienne par les non Bayésiens. Sur le plan pratique de l'approche Bayésienne, il n'existe jamais une unique loi a priori pour θ , mais plutôt un ensemble de lois compatibles avec les informations a priori disponibles et les opinions des modélisateurs. Et c'est bien pour cette cause qu'on a cherché à réduire l'influence de ce choix en retenant des lois a priori à faible contenu informatif, ou des lois dites non informatives. Une fois la loi a priori est construite, le théorème de Bayes rassemble l'information apportée par cette loi a priori avec celle apportée par les données dans une nouvelle distribution dite la distribution a posteriori notée $\pi(\theta | x)$ et qui est la base de toute inférence Bayésienne.

le développement d'une méthodologie d'estimation et d'inférence statistique pour les modèles de copule a été limité. La plupart des recherches ont porté soit sur le développement et les propriétés des fonctions de la copule, soit sur leur utilisation pour résoudre des problèmes appliqués. Moins d'attention a été accordée à la ques-

tion de savoir comment estimer de manière efficace la variété croissante des modèles de copules. À ce jour, les méthodes d'estimation les plus courantes sont la méthode de maximum de vraisemblance (Joe 2005) et la méthode des moments dans le cas des petits dimensions (Genest and Rivest 1993). Il n'y a eu que des travaux limités sur le développement d'approches Bayésiennes pour formuler et estimer des modèles de copule. Cela est surprenant, étant donné que les méthodes Bayésiennes se sont avérées efficaces à la fois pour formuler et estimer des modèles multivariés ailleurs. Le but de ce travail est d'ouvrir des voies pour l'utilisation de l'approche Bayésienne dans des modèles de copules, en illustrant un exemple où les paramètres d'une copule archimédienne sont estimés dans un contexte Bayésien par une méthode MCMC. Dans ce but, ce mémoire à été réalisé en deux parties :

Dans la première partie, nous avons introduit les notions de base de la théorie des copules bidimensionnelles ($d=2$). Tandis que dans la deuxième partie, nous avons d'abord illustrer brièvement les fondaments de la statistique Bayésienne ensuite, nous avons appliquer ces méthodes sur une copule archimédienne.

Chapitre 1

Les copules et les mesures de dépendance

1.1 La théorie des copules

1.1.1 Introduction

Les copules constituent un outil statistique assez récent permettant de modéliser la dépendance entre des variables aléatoires (vas). La fonction copule permet de relier la fonction de densité conjointe aux fonctions de densités marginales.

Dans ce chapitre, nous introduisons quelques définitions de base qui nous permettront de décrire le concept de copule et ces principales propriétés, quelques copules paramétriques les plus utilisées ainsi que leurs propriétés sont également décrites.

La construction des copules repose sur les propriétés des fonctions de répartition (fds). On rappelle ci-dessous quelques propriétés importantes des fds bivariées.

Soit (X, Y) un couple de vas, la loi du couple est caractérisée par la fd bivariée H définie comme ceci

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

On appelle lois marginales, les lois de X et de Y prises séparément. On peut exprimer les fds de ces lois marginales en fonction de H . Par exemple, pour X et Y on obtient

$$F_X(x) = \lim_{y \rightarrow \infty} H(x, y) \tag{1.1}$$

$$G_Y(y) = \lim_{x \rightarrow \infty} H(x, y) \tag{1.2}$$

On rappelle que les vas X et Y sont indépendantes ssi :

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

où F et G sont les fds de X et Y . On note par la suite :

$$F^{-1}(t) = \inf \{x \in \mathbb{R} : F(x) \geq t\}, \text{ pour } t \in [0, 1]$$

la fonction quantile qui est l'inverse de la fd F . On rappelle le lemme d'inversion qui est couramment utilisé pour générer des échantillons d'une loi dont la fonction quantile est connue : si U suit une loi uniforme sur l'intervalle $I = [0, 1]$, alors $F^{-1}(U)$ suit la même loi que X . Les fds des lois marginales sont alors des fonctions continues et les vas $F(x)$ et $G(y)$ suivent des lois uniformes sur l'intervalle I .

La définition suivante introduit la notion des copules.

1.1.2 Définition d'une copule et le théorème d'existence

Définition 1.1. Une copule bivariée est une fonction $C : I^2 \rightarrow I$ qui vérifie les conditions suivantes :

1. La copule C est attachée (grounded), c-à-d :

$$\forall (u, v) \in I^2, C(u, 0) = C(0, v) = 0 \quad (1.3)$$

2. Les marges sont uniformes, c-à-d :

$$\forall (u, v) \in I^2, C(u, 1) = u \text{ et } C(1, v) = v \quad (1.4)$$

3. La copule C est 2-croissante c-à-d : $\forall (u_1, u_2), (v_1, v_2) \in I^2$ avec $u_1 \leq u_2$ et $v_1 \leq v_2$ on a :

$$C(u_1, v_1) - C(u_1, v_2) - C(u_2, v_1) + C(u_2, v_2) \geq 0 \quad (1.5)$$

Exemple 1.2.

- **Copule produit :** La copule produit, noté Π est définie par

$$C(u, v) = \Pi(u, v) = uv, \quad \forall u, v \in I$$

Comme on observe cette copule vérifié les propriétés (1.3) et (1.4), telle que :

$$C(u, 0) = \Pi(u, 0) = u0 = 0 \text{ et } C(0, v) = \Pi(0, v) = 0v = 0,$$

$$C(u, 1) = \Pi(u, 1) = u1 = u \text{ et } C(1, v) = \Pi(1, v) = 1v = v.$$

pour la propriété (1.5), on a :

$$\Pi(u_1, v_1) - \Pi(u_1, v_2) - \Pi(u_2, v_1) + \Pi(u_2, v_2) = u_1v_1 - u_1v_2 - u_2v_1 + u_2v_2 \geq 0$$

où

$$\text{si } u_2 \geq u_1 \implies \begin{cases} u_2 v_2 \geq u_1 v_2 \\ u_2 v_1 \geq u_1 v_1 \end{cases}$$

alors

$$u_2 v_2 - u_2 v_1 \geq u_1 v_2 - u_1 v_1 \implies u_2 v_2 - u_2 v_1 - u_1 v_2 + u_1 v_1 \geq 0$$

- **Copule min** : La copule min, noté M , est définie par :

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

- **Copule max** : La copule max, noté W , est définie par :

$$W(u, v) = \max(u + v - 1, 0), \quad \forall u, v \in I.$$

Théorème d'existence (théorème de Sklar)

Ce théorème représente la base de la théorie de copules, car il valide l'existence de la copule C qui permet de relier les lois marginales de deux vas pour obtenir la distribution jointe. Ce théorème a été développé par Sklar (1959) [9] et porte son nom. La relation entre la distribution jointe, les fonctions de distribution marginales et la copule C est montré dans le théorème suivant.

Théorème 1.3. Soient X, Y deux vas de fd jointe H et des marginales F et G . Alors il existe une copule C telle que

$$\forall (x, y) \in \mathbb{R}^2 : H(x, y) = C(F(x), G(y)) \quad (1.6)$$

Démonstration. Soit C la fd du vecteur (U, V) où $U = F(X)$ et $V = G(Y)$:

$$\begin{aligned} H(x, y) &= P[X \leq x, Y \leq y] \\ &= P[F^{-1}(U) \leq x, G^{-1}(V) \leq y] \\ &= P[U \leq F(x), V \leq G(y)] \\ &= C(F(x), G(y)) \end{aligned}$$

□

Corollaire 1.4. (Inverse de théorème de Sklar)

Soit H une fonction de répartition 2-dimensionnelle de fds marginales F et G . Alors la copule C associée à H est donnée par

$$C(u, v) = H(F^{-1}(u), G^{-1}(v)), \text{ pour tout } (u, v) \in I^2$$

Démonstration. En appliquant (1.6) en posant $U = F(x)$ et $V = G(y)$, on trouve que

$$\begin{aligned} C(u, v) &= P[U \leq u, V \leq v] \\ &= p[F(x) \leq u, G(y) \leq v] \\ &= p[x \leq F^{-1}(u), y \leq G^{-1}(v)] \\ &= H(F^{-1}(u), G^{-1}(v)) \end{aligned}$$

□

Remarque 1. Si F et G sont continues, alors C est unique.

Si non, lorsque les marginales ne sont pas continues, il est toujours possible de définir une copule mais elle ci n'est plus unique.

Exemple 1.5.

Soit la distribution logistique bivariée de Gumbel [5] dont la fd conjointe est donnée par

$$H(x, y) = (1 + e^{-x} + e^{-y})^{-1}, \quad \text{pour tout } (x, y) \in \mathbb{R}^2$$

Des équations (1.1) et (1.2), nous obtenons les fds marginales

$$F(x) = (1 + e^{-x})^{-1}, \quad G(y) = (1 + e^{-y})^{-1}$$

Comme les fonctions inverses sont respectivement

$$F^{-1}(u) = -\ln\left(\frac{1}{u} - 1\right), \quad G^{-1}(v) = -\ln\left(\frac{1}{v} - 1\right)$$

En se basant sur le théorème de sklar, la copule associée à $H(x, y)$ s'écrit :

$$\begin{aligned} C(u, v) &= H(F^{-1}(u), G^{-1}(v)) \\ &= [1 + e^{\ln(\frac{1}{u}-1)} + e^{\ln(\frac{1}{v}-1)}]^{-1} \\ &= \frac{uv}{u + v - uv} \end{aligned}$$

1.1.3 Bornes de Fréchet-Hoeffding

Toute copule bivariée C est bornée par deux copules. Ces bornes sont déterminées dans le théorèmes suivant

Théorème 1.6. Soit C une copule alors :

$$\forall (u, v) \in I^2 : W(u, v) = \max(u + v - 1, 0) \leq C(u, v) \leq \min(u, v) = M(u, v)$$

telles que :

Les copules W et M sont appelées borne inférieure (respectivement borne supérieure) de Fréchet-Hoeffding.

Démonstration. Soit $(u, v) \in I^2$ on a :

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

d'où la deuxième égalité.

Pour montrer la première égalité, nous utilisons le fait que $C(u, v) \geq 0$ et l'inégalité (1.5) :

$$\begin{aligned} C(1, 1) - C(1, v) - C(u, 1) + C(u, v) &\geq 0 \\ \Rightarrow 1 - v - u + C(u, v) &\geq 0 \\ \Rightarrow C(u, v) &\geq u + v - 1 \\ \Rightarrow C(u, v) &\geq \max(u + v - 1, 0) \\ \Rightarrow C(u, v) &\geq W(u, v) \end{aligned}$$

Donc : $W(u, v) \leq C(u, v) \leq M(u, v)$ □

Comme une conséquence de théorème de Sklar, si X, Y sont des vas dont la fd jointe est H et des marginales F et G , respectivement, alors pour tout $x, y \in \mathbb{R}$

$$\max(F(x) + G(y) - 1, 0) \leq H(x, y) \leq \min(F(x), G(y)).$$

Car M et W sont des copules, les bornes ci-dessus sont des fds jointes et sont appelés les bornes de Fréchet-Hoeffding pour une fd jointe H et des marginales F et G .

1.1.4 Propriétés des copules

On note tout d'abord que les copules vérifient certaines propriétés.

1. la continuité [8]

On note que les copules sont des fonctions continues. Plus précisément, elles vérifient une condition de Lipschitz. $\forall u_1, u_2, v_1, v_2 \in I$ avec $u_1 \leq u_2$ et $v_1 \leq v_2$, on a :

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

2. L'ordre

Soit C_1, 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 $C_1 < C_2$ si

$$\forall (u, v) \in I^2, C_1(u, v) < C_2(u, v)$$

comme exemple on a La copule $W = \max(u + v - 1, 0)$ est la plus petite copule et $M = \min(u, v)$ est la plus grande copule.

3. la symétrie

Soit X, Y deux vas continues de fd jointe H et de marginales F et G , et soit C la copule associe. On dit que X, Y sont échangeables ssi : $F = G$ et $\forall (u, v) \in I^2$ si $C(u, v) = C(v, u)$ on dit que C est symétrique.

