



Faculté des Sciences Exactes et Informatique
Département de Mathématiques

N° d'ordre :

N° de séries :

Mémoire de fin d'études

Présenté pour l'obtention du diplôme de

Master

Spécialité : Mathématiques

Option : Probabilités et statistique

Thème

Inférence statistique sur les copules

Présenté par :

- Bazine Amal
- Bezzaz Rahma

Devant le jury :

Présidente	: Stihi Sara	M.C.B Université de constantine
Encadreur	: Gherda Mebrouk	M.A.A Université de Jijel
Examinatrice	: Yakoubi Fatima	M.A.A Université de Jijel

Remerciements

Avant, nous remercions Allah pour nous avoir donné la santé, la volonté, le courage et la détermination qui nous ont accompagnés tout au long de la préparation et l'élaboration de ce travail et qui nous ont permis d'achever ce modeste travail. Le présent travail est non seulement le résultat de notre courage, sacrifice, patience et endurance mais aussi une participation de plusieurs personnes qui nous sont chères.

Nous tenons d'abord à remercier très chaleureusement Mr.Mebrouk Gherda qui nous a permis de bénéficier de son encadrement. Les conseils qu'il nous a prodigué, la patience, la confiance qu'il nous a témoignés ont été déterminants dans la réalisation de notre travail.

Nos précieux remerciements vont à le président et les membres de jury :

M^{me} Yaakoubi Fatima et M^{me} Setihi Sara pour l'honneur qu'ils nos ont fait en acceptant de juger nous travail et faire partie de ce jury.

Nous n'oublions pas dans nos éloges l'ensemble du personnel du Département de mathématiques de l'Université Mohammed Seddik Ben Yahyia, et plus personnellement nos enseignants pour tous leurs efforts.

Nos remerciements vont également à tous ceux qui ont participé de près ou de loin à la réalisation de ce travail.

Dédicace

Je dédie ce travail

A la mémoire de mon défunt père

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

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

*A mon frère Mourad et sa femme Rabha ainsi que leurs enfants Behâa,
Rawnaq, Diâa et Taqwa*

A Nouredine, sa femme Hanane et leur fille Tasnime

A mes chers frères Samir et Tariq

A mes chères sœurs

Nassima, Souad, Ratiba, Wiam

A toute ma famille

A mon binôme Amal

A mes chères amies : Khadidja, Fahima, Nadjwa, Fayrouze.

Rahma

Dédicace

Je dédie ce travail

Au terme de cette étude un grand merci au Dieu, pour le courage et la force qui nous a offert pour terminer ce memoire.

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

A mes chers frères

Youcef, Mouhammed, Nasreddinne

A ma petite sœur Ikram

A mon oncle le plus précieux Idris

A la mémoire de mon oncle Mahmoud

A mon binôme Rahma

A mes chères amies : Azza, Fahima, Fayrouze.

*Finalement, à grâce aux personne les plus chères **Toutou***



Amal

Résumé

Ces dernières décennies, nous constatons à un émergence du concept de copule en modélisation statistique. Cet essor est justifié par le fait que les copules permettent de faire une analyse séparée des marges et de la structure de dépendance induite par une distribution statistique. Cette séparation facilite l'incorporation des lois non gaussiennes et la prise en compte des dépendances non linéaires entre les variables aléatoires. La finance et l'hydrologie sont deux exemples de domaines où les copules sont très utilisées. Puisqu'il existe beaucoup de familles de copules bivariées, il sera toujours possible à l'utilisateur d'en choisir une qui lui convienne. Cependant, on ne peut pas dire autant dans le cas multivarié, la gamme de ces modèles n'est pas encore assez riche pour pouvoir en choisir un qui satisfasse toutes les propriétés que l'on souhaiterait a priori.

Abstract

In the last decades, copulas have been more and more used in statistical modeling. Their popularity owes much to the fact that they allow to separate the analysis of the margins from the analysis of the dependence structure induced by the underlying distribution. This renders easier the modeling of non Gaussian distributions, and moreover, it allows to take into account non linear dependencies between random variables. Finance and hydrology are two examples of scientific fields where the use of copulas is nowadays standard. Since there exists many families of bivariate copulas, it is always possible for the user to choose one that suits his/her needs. Unfortunately, the multivariate case is not that simple. The range of these models is still not rich enough for the user to choose one that satisfies all the desired properties.

Table des matières

Introduction	1
1 Les copules et ces propriétés	3
1.1 Les mesures de dépendance	3
1.1.1 coefficient de corrélation linéaire	3
1.1.2 Les coefficients de corrélation tau de Kendall et rho de Spearman	4
1.2 Les copules	5
1.2.1 Introduction aux copules	5
1.2.2 Définition d'une copule et le théorème d'existence	6
1.3 Propriétés des copules	9
1.3.1 Propriétés principales	9
1.3.2 Les bornes de Fréchet-Hoeffding	11
1.3.3 Copules et variables aléatoires	12

1.3.4	La notion de dépendence des queues	13
1.4	Les copules paramétriques	15
1.4.1	Copule Archimédienne	15
1.4.2	les copules elliptiques	17
1.4.3	Copule des valeurs extrêmes	19
2	Inférence statistique des copules	20
2.1	Estimation	20
2.1.1	quelques méthodes d'estimation	21
2.1.2	Estimation des copules	23
2.2	Les tests	33
2.2.1	Test d'adéquation et choix de la copule	33
2.2.2	Sélection d'une copule à partir du test d'adéquation du chi-deux	34
2.2.3	Test de Cramér-Von Mises	37
3	Applications	38
3.1	Exemple d'utilisation du modèle copule en physique	38
3.2	Application en finance	41
3.2.1	Notions de base de risque	42
3.2.2	Mesure de risque	42
3.2.3	Valeur-à-Risque	43

3.2.4	La gestion de risque en utilisant la copule	46
	Conclusion	53
	Annexe A	54
	Annexe B	58
	Bibliographie	61

Introduction générale

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.

Un riche ensemble de mesures de dépendance entre les variables aléatoires a été proposé comme le coefficient de corrélation de Pearson, le tau de Kendall et le rho de Spearman. Bien que ces mesures sont simples à calculer et peuvent être facilement interprétées, elles ne sont pas en mesure de détecter toutes les formes de dépendances, donc il était indéniable de

trouver un autre moyen pour résoudre ce problème. En effet, la fonction copule a l'avantage de modéliser complètement la dépendance entre les variables.

Ce mémoire se divise en trois chapitres.

Le premier chapitre est une introduction mathématique aux copules. On y retrouve les principaux théorèmes qui serviront de base aux chapitres suivants. Dans ce chapitre nous présentons quelques familles de copules paramétriques les plus utilisées en pratique.

Le second chapitre est réparti en deux sections. Dans la première section nous allons présenter les différentes méthodes d'estimation des copules à savoir l'approche paramétrique, semi-paramétrique et non-paramétrique. La première approche consiste à estimer le paramètre de la copule par l'une des méthodes suivantes : méthode MLE (Maximum de vraisemblance), méthode IFM (fonctions d'inférence des marginales). Dans la seconde approche on utilise la méthode de maximum de vraisemblance canonique (CML) et la méthode des moments. Quand à la troisième approche on peut utiliser les méthodes suivantes : la méthode de la copule empirique, et la méthode à noyau.

