
République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

UNIVERSITÉ MOHAMED KHIDER BISKRA

Faculté des Sciences Exactes et Sciences de la Nature et de la Vie
Département de Mathématiques



MÉMOIRE

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

Par

Abderrahim GUEMMAZ

Thème

**CARACTÉRISATION DE LA COPULE DE
MARSHALL-OLKIN ET APPLICATION**

Option

Probabilités & Statistique

Soutenu publiquement le : 04 / 03 /2012

Devant le jury :

Président	: Brahim MEZERDI	Pr.	U.M.K BISKRA
Rapporteur	: Abdelhakim NECIR	Pr.	U.M.K BISKRA
Examineur	: Khaled MELKEMI	Pr.	U.M.K BISKRA
Examineur	: Djamel MERAGHNI	M.C. (A).	U.M.K BISKRA

Résumé

La copule de Marshall-Olkin a été construite à partir de la fonction conjointe H dont les distributions marginales F et G , telle que F est la fonction de la distribution de Marshall-Olkin Burr $MOB(\alpha, \beta, \gamma)$ et G est la fonction de la distribution de Marshall-Olkin Pareto $MOP(\alpha, \beta)$.

Dans ce mémoire de magister, nous faisons une synthèse sur la théorie des copules et les mesures de concordances. Particulièrement nous nous intéressons à la caractérisation de la copule de Marshall-Olkin et les mesures de concordances associées à cette famille de copule et ses applications en actuariat et en finance.

Abstract

The Marshall-Olkin copula has been built from the joint function H whose marginal distributions F and G , so that F is the Marshall–Olkin Burr distribution $MOB(\alpha, \beta, \gamma)$ function, and G is the Marshall–Olkin Pareto distribution $MOP(\alpha, \beta)$ function.

In this work, we presented some background on copula theory and measures of concordance. In particular we focus on the characterization of the Marshall-Olkin copula and measures of concordance associated with this family of copula and their applications in actuarial sciences and finance.

Introduction

Pendant très longtemps, les copules ont été utilisées comme outil pour résoudre des problèmes mathématiques sans toutefois être l'objet d'études spécifiques. En effet, le concept de copule a été introduit par Sklar en 1959 afin de résoudre un problème de probabilité énoncé par Maurice Fréchet. Ce problème concernait les espaces métriques aléatoires. Même si les copules occupent une place importante dans les travaux de Sklar et Schweizer, elles ne sont pas l'objet de leurs recherches. De façon générale, les copules n'interviennent que pour résoudre certains problèmes et ne font pas l'objet d'études spécifiques. Viennent ensuite 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. Il faut par contre attendre le milieu des années 1980 pour que les copules fassent l'objet d'études approfondies de la part de quelques statisticiens. Le point de départ de ces études est l'article *The Joy of copulas* de Genest et MacKey (1986) publié dans *The American Statistician*. Le premier livre d'introduction aux copules ne sera écrit qu'en 1999 par Nelson. Depuis, les copules sont un outil largement utilisé pour étudier la dépendance.

Alors et en toute simplicité, les copules constituent un outil statistique qui présente de nombreux avantages. La dépendance entre les séries n'a aucune raison a priori d'être linéaire, encore moins monotone. Le plus significatif est que les marchés semblent plus dépendants en période de crash qu'en période normale, traduisant ainsi une dépendance des queues. Les fonctions copules permettent de déterminer la nature de dépendance des séries qu'elle soit linéaire ou pas, monotone ou pas. Les copules autorisent une sélection plus étendue des distributions jointes des séries financières. Les fonctions copules permettent une représentation moins naïve de la dépendance statistique en finance fondée sur la mesure traditionnelle de corrélation (Embrechts et al., 1999). En plus, elles autorisent des distributions de probabilités jointes moins restrictives qui permettent de mieux prendre en compte les faits stylisés en finance (leptokurticité, asymétrie, dépendance des queues). La copule est donc une statistique exhaustive de la dépendance. Nous nous intéresserons dans ce document à une copule bivariée particulière, soit "la copule de Marshall-Olkin".

Le mémoire que je présente se décompose en deux parties :

La première partie est partagée en trois chapitres, le premier chapitre parle de la théorie des copules bivariées, notamment le théorème de Sklar, on essayera de prime à bord de donner une définition précise aux copules. Nous faisons ensuite un bref survol sur les principales propriétés des copules. Le second

comprend quelques familles de copules, et dans la troisième nous avons abordé la notion de dépendance, et les mesures de concordance qui sont des coefficients de corrélation non linéaire et non paramétrique comme (le tau de Kendall et le rho de Spearman) qui peuvent être exprimées à partir des copules.

La deuxième partie est partagée en trois chapitres, le premier comprend des généralités sur la distribution de Marshall-Olkin, et le second est consacré à la caractérisation de la copule de Marshall-Olkin et les mesures de concordance associées à cette famille de copules. Dans le troisième chapitre, nous avons terminé notre travail avec des applications en actuariat et en finance.

Table des matières

Introduction	ii
I Théorie des copules	3
1 Généralités sur les copules bivariées	4
1.1 Définition d'une copule bivariée	5
1.1.1 Propriétés immédiates des copules	6
1.2 Copules et variables aléatoires	7
1.2.1 Le théorème de Sklar et ses applications	7
1.3 Les propriétés fondamentales des copules	12
1.3.1 Copules associées à une copule	14
2 Familles paramétriques de copules usuelles bivariées	18
2.1 Les copules usuelles	18
2.1.1 La copule d'indépendance	18
2.1.2 La copule (minimale et maximale)	18
2.2 Les copules elliptiques	20
2.2.1 Caractérisation	20
2.2.2 Copules élliptiques classiques	21
2.3 Les copules aux valeurs extrêmes	24
2.3.1 Caractérisation	24
2.4 Les copules Archimédiennes	25
2.4.1 Caractérisation	25
2.4.2 Familles classiques de copules archimédiennes	26
2.5 La famille Archimax	31
2.5.1 Caractérisation	31

TABLE DES MATIÈRES **2**

2.6	La copule empirique	31
3	Mesures de concordance	33
3.1	Mesure de dépendance ou corrélation	33
3.2	Mesure de concordance	33
3.2.1	Fonction de concordance	33
3.3	Mesure de concordance	38
3.3.1	Caractérisation d'une mesure de concordance	38
3.3.2	Le coefficient de corrélation de Kendall (Le tau de Kendall)	39
3.3.3	Le coefficient de corrélation de Spearman (Le rho de Spearman)	41
3.3.4	La relation entre le tau de Kendall et le rho de Spearman .	42
3.4	Le coefficient de corrélation de Pearson	42
3.5	Dépendance de queue	43
II	Copule de Marshall-Olkin	45
4	Généralités sur la distribution de Marshall-Olkin	46
4.1	La distribution de Marshall-Olkin Burr	46
4.1.1	Introduction	46
4.1.2	La distribution de Marshall-Olkin semi-Burr	48
4.1.3	La distribution de Marshall-Olkin Burr	50
5	Caractérisation de la copule de Marshall-Olkin	53
5.1	Copule de Marshall-Olkin	53
5.1.1	La densité de la copule de Marshall-Olkin	56
5.1.2	Les mesures de concordances associées à cette famille de copules	56
6	Application	59
6.1	Applications de la distribution de Marshall-Olkin Burr dans la modélisation de série chronologique	59
6.2	Application de la distribution de Marshall-Olkin semi-Burr dans la modélisation des taux de change	65
	Bibliographie	67

Première partie
Théorie des copules

Chapitre 1

Généralités sur les copules bivariées

Les copules constituent un outil statistique qui présente de nombreux avantages, tant pour les statisticiens que pour les financiers. Outre une grande souplesse dans la mise en œuvre de l'analyse multivariée, les copules autorisent une sélection plus étendue des distributions conjointes des séries financières. Les fonctions copules permettent une représentation moins naïve de la dépendance statistique en finance fondée sur la mesure traditionnelle de corrélation qui présente des limites dans l'étude de l'interdépendance entre deux variables (cf., Embrechts et al. 1999).

En outre, elles autorisent des distributions de probabilités jointes moins restrictives, prenant mieux en compte certains faits en finance (léptokurticité, asymétrie, dépendance des queues). Elles permettent la construction de distributions multidimensionnelles assez générales et ce, indépendamment des lois marginales qui peuvent avoir des lois différentes et quelconques. Par conséquent, elles permettent de s'affranchir de certaines hypothèses peu réalistes faites dans les études empiriques. Par ailleurs, l'approche par les copules a beaucoup contribué dans l'analyse, dans la modélisation statistique multivariée.

En effet, la théorie des copules permet une décomposition de la loi multidimensionnelle en ses marginales univariées et en une fonction de dépendance, rendant possibles des extensions naturelles de certains résultats obtenus dans le cas univarié au cas multivarié. Les distributions multidimensionnelles ainsi obtenues sont d'avantage en adéquation avec la réalité surtout dans l'utilisation

financière des statistiques. Par souci de simplicité et du fait que la théorie multivariée est une extension du cas bivarié nous nous limiterons à la théorie bivariée des couples.

1.1 Définition d'une copule bivariée

Dans tout ce qui suit I désignera l'intervalle unité $I = [0, 1]$.

On appelle copule bivariée toute fonction C définie de $I^2 = [0, 1]^2$ dans I vérifiant les propriétés suivantes :

- i) $\forall u \in I, C(u, 0) = C(0, u) = 0$ et $C(u, 1) = C(1, u) = u$.
- ii) C est 2-croissante i.e.

$$C(u', v') - C(u', v) - C(u, v') + C(u, v) \geq 0, \forall u, u', v, v' \in I,$$

avec $u \leq u'$ et $v \leq v'$.

La propriété (i) traduit, en particulier, que toute copule est une distribution dont les distributions marginales sont de loi uniforme définie sur $I = [0, 1]$. En effet en considérant le vecteur aléatoire $U = (U_1, U_2)$ où U_1 et U_2 sont deux variables aléatoires uniformes sur $I = [0, 1]$ alors on a :

$$C(u_1, u_2) = P(U_1 \leq u_1, U_2 \leq u_2). \quad (1.1)$$

La propriété (ii) est la 2-croissance ou inégalité du rectangle de la distribution C . Elle traduit le fait que si C admet une densité $c(u, v)$ alors elle est positive,

$$c(u, v) = \frac{\partial^2 C(u, v)}{\partial u \partial v} \geq 0. \quad (1.2)$$

Exemples de copules

Exemple 1.1.1 (1) $\forall u, v \in I$, la fonction $M(u, v) = \min(u, v)$ définit une copule.

En effet :

$$\forall u, v \in I, \text{ on a } \begin{cases} \min(u, 0) = 0 = \min(0, u), \\ \text{et} \\ \min(u, 1) = u \text{ et } \min(1, v) = v, \end{cases}$$

alors M vérifie (i)

De même $\forall u, v, u', v' \in I$, avec $u \leq v$ et $u' \leq v'$,

$$\min(u', v') \geq \min(u', v),$$

On a : et

$$\min(u, v') \geq \min(u, v),$$

alors M vérifie (ii).

Par conséquent M est une copule.

Exemple 1.1.2 (2) De la même façon on établit que les fonctions :

$$\begin{cases} W(u, v) = \max(u + v - 1, 0), \\ \text{et} \\ \Pi(u, v) = uv, \end{cases}$$

définissent aussi des copules. M , W et Π sont des copules usuelles.

Les résultats suivants donnent les propriétés immédiates d'une copule bivariée.

1.1.1 Propriétés immédiates des copules

Théorème 1.1.1 (Bornes de Fréchet) Pour toute copule C on a :

$$\forall u, v \in I, W(u, v) \leq C(u, v) \leq M(u, v). \quad (1.3)$$

Preuve. En effet,

$$\forall u, v \in I, C(u, v) \leq C(u, 1) = u \text{ et } C(u, v) \leq C(1, v) = v.$$

Alors :

$$C(u, v) \leq \min(u, v). \quad (a)$$

La propriété (ii) $\implies \forall u, v \in I, C(u, v) \geq C(u, 1) + C(1, v) - C(1, 1)$,

alors $C(u, v) \geq u + v - 1$ ou $\forall u, v \in I, C(u, v) \geq 0$.

Donc :

$$C(u, v) \geq \max(u + v - 1, 0). \quad (b)$$

(a) et (b) implique que :

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

Les copules W et M sont appelées borne inférieure (respectivement borne supérieure) de Fréchet-Hoeffding ou copule minimale (respectivement copule maximale). Elles sont aussi notées respectivement C^- et C^+ .

Conséquence du théorème

$$\forall u, v \in I, W(u, v) \leq \Pi(u, v) \leq M(u, v).$$

Le résultat suivant montre, via la condition de Lipschitz, que toute copule est absolument continue.

Théorème 1.1.2 *Pour toute copule bivariée C on a $\forall u, v, u', v' \in I$, on a :*

$$|C(u', v') - C(u, v)| \leq |C(u', v') - C(u, v')| + |C(u, v') - C(u, v)|.$$

1.2 Copules et variables aléatoires

Un des problèmes auxquels les statisticiens se sont intéressés dans les années 50 est l'étude de la relation entre une distribution multivariée et les distributions marginales d'ordre inférieur (d'ordre 1 ou supérieur à 1). En 1959 Abel Sklar apporte une solution partielle à ce problème pour les distributions marginales univariées. Par un théorème qui porte son nom, Sklar établit que si H est une distribution conjointe de distributions marginales $F(x)$ et $G(y)$ alors il existe une copule C telle que :

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

Etant donné qu'une fonction de répartition est d'abord une distribution avant d'être continue à droite et à gauche nous donnons ici le théorème de Sklar dans sa version probabiliste.

1.2.1 Le théorème de Sklar et ses applications

L'outil fondamental de la théorie des copules est **le théorème de Sklar**. Il établit le lien entre la définition formelle ci-dessus des copules et les variables aléatoires, permettant ainsi l'application des copules dans la modélisation statistique.

Le théorème de Sklar

Théorème 1.2.1 *Soit H une distribution bivariée dont les distributions marginales sont F et G . Il existe une copule bivariée C telle que :*

$$\forall (x, y) \in \mathbb{R}^2; H(x, y) = C(F(x), G(y)). \quad (1.4)$$

- Si les distributions marginales F et G sont continues alors la copule C est unique, sinon elle est déterminée de manière unique sur $(\text{Im } F) \times (\text{Im } G)$.

- Réciproquement si C est une copule et F et G sont des distributions univariées alors la fonction H définie par (1.4) est la distribution conjointe dont les marges sont F et G .

Ce théorème permet d'associer à chaque distribution bidimensionnelle une copule. La relation (1.4) donne une représentation canonique de la distribution H en mettant en présence d'une part les marges F et G des directions unidimensionnelles et d'autre part la copule qui permet de "cimenter" ces marges.

Si la distribution bivariée H est absolument continue, alors elle admet une densité :

$$h(x, y) = \frac{\partial^2 C(F(x), G(y))}{\partial x \partial y} = c(F(x), G(y)) \cdot f(x) \cdot g(y), \quad (1.5)$$

où $c(u, v)$ est la densité de la copule C .

La copule C est dite copule des variables aléatoires X et Y ou copule de la distribution H . On la note C_{XY} ou C_H .

Application à la construction des copules

Soient X et Y deux variables aléatoires de distributions respectives F et G supposées strictement monotones. On peut, à partir de leur distribution conjointe H , construire la copule associée aux deux variables X et Y . Introduisons, pour ce faire, la définition suivante pour généraliser cette construction aux distributions pas nécessairement strictement monotones.

Définition 1.2.1 (de la pseudo-inverse ou l'inverse généralisée) Soit F une distribution. On définit la pseudo-inverse $F^{(-1)}$ de F telle que :

- 1) Si $t \in \text{Im } F$, alors $F^{(-1)}(t) = \{x \in \mathbb{R} : F(x) = t\}$.
- 2) Si $t \notin \text{Im } F$, alors $F^{(-1)}(t) = \inf \{x : F(x) \geq t\} = \sup \{x : F(x) \leq t\}$.

Remarque 1.2.1 Si la fonction F est strictement croissante alors la notion de pseudo-inverse coïncide avec la notion d'inverse ou de réciproque i.e. $F^{(-1)} = F^{-1}$. En utilisant la notion de pseudo-inverse dans le théorème de Sklar on obtient le résultat suivant :

Corollaire 1.2.1 Soit H une distribution bivariée dont les distributions marginales univariées sont F et G de pseudo-inverses respectives F^{-1} et G^{-1} . Alors il existe une copule C telle que :

$$\forall u, v \in I, C(u, v) = H(F^{-1}(u), G^{-1}(v)). \quad (1.6)$$

Le corollaire (1.2.1) donne ainsi une procédure de construction d'une copule associée à un couple de variables aléatoires connaissant la distribution conjointe et les distributions marginales continues.

Exemples de construction de copules

La copule de Galambos Pour tout réel $\theta \geq 1$, la distribution bivariée

$$H_\theta(x, y) = \exp \left\{ - \left[(x + y) - (x^{-\theta} + y^{-\theta})^{\frac{-1}{\theta}} \right] \right\},$$

définie sur $[0, +\infty[\times [0, +\infty[$ admet pour marges :

$$F(x) = G(x) = \exp(-x); \quad \forall x \in [0, +\infty[.$$

Soit $\forall u \in [0, 1[, F^{(-1)}(u) = G^{-1}(u) = -\ln u$.

La copule associée est telle que :

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

c'est la copule de Galambos.

Autre exemple de copule

Pour la distribution

$$H(x, y) = \begin{cases} \frac{(x + y)(e^y - 1)}{x + 2e^y - 1} & \text{sur } [-1, 1] \times [0, +\infty[\\ 1 - e^{-y} & \text{sur }]1, +\infty[\times [0, +\infty[\\ 0 & \text{ailleurs.} \end{cases}$$

On vérifie que $F^{-1}(u) = 2u - 1$ et $G^{-1}(v) = -\ln(1 - v)$.