4. la convexité et concavité

Soit $(a, b), (c, d) \in I^2$ et $\lambda \in I$.

Une copule C est **concave** si on a :

$$C(\lambda a + (1 - \lambda)c, \lambda b + (1 - \lambda)d) \geq \lambda C(a, b) + (1 - \lambda)C(c, d)$$

Une copule C est **convexe** si on a :

$$C(\lambda a + (1 - \lambda)c, \lambda b + (1 - \lambda)d) \leq \lambda C(a, b) + (1 - \lambda)C(c, d)$$

5. la différentiabilité [3]

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

- les dérivées partielles $\frac{\partial C(u_1, u_2)}{\partial u_j}$ existe p.s et $0 \leq \frac{\partial C(u_1, u_2)}{\partial u_j} \leq 1, \forall j = 1, 2$.

- 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.

6. Théorème d'invariance

L'un des théorèmes essentiels à la théorie des copules est celui de l'invariance par transformations strictement croissantes.

Soient deux vas continues X et Y de marges F et G et de copule C_{XY} . Si α et β sont deux fonctions strictement croissantes, alors :

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

Démonstration. On peut démontrer ce théorème facilement à l'aide de lois de probabilité, comme suit :

Soient F_1, G_1, F_2, G_2 les fds de $X, Y, \alpha(X)$ et $\beta(Y)$, respectivement. D'où α et β sont strictement croissantes, alors :

$$F_2 = P[\alpha(X) \leq x] = P[X \leq \alpha^{-1}(x)] = F_1(\alpha^{-1}(x))$$

aussi :

$$G_2 = P[\beta(Y) \leq y] = P[Y \leq \beta^{-1}(y)] = G_1(\beta^{-1}(y))$$

on a donc :

$$\begin{aligned} C_{\alpha(X), \beta(Y)}(F_2(x), G_2(y)) &= P[\alpha(X) \leq x, \beta(Y) \leq y] \\ &= P[X \leq \alpha^{-1}(x), Y \leq \beta^{-1}(y)] \\ &= C_{XY}(F_1(\alpha^{-1}(x)), G_1(\beta^{-1}(y))) \\ &= C_{XY}(F_2(x), G_2(y)) \end{aligned}$$

alors toute copule vérifie cette propriété. □

Exemple 1.7. nous avons :

$$\begin{aligned}
 C_{XY} &= C_{\ln.XY} \\
 &= C_{\ln.X \ln.Y} \\
 &= C_{X \exp(Y)} \\
 &= C_{\sqrt{X} \exp(Y)} \\
 &\cdot \\
 &\cdot \\
 &\cdot
 \end{aligned}$$

Une transformation croissante ne modifie donc pas la copule, mais seulement les lois marginales.

7. Théorème d'indépendance

Soient X et Y deux vas continues et C_{XY} la copule associée.

Alors X et Y sont indépendantes ssi $C_{XY}(u, v) = uv = \Pi(u, v)$ pour tout $u, v \in I$.

Démonstration.

- Si X et Y sont indépendantes alors : $\forall x, y \in \mathbb{R}, H(x, y) = F(x)G(y)$

$$\begin{aligned}
 \text{Or } \forall x, y \in I, C_{XY}(u, v) &= H(F^{-1}(u), G^{-1}(v)) \\
 &= F(F^{-1}(u))G(G^{-1}(v)) \\
 &= uv = \Pi(u, v)
 \end{aligned}$$

- Si $C_{XY} = \Pi$ alors :

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

\implies Les vas X et Y sont donc indépendantes. □

1.1.5 Densité de la copule

Les copules admettent des densités de probabilités. Si la densité c associée à la copule C existe, alors elle est définie par :

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

où $c : I^2 \rightarrow \mathbb{R}^+$. Si la fonction de distribution jointe H est absolument continue, en utilisant le théorème de Sklar, on peut exprimer la densité d'un couple aléatoire (X, Y) en fonction de la densité de sa copule et de ses marginales f et g par :

$$h(x, y) = c(F(x), G(y)) f(x)g(y).$$

1.1.6 Quelques types de copules paramétriques

Il existe un grand nombre de fonctions copules, parmi elles, les plus importantes sont les copules elliptiques et les copules Achimédiennes.

Les copules elliptiques

En travaillant dans le cadre des copules multivariées, on a été confronté dès le début à un problème. Toute la littérature abondante de ces dernières années sur les copules était entièrement dévolue au cas bivarié et lorsque le nombre de dimensions est supérieur à deux, les auteurs se ramènent toujours à une copule elliptique (normale ou Student) pour la simple et bonne raison que les densités de ces copules sont faciles à calculer et donc plus aisées à simuler. Les copules elliptiques sont des copules associées aux distributions elliptiques.

Définition 1.8. *On appelle copule elliptique toute copule qui s'écrit de la forme suivante :*

$$C_\rho(u, v) = \frac{1}{\sqrt{1-\rho^2}} \int_{-\infty}^{\varphi_{g,1}^{-1}(u)} \int_{-\infty}^{\varphi_{g,2}^{-1}(v)} g\left(\frac{x^2 - 2\rho xy + y^2}{\sqrt{1-\rho^2}}\right) dx dy = \varphi_\rho(\varphi_{g,1}^{-1}(u), \varphi_{g,2}^{-1}(v)) ,$$

où φ_ρ est la distribution jointe des vas X et Y , $\varphi_{g,1}^{-1}(u)$ et $\varphi_{g,2}^{-1}(v)$ sont leurs fonctions quantiles respectives et ρ leurs coefficients de corrélation ($\rho \in [-1, 1]$).

Copule gaussienne

La copule gaussienne fait partie de la famille des copules elliptiques des copules bivariées à un paramètre utiliser cette copule est conséquent avec la mesure de dépendance obtenue par le coefficient de corrélation linéaire.

Il s'agit d'une propriété contraignante lorsqu'on veut évaluer la dépendance entre les événements rares. Un des types de copules les plus utilisées dans la modélisation est la copule normale bivariée.

Définition 1.9. *soient X, Y deux vas Gaussiennes de moyenne μ et de matrice de*

covariance $\Sigma = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$ la copule Gaussienne est définie par :

$$C_\rho(u, v) = \varphi_\rho(\varphi^{-1}(u), \varphi^{-1}(v)) \\ = \int_{-\infty}^{\varphi^{-1}(u)} \int_{-\infty}^{\varphi^{-1}(v)} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{x^2 + y^2 - 2\rho xy}{2(1-\rho^2)}\right) dx dy$$

avec $x = \varphi^{-1}(u)$ et $y = \varphi^{-1}(v)$.

La densité de cette copule :

$$c(u, v) = \frac{1}{\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}(\rho^2 x^2 + \rho^2 y^2 - 2\rho xy)\right\}$$

Cette copule est paramétrée par le coefficient de corrélation linéaire ρ .

- Si $\rho = 0$ alors $C_\rho(u, v) = \Pi$.
- Si $\rho = -1$ alors $C_\rho(u, v) = W(u, v)$.
- Si $\rho = 1$ alors $C_\rho(u, v) = M(u, v)$.

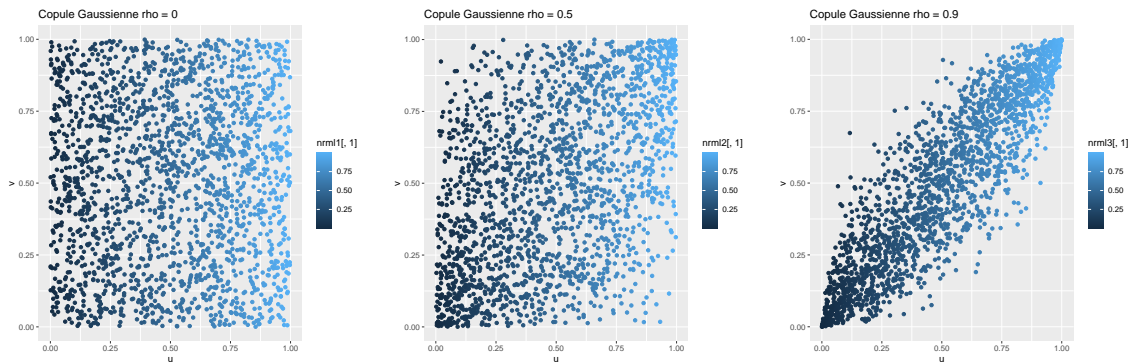


FIGURE 1.1 – Nuages de points des copules Gaussiennes selon ρ

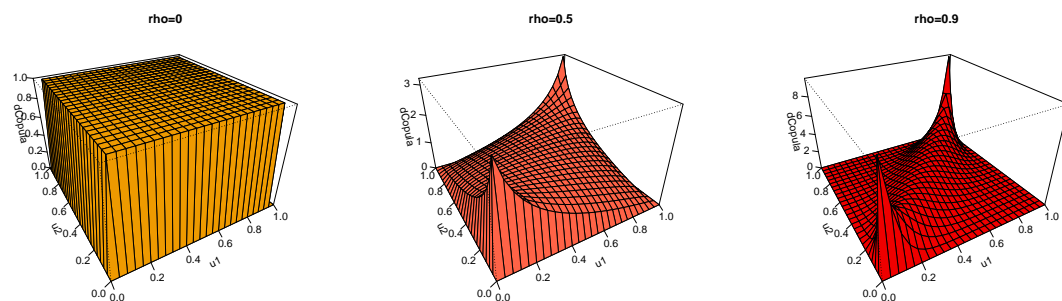


FIGURE 1.2 – Densité c_ρ des copules Gaussiennes selon ρ

Copule de Student

La copule de Student (t copula) est la copule sousjacente à une distribution bivariée de Student. Elle est définie de la même manière que la copule Gaussienne mais à partir de la distribution de Student centrée réduite. en outre lorsque le degré de liberté tend vers l'infini, la copule de Student est égale à la copule Gaussienne.

Définition 1.10. Soit $\rho \in [-1, 1]$, alors la fd de Student à degré κ de liberté est définie par :

$$t_{\rho, \kappa}(x, y) = \int_{-\infty}^x \int_{-\infty}^y \frac{1}{2\pi\sqrt{1-\rho^2}} \left(1 + \frac{s^2 + t^2 - 2\rho st}{\kappa(1-\rho^2)} \right)^{-\frac{(\kappa+2)}{2}} ds dt$$

Définition 1.11. La copule de Student est une copule paramétrique, paramétrée par le coefficient de corrélation linéaire ρ et le degré de liberté κ . Cette copule est définie par :

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

la densité de cette copule :

$$c_{\rho, \kappa}(u, v) = \frac{\kappa}{2\sqrt{1-\rho^2}} \frac{\Gamma(\kappa/2)^2}{\Gamma((\kappa+1)/2)^2} \frac{\left[1 + \frac{x^2 + y^2 - 2\rho xy}{\kappa(1-\rho^2)} \right]^{-\frac{(\kappa+2)}{2}}}{\left[1 + \frac{x^2}{\kappa} \right] \left[1 + \frac{y^2}{\kappa} \right]^{-\frac{(\kappa+2)}{2}}}$$

où Γ représente la fonction gamma, avec $x = t_{\kappa}^{-1}(u)$ et $y = t_{\kappa}^{-1}(v)$