Dans la deuxième section on donne deux tests d'adéquation : le test de khi-2 et le test de Cramér-von Mises.

Finalement nous proposons une application de l'analyse du risque dans le marché financier en mettant l'accent sur la valeur à risque (VaR), qui est une mesure capitale dans ce domaine avant l'apparition des copules, les modèles classiques présentent certaines anomalies notamment ce qui concerne l'hypothèse d'indépendance et de normalité. Les copules possèdent une structure de dépendance particulière et permettent de créer des distributions dont les marginales sont des lois quelconques et dont les propriétés de dépendances peuvent être très variées, l'implémentation se fera à l'aide du logiciel de statistique «R». Ainsi qu'une autre application concernant l'utilisation des copules dans le domaine de la physique.

Les copules et ces propriétés

Les copules sont des objets mathématiques permettant de modéliser la dépendance entre les distributions unidimensionnelles

1.1 Les mesures de dépendance

1.1.1 coefficient de corrélation linéaire

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

$$r(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X)\text{Var}(Y)}}.$$

Où $\text{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$ est la covariance entre X et Y ; $\text{Var}(X)$, $\text{Var}(Y)$ correspondent aux variances respectives des variables X , Y

La définition du coefficient de corrélation linéaire est donc subordonnée à l'existence des variances de X et Y .

Dans le cadre d'une dépendance linéaire parfaite, $Y = aX + b$, ($a \neq 0$; $b \in \mathbb{R}$), le coefficient de corrélation est égal à $+1$ ou -1 selon le signe de a .

D'autre part, ce coefficient de corrélation reste invariant par des transformations linéaires strictement croissantes des variables aléatoires. En effet : $r(aX + b, cY + d) = \text{sign}(ac) \times r(X, Y)$

Néanmoins, ce coefficient ne demeure pas constant sous l'hypothèse d'une transformation croissante non linéaire.

1.1.2 Les coefficients de corrélation tau de Kendall et rho de Spearman

Les définitions de ces deux coefficients sont intimement liées à la notion de concordance.

Ils constituent une alternative au coefficient de corrélation linéaire, qui n'est pas comme nous l'avons montré auparavant la mesure de dépendance la plus appropriée et souffre de certaines lacunes.

Définition 1.1.2. *Soient $(x_1, y_1), (x_2, y_2)$ deux observations d'un couple de variables aléatoires continus (X, Y) sont dit :*

- *concordantes si*

$$(x_1 - x_2)(y_1 - y_2) > 0$$

- *discordantes si*

$$(x_1 - x_2)(y_1 - y_2) < 0$$

Définition 1.1.3. *La fonction de concordance entre les deux vecteurs aléatoires (X_1, Y_1) et (X_2, Y_2) est définie par*

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

a/ Le tau de Kendall

Définition 1.1.4. *soient (X_1, Y_1) et (X_2, Y_2) deux vecteurs aléatoires continus iid de fonction de répartition conjointe H et de fonctions marginales F (pour X_1, X_2) et G (pour Y_1, Y_2). Le tau de Kendall noté τ est définie par*

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

Soient a et b des fonction strictement croissante, alors

$$\tau(a(X), b(Y)) = \tau(X, Y)$$

b/ Le rho de Spearman

Définition 1.1.5. *soient (X_1, Y_1) , (X_2, Y_2) et (X_3, Y_3) trois vecteurs aléatoires indépendants de même loi H . Le coefficient de corrélation de Spearman est définie par*

$$\rho = 3\mathbb{P}[(X_1 - X_2)(Y_1 - Y_3) > 0] - \mathbb{P}[(X_1 - X_2)(Y_1 - Y_3) < 0]$$

Ce coefficient est invariant sous transformation non linéaire strictement croissante c'est -à-dire si a et b sont deux fonctions strictements croissantes, alors

$$\rho(a(X), b(Y)) = \rho(X, Y)$$

1.2 Les copules

1.2.1 Introduction aux copules

Les copules représentent un outil innovant pour modéliser la structure de dépendance de plusieurs variables aléatoires, introduit par Abé Sklar en 1959. Elles sont des fonctions de répartition définie sur $[0, 1]^n$ dont les loi marginales sont égales à la loi uniforme, les copules

permet de distinguer le comportement des distributions conjointes peut s'avérer difficile lorsqu'on modélise le comportement conjoint de deux phénomènes dont les marges ne sont pas normales.

Pour faciliter la tâche nous nous intéressons plus particulièrement à l'étude des copules bivariées.

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

Dans tout la suite I désigne l'intervalle $[0,1]$.

Définition 1.2.1. *on appelle copule bivariée toute fonction C définie de $I^2 \rightarrow I$ qui possède les propriétés suivantes*

1. $\forall (u, v) \in I^2, C(u, 0) = C(0, v) = 0$
2. $\forall (u, v) \in I^2, C(u, 1) = u$ et $C(1, v) = v$
3. C est 2-croissante c-à-d : $\forall (u_1, u_2), (v_1, v_2) \in I \times I$ avec $u_1 \leq u_2$ et $v_1 \leq v_2$

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

Cette définition signifie que C est une distribution de marginales uniformes.

Exemples 1.2.2.

1. $\forall (u, v) \in I^2$, la fonction $M(u, v) = \min(u, v)$ définit une copule, en effet
 - $\forall (u, v) \in I^2, \min(u, 0) = \min(0, v) = 0 \Rightarrow M$ est vérifier (1)
 - $\forall u \in I, \min(u, 1) = \min(1, u) = u \Rightarrow M$ est vérifier (2)
 - $\forall (u_1, v_2), (u_2, v_1) \in I \times I$, avec $u_1 \leq u_2$ et $v_1 \leq v_2$
 - si $u_1 \leq u_2 < v_1 \leq v_2$ on a

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

- si $v_1 \leq v_2 < u_1 \leq u_2$ on a

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

- si $u_1 < v_1 < u_2 < v_2$ on a

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

D'où M est une copule.

2. $\forall (u, v) \in I^2$, la fonction $W(u, v) = \max(u + v - 1, 0)$ définit une copule.
3. On a aussi $\forall (u, v) \in I^2$, $\prod(u, v) = uv$ définit une copule appelée la copule d'indépendance.

Théorème d'existence (théorème de Sklar)

Ce théorème est fondamentale dans la théorie des copules, il permet d'élucider le rôle que joue la copule dans la relation conjointe et les distributions marginales.

Définition 1.2.3. Soit (X, Y) un vecteur des variables aléatoires admet une fonction de distribution jointe alors les distributions marginales de X et Y sont respectivement données par $F_X(x) = \lim_{y \rightarrow \infty} H(x, y)$, et $G_Y(y) = \lim_{x \rightarrow \infty} H(x, y)$.