Par conséquent, la copule associée est telle que :

$$C(u, v) = \frac{uv}{u + v - uv}, \quad \forall u, v \in I.$$

La copule normale Soit H_ρ une distribution normale bivariée de coefficient de corrélation ρ telle que :

$$\forall x, y \in \mathbb{R};$$

$$H(x, y) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \int_{-\infty}^x \int_{-\infty}^y \exp \left\{ \frac{-1}{2(1-\rho^2)} \left(\left(\frac{t_1 - m_1}{\sigma_1} \right)^2 - \frac{2(t_1 - m_1)(t_2 - m_2)}{\sigma_1\sigma_2} + \left(\frac{t_2 - m_2}{\sigma_2} \right)^2 \right) \right\} dt_1 dt_2.$$

Supposons que H_ρ est la loi conjointe de v.a. normales $X_1 \rightsquigarrow N(m_1, \sigma_1)$ et $X_2 \rightsquigarrow N(m_2, \sigma_2)$ (en effet, on établira par la suite que la distribution conjointe de deux lois gaussiennes n'est pas automatiquement une loi gaussienne). La copule associée à H_ρ est alors donnée par :

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

$$C_\rho(u, v) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \int_{-\infty}^{\Phi_1^{-1}(u)} \int_{-\infty}^{\Phi_2^{-1}(v)} \exp \left\{ \frac{-1}{2(1-\rho^2)} \left(\left(\frac{x - m_1}{\sigma_1} \right)^2 - \frac{2(x - m_1)(y - m_2)}{\sigma_1\sigma_2} + \left(\frac{y - m_2}{\sigma_2} \right)^2 \right) \right\} dx dy,$$

où $\Phi_i^{-1}, i = 1, 2$ est la fonction quantile de la loi $N(m_i, \sigma_i)$.

En particulier si $X_1 = X_2 \rightsquigarrow N(0, 1)$ loi normale standard de corrélation ρ on a :

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

copule de Gauss de corrélation ρ .

Comme sa distribution, la copule de Gauss est souvent caractérisée par sa densité c_ρ . Soit h_ρ la densité associée à la distribution H_ρ , de sa représentation canonique (A) on déduit que : $\forall x, y \in \mathbb{R}$;

$$h_\rho(x, y) = f(x)g(y)c_\rho(F(x), G(y)),$$

d'où

$$\begin{aligned} h_\rho(x, y) &= \frac{1}{2\pi\sqrt{1-\rho^2}} \exp \left\{ -\frac{x^2 - 2\rho xy + y^2}{2(1-\rho^2)} \right\} \\ &= \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{x^2}{2} \right\} \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{y^2}{2} \right\} c_\rho(F(x), G(y)). \end{aligned}$$

Par suite : $\forall x, y \in \mathbb{R}$;

$$c_\rho(F(x), G(y)) = \frac{1}{2\pi\sqrt{1-\rho^2}} \exp \left\{ -\frac{x^2 - 2\rho xy + y^2}{2(1-\rho^2)} + \frac{x^2 + y^2}{2} \right\}.$$

Application à la modélisation multivariée des distributions

Le corollaire précédent signifie d'autre part que si H est une distribution bivariée continue de distributions marginales F et G dont les inverses sont F^{-1} et G^{-1} alors la fonction C est définie telle que :

$$\forall u, v \in I, C(u, v) = H(F^{-1}(u), G^{-1}(v)),$$

est l'unique copule vérifiant la représentation canonique de H . En effet l'unicité résulte des propriétés suivantes :

Propriété

- i) Si X est une v.a.c. de distribution F alors la variable $Y = F(X)$ est de loi uniforme $U_{[0,1]}$.
- ii) Si F est une distribution continue d'inverse F^{-1} et X une v.a.c. telle que : $X \rightsquigarrow U_{[0,1]}$ alors $F^{-1}(X) \rightsquigarrow F$.

Rappel

On dit que X est une variable aléatoire absolument continue (v.a.c.) si sa fonction de répartition est continue et dérivable à gauche et à droite de tout point de x de \mathbb{R} .

Conséquence

Si $(x, y) \rightsquigarrow H$ de marges F et G alors $H(F^{-1}(u), G^{-1}(v)) \rightsquigarrow C$.

Réciproquement si $(U, V) \rightsquigarrow C$ alors $H(F^{-1}(u), G^{-1}(v)) \rightsquigarrow H$.

Conclusion 1.2.1 *Toute copule est une distribution indépendante de ses distributions marginales. De plus $\forall u, v \in I, C(u, 1) = C(1, u) = u$ signifie que les distributions marginales de cette distribution sont des lois uniformes.*

Remarque 1.2.2 *Certains auteurs définissent la copule comme une distribution uniforme. Ainsi, le théorème de Sklar fait des copules un outil "puissant" de l'analyse multivariée, car elles permettent de construire des modèles de distributions multivariées compatibles avec les modèles marginaux unidimensionnels, (puisque'on part de ces marges), laquelle compatibilité est souvent très importante dans la modélisation financière (modèles d'estimation de la valeur à risque). Par*

ailleurs, les copules permettent de résoudre un autre problème : l'élaboration des modèles non gaussiens. En effet, il est très difficile de construire des modèles non gaussiens. La famille des distributions non gaussiennes est non seulement réduite mais présente l'inconvénient que les marges sont identiques. Or, avec les copules on peut construire par exemple une distribution avec une marge gaussienne et une marge uniforme ou un inverse gaussien et une Beta.

Exemples de construction de distributions Partant de la copule de Gamlambos :

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

Construisons la distribution bivariée H_θ ayant une marge normale $N(0, 1)$ et une marge uniforme $U_{[0,1]}$.

Pour $\theta \geq 1$ et $\forall x, y \in \mathbb{R}$, on a :

$$H_\theta(x, y) = C_\theta(\Phi(x), y) = \frac{y}{\sqrt{2\pi}} \exp \left\{ -\frac{x^2}{2} - \left[(-\ln \Phi(x))^{-\theta} + (-\ln y)^{-\theta} \right]^{\frac{-1}{\theta}} \right\},$$

où Φ est la fonction de répartition de la loi normale standard.

De même, la distribution bivariée de marges une logistique et une Gumbel standard associée à la copule

$$C(u, v) = \frac{uv}{u + v - uv}; \forall u, v \in I,$$

est donnée, $\forall x, y \in \mathbb{R}$, par :

$$H(x, y) = C \left(\frac{1}{1 + e^{-x}}, \exp(-e^{-x}) \right) = \frac{1}{e^{-x} + \exp(-e^{-x})}.$$

1.3 Les propriétés fondamentales des copules

Comme l'a souligné Fisher dans "Encyclopedia Statistical Sciences" les copules sont d'un grand intérêt pour le statisticien pour deux raisons principales : d'une part, elles permettent de construire des familles de distributions multivariées à partir des distributions marginales univariées données (théorème de

Sklar), et d'autre part, elles constituent un outil de mesure de dépendance entre distributions univariées tout en restant invariantes sous des transformations strictement monotones de celles-ci.

Les résultats suivants résument les propriétés fondamentales des copules.

Théorème 1.3.1 (caractérisation de la copule d'indépendance) *Soient X et Y deux v.a.c. et C_{XY} la copule associée. Alors X et Y sont indépendantes si et seulement si $C_{XY} = \Pi$.*

Preuve. En effet,

1) Si X et Y sont indépendants alors, $\forall (x, y) \in \mathbb{R}^2$; $H(x, y) = F(x)G(y)$,
ou $\forall u, v \in I$, $C_{XY}(u, v) = H(F^{-1}(u), G^{-1}(v))$.

D'où $C_{XY}(u, v) = F(F^{-1}(u))G(G^{-1}(v)) = uv = \Pi(u, v)$.

2) Si $C_{XY} = \Pi$, alors $\forall u, v \in I$, $C_{XY}(u, v) = \Pi(u, v) = uv$,
alors $\forall (x, y) \in \mathbb{R}^2$;

$$\begin{aligned} H(x, y) &= C_{XY}(F(x), G(y)) \\ &= \Pi(F(x), G(y)) \\ &= F(x)G(y). \end{aligned}$$

Donc les variables X et Y sont indépendantes. ■

Théorème 1.3.2 *Soient deux v.a.c. X et Y de copule associée C_{XY} . Si α et β sont des fonctions strictement croissantes sur $\text{Im}(X)$ et $\text{Im}(Y)$ respectivement alors on a :*

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

Ce théorème révèle une propriété importante de l'outil copule : elle reste invariante sous des transformations strictement croissantes de ses distributions marginales.

Preuve. Notons, d'une part, H la distribution conjointe des v.a.c. X et Y de lois respectives F et G et, d'autre part, H' la distribution de la transformée $(\alpha(X), \beta(Y))$ de marges respectives F' et G' . on a :

$$\begin{aligned} \forall x \in \mathbb{R}; F'(x) &= P(\alpha(X) \leq x) = P(X \leq \alpha^{-1}(x)) = F(\alpha^{-1}(x)). \\ (\alpha^{-1} \text{ existe car } \alpha \text{ est strictement } \nearrow) \end{aligned}$$

De même on a : $\forall y \in \mathbb{R}; G'(y) = G(\beta^{-1}(x));$ et donc $F'^{-1}(u) = \alpha(F^{-1}(u)),$
 et $G'^{-1}(v) = \beta(G^{-1}(v)).$

Il vient que : $\forall u, v \in I,$

$$\begin{aligned} C_{\alpha(X)\beta(Y)}(u, v) &= H'(F'^{-1}(u), G'^{-1}(v)) \\ &= P[\alpha(X) < F^{-1}(u), \beta(Y) < G^{-1}(v)] \\ &= P[X \leq \alpha^{-1}[F^{-1}(u)], Y \leq \beta^{-1}[G^{-1}(v)]] \\ &= P[X \leq F^{-1}(u), Y \leq G^{-1}(v)] \\ &= H(F^{-1}(u), G^{-1}(v)) \\ &= C_{XY}(u, v). \end{aligned}$$

■

Par exemple, la copule de la distribution lognormale est la même que celle de la loi normale (en effet la première est une transformation strictement croissante ($y = \log x$) de la seconde).

Théorème 1.3.3 Soient X et Y deux variables aléatoires continues de copule C_{XY} . Si α et β sont des fonctions strictement croissantes sur $\text{Im}(X)$ et $\text{Im}(Y)$ respectivement alors :

i) Si α est strictement croissante et β strictement décroissante alors :

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

ii) Si α est strictement décroissante et β strictement croissante alors :

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

iii) Si α et β sont strictement décroissantes alors :

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

1.3.1 Copules associées à une copule

La copule de survie

Soit X une variable aléatoire de distribution F . On note $\bar{F} = 1 - F$ la fonction de survie associée à F i.e.

$$\forall x \in \mathbb{R}; \bar{F}(x) = P(X > x) = 1 - F(x).$$

Par exemple, si la v.a.c. X modélise la durée de vie d'un individu au sein d'une population alors $\bar{F}(x)$ est la probabilité que l'individu vive ou survive au delà du temps x . De même, on peut associer à la distribution conjointe H d'un couple de variables aléatoires (X, Y) une distribution de survie \bar{H} par :

$\bar{H}(x, y) = P(X > x, Y > y)$ dont les distributions marginales :

$$\bar{F}(x) = \lim_{y \rightarrow -\infty} H(x, y),$$

et

$$\bar{G}(y) = \lim_{x \rightarrow -\infty} H(x, y).$$

Problème : Comment H s'exprime en fonction de ses marges ?

$\forall (x, y) \in \mathbb{R}^2;$

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

(en effet, probabilité élémentaire) :

$$P(\bar{A} \cap \bar{B}) = P(\overline{A \cup B}) = 1 - P(A) - P(B) + P(A \cap B).$$

Alors

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

Donc $\forall (x, y) \in \mathbb{R}^2;$

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

Soit C la copule associée à la distribution H . En considérant la fonction \hat{C} telle que :

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

alors on obtient

$$\forall (x, y) \in \mathbb{R}^2; \bar{H}(x, y) = \hat{C}(\bar{F}(x), \bar{G}(y)).$$

On vérifie que \hat{C} est une copule, la copule de survie de C . Elle associe à chaque distribution conjointe de survie ses distributions marginales de survie.

Remarque 1.3.1 *Il ne faut pas confondre la copule de survie \hat{C} et la distribution de survie \bar{C} associée à C (puisque C est aussi une distribution).*

On a plutôt :

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

et on vérifie que \hat{C} et \bar{C} sont liées par la relation :

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

ou simplement $\forall u, v \in I, \hat{C}(u, v) = \bar{C}(\bar{u}, \bar{v})$,

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

Dans un raisonnement similaire on peut exprimer les probabilités :

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

La copule duale

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

En posant $\tilde{C}(F(x), G(y)) = F(x) + G(y) - C(F(x), G(y))$.

On vérifie que :

$$\forall u, v \in I, \tilde{C}(u, v) = u + v - C(u, v);$$

\tilde{C} est dite copule duale de C .

Attention : \tilde{C} n'est pas une copule car $\tilde{C}(1, 0) = 1 \neq 0$.

La co-copule

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

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

Alors $P(X > x \text{ ou } Y > y) = 1 - C(1 - \bar{F}(x), 1 - \bar{G}(y))$.

En posant $C^*(F(x), G(y)) = 1 - C(1 - \bar{F}(x), 1 - \bar{G}(y))$.

On vérifie que : $\forall u, v \in I, C^*(u, v) = 1 - C(1 - u, 1 - v)$; C^* est une co-copule de C .

La copule mixte

$$\begin{aligned} P(X > x, Y \leq y) &= F(x) - \bar{C}(F(x), G(y)) \\ &= 1 - G(y) - C(1 - \bar{F}(x), 1 - \bar{G}(y)). \end{aligned}$$

En posant $C^*(F(x), G(y)) = 1 - G(y) - C(1 - \bar{F}(x), 1 - \bar{G}(y))$.

On vérifie que : $\forall u, v \in I, \check{C}(u, v) = v - 1 + C(1 - u, 1 - v)$, copule mixte.

De même on établit que $P(X \leq x, Y > y) = \check{C}(u, v) = v - 1 + C(1 - u, 1 - v)$, la copule mixte.

A partir de chaque copule C on peut construire trois autres copules :

1) $\hat{C}(u, v) = u + v - C(1 - u, 1 - v) \longrightarrow$ la copule de survie.

2) $\tilde{C}(u, v) = u + v - C(u, v) \longrightarrow$ la copule duale.

3) $C^*(u, v) = 1 - C(1 - u, 1 - v) \longrightarrow$ la co-copule.

4) $\check{C}(u, v) = 1 - v - C(1 - u, 1 - v)$ et $\check{C}(u, v) = v - 1 + C(1 - u, 1 - v) \longrightarrow$ les copules mixtes.

Ces copules associées sont très importantes dans la modélisation financière.

Chapitre 2

Familles paramétriques de copules usuelles bivariées

Nous exposons dans la présente section les principales copules paramétriques. Ces copules présentent un intérêt particulier dans la gestion des risques (modélisation financière), de part le fait qu'elles autorisent la construction de modèles paramétriques ou semi-paramétriques.

2.1 Les copules usuelles

2.1.1 La copule d'indépendance

La copule d'indépendance a la forme suivante :

Π ; $\Pi(u, v) = uv$ (C'est à dire les v.a. X et Y sont indépendantes).

2.1.2 La copule (minimale et maximale)

La copule minimale W , $\forall u, v \in I$, $W(u, v) = \max(u + v - 1, 0)$, borne inf de Fréchet.

La copule maximale M , $M(u, v) = \min(u, v)$, borne sup de Fréchet.

Pour toute copule C on a la relation :

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

La famille de Fréchet formée par les combinaisons linéaires convexes de M , Π et W . Par exemple la copule de Mardia :

$$\forall \theta \in [-1, 1], \forall u, v \in I ;$$

$$C_{\theta}(u, v) = \frac{\theta^2 (1 + \theta)}{2} M(u, v) + (1 - \theta^2) \Pi(u, v) + \frac{\theta^2 (1 - \theta)}{2} W(u, v).$$

La borne supérieure C^+ de Fréchet sera appelée La copule comonotone , et la borne inférieure de Fréchet C^- sera appelée copule anticomonotone.

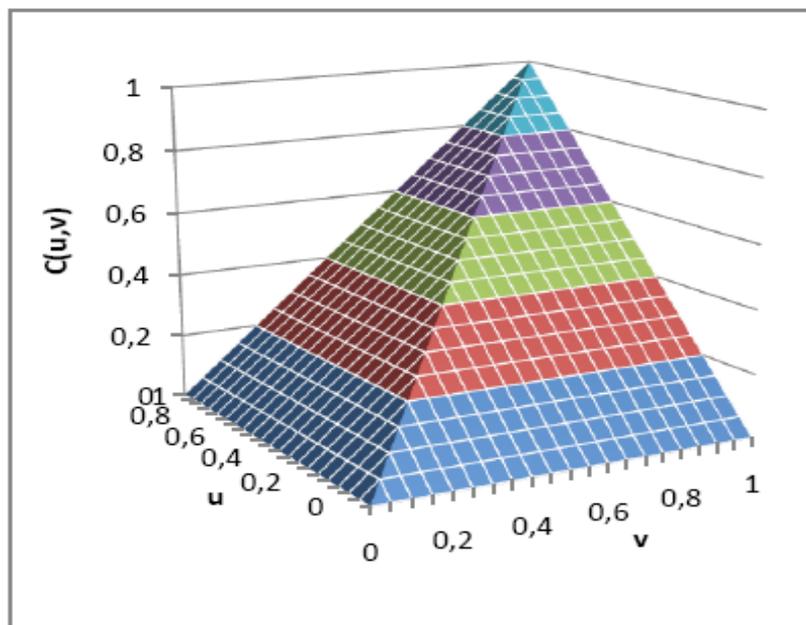


Figure 1 : Représentation de la copule comonotone

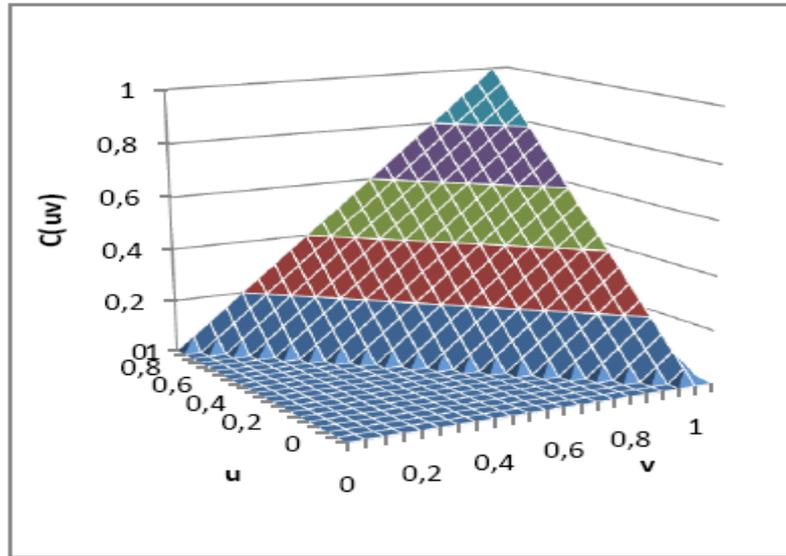


Figure 2 : Représentation de la copule anticomonotone.