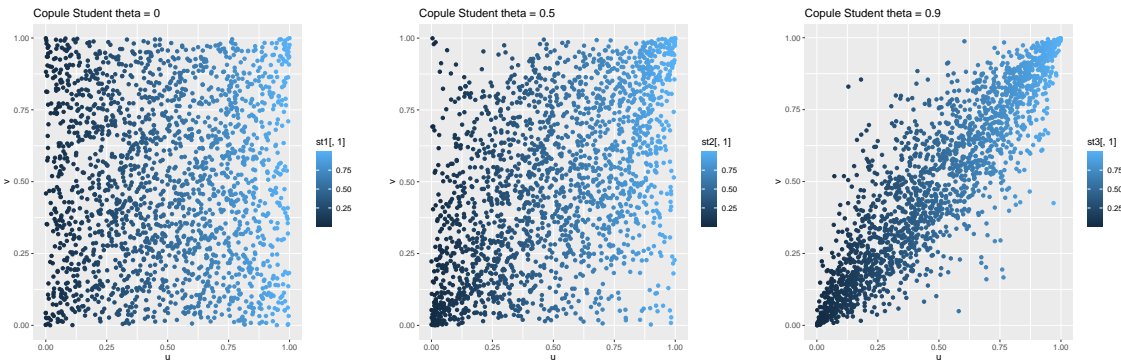


FIGURE 1.3 – Nuages de points des copules de Student selon $\kappa = 5$

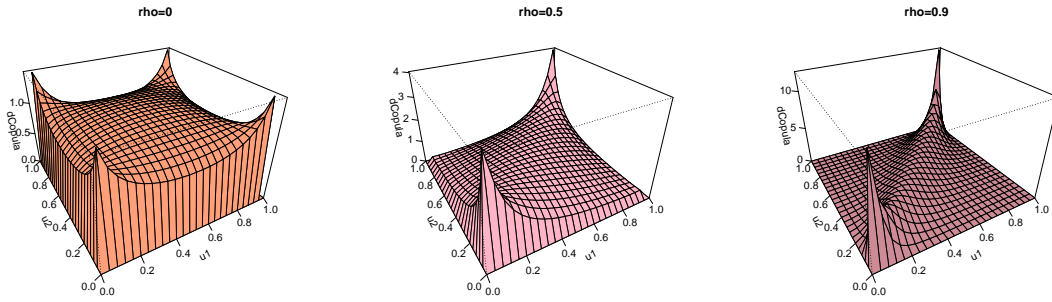


FIGURE 1.4 – Densité $c_{\rho, \kappa}$ des copules de Student selon $\kappa = 5$

Remarque 2. *Les copules elliptiques sont des copules symétriques et relativement simples d'utilisation du fait que l'on connaît bien les distributions auxquelles elles sont associées. Elles sont souvent appelées copules implicites car n'ayant pas de forme analytique explicite et s'expriment par conséquent en fonction de leurs distributions bivariées.*

Les copules Archimédiennes

La classe des copules Archimédiennes, définies par Genest et Mackay [1986] [4], joue un rôle important. D'une part, elles permettent de construire une grande variété de familles de copules, et donc de 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. En effet, contrairement aux copules gaussiennes et aux copules de Student, les copules archimédiennes ont le grand avantage de décrire des structures de dépendance.

Pour plus d'éléments sur cette famille de copules le lecteur peut se référer à l'excellent ouvrage de Nelsen [2006] [7].

Plusieurs raisons justifient l'utilisation de ce type de copules entre autres :

- Grande variété de familles paramétriques.
- Les propriétés particulières et intéressantes que cette classe possède.
- La facilité avec laquelle peuvent être construites et simulées.
- La grande variété des différentes structures de dépendance.

La définition suivante décrit la forme générale des copules Archimédiennes.

Définition 1.12. *On appelle copule Archimédienne de générateur ϕ la copule définie par :*

$$C(u, v) = \phi^{-1}(\phi(u) + \phi(v))$$

Avec $\phi : I \rightarrow [0, +\infty[$ une fonction continue, strictement décroissante et convexe,

vérifiant $\phi(1) = 0$. On définit l'inverse de ϕ par $\phi^{[-1]}$ tel que :

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

Si $\phi(0) = \infty$ alors ϕ est strictement décroissante.

ϕ est au moins deux fois continument dérivable telle que $\phi'(u) < 0$ et $\phi''(u) > 0$ pour tout $u \in I$.

Comme exemple on prend la fonction $\phi(t) = -\ln(t)$ comme générateur et le pseudo inverse $\phi^{-1}(t) = \exp(-t)$, et on construit la copule C comme suit

$$C(u, v) = \exp(-[(-\ln(u)) + (-\ln(v))]) = uv = \Pi(u, v)$$

Famille des copules archimédiennes

Copule de Gumbel

Soit la fonction de générateur $\phi(t) = (-\ln t)^\theta$, avec $\theta \geq 1$ et son pseudo-inverse $\phi^{-1}(t) = \exp\left(-t^{\frac{1}{\theta}}\right)$. On définit la famille de Gumbel comme suit :

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

- Si $\theta \rightarrow 1$ alors $C_\theta \rightarrow \Pi$.
- Si $\theta \rightarrow \infty$ alors $C_\theta \rightarrow M$.

Pour cette famille la variable latente $Z \sim \text{loi}_{\frac{1}{\theta}}$ -stable où $\theta \in [1, \infty[$.

Copule de Clayton

Soit la fonction de générateur $\phi(t) = \frac{1}{\theta}(t^{-\theta} - 1)$, avec $\theta \in [-1, 0[\cup]0, \infty[$ et son pseudo-inverse $\phi^{-1}(t) = (t + 1)^{-\frac{1}{\theta}}$. On définit la famille de Clayton comme suit :

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

- Si $\theta \rightarrow -1$ alors $C_\theta = W$.
- Si $\theta \rightarrow 0$ alors $C_\theta = \Pi$.
- Si $\theta \rightarrow \infty$ alors $C_\theta = M$.

Pour cette famille la variable latente $Z \sim \Gamma\left(\frac{1}{\theta}, 1\right)$ où $\theta \in]0, \infty[$.

Copule de Frank

Soit la fonction de générateur $\phi(t) = -\ln\left[\frac{e^{-\theta t} - 1}{e^{-\theta} - 1}\right]$, avec $\theta \in [-\infty, 0[\cup]0, \infty[$ et son pseudo-inverse $\phi^{-1}(t) = -\frac{1}{\theta} \log\left(1 - (1 - e^{-\theta})e^{-t}\right)$. On définit la famille de Frank comme suit :

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

- Si $\theta \rightarrow -\infty$ alors $C_\theta \rightarrow W$.
- Si $\theta \rightarrow +\infty$ alors $C_\theta \rightarrow M$.
- Si $\theta \rightarrow 0$ alors $C_\theta \rightarrow \Pi$.

Pour cette famille la variable latente $Z \sim \log(1 - e^{-\theta})$ où $\theta \in]0, \infty[$.

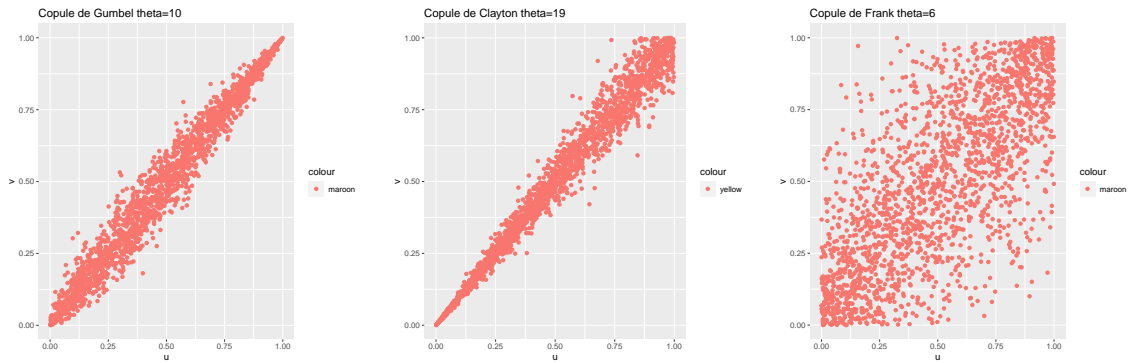


FIGURE 1.5 – Nuages de points des copules Archimédiennes de paramètres θ

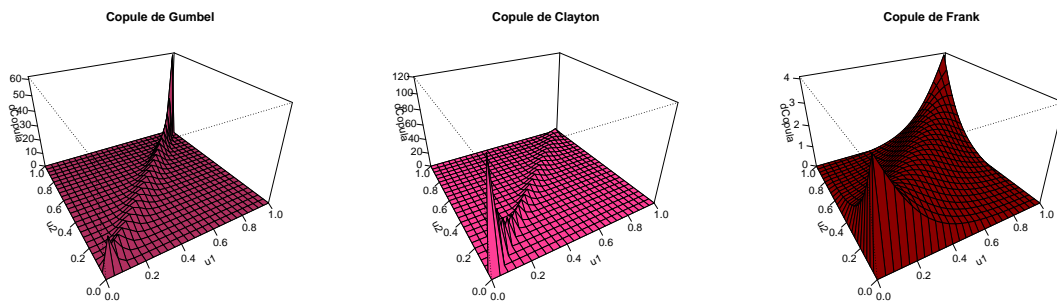


FIGURE 1.6 – Densité $c_\theta(u, v)$ des copules Archimédiennes de paramètres θ

Nous avons utiliser les packages "copula", "scatterplot3d", "ggplot2" et "grid" de langage R pour représenter les nuages de points et densités des copules.

1.2 Les mesures de dépendance

Une des premières mesures de dépendance introduite dans l'étude des statistiques est probablement le coefficient de corrélation de Pearson, une mesure de dépendance linéaire entre deux vas quantitatives. Par contre, lorsque nous travaillons avec des vas continues non normales, l'utilisation de notions basées sur la linéarité n'est pas toujours appropriée. De plus, contrairement au coefficient de corrélation de Pearson, certaines mesures sont invariantes sous transformations strictement croissantes et valent 1 pour la borne supérieure de Fréchet et -1 pour la borne inférieure de Fréchet, ce qui, en ce sens, les rendent plus intéressantes. C'est le cas entre autres

du tau de Kendall τ et du rhô Spearman ρ , deux mesures d'association bivariées qui mesurent la concordance.

1.2.1 Le coefficient de Corrélacion linéaire de Pearson

Soient X et Y deux vas. La covariance de X et Y est défini sous la forme suivante :

$$Cov(X, Y) = E(XY) - E(X)E(Y)$$

Définition 1.13. Soient X et Y deux vas continues des variances finies $var(X)$ et $var(Y)$ respectivement le coefficient de pearson entre X et Y est défini par :

$$\rho(x, y) = \frac{cov(x, y)}{\sqrt{var(X)var(Y)}}, \rho \in [-1, 1]$$

Remarque 3. Il est important de rappeler que la dépendance et la corrélation sont des notions différentes. En effet, si X et Y sont des variables indépendantes elles sont non corrélées mais la réciproque est fausse .

$$X \text{ et } Y \text{ sont indépendantes} \implies \rho = 0$$

1.2.2 Concepte de concordance

Concordance et discordance

Définition 1.14. Soient $(X_1, Y_1), (X_2, Y_2)$ deux observations d'un couple de vas continus (X, Y) sont dites que concordantes ou discordantes si :

- **Concordantes** : $(x_1 - x_2)(y_1 - y_2) > 0 \iff (x_1 < x_2 \text{ et } y_1 < y_2) \text{ ou } (x_1 > x_2 \text{ et } y_1 > y_2)$.
- **Discordantes** : $(x_1 - x_2)(y_1 - y_2) < 0 \iff (x_1 < x_2 \text{ et } y_1 > y_2) \text{ ou } (x_1 > x_2 \text{ et } y_1 < y_2)$.

Définition 1.15. (Fonction de concordance)

La fonction de concordance est la différence entre la probabilité de concordance et celle de discordance entre deux couples (X_1, Y_1) et (X_2, Y_2) : Elle est donnée par :

$$Q = P[(X_1 - X_2)(Y_1 - Y_2) > 0] - P[(X_1 - X_2)(Y_1 - Y_2) < 0]$$

Comme les vas sont continues, donc

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

Avec

$$P[(X_1 - X_2)(Y_1 - Y_2) > 0] = P[X_1 > X_2, Y_1 > Y_2] + P[X_1 < X_2, Y_1 < Y_2]$$

Le résultat suivant permet d'établir une relation entre la fonction de concordance Q et les copules associées aux deux couples.