Théorème 1.2.4. Soit H une fonction de distribution bivariable ayant comme fonction de distribution marginales F et G , alors il existe une copule C telle que pour tout $(x, y) \in \mathbb{R}^2$

$$H(x, y) = C(F(x), G(y))$$

Proposition 1.2.5. Soit X une variable aléatoire de fonction de distribution F , alors

1. Si U est uniforme sur $[0, 1]$, alors $F^{-1}(u) \xrightarrow{d} F$
2. F est continue, alors $F(x) \xrightarrow{d} \mathcal{U}_{[0,1]}$

Revenant à la preuve de ce théorème

$$\begin{aligned}
 H(x, y) &= \mathbb{P}(X \leq x, Y \leq y) \\
 &= \mathbb{P}(F_X^{-1}(U_1) \leq x, G_Y^{-1}(U_2) \leq y) \\
 &= \mathbb{P}(U_1 \leq F(x), U_2 \leq G(y)) \\
 &= C(F(x), G(y)).
 \end{aligned}$$

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

Sinon, lorsque les marginales ne sont pas continues, il est toujours possible de définir une copule mais elle ci n'est plus unique et de ce fait perd beaucoup de son intérêt, elle définit de manière unique sur $(Im(F) \times Im(G))$. En effet : on pose $F(x) = u$, $G(y) = v$ on trouve $C(u, v) = H(F^{-1}(u), G^{-1}(v))$.

Définition 1.2.6 (Inverse généralisé). *Soit F une fonction de distribution. L'inverse généralisé $F^{[-1]}$ de F est une fonction définie de la façon suivante*

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

Remarque 1.2.7. *Si la fonction F est strictement croissante, alors la notion de l'inverse généralisé coïncide avec la notion d'inverse ou de réciproque c'est-à-dire $F^{[-1]} = F^{-1}$.*

Exemple 1.2.8. *soit*

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

$$\begin{aligned}
F_X(x) &= \lim_{y \rightarrow +\infty} H(x, y) \\
&= H(x, +\infty) = \frac{(x+1)(e^{+\infty} - 1)}{x + 2e^{+\infty} - 1} \\
&= \lim_{y \rightarrow +\infty} \frac{(x+1)(e^y - 1)}{x + 2e^y - 1} \\
&= \frac{x+1}{2}
\end{aligned}$$

$$\begin{aligned}
G_Y(y) &= \lim_{x \rightarrow +\infty} H(x, y) \\
&= \lim_{x \rightarrow +\infty} 1 - e^{-y} \\
&= 1 - e^{-y}
\end{aligned}$$

alors

$$F^{-1}(u) = 2u - 1$$

$$G^{-1}(v) = \ln \frac{1}{1-v}$$

et donc

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

Remarque 1.2.9. Si C est une copule et F et G des fonction de répartition, alors la fonction H est une fonction de répartition jointe dont les marginales sont F et G .

1.3 Propriétés des copules

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

1.3.1 Propriétés principales

Théorème 1.3.1 (La continuité). Soit C une copule bivariée; $\forall u_1, u_2, v_1, v_2 \in I$, avec $u_1 \leq u_2$ et $v_1 \leq v_2$, on a

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

Théorème 1.3.2 (La différentiabilité). *Soit C une copule bivariée ; $\forall u_1, u_2 \in I$*

1. *Les dérivés partielles : $\frac{\partial C(u_1, u_2)}{\partial u_j}$ existent p.s et $0 \leq \frac{\partial C(u_1, u_2)}{\partial u_j} \leq 1, \forall j = 1, 2.$*
2. *Les fonctions de la forme $\frac{\partial C(u_1, u_2)}{\partial u_j}$ sont bien définies et décroissantes sur I ps.*

Remarque 1.3.3 (La densité). *Les copules admettent des densités de probabilité, si la densité c associée à la copule C existe, alors elle est définie par : $c(u, v) = \frac{\partial^2 C(u, v)}{\partial u \partial v}.$*

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

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

Théorème 1.3.4 (Convexité et concavité). *Soit $(a_1, a_2), (b_1, b_2) \in I^2$ et $\forall \lambda \in I$*

On dit que C est convexe si

$$C(\lambda a_1 + (1 - \lambda)b_1, \lambda a_2 + (1 - \lambda)b_2) \leq \lambda C(a_1, a_2) + (1 - \lambda)C(b_1, b_2)$$

On dit que C est concave si

$$C(\lambda a_1 + (1 - \lambda)b_1, \lambda a_2 + (1 - \lambda)b_2) \geq \lambda C(a_1, a_2) + (1 - \lambda)C(b_1, b_2)$$

La relation d'ordre partiel sur les copules

Définition 1.3.5. *Soient C_1 et C_2 deux copules, on dit que C_1 est plus petit que C_2 et on note : $C_1 \prec C_2$ si $C_1(u, v) \leq C_2(u, v), \forall u, v \in [0, 1].$*

Cette relation est une relation d'ordre partiel.

1.3.2 Les bornes de Fréchet-Hoeffding

Tout copule C admet une borne inférieure et une borne supérieure déterminée par le théorème suivant

Théorème 1.3.6 (Fréchet-Hoeffding). *Soit C une copule alors*

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

. telles que : $W(u, v) = \max(u + v - 1, 0)$ et $M(u, v) = \min(u, v)$ représentent les bornes inférieure et supérieure de Fréchet-Hoeffding.

Preuve 1.3.7. *Soit $(u, v) \in I^2$, puisque $C(u, v) \leq C(u, 1) = u$ et $C(u, v) \leq C(1, v) = v$, alors $C(u, v) \leq \min(u, v) = M(u, v)$, d'où la deuxième inégalité. Pour démentrer la première inégalité, nous utilisons le fait que $C(u, v) \geq 0$, et l'inégalité (3) de la définition(1.2.1) on a*

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

D'où la première inégalité.

Remarque 1.3.8. *Comme conséquence du théorème de Sklar, les copules W et M sont appelées respectivement les limites inférieure et supérieure de Fréchet l'inégalité $W \leq C \leq M$ peut-être réécrite comme : $\max(F(x) + G(y) - 1, 0) \leq H(x, y) \leq \min(F(x), G(y))$ est connue sous le nom de l'inégalité de Fréchet-Hoeffding.*

1.3.3 Copules et variables aléatoires

Symétrie

Soit X, Y deux variables aléatoires continues de fonction de répartition jointe H et de marginales F et G , et soit C la copule associée on dit que X, Y sont échangeables ssi $F = G$ et $\forall u, v \in I^2$ si $C(u, v) = C(v, u)$ on dit que C est symétrique.

Invariance fonctionnelle

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

Théorème 1.3.9. *Soient deux variables aléatoires continues X, Y de marges F et G et de copule $C_{X,Y}$, si α et β deux fonctions strictement croissantes alors*

$$C(\alpha(X), \beta(Y)) = C(X, Y)$$

Ainsi la copule $C_{X,Y}$ est invariante par transformations strictement croissantes des variables aléatoires.

Par exemple nous avons

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

Une transformation croissante ne modifie donc pas la copule, mais seulement les loi marginale.