- La borne inférieure de Fréchet n'est pas une copule que dans le cas où $d = 2$.
- Les bornes de Fréchet n'admettent pas de densité.
- Toutes les copules sont comprises entre ces deux bornes.

2.2 Les copules elliptiques

2.2.1 Caractérisation

Définition 2.2.1 On appelle copule elliptique toute copule de la forme :

$\forall u, v \in I :$

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

$$= H_\rho(\Phi_{g,1}^{-1}(u), \Phi_{g,2}^{-1}(v)),$$

où H_ρ est la distribution conjointe des variables X et Y , $\Phi_{g,1}^{-1}$ et $\Phi_{g,2}^{-1}$ les fonctions quantiles respectives et ρ leur coefficient de corrélation.

2.2.2 Copules élliptiques classiques

La Copule normale bivariée

Tout comme les distributions normale et binormale, un des types de copules beaucoup utilisées dans la modélisation financière est la copule normale bivariée. Si X et Y sont deux variables normales standard de corrélation ρ leur copule est donnée par :

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

où Φ^{-1} est la fonction quantile de la loi normale standard $N(0, 1)$.

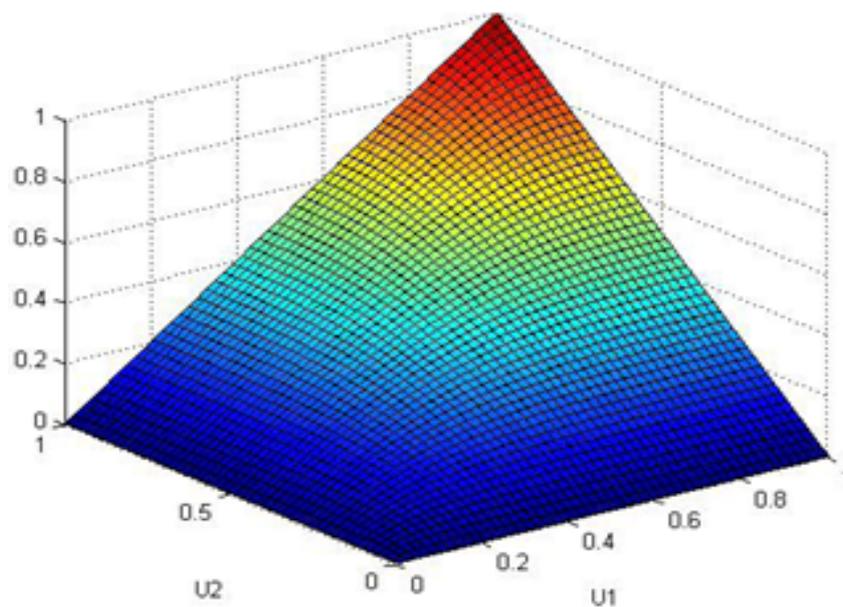


Figure 3 : Copule normale

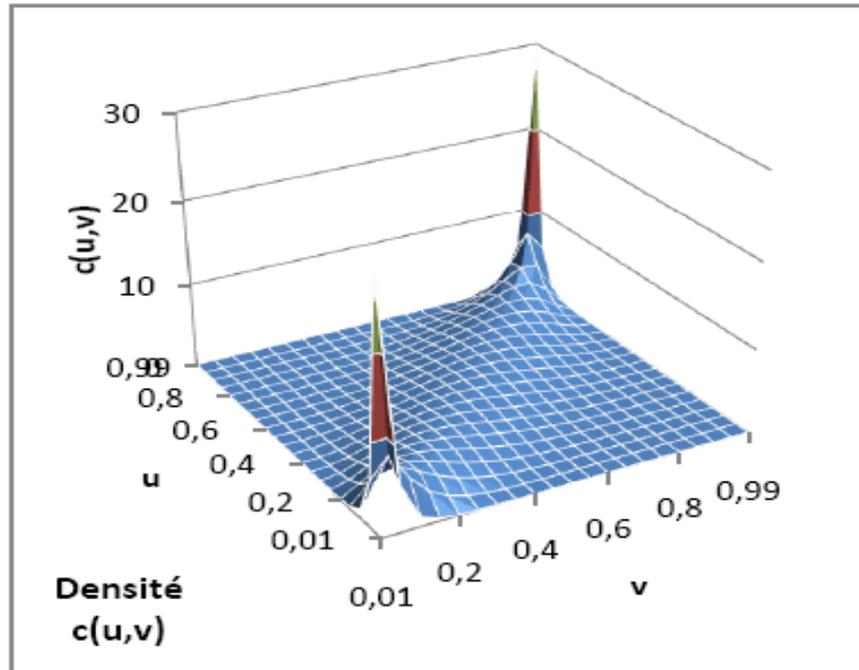


Figure 4 : *Densité de la copule normale* ($\tau = 0.7 \Leftrightarrow \rho = 0.89$)

La Copule de Student

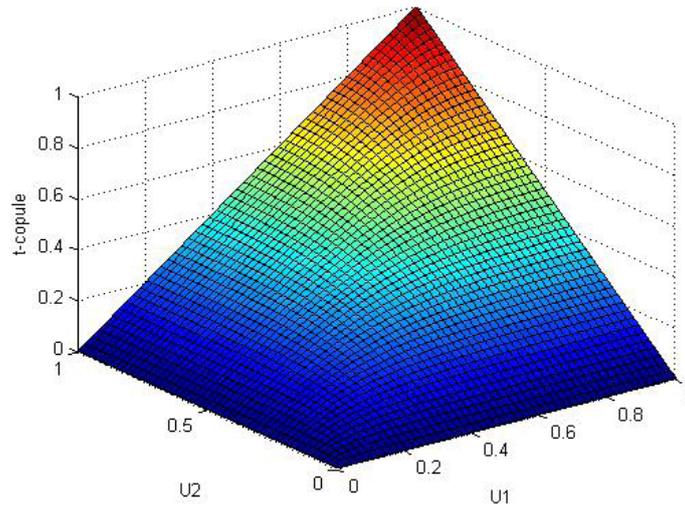
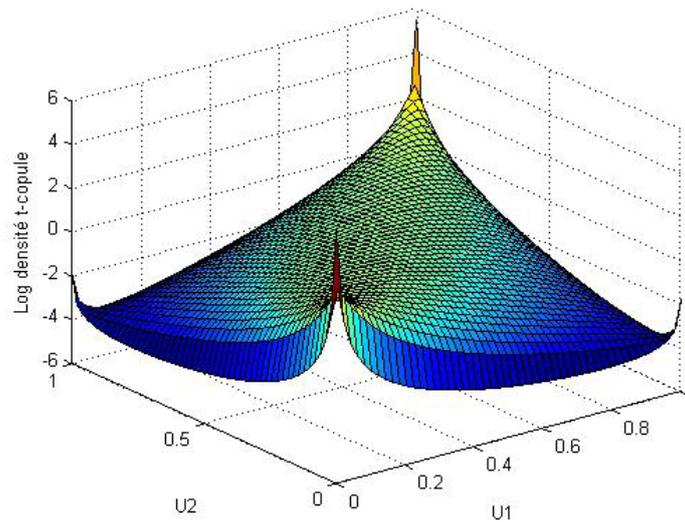
Soit une matrice diagonale définie positive avec $\text{diag } \rho = 1$ et $t_{\rho,\nu}$, la distribution de Student bivariable standard à ν degrés de liberté et matrice de corrélation.

La copule de Student associée à cette distribution est alors définie de la façon suivante :

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

où t_{ν}^{-1} est la fonction inverse de la distribution standard de Student à ν degrés de liberté.

La copule Student est extraite de la même manière que la copule Gaussienne.

Figure 5 : *La Copule de Student*Figure 6 : *Log densité de la copule de Student*

Remarque 2.2.1 *Les copules gaussienne et de Student, dites copules elliptiques, elles sont en effet moins bien adaptées en assurance car elles s'appliquent à*

des distributions symétriques. Elles 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'exprime par conséquent en fonction de distributions bivariées à eux, associées par le théorème de Sklar.

La copule mixte normale

C'est une combinaison convexe de deux copules normales. Soient (X_1, Y_1) et (X_2, Y_2) deux couples de variables de copules normales C_{δ_1} et C_{δ_2} respectivement.

Soit (X, Y) un couple de variables, égal à (X_1, Y_1) avec une probabilité P et égal à (X_2, Y_2) avec une probabilité $1 - P$.

La copule associée à (X, Y) est la copule C_P telle que $\forall u, v \in I$;

$$C_P(u, v) = PC_{\delta_1}(u, v) + (1 - P)C_{\delta_2}(u, v).$$

C_P est une copule élliptique.

2.3 Les copules aux valeurs extrêmes

2.3.1 Caractérisation

On appelle copule des valeurs extrêmes, toute copule C vérifiant la propriété suivante :

$$\forall u, v \in I, \forall t > 0; C(u^t, v^t) = C^t(u, v),$$

ou encore

$$\forall t > 0; C^{\frac{1}{t}}(u^t, v^t) = C(u, v).$$

Par la suite on tentera de justifier l'appellation de ces copules en établissant l'existence d'un lien entre la caractérisation ci-dessus et la théorie des valeurs extrêmes bivariées.

Considérons la copule :

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

(Copule de Galambos)

$\forall t > 0$, on a :

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

donc est une copule bivariée de valeurs extrêmes notée BEV.

2.4 Les copules Archimédiennes

Dans la littérature financière, plusieurs fonctions de copules ont été utilisées, les principales sont la copule Normale, la copule de Student et les copules dites archimédiennes, comme la famille de Clayton, de Frank et de Gumbel. Nelsen (1999) donne une importante liste de familles de copules dont la plus importante est la famille des copules archimédiennes auxquelles, par exemple, les copules de Clayton, de Frank et de Gumbel, de plus en plus utilisées en finance. Les copules archimédiennes ont un double avantage. Outre le fait que la plupart de ces copules ont une expression analytique, elles permettent de prendre en compte un large éventail de structures de dépendance.

2.4.1 Caractérisation

On appelle copule archimédienne toute copule de la forme :

$$C(u, v) = \begin{cases} \varphi^{-1}(\varphi(u) + \varphi(v)) & \text{si } \varphi(u) + \varphi(v) \leq 0. \\ 0 & \text{sinon.} \end{cases}$$

Avec φ une fonction de classe C^2 qui vérifie

$$\varphi(1) = 0, \varphi'(u) < 0 \text{ et } \varphi''(u) > 0 \quad \forall u, v \in I,$$

où $\varphi : [0, 1] \rightarrow [0, +\infty[$ est une fonction convexe décroissante telle que $\varphi(1) = 0$ et φ^{-1} la pseudo-inverse de φ .

On pose $\varphi(0) = +\infty$ si $\lim_{t \rightarrow +\infty} \varphi(t) = +\infty$ et $\varphi^{-1}(t) = 0$ si $t \geq \varphi(0)$.

La fonction φ est dite générateur archimédien de la copule C .

Si $\varphi(0) = +\infty$, alors la copule C est dite strictement archimédienne auquel cas φ^{-1} coïncide avec la fonction réciproque de φ .

2.4.2 Familles classiques de copules archimédiennes

Exemple 2.4.1 $\forall \theta \geq 0$ et $\forall t > 0$ la fonction $\varphi_\theta(t) = (-\ln t)^\theta$ satisfait à la caractérisation ci-dessus avec $\varphi_\theta^{-1}(t) = \exp\left(-t^{\frac{1}{\theta}}\right)$, elle définit donc un générateur archimédien et la copule C_θ associée est telle que :

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

c'est la copule de Gumbel.

Le cas particulier $\theta = 1$ correspond à la copule indépendante $\Pi(u, v) = uv$.

Quelques familles classiques de copules archimédiennes

Le tableau suivant donne quelques familles classiques de copules archimédiennes :

Famille	Générateur $\varphi(u)$	Copule $C(u, v)$
Indépendante	$-\ln u$	$\Pi(u, v) = uv$
Clayton	$u^{-\theta} - 1$	$(u^{-\theta} + v^{-\theta} - 1)^{\frac{-1}{\theta}}; \theta \geq 0$
Gumbel	$(-\ln u)^\theta; \theta \geq 1$	$\exp\left\{-\left[(-\ln u)^{-\theta} + (-\ln v)^{-\theta}\right]^{\frac{-1}{\theta}}\right\}$
Frank	$-\log\left(\frac{e^{-\theta u} - 1}{e^{-\theta} - 1}\right)$	$\frac{-1}{\theta} \log\left(1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{e^{-\theta} - 1}\right); \theta \neq 0$

Tableau 1 : Quelques familles classiques de copules archimédiennes

La copule d'indépendance

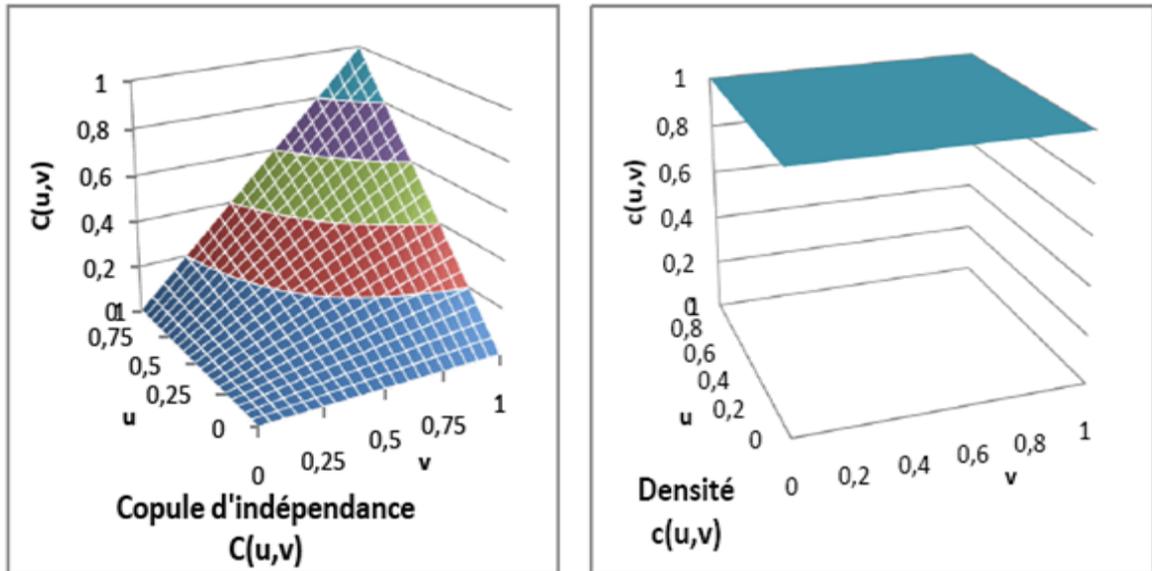


Figure 7 : Copule d'indépendance

La densité de cette copule est une constante.