Théorème 1.16. *Soient (X_1, Y_1) et (X_2, Y_2) deux vecteurs aléatoires indépendantes de vas continues de fd jointes H_1 et H_2 respectivement. Soient F et G les marges associées à X_1, X_2 et Y_1, Y_2 respectivement. Soient C_1 et C_2 les copules associées à H_1 et H_1 données par :*

$$H_1(x, y) = C_1(F(x), G(y))$$

$$H_2(x, y) = C_2(F(x), G(y))$$

alors

$$Q = Q(C_1, C_2) = 4 \int_0^1 \int_0^1 C_2(u, v) dC_1(u, v) - 1$$

1.2.3 Les coefficients de corrélation de Kendall et rho de Spearman

Le taux de Kendall et le rho de Spearman sont deux mesures de concordance bien connues en statistique. Elles donnent une mesure de la corrélation entre les rangs des observations, à la différence du coefficient de corrélation linéaire qui lui mesure la corrélation entre les valeurs des observations. Elles offrent par ailleurs l'avantage de s'exprimer simplement en fonction de la copule associée au couple de vas.

On commence par :

Le tau de kendall

Le tau de Kendall est une mesure de la dépendance entre deux vas, mais contrairement au coefficient de corrélation linéaire, il ne dépend pas des lois des deux variables étudiés mais que de leur ordre. Il est d'une importance fondamentale dans l'étude des copules, car il existe une relation analytique entre le taux de Kendall et le copule.

Définition 1.17. *Soient (X_1, Y_1) et (X_2, Y_2) deux vecteurs aléatoires continues, i.i.d de fds jointes H . Le tau de Kendall du vecteur aléatoire (X, Y) , noté τ_{XY} est défini par :*

$$\tau_{XY} = P[(X_1 - X_2)(Y_1 - Y_2) > 0] - P[(X_1 - X_2)(Y_1 - Y_2) < 0]$$

Définition 1.18. Soit (X, Y) un couple de vas continues de copule C . Si les couples sont identiquement distribués i.e $H_1 = H_2 = H$, le tau de Kendall est défini par :

$$\tau_C = Q(C, C) = 4 \int_0^1 \int_0^1 C(u, v) dC(u, v) - 1$$

Comme les couples sont uniformément distribuées sur I , alors :

$$\int_0^1 \int_0^1 C(u, v) dC(u, v) = E[C(u, v)]$$

Ainsi, on peut écrire :

$$\tau_C = 4E[C(u, v)] - 1$$

Remarque 4. Soient X et Y deux vas d'une copule Archimédienne C générée par ϕ . Alors le tau de Kendall τ_C est donne par :

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

Rho de spearman

Le coefficient de corrélation rho de Spearman représente l'une des mesures les plus connues pour quantifier le degré d'association entre deux vas. La valeur de ce coefficient dénotée par ρ est équivalente au coefficient de corrélation de Pearson. Il a été développé par Spearman. Le rho de Spearman de deux vas X et Y est égal au coefficient de corrélation entre les variables $F(X)$ et $G(Y)$ distribuées selon la loi uniforme sur l'intervalle I tel que :

$$\rho_s(X, Y) = \rho(F(X), G(Y))$$

Définition 1.19. Soient $(X_1, Y_1), (X_2, Y_2)$ et (X_3, Y_3) des copies indépendantes de vecteur aléatoire (X, Y) . Le rho de Spearman[6] noté ρ_s , est définie par :

$$\rho_s = 3(P[(X_1 - X_2)(Y_1 - Y_3) > 0] - P[(X_1 - X_2)(Y_1 - Y_3) < 0])$$

Comme le tau de Kendall, nous pouvons définir le rho de Spearman en fonction d'une copule C .

Théorème 1.20. Soient X, Y deux vas continues dont la copule est C . Le rho de Spearman de X et Y est défini par :

$$\begin{aligned} \rho_C &= 3Q(C, \Pi) = 3 \left(4 \int_0^1 \int_0^1 C(u, v) d\Pi(u, v) - 1 \right) \\ &= 12 \int_0^1 \int_0^1 C(u, v) d(uv) - 3 \\ &= 12 \int_0^1 \int_0^1 C(u, v) dudv - 3 \end{aligned}$$

Le tableau suivant présente le tau de Kendall et le rho de Spearman de quelques copules.

Copules	τ_C	ρ_C
Clayton	$\frac{\theta}{\theta+2}$	—
Gumbel	$\frac{\theta-1}{\theta}$	—
Frank	$1 - \left[\frac{4(1-D_1(\theta))}{\theta} \right]$	$1 - \left[\frac{12(D_1(\theta)-D_2(\theta))}{\theta} \right]$
Gaussienne	$\frac{2 \arcsin(\sigma)}{\pi}$	$\frac{6 \arcsin\left(\frac{\sigma}{2}\right)}{\pi}$

TABLE 1.1 – Tau de Kendall et le rho de Spearman de quelques copules

Avec $D_k(\theta)$: Fonction de Debye est égale a $\frac{k}{\theta^k} \int_0^\theta \frac{t^k}{\exp(t)-1} dt$.

Chapitre 2

Estimation des copules par l'approche Bayésienne

2.1 Fondements de la statistique Bayésienne

la statistique Bayésienne est une approche cohérente, moderne, pratique pour résoudre les problèmes d'inférence statistique et surtout adaptée aux outils informatiques de simulation et apte à répondre aux problèmes de modélisation les plus avancés dans toutes les disciplines.

L'inférence Bayésienne est une méthode qui permet de donner une crédibilité à nos croyances en s'appuyant sur nos observations et nos a priori.

Elle consiste à actualiser les croyances a priori sur un événement à la lumière des nouvelles informations disponibles pour obtenir une description quantitative des connaissances actuelles.

2.1.1 Paradigme Bayésien

On dispose d'un modèle paramétrique d'observation $x \sim f(x | \theta)$, où $\theta \in \Theta$, un espace de dimension finie. Et on dispose aussi d'une information a priori sur le paramètre θ . Cette information prend la forme d'une loi sur l'espace des paramètres Θ notée π qui s'appelle la loi a priori. Le paramètre θ n'est plus considéré comme inconnu et déterministe, mais comme une va et on note $\theta \sim \pi$. Ainsi la notion de probabilité ou densité de probabilité paramétré par θ n'a plus vraiment de sens. les notions de l'approche fréquentiste sont remplacées par des notions de probabilité d'indépendance et de densité de probabilité conditionnelles à θ .

Le concept fondamental du paradigme Bayésien est la distribution a posteriori

$\pi(\theta | x)$ qui est un résumé complet de l'information sur le paramètre θ . En effet, une fois que X est observé, l'approche Bayésienne réalise en quelque sorte l'actualisation de l'information a priori par l'observation X , au travers de $\pi(\theta | x)$.

Dans le reste de ce chapitre nous allons énoncer le théorème de Bayes, aborder la question de l'inférence Bayésienne sur un paramètre θ , citer quelque méthode de choix a priori et en fin donner le principe de fonctionnement des algorithmes MCMC : M-H, Gibbs.

2.1.2 Théorème de Bayes

Soient A et B deux événements aléatoires tels que, $P(B) \neq 0$. $P(A | B)$ et $P(B | A)$ se définissent respectivement comme étant la probabilité de A , conditionnellement à la réalisation de B et la probabilité de B , conditionnellement à la réalisation de A tels que :

$$P(A | B) = \frac{P(A, B)}{P(B)} \quad \text{et} \quad P(B | A) = \frac{P(B, A)}{P(A)}$$

où $P(A, B) = P(B, A)$ est la probabilité que les deux événements A et B aient lieu simultanément.

En remplaçant $P(A | B)$ par $P(B | A)$ qui est égale à $P(B | A) P(A)$ on déduit la relation entre les deux probabilités conditionnelles $P(A | B)$ et $P(B | A)$:

$$P(A | B) = \frac{P(A) P(B | A)}{P(B)}$$

Cette relation appelée théorème de Bayes est en fait un principe d'actualisation : elle décrit la mise à jour de la vraisemblance de $P(A)$ vers $P(A | B)$ une fois que B a été observé.

Bayes(1763) donne une version continue de ce résultat : pour deux v.a.s X et Y de distribution conditionnelle $f(x | y)$ et marginale $g(y)$, la distribution conditionnelle de y sachant x est :

$$g(y | x) = \frac{f(x | y)g(y)}{\int f(x | y)g(y)dy}$$

Ce théorème d'inversion est naturel d'un point de vue probabiliste mais Bayes et Laplace sont allés plus loin et ont considéré que l'incertitude sur le paramètre d'un modèle peut être décrite par une distribution de probabilité π sur Θ appelée distribution a priori.

L'inférence est alors fondée sur la distribution de θ conditionnelle à x , notée $\pi(\theta | x)$, et appelée distribution a posteriori et est définie par :

$$\pi(\theta | x) = \frac{f(x | \theta)\pi(\theta)}{\int_{\Theta} f(x | \theta)\pi(\theta) d\theta} = \frac{f(x | \theta)\pi(\theta)}{m(x)}$$

La quantité $m(x) = \int_{\Theta} f(x | \theta) \pi(\theta) d\theta$ est la loi marginale de x et est également appelée constante d'intégration.

Comme $m(x)$ est indépendante de θ donc on peut écrire la loi a posteriori de la manière suivante :

$$\pi(\theta | x) \propto f(x | \theta) \pi(\theta)$$

Le passage de la distribution a priori à la distribution a posteriori des paramètres du modèle statistique, exprimé par la formule de Bayes peut être interprété comme la mise à jour (actualisation) de la connaissance (information) sur la base des observations.

2.1.3 L'estimation Bayésienne

Nous présentons dans le reste de ce chapitre, les notions de base pour mener une estimation Bayésienne sur un paramètre θ , en supposant que la loi a posteriori est connue.

L'estimation ponctuelle

La loi a posteriori est une loi centrale pour l'inférence Bayésienne en ce qu'elle suffit à déterminer les procédures de décision et par extension à conduire toute inférence liée à θ . Elle joue le rôle d'un résumé exhaustif de l'information disponible sur le paramètre θ . En pratique, l'estimation d'un paramètre θ conduit à une prise de décision finale par le "décideur". La décision concerne alors le choix d'un estimateur ponctuel $\hat{\theta}$. Ce choix se fait en comparant les approximations de θ au moyen d'une fonction de coût $L(\delta, \theta)$ (à valeurs dans $[0, \infty[$), qui quantifie les conséquences de l'erreur commise en remplaçant le "vrai" paramètre θ par son approximation δ . Une fois construite la loi a posteriori, les approximations δ ont un coût moyen égal à $E^{\pi} [L(\delta, \theta) | x]$ (le coût a posteriori) qui signifie l'espérance sous $\pi(\theta | x)$.

Afin de déterminer l'estimateur $\hat{\theta}$ optimal, il suffit de minimiser cette erreur.

On définit donc l'estimateur Bayésien $\delta(x)$ comme la procédure qui à chaque observation associe la solution du problème de minimisation

$$\delta(x) = \arg \min_{\delta \in D} E^{\pi} [L(\delta, \theta) | x]$$

À noter qu'un grand nombre de fonctions de coût différentes sont possibles, et que chacune d'entre elles résulte en un estimateur ponctuel optimal différent et donc une décision optimale spécifique.

On peut prendre comme exemple la fonction de coût la plus couramment utilisée qui

est le coût (la perte) quadratique

$$L(\delta, \theta) = (\delta - \theta)^2$$

sous cette fonction de perte, l'estimateur Bayésien, s'il existe, est l'espérance de la loi a posteriori $\delta(x) = E^\pi[\theta | x]$.