L'indépendance

Si X et Y sont des variables aléatoires continues, alors : X, Y sont indépendantes ssi $C_{XY} = \Pi$ avec $\Pi(u, v) = uv$ pour tout $u, v \in I$

1.3.4 La notion de dépendance des queues

Cette notion est très importante dans l'étude de la dépendance asymptotique entre deux variables aléatoires X et Y . Ce concept sera totalement basé sur celui des copules. L'objectif est donc l'étude de la dépendance dans la queue commune de la distribution bivariée.

Le concept de dépendance de queue fournit une description de la dépendance au niveau des queues de distribution. La dépendance de queue est une mesure locale contrairement au tau de Kendall et au rho de Spearman qui mesurent la dépendance sur l'ensemble de la distribution.

1. La dépendance de queue supérieure

Prenons deux variables aléatoires continues X, Y ayant pour fonction de distribution respectives F_X et G_Y le coefficient de dépendance de queue supérieure de X et Y est défini par la limite λ_u si elle existe

$$\lambda_u = \lim_{u \rightarrow 1^-} \mathbb{P}(Y > F_Y^{-1}(u) / X > F_X^{-1}(u))$$

- Si toutefois cette limite de $\lambda_u \in [0, 1]$ existe.

La quantité λ_u est une fonction de copule et est donc invariante par transformation croissante.

- Si $\lambda_u \in]0, 1]$, on dit que X et Y sont asymptotique dépendantes au niveau supérieur de la queue de distribution.
- Si $\lambda_u = 0$, on dit que X et Y sont asymptotique indépendantes au niveau supérieur de la queue de distribution.

2. La dépendance de queue inférieure

En gardant, les mêmes notations qu'auparavant, le coefficient de dépendance de queue inférieure de X et Y est défini par la limite λ_l si elle existe

$$\lambda_l = \lim_{l \rightarrow 0^+} \mathbb{P}(Y > F_Y^{-1}(u) / X > F_X^{-1}(u))$$

- Si $\lambda_l \in]0, 1]$ on dit que X et Y sont asymptotiquement dépendantes au niveau inférieur de la queue de distribution
- Si $\lambda_l = 0$ alors, il n'y a pas de dépendance de queue au niveau inférieure de la distribution.

On peut exprimer λ_u et λ_l à l'aide de la copule C de couple (X, Y) tel que décrit dans la proposition suivante :

Proposition 1.3.10. *Etant donné deux variables aléatoires X et Y de copule $C(X, Y)$, on a*

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

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

Exemple 1.3.11. *Pour la copule W , nous obtenons*

$$\lambda_u = \lim_{u \rightarrow 1^-} \frac{1 - 2u + \max(2u - 1, 0)}{1 - u}$$

$$= 0$$

$$\lambda_l = \lim_{u \rightarrow 0^+} \frac{\max(2u - 1, 0)}{u}$$

$$= 0$$

Les mesures de dépendance en fonction de la copule

Le tau de Kendall et le rho de Spearman sont deux mesures de concordance bien connues en statistique. Elles donnent une mesure de la corrélation entre les rangs des observations, à la

différence du coefficient de corrélation linéaire qui lui mesure la corrélation entre les valeurs des observations. Elles offrent par ailleurs l'avantage de s'exprimer simplement en fonction de la copule associée au couple de variables aléatoires.

Théorème 1.3.12 (Le tau de Kendall). *Soit (X, Y) un couple de variables aléatoires continues de copule C le tau de Kendall est la différence entre la probabilité de concordance et la probabilité de discordance, son expression en terme de copule est la suivante*

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

Théorème 1.3.13 (Le rho de spearman). *Soit (X, Y) un couple de variables aléatoires continues de copule C , alors*

$$\begin{aligned}\rho(X, Y) &= 12 \int_0^1 \int_0^1 uv dC(u, v) - 3 \\ &= 12 \int_0^1 \int_0^1 C(u, v) dudv - 3.\end{aligned}$$

1.4 Les copules paramétriques

1.4.1 Copule Archimédienne

L'idée de la copule archimédienne de générateur φ est que la transformée $w(u) = \exp(-\varphi(u))$ appliquée aux marginales rend les composantes indépendantes.

Définition 1.4.1. *Soit $\varphi : [0, 1] \rightarrow \mathbb{R}_+^*$, continue, décroissante et convexe, telle que $\varphi(1) = 0$, alors φ est dite générateur. La fonction inverse généralisée de φ est définie par*

$$\varphi^{[-1]}(u) = \begin{cases} \varphi^{-1}(u) & \text{si } 0 \leq u \leq \varphi(0) \\ 0 & \text{si } \varphi(0) \leq u \leq \infty \end{cases} \quad (1.1)$$

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

Définition 1.4.2. Une copule est dit archimédienne si elle d'écrire sous la forme

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

Nous allons, dans la suite de cette partie citer quelques copules parmi les plus célèbres de cette famille.

Les familles des copules archimédiennes

1. copule Clayton

Soit la fonction génératrice $\varphi(t) = \frac{t^{-\beta} - 1}{\beta}$, et de pseudo inverse $\varphi^{[-1]}(t) = (\beta t + 1)^{-\frac{1}{\beta}}$ avec $\beta \in [-1, \infty \setminus \{0\}]$. En appliquant l'expression(1.1), on trouve la famille des copules de Clayton; sa fonction est donnée par

$$C_{\beta}(u, v) = \max((u^{-\beta} + v^{-\beta} - 1)^{-\frac{1}{\beta}}, 0)$$

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

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

2. Copule de Gumbel

Prenons comme générateur $\varphi(t) = (-\ln(t))^{\alpha}$ avec $\alpha > 1$, et son pseudo-inverse $\varphi^{-1}(t) = \exp\{-t^{\frac{1}{\alpha}}\}$. La fonction ainsi définie satisfait à les conditions du théorème sur les copules archimédiennes, ce qui permet générer la copule de Gumbel en prenant

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

3. copule de Frank

Soit $\varphi(t) = -\ln\left(\frac{e^{\theta t} - 1}{e^\theta - 1}\right)$ avec $\theta \neq 0$, l'application de la fonction (1.1) nous donne la famille de Frank qui est définie comme suit

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

1.4.2 les copules elliptiques

Les copules elliptiques sont définies à partir des lois de distribution elliptique.

Dans ce qui suit, nous donnons quelques définitions de la distribution elliptique ainsi que deux exemples classiques de cette famille de copules qui sont la copule Gaussienne et la copule de student.

Définition 1.4.3. *Un vecteur aléatoire $X = (X_1, X_2)$ est de distribution elliptique s'il admet la représentation suivante*

$$X = \mu + RAU$$

Où

- $\mu = (\mu_1, \mu_2) \in \mathbb{R}^2$
- U est un vecteur aléatoire uniforme sur la sphère unité de \mathbb{R}^2
- R est un vecteur aléatoire indépendant de U
- A est une matrice de dimension $n \times n$ telle que $\Sigma = AA^t$ est non singulière.