La copule de Gumbel

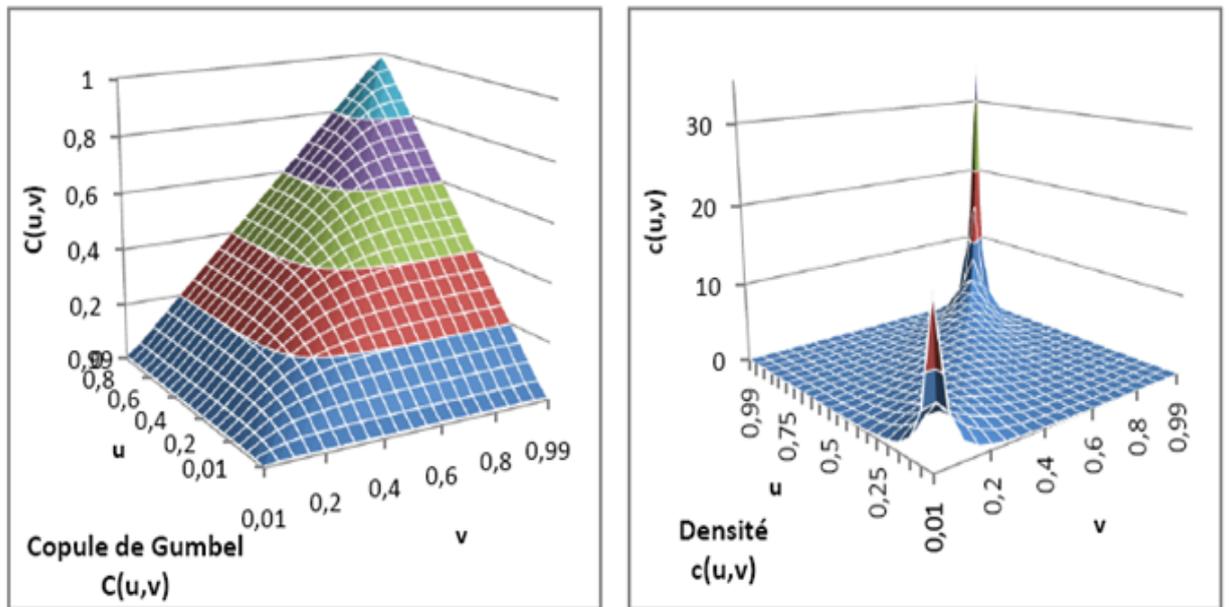
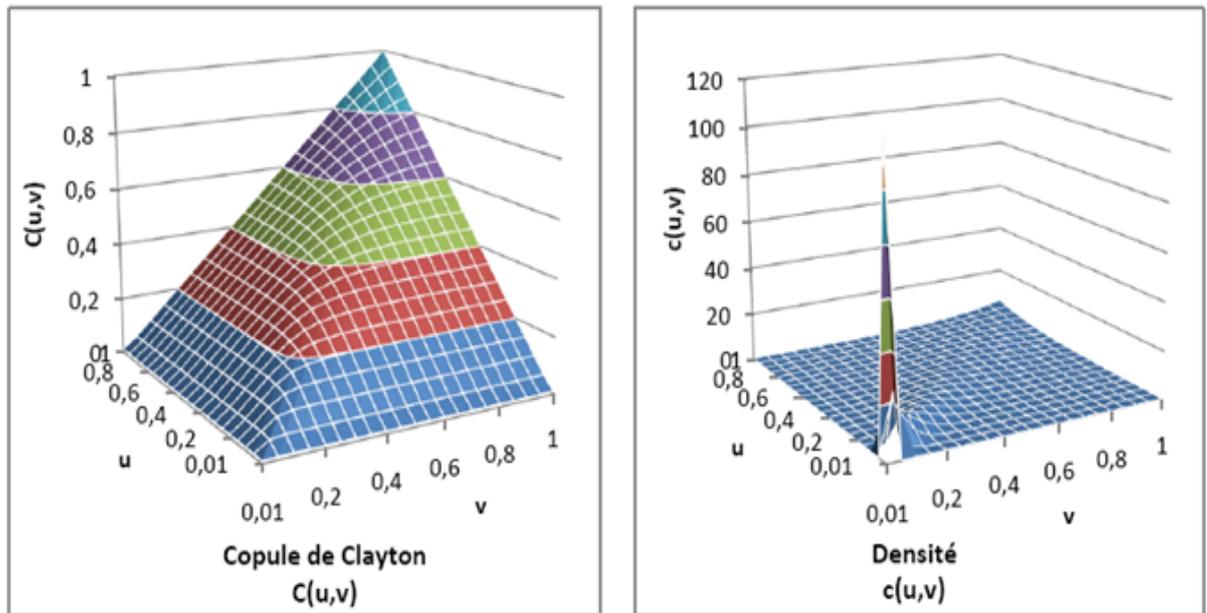


Figure 8 : Représentation de la copule de Gumbel ($\tau = 0.7 \Leftrightarrow \theta = 3.33$)

La copule de Gumbel n'appréhende que des dépendances positives et possède la caractéristique de pouvoir représenter des risques dont la structure de dépendance est plus accentuée sur la queue supérieure. Elle est à ce titre particulièrement adaptée en assurance et en finance pour étudier l'impact de la survenance d'événements de forte intensité sur la dépendance entre branches d'assurance ou actifs financiers.

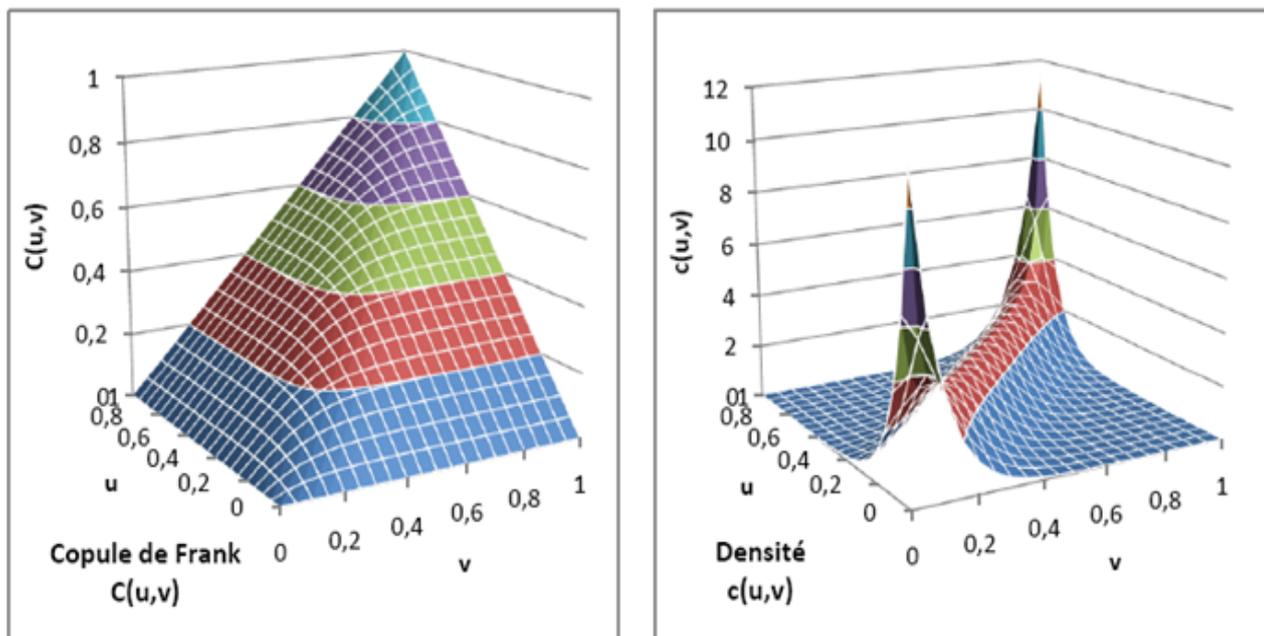
Le paramètre de la copule de Gumbel étant nécessairement supérieur à 1, celle-ci n'appréhende que les dépendances positives. Elle représente bien les risques ayant une structure de dépendance plus accentuée que la queue. C'est pourquoi elle est particulièrement adaptée en assurance et en finance, notamment lorsque l'on étudie la dépendance entre les événements de forte intensité.

La copule de Clayton

Figure 9 : Représentation de la copule de Clayton ($\tau = 0.7 \Leftrightarrow \theta = 4.67$)

Comme la copule de Gumbel, la copule de Clayton ne permet de modéliser que les dépendances positives. A l'inverse de la copule de Gumbel, elle vise à rendre compte d'une dépendance sur les événements de faible intensité.

La copule de Frank

Figure 10 : Représentation de la copule de Frank ($\tau = 0.7 \Leftrightarrow \theta = 11.4$)

La copule de Frank permet de modéliser les dépendances aussi bien positives que négatives. On note qu'il n'existe pas de dépendance de queue pour cette copule. Elle, étant la seule copule archimédienne à respecter l'équation :

$$C(u, v) = C_F(u, v),$$

celle-ci risque de ne pas être tout à fait appropriée à des applications dans le domaine de l'assurance. Elle est symétrique dans la queue inférieure et supérieure, et a donc tendance à corrélérer les petits sinistres comme les gros sinistres.

Dans les applications liées à l'assurance, il existe très certainement une plus forte dépendance entre les gros sinistres qu'entre les petits. De telles asymétries ne peuvent être mises en valeur à l'aide des copules elliptiques (les copules elliptiques vérifient $C = C_F$).

Remarque 2.4.1 Les Copules Archimédiennes sont définies de manière beaucoup plus simple que les copules elliptiques, en tout cas lorsque l'on étudie des phénomènes en deux dimensions. Les générateurs permettent également de calculer le tau de Kendall.

2.5 La famille Archimax

Cette famille de copules a l'avantage d'englober un grand nombre de copules.

2.5.1 Caractérisation

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

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

pour tout $u, v \in [0, 1]$, où

$A : [0, 1] \rightarrow [\frac{1}{2}, 1]$ telle que : $\max(t, 1-t) \leq A(t) \leq 1$ pour tout $0 \leq t \leq 1$;

$\Phi :]0, 1[\rightarrow [0, +\infty[$ est une fonction convexe, décroissante qui vérifie $\Phi(1) = 0$.

On adoptera la notation $\Phi(0) = \lim_{t \rightarrow +\infty} \Phi(t)$, et $\Phi^{-1}(s) = 0$ pour $s \geq \Phi(0)$.

Cas particuliers

Pour $\Phi(t) = -\ln t$, on obtient

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

avec $u, v \in [0, 1]$. C'est la forme générale des copules des valeurs extrêmes.

En choisissant $A = 1$ on retrouve la forme générale des copules archimédiennes :

$$C_{\Phi,A}(u, v) = C_{\Phi}(u, v) = \Phi^{-1}(\Phi(u) + \Phi(v)).$$

2.6 La copule empirique

Les copules empiriques ont été initialement introduites par Deheuvels (1979) et connue sous l'appellation Fonction empirique de dépendance.

Soient $E_n = \{(x_k, y_k), k = 1, \dots, n\}$, un échantillon de taille n de couples de variables aléatoire. La copule empirique est donc la fonction C_n définie par :

$$C_n \left(\frac{i}{n}, \frac{j}{n} \right) = \frac{\text{Nombre de paires } (x, y) \text{ dans l'échantillon tels que } x \leq x_i, y \leq y_i}{n}$$

La fonction densité empirique de la copule C notée c est donnée par :

$$c_n \left(\frac{i}{n}, \frac{j}{n} \right) = \begin{cases} \frac{1}{n} & \text{si } (x_i, y_j) \in E_n \\ 0 & \text{sinon.} \end{cases}$$

On note que la relation entre C_n et c_n est définie par :

$$C_n \left(\frac{i}{n}, \frac{j}{n} \right) = \sum_{p=1}^i \sum_{q=1}^j c_n \left(\frac{p}{n}, \frac{q}{n} \right),$$

$$c_n \left(\frac{i}{n}, \frac{j}{n} \right) = C_n \left(\frac{i}{n}, \frac{j}{n} \right) - C_n \left(\frac{i-1}{n}, \frac{j}{n} \right) - C_n \left(\frac{i}{n}, \frac{j-1}{n} \right) + C_n \left(\frac{i-1}{n}, \frac{j-1}{n} \right)$$

Chapitre 3

Mesures de concordance

3.1 Mesure de dépendance ou corrélation

L'objet de ce paragraphe est de définir ce qu'est une mesure de dépendance, et de présenter les mesures les plus utilisées. Une mesure de dépendance est une application qui associe à deux variables aléatoires ; un réel permettant de quantifier la force de la dépendance qui lie les deux variables aléatoires. Pour être jugée satisfaisante, une mesure de dépendance doit satisfaire un nombre de propriétés, entre autres être une mesure de concordance.

3.2 Mesure de concordance

La copule C d'un vecteur aléatoire continue (X, Y) est une paramétrisation, une normalisation de la distribution conjointe H après avoir éliminé les effets des marges. C'est donc une structure de dépendance entre les deux variables aléatoires X et Y connaissant leurs distributions respectives. Cette structure permet des estimations des différents moyens d'étudier cette dépendance à travers les mesures de concordance (corrélation linéaire, la concordance, tau de Kendall et rho de Spearman).

3.2.1 Fonction de concordance

Notion de concordance

Deux observations (x_1, y_1) et (x_2, y_2) d'un couple de v.a.c. (X, Y) sont dites :

a)Concordantes si :

$$(x_1 - x_2)(y_1 - y_2) > 0 \iff \begin{cases} (x_1 < x_2 \text{ et } y_1 < y_2), \\ \text{ou} \\ (x_1 > x_2 \text{ et } y_1 > y_2). \end{cases}$$

b)Discordantes si :

$$(x_1 - x_2)(y_1 - y_2) < 0 \iff \begin{cases} (x_1 < x_2 \text{ et } y_1 > y_2), \\ \text{ou} \\ (x_1 > x_2 \text{ et } y_1 < y_2). \end{cases}$$

Plus généralement, soit $(x_1, y_1); (x_2, y_2); \dots; (x_n, y_n)$ un échantillon d'observations d'un couple (X, Y) . Il existe C_n^2 paires distinctes de couples (x_i, y_i) et (x_j, y_j) qui sont, soit concordantes, soit discordantes. Soit (X_1, Y_1) et (X_2, Y_2) deux couples 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 3.2.1 *On appelle fonction de concordance entre deux couples (X_1, Y_1) et (X_2, Y_2) par :*

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

C'est la différence entre la probabilité de concordance et celle de discordance. Chaque distribution conjointe étant caractérisée par une copule unique par le théorème de Sklar, le résultat suivant permet d'établir une relation entre toute fonction de concordance Q et les copules associées aux deux couples aléatoires.

Propriétés de la fonction de concordance

Théorème 3.2.1 *Soient (X_1, Y_1) et (X_2, Y_2) deux couples de v.a. indépendants de distributions conjointes H_1 et H_2 avec des marges communes F et G respectivement. Soient C_1 et C_2 les copules associées aux distributions H_1 et H_2 respectivement. Alors :*

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

Preuve.

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

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

Par ailleurs

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

De façon similaire $P(X_1 < X_2, Y_1 < Y_2) = P(X_2 > X_1, Y_2 > Y_1)$,

donc :

$$\begin{aligned} P(X_1 < X_2, Y_1 < Y_2) &= \int_{\mathbb{R}^2} P(X_2 > x, Y_2 > y) dC_1[(F(x), G(y))] \\ &= \int_{\mathbb{R}^2} \bar{H}_2(x, y) dC_1[(F(x), G(y))], \quad (\bar{H}_2, \text{ fonction de survie de } H_2) \\ &= \int_{\mathbb{R}^2} [1 - F(x) - G(y) + C(F(x), G(y))] dC_1[(F(x), G(y))] \\ &= \int_{I^2} [1 - u - v + C_2(u, v)] dC_1(u, v) \\ &= 1 - \frac{1}{2} - \frac{1}{2} + \int_{I^2} C_2(u, v) dC_1(u, v) \\ &= \int_{I^2} C_2(u, v) dC_1(u, v). \end{aligned}$$

Par suite : $Q = 4 \int_{I^2} C_2(u, v) dC_1(u, v) - 1$ ■

Le résultat suivant résume les propriétés essentielles de la fonction Q .

Corollaire 3.2.1 Soit C_1 et C_2 et Q donnés par le théorème (3.2.1), alors :

- i) Q est symétrique ie : $Q(C_1, C_2) = Q(C_2, C_1)$
- ii) Q conserve l'ordre ie si : $C_1 \prec \hat{C}_1$ et $C_2 \prec \hat{C}_2$ alors $Q(C_1, C_2) \prec Q(\hat{C}_1, \hat{C}_2)$
- iii) Q est invariante par rapport à la survie ie : $Q(C_1, C_2) \prec Q(\hat{C}_1, \hat{C}_2)$

Preuve. ii) $Q(\hat{C}_1, \hat{C}_2) - Q(C_1, C_2) = 4 \left[\int_{I^2} \hat{C}_2(u, v) d\hat{C}_1(u, v) - \int_{I^2} C_2(u, v) dC_1(u, v) \right],$

où $\frac{\partial C_1}{\partial x}$ et $\frac{\partial \hat{C}_1}{\partial x}$ sont croissantes par rapport à chaque variable (2-croissance).

Alors $\frac{\partial^2 C_1}{\partial u \partial v} < \frac{\partial^2 \hat{C}_1}{\partial u \partial v}$ donc $dC_1(u, v) < d\hat{C}_1(u, v)$.

D'où $Q(\hat{C}_1, \hat{C}_2) - Q(C_1, C_2) > \int_{I^2} [\hat{C}_2(u, v) - C_2(u, v)] dC_1(u, v) > 0.$

Pour toute copule C , on a : $dC(u, v) = \frac{\partial^2 C}{\partial u \partial v}(u, v) du dv,$

et $\hat{C}(u, v) = 1 - u - v + C(u, v)$. Il s'en suit que : $dC(u, v) = d\hat{C}(u, v)$.

$$\begin{aligned} Q(\hat{C}_1, \hat{C}_2) &= 4 \int_{I^2} \hat{C}_2(u, v) d\hat{C}_1(u, v) - 1 \\ &= 4 \int_{I^2} [1 - u - v + C_2(u, v)] dC_1(u, v) - 1 \\ &= 4 \int_{I^2} (1 - u - v) dC_1(u, v) + 4 \int_{I^2} C_2(u, v) dC_1(u, v) - 1, \end{aligned}$$

où $4 \int_{I^2} (1 - u - v) dC_1(u, v) = 1 - \frac{1}{2} - \frac{1}{2} = 0.$

Alors :

$$Q(\hat{C}_1, \hat{C}_2) = 4 \int_{I^2} C_2(u, v) dC_1(u, v) - 1 = Q(C_1, C_2).$$

■

Le résultat suivant donne les fonctions de dépendance des copules usuelles W , et M pris deux à deux.