En effet soit :

$$\begin{aligned}\rho(\pi, \delta | x) &= E^\pi[L(\delta, \theta) | x] \\ &= \int_{\Theta} L(\delta, \theta) \pi(\theta | x) d\theta\end{aligned}$$

On remplace $L(\delta, \theta)$ par sa valeur :

$$\rho(\pi, \delta | x) = \int_{\Theta} (\delta - \theta)^2 \pi(\theta | x) d\theta$$

Puis on minimise :

$$\begin{aligned}\frac{\partial \rho(\pi, \delta | x)}{\partial \delta} &= \int_{\Theta} \frac{\partial}{\partial \delta} (\delta - \theta)^2 \pi(\theta | x) d\theta = 0 \\ \Rightarrow \int_{\Theta} -2(\delta - \theta) \pi(\theta | x) d\theta &= 0 \\ \Rightarrow -2 \int_{\Theta} \theta \pi(\theta | x) d\theta + 2\delta \int_{\Theta} \pi(\theta | x) d\theta &= 0 \\ \Rightarrow -2E^\pi(\theta | x) + 2\delta &= 0 \\ \Rightarrow \delta &= E^\pi(\theta | x)\end{aligned}$$

Le minimum du coût a posteriori est effectivement atteint par :

$$\delta(x) = E^\pi(\theta | x)$$

Exemple 2.1.

Si X est une va suivant une loi de Pareto de paramètre θ , $X \sim P(\theta)$, et si $\pi(\theta)$ est la loi a priori qui est une loi Gamma de paramètres (α, β) , alors la loi a posteriori $\pi(\theta | x)$ est aussi une loi gamma de paramètres $(x + \alpha, \beta + 1)$.

Sous l'hypothèse d'un coût quadratique, un estimateur de Bayes $\delta(x)$ de δ sera l'espérance a posteriori de δ . Puisque la loi a posteriori est une loi Gamma, l'espérance est le rapport des paramètres et on a :

$$\delta(x) = \frac{x + \alpha}{\beta + 1}$$

Estimation par Intervalles

La connaissance de la loi a posteriori peut aussi conduire à déterminer un ensemble de valeurs ayant une forte probabilité a posteriori. Un tel ensemble est appelé région de crédibilité c-à-d des régions de la forme $\{\theta, \pi(\theta | x) \geq k\}$.

Si la loi a posteriori est unimodale ; cette région est un intervalle. Par exemple un intervalle de crédibilité à 95% est un intervalle $[t_{\text{inf}}, t_{\text{sup}}]$ tel que $\int_{t_{\text{inf}}}^{t_{\text{sup}}} \pi(\theta | x) d\theta = 0.95$. En général on est intéressé par l'intervalle de crédibilité à 95% le plus étroit possible.

L'estimateur MAP

On appelle estimateur MAP (estimateur de maximum a posteriori) tout estimateur $\delta(x)$ qui maximise l'information sur θ représentée par sa loi a posteriori, c'est-à-dire tout estimateur

$$\delta(x) \in \arg \max_{\theta} \pi(\theta | x).$$

$\delta(x)$ doit donc être le mode de la distribution a posteriori.

Le grand avantage de cet estimateur est qu'il ne dépend pas d'une fonction de perte, et donc il ne requiert aucun calcul d'intégral, mais une simple maximisation de $f(x | \theta) \cdot \pi(\theta)$, est utile pour les approches théoriques. L'estimateur MAP est le pendant Bayésien de l'estimateur de maximum de vraisemblance, de ce fait ils partagent les mêmes inconvénients comme : la non unicité, l'instabilité (dus aux calculs d'optimisation), ..., etc.

2.1.4 Choix de la distribution a priori

D'un point de vue pratique, le choix de la loi a priori est souvent perçu comme une difficulté majeure de l'approche Bayésienne en ce que l'interprétation de l'information a priori disponible est rarement assez précise pour conduire à la détermination d'une seule et unique loi de probabilité.

En pratique, l'information a priori peut être codée selon une des trois façons suivantes :

- Choisir une loi a priori non-informative.
- Choisir une loi a priori informative.
- Prendre une loi a priori impropre.

Loi a priori non informative /vague

$\pi(\theta)$ est non-informative si son impact sur la distribution a posteriori $\pi(\theta | x)$ est minimal. Graphiquement $\pi(\theta)$ sera plus plate que la vraisemblance $f(x | \theta)$; est dominée par celle-ci; $\text{beta}(1, 1)$, $N(\mu, \infty)$. On peut citer quelques types de cette lois :

Loi a priori de Jeffreys : À fin d'éviter le besoin de prendre en compte une structure d'invariance potentielle, Jeffreys (1946, 1961) propose des lois a priori non informatives fondées sur l'information de Fisher, donnée par :

$$I(\theta) = E_{\theta} \left[\left(\frac{\partial \log f(x | \theta)}{\partial \theta} \right)^2 \right]$$

Dans le cas unidimensionnel, sous certaines conditions de régularité, cette information est aussi égale à :

$$I(\theta) = -E_{\theta} \left[\left(\frac{\partial^2 \log f(x | \theta)}{\partial \theta^2} \right) \right],$$

La loi a priori de Jeffreys est $\pi(\theta) \propto \sqrt{I(\theta)}$, elle définit un coefficient de normalisation près quand π est propre.

Ce choix, dépendant de l'information de Fisher se justifie par le fait que $I(\theta)$ est largement accepté comme un indicateur de la quantité d'information apportée par le modèle sur θ , (Fisher 1956).

Exemple 2.2. Soit X une va telle que : $X \sim N(0, \theta)$, On calcule l'information de Fisher tel que :

$$\begin{aligned} f(x | \theta) &= \left(\frac{1}{2\pi\theta} \right)^{\frac{1}{2}} e^{-\frac{x^2}{2\theta}} \\ \log f(x | \theta) &= -\frac{1}{2} \log 2\pi - \frac{1}{2} \log \theta - \frac{x^2}{2\theta} \\ \frac{\partial \log f(x | \theta)}{\partial \theta} &= -\frac{1}{2\theta} + \frac{x^2}{2\theta^2} \\ \frac{\partial^2 \log f(x | \theta)}{\partial \theta^2} &= \frac{1}{2\theta^2} - \frac{x^2}{\theta^3} \end{aligned}$$

calculons la loi a priori $\pi(\theta)$ de Jeffreys :

on a d'après le calcul :

$$\begin{aligned} I(\theta) &= -E \left(\frac{\partial^2 \log f(x | \theta)}{\partial \theta^2} \right) \\ &= -\frac{1}{2\theta^2} \end{aligned}$$

d'ou :

$$\pi(\theta) \propto \sqrt{I(\theta)} \Rightarrow \pi(\theta) \propto \frac{1}{\theta}$$

Loi a priori de référence : Cette technique a été mise au point par Bernardo (1979), c'est une modification de l'approche de Jeffreys dans le cas unidimensionnel, appelée approche de la loi de référence.

La différence qui caractérise cette méthode est que la loi a priori résultante par la méthode de référence ne dépend pas seulement de la loi d'échantillonnage, mais aussi du problème inférentiel considéré.

Quand $x \sim f(x | \theta)$ et $\theta = (\theta_1, \theta_2)$ est le paramètre d'intérêt, la loi de référence est obtenue en définissant d'abord $\pi(\theta_2 | \theta_1)$ comme la loi de Jeffreys associée à $f(x | \theta)$ pour θ_1 fixé, puis en calculant la loi marginale

$$\tilde{f}(x | \theta_1) = \int f(x | \theta_1, \theta_2) \pi(\theta_2 | \theta_1) d\theta_2$$

et la loi de Jeffreys $\pi(\theta_1)$ associée à $\tilde{f}(x | \theta_1)$.

Le principe sous-jacent à la loi de référence est donc d'éliminer le paramètre de nuisance en utilisant la loi de Jeffreys correspondant au cas où le paramètre d'intérêt reste fixé.

loi a priori invariante : le fait de formaliser l'absence de l'information a priori par une propriété d'invariance est naturel au sens où seules les paramètres de la distribution de θ changent lorsqu'on effectue une transformation de θ .

Par exemple les distributions de θ et $\theta - \theta_0$ en réalité ne sont pas les mêmes, mais dire qu'elles sont les mêmes c-à-d : $\pi(\theta) = \pi(\theta - \theta_0)$.

Pour tout θ_0 exprime certainement une ignorance sur la valeur de θ . On dit dans ce cas que la loi a priori π est invariante par la translation, et on prend $\pi(\theta) = c$ la loi uniforme sur Θ .

Loi a priori informative

Un a priori informatif n'est pas dominé par la vraisemblance et impacte $\pi(\theta | x)$; L'a posteriori d'une étude précédente peut être utilisée comme l'a priori $\pi(\theta)$ de l'étude courante.

loi a priori Conjuguée : Un a priori est dit conjugué à une famille de distributions si $\pi(\theta)$ et $\pi(\theta | x)$ appartiennent à la même famille.

Définition 2.3. Une famille F distribution de probabilité sur Θ est dite conjuguée (ou fermée par échantillonnage) par une fonction de vraisemblance $f(x | \theta)$ si, pour tout $\pi \in F$, la distribution a posteriori $\pi(\cdot | \theta)$ appartient également à F .

Cette définition est formellement vague puisque si l'on choisit F comme la famille de toutes les densités de probabilité, la famille de la densité a priori F est toujours conjuguée par la fonction de vraisemblance. Il est plus intéressant d'étudier des familles de densité a posteriori se réduit à un changement de paramètres et non à une modification de la forme fonctionnelle de la famille correspondante. Les densités a posteriori sont donc toujours calculables et l'évaluation des statistiques a posteriori est grandement simplifiée.

Considérons l'exemple des familles de densités exponentielles.

Définition 2.4. Soient Θ l'espace des paramètres et Ω l'espace des observations. On définit C et h respectivement fonctions de Θ et Ω dans \mathbb{R}^+ et \mathbb{R} , et T fonction de Θ et Ω dans \mathbb{R}^k .

La famille tel que les densités soient de la forme :

$$f(x | \theta) = h(x) C(\theta) \exp(R(\theta) T(x))$$

est dite famille exponentielle de dimension k .

Dans le cas particulier où $\Theta \subset \mathbb{R}^k$ et

$$f(x | \theta) = h(x) C(\theta) \exp(\theta x)$$

La famille est dite naturelle.

la fonction de vraisemblance du vecteur x tel que les $x_i, i = \overline{1, N}$, sont i.i.d. est définie par :

$$f(x_1, \dots, x_n | \theta) = \left[\prod_{i=1}^N h(x_i) \right] C(\theta)^N \exp(R(\theta) \sum_{i=1}^N T(x_i))$$

Supposons que la densité a priori conjuguée s'écrive de la manière suivante :

$$\pi(\theta) \propto C(\theta)^\eta \exp(R(\theta) \nu)$$

Notons que le symbole de proportionnalité (\propto) s'entend en termes de fonction de θ (et non de x).

Où les deux paramètres (ou hyper-paramètres) η et ν doivent être spécifiés pour mener à de terme l'analyse Bayésienne. La densité a posteriori appartient, donc, à la famille exponentielle :

$$\pi(\theta | x) \propto C(\theta)^{\eta+N} \exp\left(R(\theta) \left(\nu + \sum_{i=1}^N T(x_i)\right)\right)$$

Le choix d'un a priori conjugué est toujours un choix particulier et influence donc, dans une certaine mesure, l'inférence résultante. De plus, il peut obliger à ignorer une partie de l'information a priori si cette dernière n'est pas complètement compatible avec la structure de la loi a priori conjuguée.