Définition 1.4.4. *La fonction de densité d'une distribution elliptique (si elle existe) est donnée par*

$$f(x) = |\Sigma|^{-\frac{1}{2}} g((X - \mu)^t(X - \mu))$$

où g est une fonction définie de \mathbb{R}^+ dans \mathbb{R} , dite génératrice de densité, uniquement déterminée par la fonction de distribution de R .

1. La copule Gaussienne

La copule gaussienne fait partie de la famille des copules elliptiques. L'importance de cette copule réside dans le fait qu'elle est sous-jacente à la distribution normale multivariée.

Définition 1.4.5. Soit $X = (X_1, X_2)$ un vecteur aléatoire gaussien ($X \sim \mathcal{N}(0, \Sigma)$), avec Σ une matrice de covariance. Notons par ϕ_Σ et ϕ les fonction de distribution X et X_i , $i = 1, 2$ respectivement. Le vecteur aléatoire Gaussien possède une copule appelée copule Gaussienne donnée par

$$\begin{aligned} C(u, v) &= \phi_\Sigma(\phi^{-1}(u), \phi^{-1}(v)) \\ &= \int_{-\infty}^{\phi^{-1}(u)} \int_{-\infty}^{\phi^{-1}(v)} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left(-\frac{x_1^2 + x_2^2 - 2x_1x_2}{2(1-\rho^2)}\right) dx_1 dx_2 \\ &= \int_0^u \int_0^v \frac{1}{\sqrt{1-\rho^2}} \exp\left(-\frac{x_1^2 + x_2^2 - 2x_1x_2}{2(1-\rho^2)} + \frac{x_1^2 + x_2^2}{2}\right) dx_1 dx_2 \end{aligned}$$

avec $x_1 = \phi^{-1}(u)$, $x_2 = \phi^{-1}(v)$

et ρ est le coefficient de corrélation ($\rho \in [-1, 1]$)

2. La copule de student

Définition 1.4.6. Soit $\rho \in [0, 1]$, la fonction de répartition de Student à v degré de liberté est définie par

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

Définition 1.4.7. La copule de Student est une copule paramétrique, paramétrée par le coefficient de corrélation linéaire ρ et de degré de liberté v . Cette copule est définie par

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

Si $v \rightarrow \infty$, alors la copule de student converge vers la copule gaussienne et dans ce cas très difficile de différencier entre ces deux copules.

1.4.3 Copule des valeurs extrêmes

Une autre classe particulière des copules est celle des valeurs extrêmes. Le nom "extreme value copula" suggère un lien entre la théorie des valeurs extrêmes et ses copules.

Définition 1.4.8. *On appelle copule des valeurs extrêmes, toute copule C vérifiant la propriété suivante*

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

Où

$$C^{\frac{1}{t}}(u^t, v^t) = C(u, v)$$

Exemple 1.4.9. *La copule de Gumbel est une copule de valeurs extrêmes, en effet*

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

Inférence statistique des copules

La fonction de copule reste inconnue en pratique ainsi que ces paramètres, d'où le besoin de les estimer. Dans ce chapitre nous allons présenter quelques méthodes d'estimation et deux tests d'adéquation pour le choix de bon copule.

2.1 Estimation

Définition 2.1.1. *L'estimation est le procédé par lequel on détermine les valeurs inconnues des paramètres des lois de probabilité ou les lois seront implicites de la population, à partir des données de l'échantillon. Pour cela on utilise des méthodes statistique paramétrique ou non paramétrique selon le cas.*

Définition 2.1.2. *Un estimateur est une statistique permettant d'évaluer un paramètre ou une loi inconnu relatif à une loi de probabilité. La qualité des estimateurs s'exprime par leur convergence, leur biais, leur efficacité et leur robustesse.*

2.1.1 quelques méthodes d'estimation

Les diverses méthodes permettent d'obtenir des estimateurs de qualités différentes

- **La méthode de maximum de vraisemblance**

Définition 2.1.3. La statistique $w \mapsto \arg \max(\theta \mapsto \prod_{i=1}^n f_{\theta}(X_i(w)))$ s'appelle l'estimateur de maximum de vraisemblance de θ .

$L : \theta \mapsto \prod_{i=1}^n f_{\theta}(x_i)$ s'appelle la fonction vraisemblance du modèle.

$l : \theta \mapsto \sum_{i=1}^n \log f_{\theta}(x_i)$ s'appelle la fonction log-vraisemblance du modèle.

En pratique, on fait l'étude de l'une des fonctions L ou l . Il n'y a pas forcément unicité.

Ces fonctions ne sont pas nécessairement dérivables ce qui annule le gradient ne réalise pas forcément un maximum .

Remarque 2.1.4. L'estimateur de maximum de vraisemblance n'existe pas toujours et n'est pas toujours unique.

Exemple 2.1.5. Le modèle de la loi exponentielle

$\Theta = \mathbb{R}^+$, $f_{\theta}(x) = \theta e^{-\theta x}$ on a

$$L(\theta) = \prod_{i=1}^n f_{\theta}(x_i) = \theta^n e^{-\theta \sum_{i=1}^n x_i}$$

$$l(\theta) = \log L(\theta) = n \log \theta - \theta \sum_{i=1}^n x_i$$

$$\frac{\partial l(\theta)}{\partial \theta} = \frac{n}{\theta} - \sum_{i=1}^n x_i = 0 \iff \hat{\theta} = \frac{n}{\sum_{i=1}^n x_i} = \frac{1}{\bar{X}_n}$$

L est une application concave car on a

$$\frac{\partial^2 L(\theta)}{\partial^2 \theta} = -\frac{n}{\theta^2}$$

Donc, $\hat{\theta} = \frac{1}{\bar{X}}$ est l'estimateur de maximum de vraisemblance dans le cas d'un modèle de la loi exponentielle .

- **La méthode des moments**

L'idée de base est d'estimer une espérance mathématique par une moyenne empirique, une variance par une variance empirique, etc...

Si la loi des X_i a deux paramètres θ_1 et θ_2 tels que $(\mathbb{E}(X), \text{Var}(X)) = \varphi(\theta_1, \theta_2)$, où φ est une fonction inversible, alors les estimateurs de θ_1 et θ_2 par la méthode des moments sont : $(\hat{\theta}_{1n}, \hat{\theta}_{2n}) = \varphi^{-1}(\bar{X}_n, S_n^2)$.

Ce principe peut naturellement se généraliser aux moments de tous ordres, centrés ou non centrés : $\mathbb{E}[(X - \mathbb{E}(X))^k]$, et $\mathbb{E}(X^k)$, $k \geq 1$.