Propriété

$Q(M, M) = 1$; $Q(M, \Pi) = \frac{1}{3} = Q(W, \Pi)$; $Q(M, W) = 0 = Q(\Pi, \Pi)$ et $Q(W, W) = -1.$

Pour toute copule C on a :

$$0 \leq Q(C, M) \leq 1, \quad -1 \leq Q(C, W) \leq 1 \quad \text{et} \quad \frac{-1}{3} \leq Q(C, M) \leq \frac{1}{3}.$$

En effet,

1) Le support de la copule usuelle M est l'ensemble $D_M = \{(u, v) \in I^2 : u = v\}$.

$$\begin{aligned} Q(M, M) &= 4 \int_{I^2} M(u, v) dM(u, v) - 1 \\ &= 4 \int_{I^2} M(u, u) dM(u, u) - 1 \quad \text{car } D_M = \{(u, v) \in I^2 : u = v\} \\ &= 4 \int_0^1 u du - 1 = 4 \left[\frac{1}{2} u^2 \right] du = 1. \end{aligned}$$

$$\begin{aligned} Q(M, \Pi) &= 4 \int_{I^2} \Pi(u, v) dM(u, v) - 1 \\ &= 4 \int_{I^2} u^2 du - 1 = \frac{1}{3}. \end{aligned}$$

$$\begin{aligned} Q(M, W) &= 4 \int_{I^2} W(u, v) dM(u, v) - 1 \\ &= 4 \int_{I^2} (2u - 1) du = 0. \end{aligned}$$

De même, le support de la copule usuelle W est $D_W = \{(u, v) \in I^2 : u = 1 - v\}$.

$$\begin{aligned} Q(W, \Pi) &= 4 \int_{I^2} \Pi(u, v) dM(u, v) - 1 \\ &= 4 \int_{I^2} u^2 du - 1 = \frac{1}{3} \end{aligned}$$

$$\begin{aligned} Q(W, W) &= 4 \int_{I^2} W(u, v) dM(u, v) - 1 \\ &= 4 \int_{I^2} (2u - 1) du - 1 = 0 \end{aligned}$$

- Finalement, puisque $d\Pi(u, v) = dudv$ on a :

$$\begin{aligned} Q(\Pi, \Pi) &= 4 \int_{I^2} \Pi(u, v) d\Pi(u, v) - 1 \\ &= 4 \int_0^1 \int_0^1 uv dudv - 1 = 0. \end{aligned}$$

- 2) Pour toute copule quelconque C on a :

$$W \leq C \leq M.$$

Alors

$$Q(M, W) \leq Q(C, M) \leq Q(M, M).$$

Donc

$$0 \leq Q(C, M) \leq 1.$$

De la même manière, on établit que :

$$-1 \leq Q(C, W) \leq 1 \text{ et } -\frac{1}{3} \leq Q(C, M) \leq \frac{1}{3}.$$

3.3 Mesure de concordance

3.3.1 Caractérisation d'une mesure de concordance

Définition 3.3.1 (d'une mesure de concordance) Une mesure d'association K entre deux variables aléatoires continues X et Y de copule C est une mesure de concordance si elle vérifie les propriétés suivantes :

- 1) K est définie pour tout couple (X, Y) de variables aléatoires continues.
- 2) $-1 = K_{X, -X} \leq K_{Y, X} \leq K_{X, X} = 1$.
- 3) $K_{X, Y} = K_{Y, X}$.
- 4) Si X et Y sont indépendantes alors $K_{X, Y} = 0$.
- 5) $K_{-X, Y} = K_{X, -Y} = -K_{X, Y}$.
- 6) Si C_1 et C_2 sont deux copules telles que $C_1 < C_2$ alors on a : $K_{C_1} < K_{C_2}$.
- 7) Si $\{(X_n, Y_n)\}$ est une suite de couples aléatoires dont la copule est C_n et si $\{C_n\}$ converge vers C alors K_{C_n} converge vers K_C .

Deux des mesures de concordance sont plus connues et jouent un rôle très important en statistique non paramétriques : le tau de Kendall et le rho de Spearman.

3.3.2 Le coefficient de corrélation de Kendall (Le tau de Kendall)

Le tau de Kendall est une mesure de concordance bien connue en statistique. Elle donne une mesure de la corrélation entre les rangs des observations, à la différence du coefficient de corrélation linéaire, qui lui apprécie la corrélation entre les valeurs des observations. Elle offre par ailleurs l'avantage de s'exprimer simplement en fonction de la copule associée au couple de variables aléatoires.

Soit $(x_1, y_1); (x_2, y_2); \dots; (x_n, y_n)$ un échantillon de n observations d'un couple (X, Y) . On définit le tau de Kendall (version échantillon) par :

$$\tau = \frac{[\text{nombre de paires concordantes}] - [\text{nombre de paires discordantes}]}{\text{nombre total de paires}}.$$

Soient (X_1, X_2) et (Y_1, Y_2) deux observations d'un couple aléatoire (X, Y) . Si (X_1, X_2) et (Y_1, Y_2) sont i.i.d. de la loi H ; on définit le tau de Kendall par :

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

Remarque 3.3.1 *Si les couples sont identiquement distribués i.e.*

$H_1 = H_2 = H$, alors :

$\tau_{XY} = Q_{XY}$ si C est la copule associée au couple (X, Y) .

Le résultat suivant donne la conséquence immédiate du théorème (3.2.1).

Théorème 3.3.1 *Soit X et Y deux variables aléatoires continues de copule C . Alors le tau de Kendall de X et Y est donné par :*

$$\tau_{XY} = \tau_C = Q(C, C) = 4 \int_{I^2} C(u, v) dC(u, v) - 1.$$

Soient U et V deux variables aléatoires ; de loi jointe C uniformément distribuées sur I ; donc

$$E[C(U, V)] = \int_{I^2} C(u, v) dC(u, v).$$

Ainsi, on peut écrire :

$$\begin{aligned}
\tau_C &= 4E[C(U, V)] \\
&= 4 \int_0^1 t dF_C(t) - 1 \\
&= 4 [tF_C(t)]_0^1 - 4 \int_0^1 F_C(t) dt - 1 \\
&= 4 - 4 \int_0^1 F_C(t) dt - 1 \\
&= 3 - 4 \int_0^1 F_C(t) dt,
\end{aligned}$$

où F_C est la distribution de la variable aléatoire $C(U, V)$.

Exemple 3.3.1 Pour toute copule de la famille de Fréchet

$$C_{\alpha, \beta} = \alpha M + (1 - \alpha - \beta) \Pi + \beta W.$$

Par suite on a :

$$dC_{\alpha, \beta} = \alpha dM + (1 - \alpha - \beta) d\Pi + \beta dW.$$

Et

$$\begin{aligned}
C_{\alpha, \beta}(dC_{\alpha, \beta}) &= \alpha^2 M dM + \alpha(1 - \alpha - \beta) M d\Pi + \alpha\beta M dW + \alpha(1 - \alpha - \beta) \Pi dM \\
&+ (1 - \alpha - \beta)^2 \Pi d\Pi + \beta(1 - \alpha - \beta) \Pi dW + \alpha\beta W dM \\
&+ \beta(1 - \alpha - \beta) W d\Pi + \beta^2 W dW.
\end{aligned}$$

Alors :

$$\begin{aligned}
\int_{I^2} C_{\alpha, \beta}(u, v) dC_{\alpha, \beta}(u, v) &= \alpha^2 Q(M, M) + \alpha(1 - \alpha - \beta) Q(M, \Pi) + \alpha\beta Q(M, W) \\
&+ \alpha(1 - \alpha - \beta) Q(\Pi, M) + (1 - \alpha - \beta)^2 Q(\Pi, \Pi) \\
&+ \beta(1 - \alpha - \beta) Q(\Pi, W) + \alpha\beta Q(W, M) \\
&+ \beta(1 - \alpha - \beta) Q(W, \Pi) + \beta^2 Q(W, W) \\
&= \alpha^2 Q(M, M) + (1 - \alpha - \beta)^2 Q(\Pi, \Pi) + \beta^2 Q(W, W) \\
&+ 2[\alpha(1 - \alpha - \beta) Q(M, \Pi) + \alpha\beta Q(M, W)].
\end{aligned}$$

Donc :

$$\begin{aligned} \int_{I^2} C_{\alpha,\beta}(u,v) dC_{\alpha,\beta}(u,v) &= \alpha^2 - \beta^2 + 2 \left[\frac{1}{3}\alpha(1-\alpha-\beta) - \frac{1}{3}\beta(1-\alpha-\beta) \right] \\ \int_{I^2} C_{\alpha,\beta}(u,v) dC_{\alpha,\beta}(u,v) &= (\alpha - \beta)(\alpha + \beta) + (\alpha - \beta) \left[\frac{2}{3} - \frac{2}{3}\alpha - \frac{2}{3}\beta \right] \\ \int_{I^2} C_{\alpha,\beta}(u,v) dC_{\alpha,\beta}(u,v) &= \frac{(\alpha - \beta)(\alpha + \beta + 2)}{3} = \tau_{\alpha,\beta}. \end{aligned}$$

3.3.3 Le coefficient de corrélation de Spearman (Le rho de Spearman)

Tout comme le tau de Kendall, le rho de Spearman est une mesure de dépendance basée sur la notion de concordance. Soient (X_1, Y_1) , (X_2, Y_2) et (X_3, Y_3) trois vecteurs aléatoires indépendants de même distribution H , dont les distributions marginales sont F et G et dont la copule associée est C .

Définition 3.3.2 *La version population du rho de Spearman est définie comme étant proportionnelle à la différence de la probabilité de concordance et celle de discordance des couples aléatoires (X_1, Y_1) , (X_2, Y_3) :*

$$\rho = \rho_{X,Y} = \rho_C = 3([P(X_1, Y_1)(X_2, Y_3) > 0] - [P(X_1, Y_1)(X_2, Y_3) < 0]).$$

Propriété

La distribution de (X_1, Y_1) étant H et celle de (X_2, Y_3) étant Π (car les variables X_2 et Y_3 sont indépendantes).

Alors d'après ce qui précède on a le théorème suivant :

Théorème 3.3.2 *Soit (X, Y) un couple de variables continues de copule. Le rho de Spearman pour X et Y est donnée par :*

$$\rho_{X,Y} = \rho_C = 3Q(C, \Pi) = 12 \int_{I^2} uv dC(u,v) - 3 = 12 \int_{I^2} C(u,v) dudv - 3.$$

Le tableau suivant donne des mesures de concordance de quelques familles de copules :

Famille de copule	Le tau de Kendall	Le rho de Spearman
Independence	0	0
Marshall-Olkin	$\frac{\alpha\beta}{\alpha - \alpha\beta + \beta}$	$\frac{3\alpha\beta}{2\alpha - \alpha\beta + 2\beta}$
Gumbel	$1 - \alpha^{-1}$	pas de forme fermée
Clayton	$\frac{\alpha}{\alpha + 2}$	expression compliquée
Frank	$1 + 4\frac{D_1(\alpha) - 1}{\alpha}$	$1 - 12\frac{D_2(-\alpha) - D_1(-\alpha)}{\alpha}$

Tableau 2 : Mesures de concordance de quelques familles de copules

Notation 1 $D_k(x) = \frac{k}{x^k} \int_0^x \frac{t^k}{e^t - 1} dt$ est dite la fonction de "Debye".

3.3.4 La relation entre le tau de Kendall et le rho de Spearman

Théorème 3.3.3 Soit X et Y des variables aléatoires continues, τ et ρ désignent le tau de Kendall et le rho de Spearman, respectivement. Alors :

$$-1 \leq 3\tau - 2\rho \leq 1.$$

Et

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

Preuve. Voir Nilson page 176 ; 177. ■

3.4 Le coefficient de corrélation de Pearson

Le coefficient de corrélation de Pearson, également appelé le coefficient de corrélation linéaire, est la première mesure de dépendance à avoir été utilisée. La covariance est donnée par :

$$\text{COV}(X, Y) = E(XY) - E(X)E(Y).$$

Elle mesure le surcroît de variabilité (ou la diminution) de la somme de deux variables aléatoires par rapport à la somme de leur variance. Elle nous renseigne donc par son signe sur le sens de la covariation entre X et Y .

Le coefficient de corrélation linéaire en est une version normée, et a ainsi l'avantage d'être sans dimension. Le coefficient de corrélation de Pearson est défini par :

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{V(X)V(Y)}}.$$

Où

$$\text{Cov}(X, Y) = E[XY] - E[X]E[Y].$$

Plus ce coefficient est grand en valeur absolue, plus la dépendance entre les variables est forte. Par contre, ce n'est pas une mesure de concordance. Il est important de rappeler que la dépendance et la corrélation sont des notions différentes.

En effet, on a X et Y indépendantes ; alors X et Y non corrélées ou $\rho(X, Y) = 0$, mais la réciproque est fautive sauf dans le cas où les variables sont gaussiennes car la dépendance est alors entièrement caractérisée par le coefficient de corrélation. Le contre-exemple le plus connu dans la littérature est le suivant : soit $X \rightsquigarrow N(0, 1)$ et $Y = X^2$, alors

$$\text{Cov}(X, Y) = E(X^3) = 0.$$

3.5 Dépendance de queue

Définissons enfin un concept essentiel : la dépendance de queue pour une copule bivariée. Cette dernière mesure la probabilité de réalisations extrêmes simultanées. L'indicateur de dépendance de queue d'une copule se déduit des probabilités conditionnelles suivantes :

$$\begin{aligned}
P[U_1 \leq u_1 / U_2 \leq u_2] &= \frac{P[U_1 \leq u_1, U_2 \leq u_2]}{P[U_2 \leq u_2]} \\
&= \frac{C(u_1, u_2)}{u_2} \\
&= \frac{P[U_1 > u_1, U_2 > u_2]}{P[U_2 > u_2]} \\
&= \frac{P[U_1 \leq 1, U_2 > u_2] - P[U_1 \leq u_1, U_2 > u_2]}{P[U_2 \leq 1] - P[U_2 \leq u_2]} \\
&= \frac{P[U_1 \leq 1, U_2 \leq 1] - P[U_1 \leq 1, U_2 \leq u_2] - P[U_1 \leq u_1, U_2 > u_2]}{1 - u_2 - P[U_1 \leq u_1, U_2 \leq 1] + P[U_1 \leq u_1, U_2 \leq u_2]} \\
&= \frac{1 - u_1 - u_2 + C(u_1, u_2)}{1 - u_2}
\end{aligned}$$

Nous pouvons alors définir les dépendances de queue à gauche et à droite.

a) Une copule C a une dépendance de queue à gauche (lower tail dependence)

si :

$$\lambda_L = \lim_{u \rightarrow 0^+} \frac{C(u, u)}{u},$$

existe et $\lambda_L \in (0, 1]$. Si $\lambda_L = 0$ alors elle n'a pas de dépendance de queue à gauche.

b) Une copule C a une dépendance de queue à droite (upper tail dependence)

si :

$$\lambda_U = \lim_{u \rightarrow 1^-} \frac{1 - 2u + C(u, u)}{1 - u},$$

existe et $\lambda_U \in (0, 1]$. Si $\lambda_U = 0$ alors elle n'a pas de dépendance de queue à droite.

Deuxième partie

Copule de Marshall-Olkin

Chapitre 4

Généralités sur la distribution de Marshall-Olkin

4.1 La distribution de Marshall-Olkin Burr

4.1.1 Introduction

Par diverses méthodes, de nouveaux paramètres peuvent être présentés pour développer les familles de distributions, pour plus de flexibilité ou de construire le modèle covariable. L'introduction d'un paramètre d'échelle mène pour accélérer le modèle de vie et les pouvoirs de prise d'une fonction de survie présente un paramètre qui donne le modèle des risques proportionnels. Une méthode d'ajouter un paramètre à une famille de distribution, a été proposée par Marshall et Olkin (1997). Nous proposons ici une méthode pour présenter deux paramètres dans une famille de distribution. Ceci peut être vu comme une généralisation à la méthode suggérée par Marshall et Olkin (1997). On commençant par une fonction de survie \bar{F} , et la fonction de densité f , la famille de deux-paramètre de la fonction de survie est proposée comme suit :

$$\bar{G}(x) = \left[\frac{\alpha \bar{F}(x)}{1 - \alpha \bar{F}(x)} \right]^\gamma ; \quad -\infty < x < \infty, 0 < \alpha < \infty, 0 < \gamma < \infty. \quad (4.1)$$

Quand $\alpha = 1$ nous obtenons $\bar{G}_{1,\gamma}(x) = [\bar{F}(x)]^\gamma$,
et en particulier quand $\alpha = \gamma = 1$, nous obtenons : $\bar{G}_{1,1}(x) = \bar{F}(x)$.

$$g_{\alpha,\gamma}(x) = \gamma \left[\frac{\alpha \bar{F}(x)}{1 - \alpha \bar{F}(x)} \right]^{\gamma-1} \frac{\alpha f(x)}{[1 - \alpha \bar{F}(x)]^2}, \quad (4.2)$$

où \bar{G} et g sont : la fonction de survie, et la fonction de densité de la nouvelle famille de la distribution.

La fonction de taux de risque est :

$$r_{\alpha,\gamma}(x) = \frac{g_{\alpha,\gamma}(x)}{\bar{G}_{\alpha,\gamma}(x)} = \frac{\gamma f(x)}{\bar{F}(x) [1 - \alpha \bar{F}(x)]}. \quad (4.3)$$

Burr (1942) a présenté 12 familles de distributions qui pourraient prendre une variété de formes et docile pour travailler avec ces derniers. La distribution de Burr (type XII) a reçu la plupart d'attention dans la littérature. Elle a été appliquée dans une variété de domaines. En effet, la distribution de Burr (type XII) souvent s'appelle simplement la distribution de Burr en littérature. Puisque c'est des applications de distribution particulièrement flexibles se sont avérés être beaucoup plus large. Les applications peuvent être trouvées dans les domaines de la qualité, le contrôle, la durée de la modélisation du temps d'échec, modélisation de la distribution de revenu, biologique et le test d'hypothèse.

Rodriguez (1977) a consacré une particulière attention au type XII dont la fonction de la distribution est donnée par :

$$\bar{F}(x) = \left(\frac{1}{1 + x^\beta} \right)^\gamma ; 0 < x < \infty, 0 < \beta < \infty, 0 < \gamma < \infty. \quad (4.4)$$

La fonction de densité est unimodale avec le mode à $x = \left(\frac{\beta-1}{\gamma\beta+1} \right)^{\frac{1}{\beta}}$ si $\beta > 1$ et 0 si $\beta < 1$

On peut voir qu'un certain nombre de données que nous rencontrons par hasard en pratique peuvent exposer quelques mouvements périodiques, et avoir plusieurs maximums locaux. Pour modéliser de telles situations, la distribution de Pareto semble être insuffisante, et la recherche se termina par la distribution que l'on appelle semi-Pareto, qui reçoit Pareto et en attendant l'exhibition de mouvements périodiques.