$f(x \theta)$	$\pi(\theta)$	$\pi(\theta x)$
Normale $N(0, \delta^2)$	Normale $N(\mu, \tau^2)$	$N\left(\frac{\mu\delta^2 + \tau^2 x}{\delta^2 + \tau^2}, \frac{\delta^2 \tau^2}{\delta^2 + \tau^2}\right)$
Normale $N\left(\mu, \frac{1}{\theta}\right)$	Gamma $\Gamma(\alpha, \beta)$	$\Gamma\left(\alpha + \frac{1}{2}, \beta + \frac{(\mu - x)^2}{2}\right)$
Poisson $P(\theta)$	Gamma $\Gamma(\alpha, \beta)$	$\Gamma(\alpha + x, \beta + 1)$
Gamma $\Gamma(\nu, \theta)$	Gamma $\Gamma(\alpha, \beta)$	$G(\alpha + \nu, \beta + x)$
Binomiale $B(n, \theta)$	Beta $Be(\alpha, \beta)$	$Be(\alpha + x, \beta + n - x)$
Binomiale Négative $Bneg(m, \theta)$	Beta $Be(\alpha, \beta)$	$Be(\alpha + m, \beta + x)$
Multinomiale $M_k(\theta_1, \dots, \theta_k)$	Dirichlet $D(\alpha_1, \dots, \alpha_k)$	$D(\alpha_1 + x_1, \dots, \alpha_k + x_k)$

TABLE 2.1 – Lois a priori conjuguées pour quelques familles exponentielles

Exemple 2.5. Modèle exponentiel

Soit x un n - échantillon de va i.i.d par une exponentielle paramétrée par θ dont la vraisemblance est donnée par :

$$f(x | \theta) = \theta^n \exp(-n\bar{x}\theta)$$

le prior conjugué est une distribution Gamma de paramètres a et b :

$$\pi(\theta | a, b) = \frac{b^a}{\Gamma(a)} \theta^{a-1} \exp(-b\theta)$$

la règle de Bayes donne la distribution a posteriori de suivante :

$$\pi(\theta | x, a, b) \propto \theta^{n+a-1} \exp(-\theta(n\bar{x} + b))$$

On reconnaît la forme analytique d'une nouvelle distribution Gamma.

Si la taille de l'échantillon ($n \rightarrow \infty$) alors :

$$n + a - 1 \approx n \quad \text{et} \quad n\bar{x} + b \approx n\bar{x}$$

Ainsi, le posterior et la vraisemblance ont la même forme analytique.

- **A prior vague** : La vraisemblance fait aussi le posterior quand l'état de connaissance disponible n'autorise qu'un prior vague (peu informatif).

Dans l'exemple précédent, un prior vague est obtenu pour $(a, b) \rightarrow (0, 0)$.

La forme analytique de la densité a posteriori devient :

$$\pi(\theta | x) \propto \theta^{n-1} \exp(-(n\bar{x})\theta)$$

Quand n est assez grand, n et $n - 1$ possèdent le même ordre de grandeur.

De même, le posterior et la vraisemblance ont la même forme analytique.

Loi a priori impropre

Dans de nombreux cas, la distribution a priori est déterminée par des critères subjectifs et théoriques qui conduisent à une mesure infinie sur l'espace des paramètres Θ plutôt qu'à une mesure de probabilité, c-à-d à une mesure π telle que :

$$\int_{\Theta} \pi(\theta) d\theta = +\infty$$

Dans de tels cas, on dit que la distribution a priori $\pi(\theta)$ est impropre. Cette terminologie est bien-sûr un abus de langage puisque $\pi(\theta)$ n'est pas une densité de probabilité.

Ce type de lois n'est utile que si la loi a posteriori $\pi(\theta | x)$ existe, aussi on se limite aux lois impropres telle que :

$$\int_{\Theta} f(x | \theta) \pi(\theta) d\theta < \infty$$

Vérifier que $\int_{\Theta} f(x | \theta) \pi(\theta) d\theta < \infty$ est une difficulté pratique dans l'utilisation de ces lois, cependant cette difficulté ne doit pas être considérée comme un inconvénient car les nouvelles techniques de calcul Bayésien comme les algorithmes MCMC (Méthodes Monte-Carlo par chaîne de Markov) ne nécessitent pas dans la pratique de vérifier cette condition.

2.1.5 Méthodes de Monte-Carlo par Chaîne de Markov

Les méthodes de Monte-Carlo par chaînes de Markov, ou méthodes MCMC, sont une classe de méthodes d'échantillonnage à partir de distributions de probabilité. Ces méthodes de Monte-Carlo se basent sur le parcours de chaînes de Markov qui ont pour lois stationnaires les distributions à échantillonner.

Certaines méthodes utilisent des marches aléatoires sur les chaînes de Markov (algorithme de Metropolis-Hastings, échantillonnage de Gibbs), alors que d'autres algorithmes, plus complexes, introduisent des contraintes sur les parcours pour essayer d'accélérer la convergence. Ces méthodes sont notamment appliquées dans le cadre de l'inférence Bayésienne.

Chaîne de Markov

les chaînes de markov est un processus aléatoire portant sur un nombre fini d'états, avec des probabilités de transition sans mémoire.

Définition 2.6. On appelle processus markovien un processus stochastique possédant la propriété de Markov. Si de plus, l'ensemble des états E du phénomène est discret, alors c'est une chaîne de Markov.

Une suite $(X_n)_{n \geq 0}$ de va à valeurs dans un ensemble au plus dénombrable E est une chaîne de Markov d'espace d'états E ssi pour tout $n \in \mathbb{N}, i_n \in E$ tels que

$$P[X_{n+1} = j \mid X_0 = i_0, X_1 = i_1, \dots, X_n = i_n] = P[X_{n+1} = j \mid X_n = i_n]$$

Si de plus le terme de droite est indépendant de n , alors la chaîne de Markov est dite homogène, c-à-d, $\forall n \in \mathbb{N}$, et $i, j \in E$:

$$P[X_{n+1} = j \mid X_n = i] = P[X_1 = j \mid X_0 = i]$$

Matrice de transition

Définition 2.7. On appelle probabilité de transition de i vers j la probabilité de passer de l'état i à l'instant n vers l'état j à l'instant $n + 1$, et on la dénote p_{ij}

$$p_{ij} = P[X_{n+1} = j \mid X_n = i]$$

Définition 2.8. On appelle matrice de transition ou de passage (noyau), la matrice $p = (p_{ij})_{(i,j \in E)}$:

$$P = \begin{bmatrix} p_{11} & \cdot & \cdot & \cdot & p_{1j} \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ p_{i1} & & & & p_{ij} \end{bmatrix}$$

On aura alors :

1. $\sum_{j \in \mathbb{N}} p_{ij} = 1$
2. $\forall (i, j) \in \mathbb{N}^2, p_{ij} \geq 0$.

P est donc une matrice stochastique.

Chaîne irréductible

Définition 2.9.

- On dit que l'état j est accessible à partir de l'état i , et on note $i \rightsquigarrow j$, s'il existe un $n \geq 0$ tel que $P_{ij}^{(n)} \geq 0$.

- On dit aussi que deux états i et j communiquent, si on a : $i \rightsquigarrow j$ et $j \rightsquigarrow i$; On écrit alors $i \leftrightarrow j$.

- On dit qu'une chaîne de Markov est irréductible si tous ses états communiquent entre eux, c-à-d $\forall i, j \in E, i \leftrightarrow j$.

On peut formuler l'irréductibilité comme tel :

$$\forall i, j \in E, \forall n \geq 0, \exists n_0 \in \mathbb{N} / P(X_{n+n_0} = j \mid X_n = i) > 0$$

et n_0 dépend des états i et j .

Chaîne Périodique

L'étude ici se focalise sur les conditions de temps qui sépare deux retours au même état est ou n'est pas multiple d'un temps minimum. Pour cela, on introduit la notion de période.

Définition 2.10. Soit $i \in E$. On appelle période de i , et on le note $d(i)$, le pgcd de tous les entiers $n \geq 1$, tels que $P_{ii}^{(n)} > 0$.

$$d(i) = \text{pgcd}(n \geq 1, P_{ii}^{(n)} > 0)$$

- Si $d(i) > 1$ alors on dit que i est périodique de période $d(i)$.
- Si $d(i) = 1$ alors on dit que i est apériodique. On dit qu'une chaîne de Markov est apériodique si tous ses états sont apériodiques.

Chaîne récurrente

$\forall i \in E$, on désigne par T_i le temps d'atteinte de l'état i à partir de l'instant 1 :

$$T_i = \inf_{n \geq 1} \{X_n = i\}$$

On notera par N_i le nombre de fois que la chaîne visite l'état i en comptant le point de départ :

$$N_i = \sum_{n \in \mathbb{N}} 1_{X_n = i}$$

Définition 2.11. On dit qu'un état $i \in E$ est récurrent(ou persistant) si, partant de cet état, la probabilité que la chaîne de Markov retourne à son état initial en un temps fini est égale à 1 :

$$P_i(T_i < \infty) = P(T_i < \infty \mid X_0 = i) = 1$$

Sinon, s'il n'est pas récurrent alors on dit qu'il est transient(ou transitoire), c-à-d :

$$P_i(T_i < \infty) = P(T_i < \infty \mid X_0 = i) < 1$$

Nous pouvons tout de suite remarquer que dès que la chaîne parvient à revenir à son état initial, $T_i < \infty$ et $N_i > 1$. En d'autre terme

$$P_i(T_i < \infty) = P(T_i < \infty \mid X_0 = i) = P(N_i > 1 \mid X_0 = i) = P_i(N_i > 1)$$

Une chaîne de Markov est dite transiente (respectivement récurrente) si tous ses états sont transients (respectivement récurrents).

Proposition 2.12. *Les propriétés suivantes permettent de caractériser les états récurrents et transients :*

- Si i est transient alors $P_i(N_i < \infty) = 1$, $\sum p_{ii}^{(n)} < \infty$, et conditionnellement à $\{X_0 = i\}$.

- Si i est récurrent alors $P_i(N_i = \infty) = 1$ et $\sum p_{ii}^{(n)} = \infty$.

Théorème 2.13. *Théorème ergodique :*

Soient $\theta^{(1)}, \theta^{(2)}, \dots, \theta^{(N)}$, N valeurs d'une chaîne de Markov ergodique de distribution stationnaire π , et soit f une fonction telle que $E^\pi[f(\theta)] < \infty$ alors :

$$\frac{1}{N} \sum_{n=1}^N f(\theta^{(n)}) \xrightarrow[n \rightarrow \infty]{} \int_{\Theta} f(\theta) \pi(\theta | x) d\theta < \infty$$

2.1.6 Principe général des méthodes MCMC

le principe générale des algorithmes MCMC est le suivant : pour produire une approximation acceptable d'une intégrale ou d'une autre fonctionnelle d'une distribution d'intérêt (i.e. la loi a posteriori), il suffit de générer une chaîne de Markov dont la distribution limite est la distribution d'intérêt, puis d'y appliquer la méthode de Monte-Carlo.

Il faut donc avoir une double convergence.

1- la convergence de la Chaîne de Markov vers sa distribution stationnaire π .

2- La convergence de Monte-Carlo, une fois la distribution stationnaire est atteinte :

$$\frac{1}{N} \sum_{n=1}^N f(\theta^n) \xrightarrow[N \rightarrow +\infty]{} E[f(\theta)]$$

L'échantillonnage de Gibbs

Cette méthode a été utilisée par Geman, S. et Geman, D. (1984) pour générer des observations à partir d'une distribution de Gibbs (distribution de Boltzmann). Il s'agit d'une forme particulière de la méthode de Monte-Carlo par chaîne de Markov qui, du fait de son efficacité, est largement utilisée dans de nombreux domaines d'analyse statistique Bayésienne.