Exemple 2.1.6. *La loi Gamma*

Si X_1, \dots, X_n sont indépendantes et de même loi gamma $G(\alpha, \lambda)$, $\mathbb{E}(X) = \frac{\alpha}{\lambda}$ et $\text{Var}(X) = \frac{\alpha^2}{\lambda}$. On en déduit facilement que

$$\lambda = \frac{\mathbb{E}(X)}{\text{Var}(X)} \text{ et } \alpha = \frac{[\mathbb{E}(X)]^2}{\text{Var}(X)}$$

Donc les EMM de α et λ sont

$$\hat{\lambda}_n = \frac{\bar{X}_n}{S_n^2} \text{ et } \hat{\alpha}_n = \frac{\bar{X}_n^2}{S_n^2}$$

Remarque 2.1.7. *Dans certains cas, l'estimation par la méthode des moments est moins bonne que l'estimation par maximum de vraisemblance. Néanmoins, dans le cas de la loi Gamma par exemple, le calcul de la fonction de vraisemblance peut poser des problèmes (l'utilisation de l'ordinateur et d'algorithmes numériques est indispensable) tandis que l'estimation des moments est très facilement accessible.*

Lorsque la taille de l'échantillon n'est pas suffisamment grande, la loi des grands nombres ne s'applique pas et par conséquent, les moments empiriques n'approchent pas suffisamment les moments théoriques.

2.1.2 Estimation des copules

A- Estimation paramétrique

Lorsqu'il s'agit d'une estimation paramétrique d'une copule, on impose un modèle paramétrique pour la copule et pour les distributions marginales. Pour estimer ces paramètres il existe différentes méthodes. Dans la suite, nous nous intéressons qu'aux deux méthodes d'estimation les plus usuelles, la méthode maximum de vraisemblance (MLE), et la méthode fonctions d'inférence des marginales (IFM).

1. la methode de maximum de vraisemblance

Soit C_θ une copule paramétrique multivariée de paramètre θ . On estime sous les deux hypothèses suivantes

$$H_0 : C \in C_0$$

telle que $C_0 = \{C_\theta : \theta \in \Theta\}$; où Θ est un sous ensemble de \mathbb{R}^p pour tout entier $p \geq 1$. Et

$$H'_0 : F_1 \in \mathcal{F}_1, \dots, F_d \in \mathcal{F}_d$$

telle que F_j pour $j = 1, \dots, d$ sont les lois marginales de X_1, \dots, X_d est $\mathcal{F}_j = \{\mathcal{F}_{\gamma_j} : \gamma_j \in \Gamma_j\}$, où Γ_j est un sous ensemble de \mathbb{R} et en maximisant la fonction log-vraisemblance $L(\theta)$ définie par

$$L(\theta) = \sum_{j=1}^d \ln f(x_j) \tag{2.1}$$

telle que $f(x_1, \dots, x_d)$ est la fonction de densité jointe de X_1, \dots, X_d donnée par

$$f(x_1, \dots, x_d) = c(F_1(x_1), \dots, F_d(x_d)) \prod_{j=1}^d f_j(x_j)$$

en peut voir que la fonction $L(\theta)$ de l'échantillon $(x_1^i, \dots, x_d^i)_{1 \leq i \leq n}$ peut s'écrire

$$\begin{aligned}
L(\theta) &= \sum_{i=1}^n \ln(c_\theta(F_1(x_{i1}), \dots, F_d(x_{id}))) \prod_{j=1}^d f_j(x_j) \\
&= \sum_{i=1}^n \ln(c_\theta(F_1(x_{i1}), \dots, F_d(x_{id}))) + \sum_{i=1}^n \sum_{j=1}^d \ln f_j(x_{ij}).
\end{aligned}$$

Où c représente la densité de la copule C_θ , alors l'estimateur de θ , noté $\hat{\theta}_{MLE}$ est

$$\hat{\theta}_{MLE} = \mathbf{argmax} L(\theta)$$

Cet estimateur est consistant est vérifier la propriété de normalité asymptotique

$$\sqrt{n}(\hat{\theta}_{MLE} - \theta) \longrightarrow \mathcal{N}(0, I^{-1}(\theta))$$

telle que $I(\theta)$ est la matrice d'information de Fisher, cette matrice est estimée par l'inverse de la matrice Hessian de la fonction de vraisemblance.

2. Méthode des fonctions d'inférence des marginales

La méthode d'estimation MLE estime conjointement les paramètres des distributions marginales et ceux de la fonction de copule, donc elle exige des calculs intensifs parce que l'estimation des paramètres des lois marginales et des paramètres de la copule se fait d'une façon simultanée ce qui engendre des temps de calcul très longs dans le cas d'une grande dimension. Joe et Xu [1996] ont proposé la méthode IFM (Inference Functions for Margins) et ont estimé les paramètres séparément. Cette méthode repose sur le fait que la représentation en copule permet de séparer les paramètres spécifiques des distributions marginales de ceux de la structure de dépendance, cette méthode se compose de deux étapes :

- Estimer les paramètres de distribution marginales $\gamma_1, \dots, \gamma_d$ qui sont définie par

$$\hat{\gamma}_j = \mathbf{argmax} L_j(\gamma_j) \tag{2.2}$$

Telle que

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

Où f_j est la densité de F_j

- l'estimation du paramètre θ de la copule. En utilisant les estimateurs $\hat{\gamma}_j$ obtenues par (1.2), nous avons alors

$$\hat{\theta}_{IFM} = \mathbf{argmax} L(\theta)$$

Où

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

L'estimateur $\hat{\theta}_{IFM}$ vérifie aussi la propriété de normalité asymptotique démontrée par Joe (1997)

$$\sqrt{n}(\hat{\theta}_{IFM} - \theta) \longrightarrow \mathcal{N}(0, V^{-1}(\theta))$$

Avec $V(\theta)$ est la matrice d'information de Godambe, définie par

$$V(\theta) = D^{-1}M(D^{-1})^t$$

Où

$$D = \mathbb{E}\left[\frac{\partial}{\partial \theta}(g(\theta)^t)\right], \quad M = \mathbb{E}[g(\theta)^t g(\theta)], \quad \text{et } g(\theta) = \left(\frac{\partial}{\partial \gamma_1} L_1, \dots, \frac{\partial}{\partial \gamma_d} L_d\right)$$

B- Estimation semi-paramétrique

L'approche semi-paramétrique suppose un modèle paramétrique pour la copule C , et non-paramétrique pour les distributions marginales. Par exemple la méthode de maximum de vraisemblance canonique (CML) et la méthode de moments

1. la méthode de maximum de vraisemblance canonique (pseudo-maximum de vraisemblance)

Cette méthode a été proposé dans le cas où les marginales F_1, \dots, F_d associées aux X_1, \dots, X_d sont inconnues, elle compose de deux étapes :

- On remplace les marginales F_1, \dots, F_d par leurs estimations naturelles (estimateur empirique), définies par

$$\hat{F}_{X_j}(x) = \frac{1}{n} \sum_{i=1}^n 1_{\{X_{ij} \leq x\}}$$

Avec $1_{\{X_{ij} \leq x\}}$ est la fonction indicatrice.

- Estimer le paramètre de la copule paramétrique par la méthode MLE

$$\hat{\theta}_{CML} = \mathbf{argmax} L(\theta)$$

Telle que

$$L(\theta) = \sum_{i=1}^n \ln c_{\theta}(\hat{F}_1(x_{i1}), \dots, \hat{F}_d(x_{id}))$$

Cette méthode génère des temps de calcul limités et elle procède à une estimation paramétrique de la copule indépendamment de la forme paramétrique des lois marginales.