Définition 4.1.1 *On dit que la variable aléatoire X , de support positif, suit la distribution semi-Pareto notée par $SP(\beta)$, si sa fonction de survie est de la*

forme :

$$\bar{F}(x) = \frac{1}{1 + \Psi(x)}, \quad (4.5)$$

où $\Psi(x)$ satisfait l'équation fonctionnelle

$$p\Psi(x) = \left(p^{\frac{1}{\beta}}x\right); \beta > 0, 0 < p < 1. \quad (4.6)$$

La solution de l'équation fonctionnelle est :

$$\Psi(x) = x^\beta h(x),$$

où $h(x)$ est périodique au $\ln x$ avec le périodicité $\left(-\frac{2\pi\beta}{\ln p}\right)$. (Pour la preuve voir Kagan et al., 1973, p. 163).

Par exemple, si $h(x) = e^{\theta \cos(\beta \ln x)}$ il satisfait l'équation fonctionnelle avec $p = e^{-2\pi}$ et $\Psi(x)$ monotone croissante avec $0 < \theta < 1$.

Alice et Jose (2003) ont utilisés la méthode introduite par Marshall et Olkin (1997) pour définir la distribution de Marshall-Olkin semi-Pareto. La distribution de Burr (type XII), ce qui donne un large éventail de valeurs de skewness et kurtosis, peut être utilisé pour s'adapter à presque n'importe quel ensemble de données unimodales. Parfois, nous rencontrons des données qui présentent la nature périodique d'objet exposé, et en même temps ne peut être modélisée par la distribution semi-Pareto. Dans de telles situations, il devient nécessaire d'introduire une classe plus générale de distribution, qui inclut la distribution semi-Pareto.

Définition 4.1.2 *On dit qu'une variable aléatoire X , de support positif, suit la distribution de semi-Burr notée par $SB(\beta, \gamma)$, si sa fonction de survie est de la forme :*

$$\bar{F}(x) = \left(\frac{1}{1 + \Psi(x)}\right)^\gamma; \gamma > 0, \quad (4.7)$$

où $\Psi(x)$ satisfait l'équation fonctionnelle (4.6).

4.1.2 La distribution de Marshall-Olkin semi-Burr

En substituant l'équation (4.5) dans l'équation (4.1), nous obtenons la distribution de Marshall-Olkin semi-Burr ($MOSB(\alpha, \beta, \gamma)$) dont la fonction de survie est donnée par :

$$\begin{aligned}\bar{G}(x) &= \left(\frac{\alpha}{\alpha + \Psi(x)} \right)^\gamma; \quad \alpha, \gamma, \beta > 0 \\ &= \left(\frac{1}{1 + \frac{1}{\alpha}\Psi(x)} \right)^\gamma,\end{aligned}\tag{4.8}$$

où $\Psi(x)$ satisfait l'équation fonctionnelle. (4.6).

Notez qu'il s'avère être une distribution semi-Burr de trois paramètres définie en (4.8).

Maintenant la fonction de densité de probabilité de $MOSB(\alpha, \beta, \gamma)$ est donnée par :

$$g(x) = \frac{\gamma}{\alpha} \left(\frac{\alpha}{\alpha + \Psi(x)} \right)^{\gamma+1} \Psi'(x); \quad \alpha, \beta, \gamma > 0.\tag{4.9}$$

Ici, nous étudions le cas particulier où $\Psi(x) = x^\beta h(x)$ où $h(x) = e^{\theta \cos(\beta \ln x)}$.

$$g(x) = \left(\frac{1}{1 + \frac{1}{\alpha}x^\beta e^{\theta \cos(\beta \ln x)}} \right) \frac{\gamma \beta x^{\beta-1} e^{\theta \cos(\beta \ln x)} (1 - \theta \sin(\beta \ln x))}{1 + \frac{1}{\alpha}x^\beta e^{\theta \cos(\beta \ln x)}}.\tag{4.10}$$

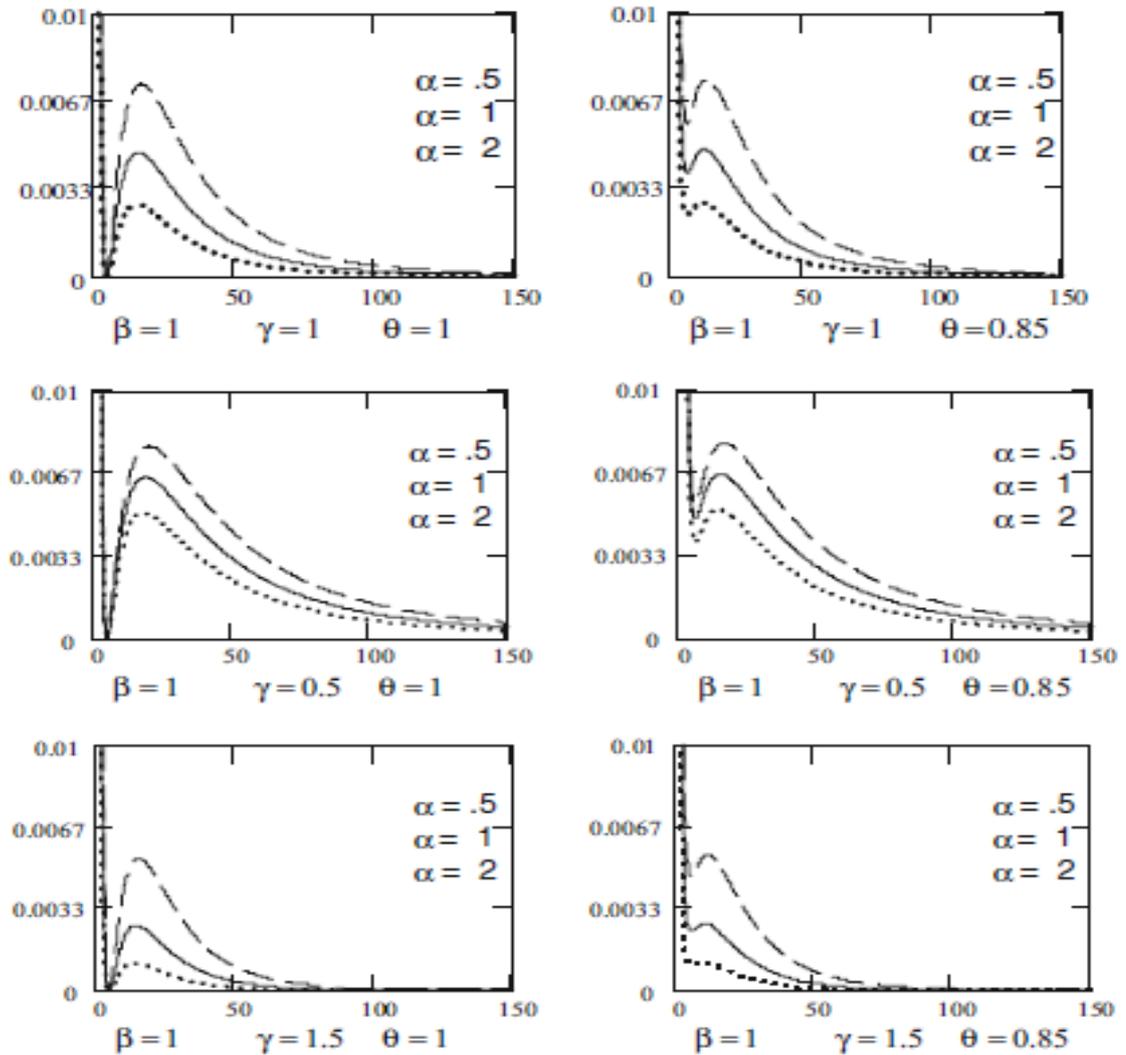


Figure 11 : La distribution de $MOSB(\alpha, \beta, \gamma)$
pour différentes valeurs de α, β, γ et θ .

4.1.3 La distribution de Marshall-Olkin Burr

Comme un cas spécial de la distribution de semi-Burr de Marshall-Olkin, nous présentons la distribution de Burr de Marshall-Olkin.

Définition 4.1.3 On dit qu'une variable aléatoire X , de support positif, suit la

distribution de Pareto notée par $P(\beta)$, si sa fonction de survie est de la forme :

$$\bar{F}(x) = \frac{1}{1+x^\beta}, \quad \beta > 0. \quad (4.11)$$

Lorsque $\bar{F}(x) = \frac{1}{1+x^\beta}$, l'équation (4.1) devient :

$$\begin{aligned} \bar{G}(x) &= \left(\frac{\alpha}{\alpha+x^\beta} \right)^\gamma, \quad 0 < \alpha < \infty, 0 < \beta < \infty, 0 < \gamma < \infty \\ &= \left(\frac{1}{1+\frac{1}{\alpha}x^\beta} \right)^\gamma. \end{aligned} \quad (4.12)$$

La distribution avec la fonction de survie (4.12) s'appelle la distribution de Marshall-Olkin Burr notée par $MOB(\alpha, \beta, \gamma)$.

La fonction correspondante de densité est :

$$g(x) = \frac{\gamma\beta}{\alpha} \left(\frac{\alpha}{\alpha+x^\beta} \right)^{\gamma+1} x^{\beta-1}. \quad (4.13)$$

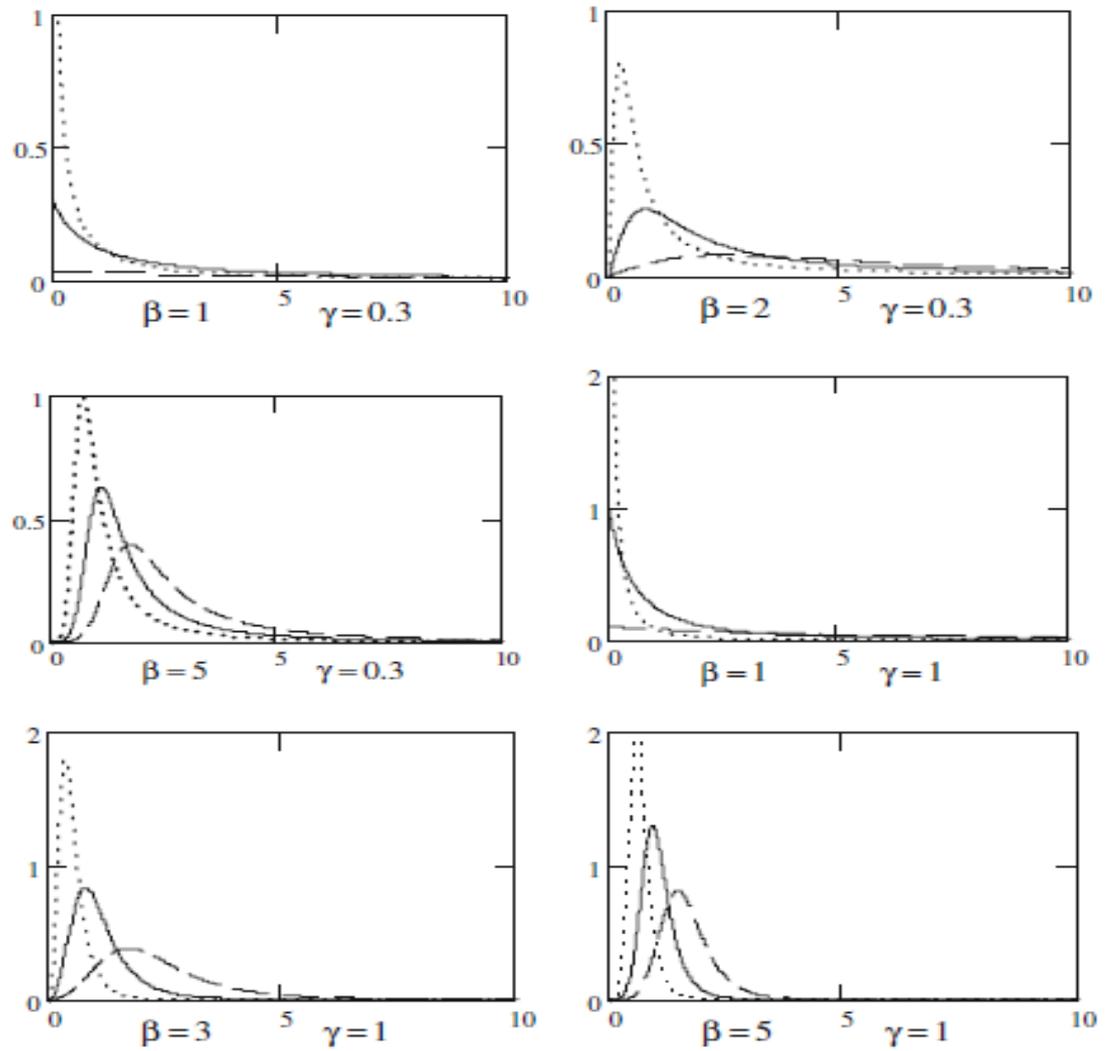


Figure 12 : La densité de $MOB(\alpha, \beta, \gamma)$ pour différentes valeurs de α, β et γ .

Chapitre 5

Caractérisation de la copule de Marshall-Olkin

5.1 Copule de Marshall-Olkin

En considérant la distribution $MOB(\alpha, \beta, \gamma)$ avec la fonction de survie :

$$\bar{F}(x) = \left(\frac{1}{1 + \frac{1}{\alpha}x^\beta} \right)^{\gamma+1}, \text{ et la distribution } MOP(\alpha, \beta) \text{ avec la fonction de}$$

survie :

$$\bar{G}(y) = \frac{1}{1 + \frac{1}{\alpha}y^\beta}.$$

Lemme 5.1.1 Soient X et Y deux variables aléatoires telle que : $X \rightsquigarrow MOB(\alpha, \beta, \gamma)$ et $Y \rightsquigarrow MOP(\alpha, \beta)$, alors

d'après (Jayakumar et Thomas, 2002.) la fonction de survie conjointe de (X, Y) est de la forme :

$$\bar{H}_{XY}(x, y) = \frac{1}{\left(1 + \frac{1}{\alpha}y^\beta\right)} \int_0^1 \frac{1}{\left(1 + \frac{1}{\alpha} \max(x^\beta, v^\beta y^\beta)\right)^{\gamma+1}} \beta \gamma v^{\beta\gamma-1} dv.$$

Preuve. (voir Jayakumar et Gupta, 2006, page 205).

Pour plus de clarification voir applications (dernier chapitre). ■

On va trouver maintenant la copule de Marshall-Olkin :

Soit X une variable aléatoire qui suit la distribution $MOB(\alpha, \beta, \gamma)$, alors :

$$\bar{F}_X(x) = \left(\frac{1}{1 + \frac{1}{\alpha}x^\beta} \right)^{\gamma+1}; \quad 0 < \alpha < \infty, \quad 0 < \beta < \infty, \quad 0 < \gamma < \infty, \quad 0 < x < \infty.$$

Et soit Y une variable aléatoire qui suit la distribution $MOP(\alpha, \beta)$, alors :

$$\bar{G}_Y(y) = \frac{1}{1 + \frac{1}{\alpha}y^\beta}; \quad 0 < \alpha < \infty, \quad 0 < \beta < \infty, \quad 0 < x < \infty.$$

D'après le lemme (5.1.1) La fonction de survie conjointe de (X, Y) est de la forme :

$$\bar{H}(x, y) = \frac{1}{\left(1 + \frac{1}{\alpha}y^\beta\right)} \int_0^1 \frac{1}{\left(1 + \frac{1}{\alpha} \max(x^\beta, v^\beta y^\beta)\right)^{\gamma+1}} \beta \gamma v^{\beta\gamma-1} dv,$$

alors :

$$\bar{H}(x, y) = \begin{cases} \frac{1}{\left(1 + \frac{1}{\alpha}y^\beta\right)} \int_0^1 \frac{1}{\left(1 + \frac{1}{\alpha}x^\beta\right)^{\gamma+1}} \beta \gamma v^{\beta\gamma-1} dv & \text{si } x^\beta > v^\beta y^\beta \\ \frac{1}{\left(1 + \frac{1}{\alpha}y^\beta\right)} \int_0^1 \frac{1}{\left(1 + \frac{1}{\alpha}v^\beta y^\beta\right)^{\gamma+1}} \beta \gamma v^{\beta\gamma-1} dv & \text{si } x^\beta < v^\beta y^\beta \\ \frac{1}{\left(1 + \frac{1}{\alpha}y^\beta\right)} \times \frac{1}{\left(1 + \frac{1}{\alpha}x^\beta\right)^{\gamma+1}} & \text{si } x^\beta > v^\beta y^\beta \\ \frac{1}{\left(1 + \frac{1}{\alpha}y^\beta\right)^{\gamma+1}} & \text{si } x^\beta < v^\beta y^\beta \end{cases}$$

Soit :

$$u = F(x) = 1 - \bar{F}(x) \text{ et } v = G(y) \implies v = 1 - \bar{G}(y),$$

alors

$$u = 1 - \left(\frac{1}{1 + \frac{1}{\alpha}x^\beta}\right)^{\gamma+1} \text{ et } v = 1 - \frac{1}{1 + \frac{1}{\alpha}y^\beta}.$$

Donc

$$x = \left[\frac{\alpha}{(1-u)^{\frac{1}{\gamma+1}}} - \alpha\right]^{\frac{1}{\beta}} \text{ et } y = \left[\frac{\alpha}{1-v} - \alpha\right]^{\frac{1}{\beta}}.$$

Et on a :

$$\begin{aligned} C(u, v) &= H(F^{-1}(u), G^{-1}(v)) \\ &= \bar{H}(F^{-1}(u), G^{-1}(v)) - 1 + u + v. \end{aligned}$$