Dans la méthode de Gibbs, après avoir choisi un point de départ, les d composantes

du vecteur de covariables (θ) sont générées les unes après les autres conditionnellement à toutes les autres composantes. Si $\pi(\theta/x)$ est la densité des d composantes du vecteur θ , conditionnellement aux données observées (x) , nous utilisons alors les densités conditionnelles $\pi(\theta_1/\theta_2, \theta_3, \dots, \theta_d, x)$, $\pi(\theta_2/\theta_3, \theta_4, \dots, \theta_d, x)$ et ainsi de suite. À chaque k^e étape, la distribution conditionnelle utilise les valeurs générées les plus récentes parmi toutes les autres composantes. Par la théorie des chaînes de Markov, on a que lorsque $k \rightarrow \infty$ la densité des réalisations obtenues converge vers $\pi(\theta/x)$. Le schéma de simulation est décrit ci-dessous :

Algorithme d'échantillonnage de Gibbs

1. Initialisation $\theta^{(0)} = (\theta_1^{(0)}, \dots, \theta_d^{(0)})$
2. À l'itération k , ($k = 1, \dots, K$)
 - a- Générer $\theta_1^{(k)}$ à partir de $\pi(\theta_1/\theta_2^{(k-1)}, \theta_3^{(k-1)}, \dots, \theta_d^{(k-1)}, x)$
 - b- Générer $\theta_2^{(k)}$ à partir de $\pi(\theta_2/\theta_1^{(k)}, \theta_3^{(k-1)}, \dots, \theta_d^{(k-1)}, x)$
 - c- ...
 - d- Générer $\theta_{d-1}^{(k)}$ à partir de $\pi(\theta_{d-1}/\theta_1^{(k)}, \theta_2^{(k)}, \dots, \theta_d^{(k-1)}, x)$
 - e- Générer $\theta_d^{(k)}$ à partir de $\pi(\theta_d/\theta_1^{(k)}, \theta_2^{(k)}, \dots, \theta_{d-1}^{(k)}, x)$

Le schéma de simulation est d'autant plus efficace lorsque les distributions conditionnelles $\pi_i(\theta_i/\theta_1, \theta_2, \dots, \theta_{(i-1)}, \theta_{(i+1)}, \dots, \theta_d)$ sont simulables rapidement.

Remarque 5.

- *Un des points remarquables de l'algorithme de Gibbs est qu'il rend possible la simulation de la distribution jointe $\pi(\theta) = \pi(\theta_1, \dots, \theta_d)$ à partir des distributions conditionnelles $\pi_i(\theta_i/\theta_{j \neq i})$.*
- *Lorsqu'il est difficile de simuler selon l'une des distributions conditionnelles $\pi_i(\theta_i/\theta_{j \neq i})$. On peut remplacer cette étape par une étape de Métropolis-Hastings dont la distribution cible est $\pi_i(\theta_i/\theta_{j \neq i})$. L'introduction d'une étape de Métropolis-Hastings ne modifie pas la distribution stationnaire de la chaîne et est donc valide.*

Algorithme de Metropolis-Hastings

La technique de Metropolis-Hastings a été développée par Metropolis et al (1953), au départ pour la physique particulaire, et généralisée par Hastings (1970) dans un cadre plus statistique.

L'algorithme de Metropolis-Hastings peut être vu comme une alternative à l'algorithme de Gibbs dans le cas où on ne peut pas simuler facilement dans les distributions conditionnelles a posteriori, par exemple lorsque les distributions ne sont pas conjuguées ou lorsque les expressions ne peuvent pas s'exprimer sous forme analytique.

L'algorithme de Metropolis-Hastings est un algorithme d'acceptation/rejet. L'idée de l'algorithme est de simuler selon une autre distribution, plus simple à simuler, appelée la loi de proposition, et d'accepter la valeur simulée avec une certaine probabilité d'être effectivement un tirage selon la loi cible.

Sous les conditions de convergence, les valeurs successivement acceptées de la loi de proposition forment une chaîne de Markov convergeant vers la distribution cible.

La différence avec un algorithme acceptation/rejet classique est que tant qu'une nouvelle valeur n'est pas acceptée l'algorithme retourne la dernière valeur acceptée comme nouvelle valeur de la chaîne.

Le schéma de simulation est décrit ci-dessous :

Algorithme de Metropolis-Hastings

1. initialiser $\theta^{(0)}$.
2. À l'itération k , ($k = 1, \dots, K$)
 - a- proposer un candidat $\theta^* \sim q(\theta^* | \theta^{(k-1)})$
 - b- calculer la probabilité d'acceptation :

$$\alpha^{(k)} = \min \left\{ 1, \frac{\pi(\theta^*) q(\theta^{(k-1)} | \theta^*)}{\pi(\theta^{(k-1)}) q(\theta^* | \theta^{(k-1)})} \right\}$$

- c- générer une valeur $u^{(k)} \sim U_{[0,1]}$ et prendre :

$$\theta^{(k)} = \begin{cases} \theta^* & \text{si } u^{(k)} \leq \alpha^{(k)} \\ \theta^{(k-1)} & \text{sinon} \end{cases}$$

2.2 Approche Bayésienne et copules

Dans cette section nous allons mettre en exergue l'utilisation de l'approche Bayésienne dans la formulation et l'estimation des modèles de copules.

Tandis qu'avant les années 1990, l'inférence Bayésienne a été au mieux une entreprise difficile dans la pratique réservée à un nombre restreint de chercheurs spécialisés

et limitée à un nombre plutôt restreint de modèles et ça du au fait que la statistique Bayésienne nécessite souvent des calculs potentiellement lourds ou infaisable lorsque l'on sort des exemples simples. Elle est devenue au premier plan et un procédé très performant qui apte à résoudre les problèmes les plus avancés dans tous les domaines ; grâce aux méthodes de calcul numériques qui permettent d'obtenir des approximations numériques en des temps raisonnables.

L'approche Bayésienne fait face aux problèmes rencontrés dans les modèles des copules telle qu'elle combine les informations apportées par les données avec celle de la loi a priori.

Malgré le nombre de travaux qui utilisent l'approche Bayésienne pour résoudre les problèmes liés aux copules est limité, elle a pu accru sa cohérence et sa performance dans ce domaine aussi. Plusieurs chercheurs ont travaillé sur le développement des outils Bayésiennes dans des problèmes liés aux copules, à savoir :

Michael Stanly (2011) a proposé une modélisation des copules par l'approche Bayésienne, Simon Guillotte (2008) a proposé également une approche Bayésienne pour l'estimation simultanée de la fonction de dépendance et des lois marginales. Un modèle non paramétrique est construit, et un algorithme MCMC à sans réversible pour le calcul de l'estimateur Bayésien est proposé, Silva and Lopes (2008) ont utilisé une méthode MCMC pour estimer les fonctions de copules de petites dimensions, Hoff (2007) and Denaher and Smith (2011) ont estimé une copule gaussienne d'un modèle de régression en utilisant une méthode MCMC, Juan Wu. XueWang. Stephen G. Walker (2013) ont proposé une approche Bayésienne non paramétrique pour l'estimation d'une copule en utilisant une méthode MCMC, Ausien et Lopes (2010) ont considéré une estimation Bayésienne pour une série temporelle multivariée avec une copule basée sur time varying cross-sectional dépendance.

2.2.1 Exemple d'application

L'objectif de cette section est de présenter une méthode Bayésienne pour estimer le paramètre θ d'une copule archimédienne $C(u_1, u_2) = \phi^{-1}(\phi(u_1) + \phi(u_2))$ avec $\phi : I \rightarrow [0, \infty[$.

Nous avons pris comme exemple le travail présenté par Maxime Dion (2014) dans journal du colloque des étudiants de premier cycle en mathématique de l'université Laval dont une méthode MCMC qui est l'algorithme de Metropolis-Hastings pour estimer les paramètres d'une copules archimédienne est présentée.

Considérons le modèle de copule archimédienne suivant :

$$C(u_1, u_2) = \phi^{-1} \left(\sum_{j=1}^2 \phi(u_j) \right)$$

La densité de cette copule

$$c(u_1, u_2) = (\phi^{-1})^{(2)} \left(\sum_{j=1}^2 \phi(u_j) \right) \prod_{j=1}^2 (\phi)'(u_j),$$

où $(\phi)'$ est la dérivée du générateur, $(\phi^{-1})^{(2)}$ est la 2-ième dérivée de l'inverse générateur.

Les copules archimédienne sont caractérisées par l'intervention d'une va latente Z qui se diffère d'une famille à l'autre. Ainsi la variable U_k peut être exprimée en termes de la variable latente Z et de distributions exponentielles indépendante comme suit : $(U_1, U_2) = (\phi_\theta^{-1}(\frac{E_1}{Z}), \phi_\theta^{-1}(\frac{E_2}{Z}))$ où $E_k \sim \text{Exp}(1)$ pour $k = 1, 2$.

D'où la fonction de répartition de U_k sachant θ et Z est donnée par :

$$\begin{aligned} F_{U_k}(u_k | \theta, Z = z) &= P[U_k \leq u_k | \theta, Z = z] \\ &= P \left[\phi_\theta^{-1} \left(\frac{E_k}{z} \right) \leq u_k | \theta, Z = z \right] \\ &= P [E_k \geq z\phi_\theta(u_k) | \theta, Z = z] \\ &= \exp(-z\phi_\theta(u_k)). \end{aligned}$$

Ainsi, avec n échantillons indépendants mis dans une matrice $n * d$ ($d = 2$) notée U , et avec respectivement un Z_k où $k \in \{1, \dots, n\}$ pour chacun des échantillons, la de vraisemblance est $F(U|\theta, Z = z) = \exp(-\sum_{k=1}^n z_k \sum_{j=1}^2 \phi_\theta(u_{k,j}))$.

Alors la fonction de densité de vraisemblance est :

$$\begin{aligned} f(U | \theta, Z = z) &= (-1)^{2*n} \exp \left(- \sum_{k=1}^n z_k \sum_{j=1}^2 \phi_\theta(u_{k,j}) \right) \prod_{k=1}^n z_k^2 \prod_{j=1}^2 (\phi_\theta)'(u_{k,j}) \\ &= \exp \left(- \sum_{k=1}^n z_k [\phi_\theta(u_{k,1}) + \phi_\theta(u_{k,2})] \right) \prod_{k=1}^n z_k^2 [(\phi_\theta)'(u_{k,1}) (\phi_\theta)'(u_{k,2})] \end{aligned} \quad (2.1)$$

Comme nous avons vu précédemment, toute inférence Bayésienne ne peut être conduite qu'après le calcul de la loi a posteriori qui est le produit de la vraisemblance (2.1) et des lois a priori des paramètres à estimer.

Dans notre cas, nous avons choisi la copule archimédienne de Clayton de générateur $\phi(t) = \frac{1}{\theta} (t^{-\theta} - 1)$, avec $\theta \in [-1, 0[\cup]0, \infty[$.

Nous avons pris les mêmes lois a priori choisi par Maxime Dion : la loi normale tronquée sur $(0, \infty)$ comme loi a priori de θ ($\theta \sim N^+(0, \sigma_\theta^2)$), et la loi de Gamma comme loi a priori de Z ($Z \sim \Gamma(\frac{1}{\theta}, 1)$).

Ainsi, le calcul de la loi a posteriori via le théorème de Bayes a donné :

$$\begin{aligned}
 \pi(\theta, Z | U) &\propto f(U | \theta, Z) * \pi_\theta(z) \\
 &= \exp\left(-\sum_{k=1}^n z_k \left[\frac{1}{\theta} (u_{k,1}^{-\theta} + u_{k,2}^{-\theta} - 2)\right]\right) \prod_{k=1}^n z_k^2 \left[(u_{k,1} u_{k,2})^{-\theta-1}\right] \\
 &* \frac{1}{\Gamma\left(\frac{1}{\theta}\right)} z^{\frac{1}{\theta}-1} \exp(-z) 1_{\mathbb{R}_+^*}(z) \\
 &\propto z^{\frac{1}{\theta}-1} \exp\left(-z - \sum_{k=1}^n z_k \left[\frac{1}{\theta} (u_{k,1}^{-\theta} + u_{k,2}^{-\theta} - 2)\right]\right) \prod_{k=1}^n z_k^2 \left[(u_{k,1} u_{k,2})^{-\theta-1}\right]
 \end{aligned}$$

Nous visons à estimer θ et Z dans un contexte Bayésien, il est clair qu'il n'est pas facile de le faire analytiquement. Pour cela, nous avons appliqué l'algorithme M-H pour estimer θ et Z .

Comme nous avons mentionné précédemment l'algorithme M-H repose sur le choix d'une loi de proposition. Dans notre cas, nous avons choisi les deux lois de proposition suivantes : $\theta^* \sim N^+(0, \sigma_\theta^2)$ et $z^* \sim N^+(0, \sigma_z^2)$.