Remarque 2.1.8. *Nous avons pour tout $i \in \{1, \dots, n\}$ et $j \in \{1, \dots, d\}$, $R_{ij} = n\hat{F}_j(x_{ij})$, où R_{ij} est le rang de X_{ij} . Donc l'estimateur tiré de cette méthode est basé sur les rangs des observations.*

Exemple 2.1.9. *Dans le cas bivarié*

Soit X, Y deux variables aléatoires dont les marginales sont F et G respectivement. Soit C_{θ} la copule associée à X et Y de paramètre θ . La pseudo log-vraisemblance est définie à partir de l'équation (1.1) par

$$L(\theta) = \sum_{i=1}^n \ln c_{\theta}(F_n(x_i), G_n(y_i)),$$

telle que :

$$F_n(x_i) = \frac{1}{n} \sum_{i=1}^n 1_{\{X_1 \leq x\}} \text{ et } G_n(y_i) = \frac{1}{n} \sum_{i=1}^n 1_{\{Y_1 \leq y\}}.$$

Comme $nF_n = R_i$ et $nG_n = S_i$; où R_i et S_i sont les rangs de X_i et Y_i dans leurs échantillons respectivement, donc $L(\theta)$ devient

$$L(\theta) = \sum_{i=1}^n \ln c_{\theta}\left(\frac{R_i}{n}, \frac{S_i}{n}\right)$$

2. La méthode des moments

Cette méthode consiste à estimer les paramètres γ_j , $j = 1, \dots, d$ des lois marginales et le paramètre θ de la copule par la méthode des moments, i.e.,

(a) Résoudre le système des d équations à d inconnues

$$\left\{ \begin{array}{l} \overline{X}_t = f(\gamma_1, \dots, \gamma_d) \\ S_t^2 = g(\gamma_1, \dots, \gamma_d) \\ \mu_{3,t} = h(\gamma_1, \dots, \gamma_d) \\ \cdot \\ \cdot \\ \cdot \end{array} \right.$$

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

(b) Inverser le tau de Kendall ou le rho de Spearman pour obtenir le paramètre θ de la copule, car il existe une relation entre ces mesures et le paramètre de dépendance de la copule.

Soit (X, Y) un couple de variables aléatoires dont la copule est C_θ , de paramètre θ , telle que $\theta \in \Theta \subset \mathbb{R}$.

- **L'estimateur de θ basé sur le tau de Kendall**

Définition 2.1.10. Soit un échantillon de taille n de données bivariées $\{x_i, y_i\}_{i=1}^n$. Définissons les nombres c et d comme étant respectivement le nombre de paires concordantes et discordantes dans cet échantillon. On définit la version empirique du tau de Kendall par

$$\tau_n = \frac{c - d}{c + d} = \frac{c - d}{C_2^n},$$

avec

$$C_2^n = \frac{n!}{2!(n-2)!}$$

Supposant qu'il existe une relation entre le tau et le paramètre θ , définie par l'égalité suivante

$$\tau(X, Y) = g_1(\theta),$$

où g est une fonction continue et dérivable, alors un estimateur $\hat{\theta}_{TK}$, de θ est défini par

$$\hat{\theta}_{TK} = g_1^{-1}(\tau_n),$$

telle que τ_n est l'estimateur empirique du tau de Kendall .

Cet estimateur est asymptotiquement normale

$$\sqrt{n}(\hat{\theta}_{TK} - \theta) \longrightarrow \mathcal{N}(0, \hat{\sigma}_\tau^2),$$

telle que $\hat{\sigma}_\tau^2$ est la variance empirique de σ_τ^2 définie par

$$\hat{\sigma}_\tau^2 = 4S'g(\tau_n)^2,$$

où

$$S'^2 = \frac{1}{n} \sum_{i=1}^n (W_i + \tilde{W}_i - 2\bar{W})^2$$

et

$$W_i = \frac{1}{n} \sum_{j=1}^n 1_{\{X_j \leq X_i, Y_j \leq Y_i\}}$$

$$\tilde{W}_i = \frac{1}{n} \sum_{i=1}^n 1_{\{X_i \leq X_j, Y_i \leq Y_j\}}$$

- **L'estimateur de θ basé sur le rho de Spearman**

Définition 2.1.11. Soit un échantillon de taille n de données bivariées $\{x_i, y_i\}_{i=1}^n$ la version empirique du rho de Spearman est définie par

$$\rho_n = \frac{12}{n(n^2 - 1)} \sum_{i=1}^n R_i S_i - 3 \left(\frac{n+1}{n-1} \right) \quad (2.3)$$

telle que, R_i est le rang de X_i parmi les X_1, \dots, X_n et S_i est le rang de Y_i parmi les Y_1, \dots, Y_n .

De la même manière, si on suppose que le rho de Spearman se définit en fonction de θ par la relation suivante

$$\rho(X, Y) = g_2(\theta),$$

où g_2 est une fonction continue et dérivable, alors l'estimateur $\hat{\theta}_{RS}$ de θ est défini par

$$\hat{\theta}_{RS} = g_2^{-1}(\rho_n),$$

telle que ρ_n est l'estimateur empirique de ρ .

Cette estimateur est asymptotiquement normale

$$\sqrt{n}(\hat{\theta}_{RS} - \theta) \longrightarrow \mathcal{N}(0, (\sigma_n g_2^{-1}(\rho_n)^2)^2),$$

telle que σ_n^2 est l'estimateur de σ^2 définie par

$$\sigma_n^2 = 144(-9A_n^2 + B_n + 2C_n + D_n + 2\epsilon_n),$$

où

$$A_n = \frac{1}{n} \sum_{i=1}^n \frac{R_i}{n+1} \frac{S_i}{n+1}$$

$$B_n = \frac{1}{n} \sum_{i=1}^n \left(\frac{R_i}{n+1}\right)^2 \left(\frac{S_i}{n+1}\right)^2$$

$$C_n = \frac{1}{n^3} \sum_{i=1}^n \sum_{j=1}^n \sum_{k=1}^n \frac{R_i}{n+1} \frac{S_i}{n+1} 1_{\{R_k \leq R_i, S_k \leq S_i\}} + \frac{1}{4} - A_n$$

$$D_i = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \frac{R_i}{n+1} \frac{S_i}{n+1} \max\left(\frac{R_i}{n+1}, \frac{R_j}{n+1}\right)$$

$$\epsilon_n = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \frac{R_i}{n+1} \frac{S_i}{n+1} \max\left(\frac{S_i}{n+1}, \frac{S_j}{n+1}\right)$$

Exemple 2.1.12. Pour la copule de Gumbel de paramètre θ on a

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

Nous en déduisons que

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

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

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

Dans le cas général, on a l'estimateur non paramétrique du tau de Kendall est donné par

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

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

C- Estimation non paramétrique

Si le modèle paramétrique de la copule est mal spécifié, l'approche paramétrique et semi-paramétrique conduisent généralement à un mauvais ajustement aux données. Dans ce cas les méthodes d'estimation non paramétriques peuvent être une bonne alternative.