Lorsque $x = \left[\frac{\alpha}{(1-u)^{\frac{1}{\gamma+1}}} - \alpha\right]^{\frac{1}{\beta}}$ et $y = \left[\frac{\alpha}{1-v} - \alpha\right]^{\frac{1}{\beta}}$, (5.6) devient :

$$\begin{aligned}
C(u, v) &= \left\{ \begin{array}{l} \frac{1}{\left(1 + \frac{1}{\alpha} \left[\frac{\alpha}{1-v} - \alpha\right]^{\frac{1}{\beta}\beta}\right)} \times \frac{1}{\left(1 + \frac{1}{\alpha} \left[\frac{\alpha}{(1-u)^{\frac{1}{\gamma+1}}} - \alpha\right]^{\frac{1}{\beta}\beta}\right)^{\gamma+1}} - 1 + u + v \\ \text{si } \left[\frac{\alpha}{(1-u)^{\frac{1}{\gamma+1}}} - \alpha\right]^{\frac{1}{\beta}\beta} > v^\beta \left[\frac{\alpha}{1-v} - \alpha\right]^{\frac{1}{\beta}\beta}, \\ \text{et} \\ \frac{1}{\left(1 + \frac{1}{\alpha} \left[\frac{\alpha}{1-v} - \alpha\right]^{\frac{1}{\beta}\beta}\right)^{\gamma+1}} - 1 + u + v \\ \text{si } \left[\frac{\alpha}{(1-u)^{\frac{1}{\gamma+1}}} - \alpha\right]^{\frac{1}{\beta}\beta} < v^\beta \left[\frac{\alpha}{1-v} - \alpha\right]^{\frac{1}{\beta}\beta} \end{array} \right. , \\
&= \left\{ \begin{array}{l} \frac{1}{\left(\frac{1}{1-v}\right)} \times \frac{1}{\left(\frac{1}{(1-u)^{\frac{1}{\gamma+1}}}\right)^{\gamma+1}} - 1 + u + v \\ \text{si } \left[\frac{1}{(1-u)^{\frac{1}{\gamma+1}}} - 1\right] > v^\beta \left[\frac{1}{1-v} - 1\right]. \\ \text{et} \\ \frac{1}{\left(\frac{1}{1-v}\right)^{\gamma+1}} - 1 + u + v \\ \text{si } \left[\frac{1}{(1-u)^{\frac{1}{\gamma+1}}} - 1\right] < v^\beta \left[\frac{1}{1-v} - 1\right]. \end{array} \right. \\
&= \left\{ \begin{array}{l} (1-v)(1-u) - 1 + u + v \\ (1-v)^{\gamma+1} - 1 + u + v \end{array} \right. \text{ si } \left\{ \begin{array}{l} \left[\frac{1}{(1-u)^{\frac{1}{\gamma+1}}} - 1\right] > v^\beta \left[\frac{1}{1-v} - 1\right] \\ \left[\frac{1}{(1-u)^{\frac{1}{\gamma+1}}} - 1\right] < v^\beta \left[\frac{1}{1-v} - 1\right] \end{array} \right.
\end{aligned}$$

Donc

$$C(u, v) = \begin{cases} uv & \text{si } u > 1 - \left(\frac{1-v}{v^{\beta+1}+1-v}\right)^{\gamma+1} \\ (1-v)^{\gamma+1} - 1 + u + v & \text{si } u < 1 - \left(\frac{1-v}{v^{\beta+1}+1-v}\right)^{\gamma+1} \end{cases} . \quad (5.1)$$

5.1.1 La densité de la copule de Marshall-Olkin

On a :

$$C(u, v) = \begin{cases} uv & \text{si } u > 1 - \left(\frac{1-v}{v^{\beta+1} + 1 - v} \right)^{\gamma+1} \\ (1-v)^{\gamma+1} - 1 + u + v & \text{si } u < 1 - \left(\frac{1-v}{v^{\beta+1} + 1 - v} \right)^{\gamma+1} \end{cases}, \quad (5.2)$$

et :

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

alors :

$$c(u, v) = \begin{cases} 1 & \text{si } u > 1 - \left(\frac{1-v}{v^{\beta+1} + 1 - v} \right)^{\gamma+1} \\ 0 & \text{si } u < 1 - \left(\frac{1-v}{v^{\beta+1} + 1 - v} \right)^{\gamma+1} \end{cases} \quad (5.3)$$

5.1.2 Les mesures de concordances associées à cette famille de copules

Le tau de Kendall

$$\tau = 4 \int_{I^2} C(u, v) dC(u, v) - 1.$$

$$\text{On a } C(u, v) = \begin{cases} uv & \text{si } u > 1 - \left(\frac{1-v}{v^{\beta+1} + 1 - v} \right)^{\gamma+1} \\ (1-v)^{\gamma+1} - 1 + u + v & \text{si } u < 1 - \left(\frac{1-v}{v^{\beta+1} + 1 - v} \right)^{\gamma+1} \end{cases},$$

$$\text{alors } dC(u, v) = \begin{cases} 1 dudv & \text{si } u > 1 - \left(\frac{1-v}{v^{\beta+1} + 1 - v} \right)^{\gamma+1} \\ 0 & \text{si } u < 1 - \left(\frac{1-v}{v^{\beta+1} + 1 - v} \right)^{\gamma+1} \end{cases}$$

$$\text{Donc } C(u, v) dC(u, v) = \begin{cases} uv dudv & \text{si } u > 1 - \left(\frac{1-v}{v^{\beta+1} + 1 - v} \right)^{\gamma+1} \\ 0 & \text{si } u < 1 - \left(\frac{1-v}{v^{\beta+1} + 1 - v} \right)^{\gamma+1} \end{cases}$$

alors :

$$\begin{aligned}
\tau &= 4 \int_0^1 \int_{\omega}^1 uv du dv - 1 \\
&= 4 \int_0^1 \left(\int_{\omega}^1 uv du \right) dv - 1; \omega = 1 - \left(\frac{1-v}{v^{\beta+1} + 1 - v} \right)^{\gamma+1} \\
&= 4 \int_0^1 \left[\frac{u^2 v}{2} \right]_{\omega}^1 dv - 1 \\
&= 4 \int_0^1 \frac{v}{2} dv - 4 \int_0^1 \left[\frac{v}{2} \omega^2 \right] dv - 1 \\
&= 4 \int_0^1 \frac{v}{2} dv - 4 \int_0^1 \frac{v}{2} \left[1 - \left(\frac{1-v}{v^{\beta+1} + 1 - v} \right)^{\gamma+1} \right]^2 dv - 1 \\
&= 4 \int_0^1 \frac{v}{2} dv - 4 \int_0^1 \frac{v}{2} dv - 4 \int_0^1 \frac{v}{2} \left(\frac{1-v}{v^{\beta+1} + 1 - v} \right)^{2\gamma+2} dv + 4 \int_0^1 v \left(\frac{1-v}{v^{\beta+1} + 1 - v} \right)^{\gamma+1} dv - 1 \\
&= 2 \int_0^1 \left[2 - \left(\frac{1-v}{v^{\beta+1} + 1 - v} \right)^{\gamma+1} \right] v \left(\frac{1-v}{v^{\beta+1} + 1 - v} \right)^{\gamma+1} dv - 1.
\end{aligned}$$

Le rho de Spearman

$$\rho_{X,Y} = \rho_C = 3Q(C, \Pi) = 12 \iint_{I^2} uv dC(u, v) - 3 = 12 \iint_{I^2} C(u, v) dudv - 3.$$

$$\text{On a } dC(u, v) = \begin{cases} 1 dudv & \text{si } u > 1 - \left(\frac{1-v}{v^{\beta+1} + 1 - v} \right)^{\gamma+1} \\ 0 & \text{si } u < 1 - \left(\frac{1-v}{v^{\beta+1} + 1 - v} \right)^{\gamma+1} \end{cases}.$$

Alors :

$$\begin{aligned}
\rho_{X,Y} &= 12 \int_{I^2} uv dC(u, v) - 3 \\
&= 12 \int_0^1 \left[\int_{\omega}^1 uv du \right] dv - 3; \omega = 1 - \left(\frac{1-v}{v^{\beta+1} + 1 - v} \right)^{\gamma+1} \\
&= 12 \int_0^1 \left[v \frac{u^2}{2} \right]_{\omega}^1 dv - 3 \\
&= 12 \int_0^1 \left[\frac{v}{2} - \frac{v}{2} \omega^2 \right] dv - 3 \\
&= 12 \int_0^1 \left[\frac{v}{2} - \frac{v}{2} \left[1 - \left(\frac{1-v}{v^{\beta+1} + 1 - v} \right)^{\gamma+1} \right]^2 \right] dv - 3 \\
&= 6 \int_0^1 v dv - 6 \int_0^1 v \left[1 - \left(\frac{1-v}{v^{\beta+1} + 1 - v} \right)^{\gamma+1} \right]^2 dv - 3 \\
&= 6 \int_0^1 v dv - 6 \int_0^1 v dv + 6 \int_0^1 v \left[2 - \left(\frac{1-v}{v^{\beta+1} + 1 - v} \right)^{\gamma+1} \right] \left(\frac{1-v}{v^{\beta+1} + 1 - v} \right)^{\gamma+1} dv - 3 \\
&= 6 \int_0^1 \left[2 - \left(\frac{1-v}{v^{\beta+1} + 1 - v} \right)^{\gamma+1} \right] v \left(\frac{1-v}{v^{\beta+1} + 1 - v} \right)^{\gamma+1} dv - 3.
\end{aligned}$$

Chapitre 6

Application

Nous avons présenté dans ce chapitre l'application de la distribution de Marshall-Olkin Burr dans la construction de la modélisation de séries chronologiques, la distribution de Marshall-Olkin semi- Burr est utilisée pour modéliser le taux de change quotidien du yuan chinois avec le dollar américain. La distribution de Burr (type XII) est largement utilisée dans des domaines comme les affaires, le génie, la fiabilité, l'hydrologie et la minéralogie comme le modèle d'échec et ses propriétés a été étudiée par Burr et Cislak (1968) et Tadikamalla (1980).

6.1 Applications de la distribution de Marshall-Olkin Burr dans la modélisation de série chronologique

L'analyse de la série chronologique dans l'installation classique est fondée sur l'hypothèse qu'une série observée soit une réalisation d'une séquence gaussienne. Cependant, il y a beaucoup de situations où les données naturellement arrivantes montrent une tendance de suivre des distributions asymétriques et à queues lourdes, qui ne peuvent pas être modélisées par des distributions gaussiennes. Les techniques habituelles de transférer des données pour utiliser le modèle gaussien, échoue aussi dans certaines situations (voir Lawrance, 1991). Ainsi un certain nombre de modèles non gaussiens de séries chronologiques ont été présentés par plusieurs chercheurs au cours dernières années (voir par exemple, Balakrishnan et Jayakumar, 1997; Jayakumar, 1995 et 1997; Jayakumar et Pillai,

1993; Jayakumar et Thomas, 2002).

L'étude sur le processus autorégressif de minification a commencé par le travail innovateur de Tavares (1980). Dans son travail, les observations sont générées par l'équation :

$$X_n = k \min(X_{n-1}, \varepsilon_n); \quad n \geq 0, \quad (6.1)$$

où $k > 1$ est une constante et $\{\varepsilon_n\}$ est un processus novateur des variables aléatoires i.i.d. choisies pour assurer que $\{X_n\}$ est un processus de Markov stationnaire avec la fonction de distribution marginale $F_{X_0}(x)$. A cause de la structure de (6.1), le processus $\{X_n\}$ est appelé processus de minification. Même si la distribution de Burr (type XII) peut être utilisée pour s'adapter à presque n'importe quel ensemble de données unimodales, pas suffisamment d'études ont été faites sur des modèles de séries chronologiques avec la distribution marginales de Burr (type XII).

Jayakumar et Thomas (2002) ont présenté un processus autorégressif du premier ordre de minification avec la distribution marginale de Burr (type XII). Ce modèle peut être étendu à définir un processus autorégressif de premier ordre de minification avec la distribution de Marshall-Olkin Burr en tant que marginale.

En considérant la distribution de Marshall-Olkin Burr ($MOB(\alpha, \beta, \gamma)$) avec la fonction de survie :

$$\bar{F}(x) = \left(\frac{1}{1 + \frac{1}{\alpha}x^\beta} \right)^{\gamma+1}; \quad 0 < \alpha < \infty, \quad 0 < \beta < \infty, \quad 0 < \gamma < \infty, \quad 0 < x < \infty,$$

et la distribution de Marshall-Olkin Pareto ($MOP(\alpha, \beta)$) (voir Alice et Jose, 2003) avec la fonction de survie :

$$\bar{F}(x) = \frac{1}{1 + \frac{1}{\alpha}x^\beta}; \quad 0 < \alpha < \infty, \quad 0 < \beta < \infty, \quad 0 < x < \infty.$$

Théorème 6.1.1 Soit le processus $\{X_n\}$ être définie comme :

$$X_n = \min(V_n^{-1}X_{n-1}, \varepsilon_n), \quad n = 1, 2, \dots, \quad (6.2)$$

où $\{V_n\}$ et $\{\varepsilon_n\}$ sont deux suites de variables aléatoires i.i.d. indépendantes telle que $\{V_n\}$ a la fonction de distribution $F_{V_n}(v) = v^{\beta\gamma}$, $\beta, \gamma > 0$ et $0 < v < 1$.

Supposons que le processus $\{X_n\}$ est stationnaire.

Alors :

$$X_n \stackrel{d}{=} MOB(\alpha, \beta, \gamma) \text{ si et seulement si } \varepsilon_n \stackrel{d}{=} MOP(\alpha, \beta).$$

Preuve. On note par $\bar{F}_{X_n}(x)$ et $\bar{F}_{\varepsilon_n}(x)$ la fonction de survie de X_n et ε_n respectivement. le processus (6.2) en termes de fonction de survie est :

$$\bar{F}_{X_n}(x) = \bar{F}_{\varepsilon_n}(x) \int_0^1 \bar{F}_{X_{n-1}}(xv) f_{V_n}(v) dv,$$

où $f_{V_n}(v) = \beta\gamma v^{\beta\gamma-1}$ est la fonction de densité de probabilité de V_n . C'est,

$$\bar{F}_{X_n}(x) = \bar{F}_{\varepsilon_n}(x) \int_0^1 \bar{F}_{X_{n-1}}(xv) \beta\gamma v^{\beta\gamma-1} dv.$$

Alors

$$\begin{aligned} \bar{F}_{X_n}(x) &= \frac{1}{\left(1 + \frac{1}{\alpha}x^\beta\right)} \int_0^1 \frac{1}{\left(1 + \frac{1}{\alpha}(xv)^\beta\right)^{\gamma+1}} \beta\gamma v^{\beta\gamma-1} dv \\ &= \frac{1}{\left(1 + \frac{1}{\alpha}x^\beta\right)} \int_0^1 \frac{1}{\left(1 + v^\beta \left(\frac{1}{\alpha}x^\beta\right)\right)^{\gamma+1}} \beta\gamma v^{\beta\gamma-1} dv \\ &= \bar{F}_X(x) \\ &= \left(\frac{1}{1 + \frac{1}{\alpha}x^\beta}\right)^{\gamma+1}. \end{aligned}$$

(voir Jayakumar et Gupta, 2006, page 204)

Réciproquement, si $\{X_n\}$ est stationnaire avec la distribution $MOB(\alpha, \beta, \gamma)$ comme marginal, alors $\{\varepsilon_n\}$ est la $MOP(\alpha, \beta)$.

$$\begin{aligned} \frac{1}{\bar{F}_{\varepsilon_n}(x)} &= \int_0^1 \frac{\bar{F}_X(xv) \beta\gamma v^{\beta\gamma-1}}{\bar{F}_X(x)} dv \\ &= \int_0^1 \left(\frac{1 + \frac{1}{\alpha}x^\alpha}{1 + \frac{1}{\alpha}v^\alpha x^\alpha}\right)^{\gamma+1} \beta\gamma v^{\beta\gamma-1} dv \\ &= 1 + \frac{1}{\alpha}x^\beta \\ \bar{F}_{\varepsilon_n}(x) &= \frac{1}{1 + \frac{1}{\alpha}x^\beta}. \end{aligned}$$

Par conséquent $\varepsilon_n \stackrel{d}{=} MOP(\alpha, \beta)$. Ceci accomplit la preuve. ■

Sur la base de ce résultat nous définissons le processus autorégressif du premier ordre de Burr comme suit :

Soit $X_0 = MOB(\alpha, \beta, \gamma)$ et $X_n = \min(V_n^{-1}X_{n-1}, \varepsilon_n)$ pour $n = 1, 2, \dots$, où $\{V_n\}$ est une suite de variables aléatoires i.i.d. de fonction de puissance avec une fonction de distribution $F_{V_n}(v) = v^{\beta\gamma}$, $\beta, \gamma > 0$ et $0 < v < 1$ et $\{\varepsilon_n\}$ est une suite de variables aléatoires i.i.d. de $MOP(\alpha, \beta)$ indépendantes de $\{V_n\}$. Le processus $\{X_n\}$ est stationnaire avec la distribution marginale de $MOB(\alpha, \beta, \gamma)$.