Les étapes de l'algorithme de M-H

- Entrer la taille de l'échantillon : $n = 100$.
- Charger le package TruncatedNormal.
- Entrer $\theta \sim N^+(0, \sigma_\theta^2)$ et $Z \sim \Gamma\left(\frac{1}{\theta}, 1\right)$.
- Générer deux variables suivant la loi uniforme :
 $u_1 \sim U_{[0,1]}$, $u_2 \sim U_{[0,1]}$.
- Entrer la distribution a posteriori $p(\theta, Z = z|U)$ et la loi de proposition $q(z)$.
- Générer des nouveaux candidats $\theta^* \sim N^+(0, \sigma_\theta^2)$ et $z^* \sim N^+(0, \sigma_z^2)$.
- Initialiser la chaîne : $X[1, 1] = \theta[1]$, $X[1, 2] = z[1]$
- Calculer la probabilité d'acceptation :

$$\rho = \min \left\{ 1, \frac{p(\theta^*, z^*) * q(X[i-1, 1], X[i-1, 2])}{p(X[i-1, 1], X[i-1, 2]) * q(\theta^*, z^*)} \right\}$$

- Générer u Uniforme et prendre :

$$X_i = \begin{cases} \theta^*, z^* & , u \leq \rho \\ X_{i-1} & \text{sinon} \end{cases}$$

Algorithme sous R

```

n=100
theta=rtnorm(n,0,1,lb=0,ub=Inf)
z=rgamma(n,1/theta,1)
u=runif(n,0,1)
v=runif(n,0,1)
for (k in 1:n){
s=sum(z[k]*((1/theta[k])*(u[k]^(-theta[k])+v[k]^(-theta[k])-2)))
pro=prod(z[k]^2*(u[k]^(-theta[k]-1)*v[k]^(-theta[k]-1)))
}
p=function(x,y){(y^((1/x)-1))*exp(-(y+s))*pro}
q=function(x,y){exp(-0.5*t(y)%*%y)}
M=100 ; ind=2
X=matrix(rep(0,ind),ncol=ind,nrow=M)
X[1,1]=theta[1] ; X[1,2]=z[1]
for(i in 2 :M) {
thetac=rtnorm(1,0,1,lb=0,ub=Inf)
zc=rtnorm(1,0,0.5,lb=0,ub=Inf)
Y=c(thetac,zc)
val=(p(Y[1],Y[2])*q(X[i-1,1],X[i-1,2]))/(p(X[i-1,1],X[i-1,2])*q(Y[1],Y[2]))
theta0=min(1,val)
u=runif(n,0,1)
if(u<theta0){X[i,1]=Y[1];X[i,2]=Y[2]}else{X[i,1]=X[i-1,1];X[i,2]=X[i-1,2]}
}
theta0
thetaB=mean(X[,1])
zB=mean(X[,2])
thetaB;zB

```

Le tableau suivant représente les valeurs de les estimateurs θ_B et z_B pour différentes valeurs de n :

Estimateur	n		
	$n = 1000$	$n = 2000$	$n = 3000$
θ_B	0.7982544	1.656317	0.2624536
z_B	0.7008208	1.881934	1.082474
Taux de kandell	0.2852687	0.4530015	0.116004

TABLE 2.2 – Les valeurs de l'estimateur θ_B et Z_B .

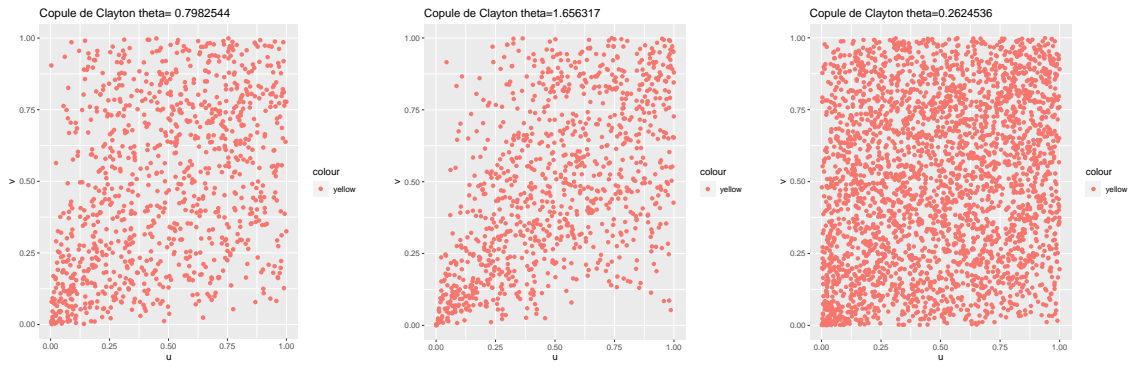


FIGURE 2.1 – Nuage des points de la copule de Clayton selon θ

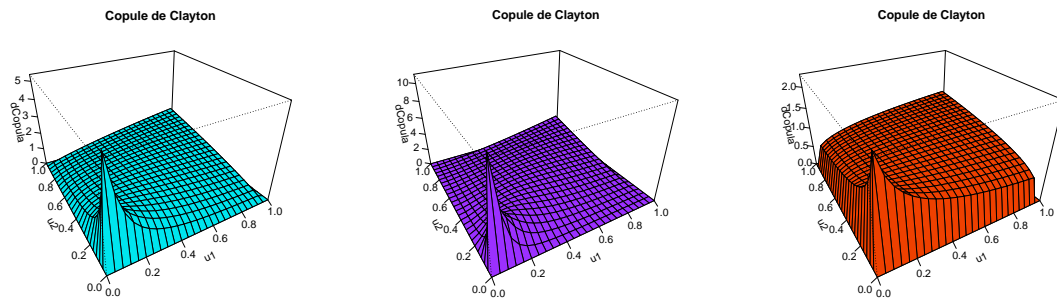


FIGURE 2.2 – Les densités de la copule de Clayton selon θ

Conclusion et perspectives

Comme les méthodes MCMC ont rendu la statistique Bayésienne applicable à presque tout type de modèles, nous nous sommes intéressé dans notre travail à établir une estimation Bayésienne des paramètres d'une copule archimédienne comme un type d'une copule paramétrique, en appliquant une méthode MCMC (M-H).

Nous pouvons constater une autre fois ; à travers le résultat de ce travail ; l'importance et la performance de l'approche Bayésienne pour les modèles de copules. Nous avons pu développer un algorithme M-H pour estimer le paramètre θ de la copule archimédienne de Clayton, en utilisant le langage R.

En perspective, il serait intéressant de généraliser ce travail en proposant d'autres lois a priori qui permettent de meilleurs estimation. Aussi il serait important de développer cet algorithme pour les autres familles archimédiennes ou même pour d'autre copules paramétriques et non paramétriques.

Annexe : Abréviations et Notations

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

Symbole	Signification
\mathbb{R}	Ensemble des nombres réels
I	Intervalle $[0, 1]$
C	Distribution de la copule
c	Densité de la copule
i.i.d	Indépendante et identiquement distribuées
i.e	c'est-à-dire
U, V	vas qui suivent la loi uniforme
$E(x)$	Espérance mathématique de la va X
$var(x)$	Variance mathématique de la va X
$cov(X, Y)$	covariance mathématique du couple (X, Y)
F, G	fd des lois marginales
H	Distribution jointe
f, g	Densité des lois marginales
h	Densité de distribution jointe
ϕ	Fonction génératrice de la copule archimédienne
ϕ^{-1}	Fonction quantile de ϕ
τ	Tau de Kendall
ρ	Rho de Spearman
Q	Fonction de concordance
Π, W, M	copules produit, maximum et minimum
Z	va latente
Θ	Espace des paramètres
θ	paramètre de la copule et Bayésien
$f(x \theta)$	Fonction de vraisemblance
$\pi(\theta)$	La loi a priori

$\pi(\theta x)$	La loi a posteriori
$L(\delta, \theta)$	Fonction de coût
$\hat{\theta}$	Estimateur ponctuel
$\sigma(x)$	Estimateur Bayésien
E	Espace d'état
X_n	suite de va
q	la loi de proposition

Bibliographie

- [1] **Christian P. Robert** : Le choix bayésien Principes et pratique. Springer-Verlag France, Paris, 2006.
- [2] **Christian P. Robert** et **George Casella** : Méthodes de Monte-Carlo avec R. Springer-Verlag France, 2011.
- [3] **Denuit, M. Dhaene, J. Goovaerts, M & Kaas, R.** Actuarial Theory for Dependent Risks Measures, Wiley , England, (2005).
- [4] **Genest, C. , Mackay, J.**. The joy of copulas : Bivariate distributions with uniform marginals. The American Statistician 40, 280-285, 1986.
- [5] **Gumbel, E. J** : Bivariate logistic distributions, 56 : 335-349, 1961.
- [6] **Lehmann EL.**. Some concepts of dependence. Ann Math Statist. 37,1137-1153, 1966.
- [7] **Nelsen, R. B.**, An Introduction to copula. Springer Verlag, New York.
- [8] **Nelsen, R. B.** An Introduction to Copulas, 2nd ed. New York : Springer (2006).
- [9] **Sklar, A.**. Fonctions de répartition à n dimensions et leurs marges, Publ. Inst. Univ. Paris. 8, 229-231, 1959.
- [10] **BENELMIR Imane** : Modélisation de la Dépendance par les Copules, Université Mohamed Khider, Biskra.
- [11] **CHINE Amel** : la statistique de copules, Université Mohamed Khider, Biskra.
- [12] **Nabil KADI** : Estimation non-paramétrique de la distribution et densité de copules, Université Sherbrooke, Québec, Canada, avril 2014.
- [13] **TARIK BAHRAOUI** : Fonction de C-puissance et application à l'inférence de copules, Université du Québec, Novembre 2012.
- [14] **ABDERRAHMANE Nour El Houda** Et **:ZEGHMAR Radia** : Modélisation De La Dépendance Par Les Copules, Université M'hamed Bougara, Boumerdes, 2016.
- [15] **ALLICHE Fadhila** et **BRAHIMI Kenza** : Segmentation d'images couleurs basée sur les copules, Université Mouloud Mammeri, Tizi-Ouzou.

- [16] **Lounas Fadhila** : Modélisation de la dépendance par les copules et applications, Université de Mouloud Mammeri, Tizi-Ouzou.
- [17] **Slimani** : Estimation des paramètres d'une copules, Université Mohamed Khider, Biskra.
- [18] **ISMAËL COULIBALY** : Conception d'algorithmes probabilistes pour l'estimation des génotypes d'un corpus de généalogie par chaîne de Markov, Université du Québec à Chicoutimi, 29 mars 2009.
- [19] **LOTFI KHRIBI** : L'échantillonnage de Gibbs pour l'estimation Bayésienne dans l'analyse de survie, Université du Québec, Montréal, Février 2007.
- [20] **Paméla ElHajj** : Méthodes d'aide à la décision thérapeutique dans les cas des maladies rares : intérêt des méthodes bayésiennes et application à la maladie de Horton, Université de Montpellier, 29 Septembre 2017.
- [21] **ILLA AMADOU Oumoul Hair** : Approche Bayésienne des tests sur les modèles autorégressifs vectoriels, Université Mouloud Mammeri, Tizi-Ouzou, 03 Juillet 2013.
- [22] **Kamilia HAMROUN** :Eléments de robustesse Bayésienne, Université de Mouloud Mammeri, Tizi-Ouzou.
- [23] **RATISARIJAONA Mickaël David** :Pratique de la methode MCMC en estimation Bayésienne, Université d'Antananarivo, 23 Janvier 2018.
- [24] **TOUAM ABDERRAZAK** :Méthodes Monte-Carlo, application Bayésienne, Université Mouloud Mammeri, Tizi Ouzou, 27 septembre 2018.
- [25] **Somia Ayed et Wafaa Ogbi** : Monté Carlo dans l'estimation de la densité de probabilité par méthode du noyau associé. Université Djilali Liabes, Sidi Bel Abbes, 2017.