Lemme 6.1.1 *La fonction de survie conjointe de (X_n, X_{n+1}) est de la forme :*

$$\bar{F}_{X_n X_{n+1}}(x, y) = \frac{1}{\left(1 + \frac{1}{\alpha} y^\beta\right)} \int_0^1 \frac{1}{\left(1 + \frac{1}{\alpha} \max(x^\beta, v^\beta y^\beta)\right)^{\gamma+1}} \beta \gamma v^{\beta\gamma-1} dv. \quad (6.3)$$

Preuve. En effet :

$$\begin{aligned} \bar{F}_{X_n X_{n+1}}(x, y) &= P(X_n > x, Y_n > y) \\ &= P(\varepsilon_{n+1} > y) \int_0^1 P[X_n > \max(x, vy)] \beta \gamma v^{\beta\gamma-1} dv \\ &= \frac{1}{\left(1 + \frac{1}{\alpha} y^\beta\right)} \int_0^1 \frac{1}{\left[1 + \frac{1}{\alpha} (\max(x, vy)^\beta)\right]^{\gamma+1}} \beta \gamma v^{\beta\gamma-1} dv \\ &= \frac{1}{\left(1 + \frac{1}{\alpha} y^\beta\right)} \int_0^1 \frac{1}{\left(1 + \frac{1}{\alpha} \max(x^\beta, v^\beta y^\beta)\right)^{\gamma+1}} \beta \gamma v^{\beta\gamma-1} dv. \end{aligned}$$

(voir Jayakumar et Gupta, 2006, page 205). ■

Lemme 6.1.2 *Un calcul élémentaire montre que :*

$$P(X_{n+1} > X_n) = \frac{\gamma + 1}{\gamma + 2}. \quad (6.4)$$

Preuve. En effet :

$$\begin{aligned}
 P(X_{n+1} > X_n) &= \int_0^{\infty} \left(\frac{1}{1 + \Psi(x)} \right) \frac{(1 + \gamma) \Psi'(x)}{(1 + \Psi(x))^{\gamma+2}} dx, \Psi(x) = \frac{1}{\alpha} x^\beta \\
 &= \int_0^{\infty} \frac{(1 + \gamma) \Psi'(x)}{(1 + \Psi(x))^{\gamma+3}} dx \\
 &= \int_0^{\infty} \frac{(1 + \gamma) \frac{\beta}{\alpha} x^{\beta-1}}{\left(1 + \frac{1}{\alpha} x^\beta\right)^{\gamma+3}} dx \\
 &= (\gamma + 1) \int_0^{\infty} \frac{1}{(1 + y)^{\gamma+3}} dy, y = \frac{1}{\alpha} x^\beta \\
 &= (\gamma + 1) \left[\frac{(1 + y)^{-\gamma-2}}{-\gamma - 2} \right]_0^{\infty} \\
 &= (\gamma + 1) \left[0 - \frac{1}{-\gamma - 2} \right] \\
 &= \frac{\gamma + 1}{\gamma + 2}.
 \end{aligned}$$

(voir Jayakumar et Gupta (2006), page 207). ■

Le comportement trajet de l'échantillon du processus de Marshall-Olkin Burr est donné dans La figure : 11.

Le tableau 03 donne l'autocorrélation d'ordre jusqu'à 13 pour différentes valeurs de γ avec $\beta = 2$ et $\alpha = 1$. La première colonne donne les valeurs de γ et la première rangée donne l'ordre de corrélation.

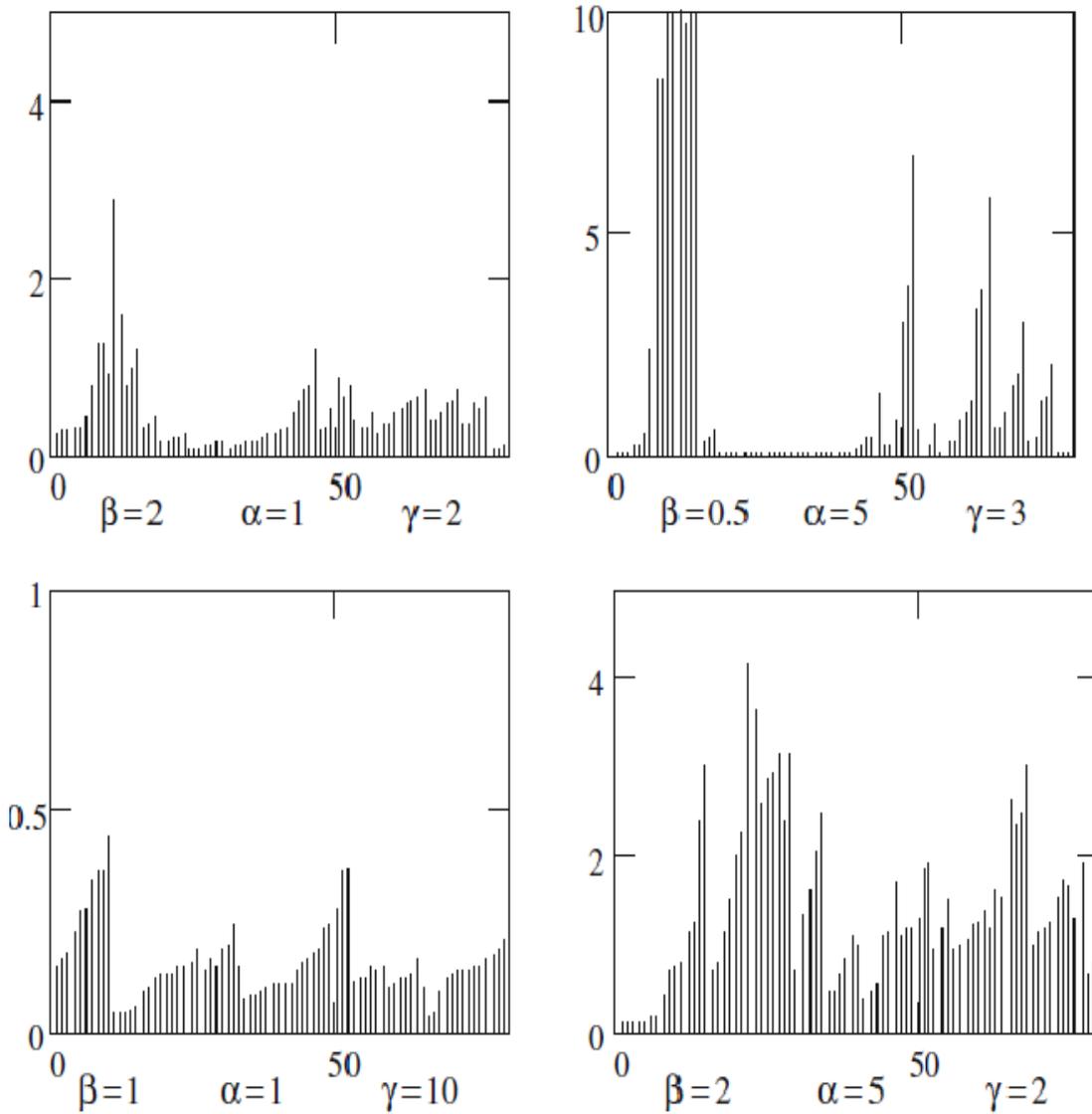


Figure 13 : *Le comportement chemin d'échantillons de processus de $MOB(\alpha, \beta, \gamma)$*

Le tableau suivant donne l'autocorrélation d'ordre jusqu'à 13 pour $\beta = 2$ et $\alpha = 1$

$\gamma \setminus r$	1	2	3	4	5	6	7	8	9	10	11	12	13
0.1	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.01	0.00
0.9	0.36	0.17	0.08	0.03	0.02	0.00	0.01	0.03	0.02	0.01	0.00	0.00	0.00
1.7	0.55	0.34	0.21	0.12	0.08	0.04	0.04	0.03	0.03	0.03	0.02	0.00	0.00
2.5	0.67	0.48	0.33	0.23	0.17	0.12	0.09	0.07	0.06	0.05	0.03	0.01	0.01
3.9	0.79	0.63	0.50	0.40	0.32	0.26	0.21	0.17	0.14	0.12	0.09	0.07	0.05
4.7	0.82	0.68	0.56	0.47	0.39	0.32	0.27	0.23	0.19	0.16	0.13	0.11	0.09
5.5	0.85	0.78	0.61	0.52	0.44	0.38	0.32	0.28	0.24	0.20	0.17	0.14	0.12
6.5	0.87	0.76	0.66	0.58	0.50	0.44	0.38	0.34	0.29	0.26	0.22	0.19	0.16
8.1	0.90	0.81	0.72	0.65	0.58	0.52	0.47	0.42	0.38	0.34	0.30	0.27	0.24
10	0.92	0.84	0.77	0.71	0.65	0.60	0.55	0.50	0.46	0.42	0.39	0.36	0.33
14	0.94	0.89	0.83	0.78	0.74	0.69	0.65	0.62	0.58	0.55	0.51	0.48	0.46
20	0.96	0.92	0.88	0.85	0.81	0.78	0.75	0.72	0.69	0.67	0.64	0.62	0.59
24	0.97	0.94	0.90	0.87	0.85	0.82	0.79	0.76	0.74	0.72	0.69	0.67	0.65
30	0.97	0.95	0.92	0.90	0.87	0.85	0.83	0.80	0.78	0.76	0.74	0.72	0.70
35	0.98	0.95	0.93	0.91	0.89	0.87	0.85	0.83	0.81	0.79	0.77	0.75	0.74
40	0.98	0.96	0.94	0.92	0.90	0.88	0.86	0.84	0.83	0.81	0.79	0.78	0.76
50	0.98	0.97	0.95	0.93	0.92	0.90	0.89	0.87	0.86	0.84	0.83	0.81	0.80
80	0.99	0.98	0.97	0.96	0.95	0.94	0.93	0.92	0.91	0.90	0.89	0.89	0.88
100	0.99	0.98	0.98	0.97	0.96	0.95	0.94	0.94	0.93	0.92	0.91	0.91	0.90
150	0.99	0.99	0.98	0.98	0.97	0.97	0.96	0.95	0.95	0.94	0.94	0.93	0.93

Tableau 03 : Autocorrélation d'ordre jusqu'à 13 pour $\beta = 2$ et $\alpha = 1$

6.2 Application de la distribution de Marshall-Olkin semi-Burr dans la modélisation des taux de change

Des observations quotidiennes du taux de change de la Chine-États-Unis sont pris en compte. Les données se composent de 1024 observations à partir du 2 janvier 1981 au 11 janvier 1985. Les données sont rassemblées sur le site web du

conseil de gouvernement de Système Fédérale de Réserve Etats-Unis. Tracer des séries chronologiques de données est fournie dans la figure 14.

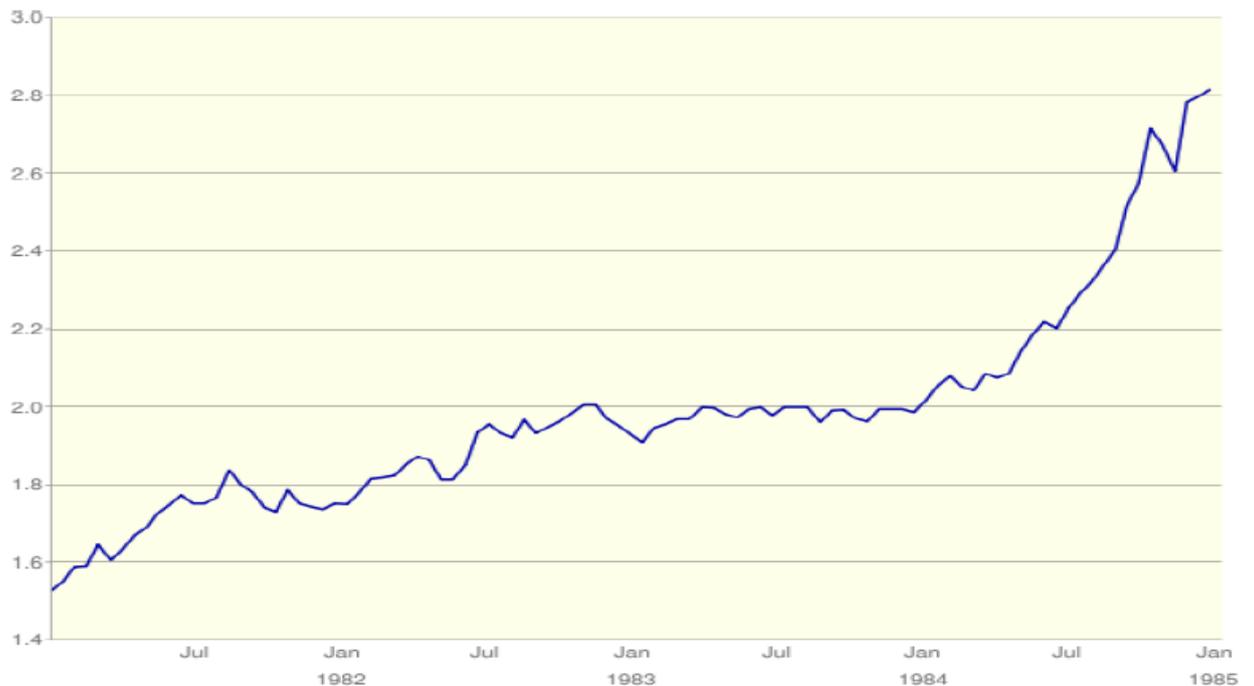


Figure 14 : Sries chronologiques des taux de change quotidien du yuan chinois en dollars amricains.

L'autocorrélation du premier ordre de la série $\{X_n\}$ s'avère 0.9993. L'autocorrélation de premier ordre de différenciation est prise pour rendre la série stationnaire. La série qui en résulte est obtenue par l'équation :

$$Y_n = X_n - r_0 X_{n-1},$$

où r_0 est l'autocorrélation du premier ordre. En soustrayant la moyenne et en divisant par l'écart-type d'obtenir la série normalisée. La série qui en résulte est rendue positive en prenant le module. L'autocorrélation de la série qui en résulte est jugée insignifiante. Chaque observation dans la série est multipliée par 10. La valeur maximale de la série est trouvée à 66,274. Les observations sont classées dans les classes à 49 de l'égalité (1,333) de largeur. L'histogramme est construit avec une valeur moyenne de classes sur le long de l'axe des x et de la fréquence

sur le long de l'axe des y . Le tracé de l'histogramme et la courbe de fréquence cumulative sont présentés dans la figure : 15 a,b.

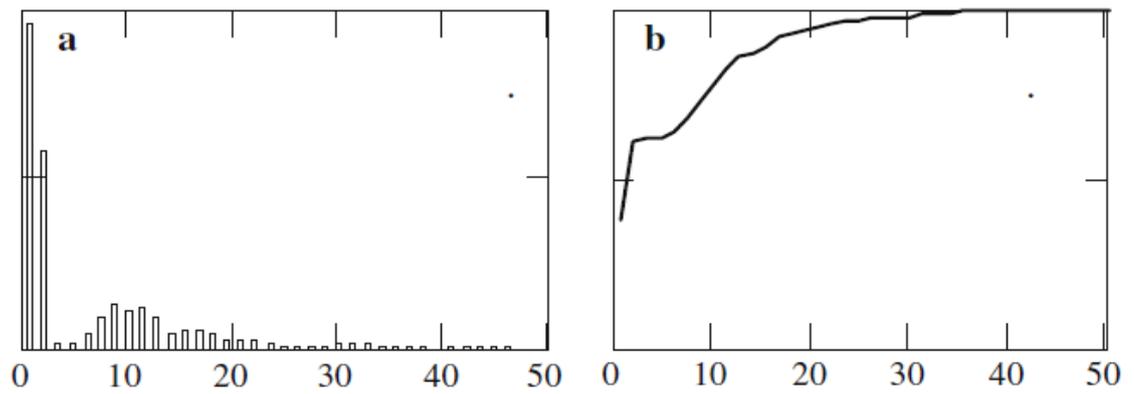


Figure 15 a,b : *L'histogramme et la courbe de fréquence cumulative de la série observée*

L'histogramme rapproche de la forme de fonction de la distribution semi-Burr présentée dans la figure 11.

Bibliographie

- [1] A. Fredricks, B. Nelsen (2007). On the relationship between Spearman's rho and Kendall's tau for pairs of continuous random variables. *Journal of Statistical Planning and Inference* 137 2143 – 2150.
- [2] A. Ferreira, L. de Hann (2006). *Extreme value theory, An introduction*. Springer.
- [3] A. Khodraji, L. Rivest, et K. Ghoudi(1998). Proprietes statistiques des couples de valeurs extremes bidimensionnelles. *The Canadian Journal of Statistics* Vol. 26, No. 1 : 187-197.
- [4] B. Nelsen (2006). *An Introduction to Copulas*. Springer Science+Business Media, Inc.
- [5] C. Robert (2001). *Gestion de risques multiples*. Journées Statistiques et Assurances Dommages, SFdS.
- [6] D.Michele, K. Jayakumar, and R.Gupta (2006). On generalized to semi-Pareto and semi-Burr distributions and random coefficient minification processes. *Statistical Papers* 51 : 193-208.
- [7] E. Luciano, U. Cherubini, and W. Vecchiato (2004). *Copula Methods in Finance*. Wiley Finance, England.
- [8] F. Bretz, G. Alan (2009). *Computation of Multivariate Normal and t Probabilities*. Springer.
- [9] G. Melfi, Y. Dodge (2008). *Premiers pas en simulation*. Springer.
- [10] G. Salvadori (2007). *Extremes in nature*. Springer
- [11] J. Schiller, R. Spiegel, and R.Alu (2001). *Probability and statistics*. Schauum's Easy outlines.
- [12] K. Jayakumar, T. Mathew (2008). On a generalization to Marshall–Olkin scheme and its application to Burr type XII distribution. *Statistical Papers* 49 : 421–439.

-
- [13] K. Rakotomalala (2009). Les fonctions copule en finance, Paris : Presses de la Sorbonne.
 - [14] K. Trivedi, M. Zimmer (2007). Copula Modeling : An Introduction for Practitioners. Foundations and Trends in Econometrics Vol. 1, No 1, 1–111.
 - [15] L.Mario (2005). Applied probability and statistics. Springer.
 - [16] O. Okhrin W. Härdle and Y. Okhrin (2008). Modeling Dependencies in Finance using Copulae. SFB 649 Discussion Paper 043.
 - [17] R. Russo (2003). Statistics for the Behavioural Sciences, An Introduction. ISBN 0-203-64157-4 Master e-book ISBN.