1. La fonction empirique :

Le cadre général de l'estimation non paramétrique d'une loi marginale s'appuie sur la fonction de répartition empirique, définie par

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

pour (x_1, \dots, x_n) un échantillon de taille n de la loi F .

Si on se donne $(x_1^i, \dots, x_d^i)_{1 \leq i \leq n}$ un n -échantillon du vecteur X (de dimension d), on peut généraliser l'expression de la fonction de répartition empirique en posant

$$F_n(x_1, \dots, x_d) = \frac{1}{n} \sum_{i=1}^n 1_{\{x_{i1} \leq x_1, \dots, x_{id} \leq x_d\}}$$

Cet estimateur conduit à un estimateur non paramétrique naturel d'une copule.

Propriétés asymptotiques

- F_n est un estimateur sans biais

$$\mathbb{E}(F_n(x)) = F(x).$$

- Par la loi forte des grands nombres

$$F_n(x) \xrightarrow[p.s.]{} F(x)$$

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

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

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

$$\|F_n - F\|_\infty \xrightarrow[p.s.]{} 0$$

2. Copule empirique :

Soit comme ci-dessus $(x_1^i, \dots, x_d^i)_{1 \leq i \leq n}$ un n-échantillon du vecteur X (de dimension d). En observant que : $C(u_1, \dots, u_d) = F(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d))$, on introduit la notion de copule empirique à partir des versions empiriques des fonctions de répartition de cette expression

$$C_n(u_1, \dots, u_d) = F_n(F_{1,n}^{-1}(u_1), \dots, F_{d,n}^{-1}(u_d)),$$

avec $F_{j,n}^{-1}(u) = \inf\{t/F_{j,n}(t) \geq u\}$, $j = 1, d$. Cette notion a été introduite par Deheuvels [1979].

Remarque 2.1.13. C est la fonction de répartition d'un vecteur dont les marginales sont uniformes sur $[0, 1]$; cette propriété n'est plus vraie pour C_n qui est conditionnellement aux observations, la distribution d'un vecteur dont les marges sont réparties uniformément sur l'ensemble discret $\{0, \frac{1}{n}, \dots, \frac{n-1}{n}\}$. Il suffit donc de connaître les valeurs de la copule empirique en ces points discrets; on a en particulier le lien avec les statistiques d'ordre suivant

$$C_n\left(\frac{t_1}{n}, \dots, \frac{t_d}{n}\right) = \frac{1}{n} \sum_{t=1}^n 1_{\{x_1^t \leq x_1^{t_1}, \dots, x_d^t \leq x_d^{t_d}\}}$$

Cette expression peut être également écrite à partir des rangs des observations (ce qui est intuitif, puisque la copule est invariante par toute transformation croissante des marginales)

$$C_n\left(\frac{t_1}{n}, \dots, \frac{t_d}{n}\right) = \frac{1}{n} \sum_{t=1}^n 1_{\{r_1^t \leq t_1, \dots, r_d^t \leq t_d\}}$$

Deheuvels [1979] a obtenu des résultats asymptotiques sur C_n et $n^{-\frac{1}{2}}(C_n - C)$. Les copules empiriques sont notamment utiles pour fournir des estimateurs non paramétriques de mesures de dépendance. Par exemple on peut ainsi proposer l'estimateur suivant du ρ de Spearman

$$\hat{\rho} = \frac{12}{n^2 - 1} \sum_{t_1=1}^n \sum_{t_2=1}^n \left(C_n\left(\frac{t_1}{n}, \frac{t_2}{n}\right) - \frac{t_1 t_2}{n^2} \right)$$

3. Estimation de la copule C par la méthode à noyau :

Malgré les qualités de C_n , cet estimateur présente des discontinuités. Dans cette partie, nous présentons les estimateurs à noyau les plus utilisés en pratique et qui résolvent ce problème. Fermanian, Radulovic et Wegkamp [2004] ont proposé une version lisse notée SE (Smooth version of empirical copula) pour la copule empirique, à savoir

$$\hat{C}_{SE}(u, v) = \hat{H}_n(\hat{F}_n^{-1}(u), \hat{G}_n^{-1}(v)),$$

où les quantités $\hat{H}_n, \hat{F}_n, \hat{G}_n$ sont données par

$$\hat{H}_n(x, y) = \frac{1}{n} \sum_{i=1}^n K_n(x - X_i, y - Y_i),$$

$$\hat{F}_n(x) = \hat{H}_n(x, +\infty), \hat{G}_n(y) = \hat{H}_n(+\infty, y),$$

avec

$$K_n(x, y) = K\left(\frac{x}{b_n}, \frac{y}{b_n}\right), \quad K(x, y) = \int_{-\infty}^x \int_{-\infty}^y k(s, t) ds dt$$

où $k(s, t)$ est la fonction densité du noyau bidimensionnelle, et b_n est un paramètre de lissage qui tend vers zéro quand n converge vers l'infini.

Cependant cet estimateur souffre du problème du biais, ce biais typique associé à l'estimation à noyaux est présent puisque le support d'une copule est borné sur $[0, 1]^2$. L'estimateur \hat{C}_{SE} n'est pas convergent. Pour trouver solution à ce problème, plusieurs méthodes sont proposées. La

première méthode est basée sur la transformation des données proposée par Gijbels, Omelka et Veraverbeke [2009]. La seconde méthode est l'estimateur de Chen et Huang [2007], qui ont proposé d'utiliser un noyau linéaire local afin de faire face aux biais près des frontières du carré unité.

2.2 Les tests

Diverses familles de copules peuvent représenter des structures de dépendance très différentes. Lorsqu'il s'agit de modéliser un jeu de données réel, certaines classes paramétriques peuvent se révéler inadaptées, i.e, incapables de capter la spécificité de corrélation entre les variables. Les tests d'adéquation permettent de valider un modèle de copule adéquat et d'éliminer ceux qui ne correspondent pas à la dépendance étudiée.

2.2.1 Test d'adéquation et choix de la copule

Soit un échantillon de vecteurs aléatoires iid $x_i \in \mathbb{R}$ avec $X_i = (X_{1,i}, \dots, X_{d,i}) \sim X$ iid. Soit H la fonction de répartition de X et soit C la copule de X donc

$$H(x_1, \dots, x_d) = C(F_1(x_1), \dots, F_d(x_d))$$

Généralement, un test statistique distingue entre deux hypothèses

- Hypothèse nulle $H = H_0$ ou bien $C = C_0$
- Hypothèse alternative $H \neq H_0$ ou bien $C \neq C_0$

Une première solution peu puissante consiste à comparer graphiquement la copule empirique définie par

$$\hat{C}\left(\frac{k_1}{n}, \dots, \frac{k_p}{n}\right) = \frac{1}{n} \text{card}\{i \mid R_{1,i} \leq k_1, \dots, R_{p,i} \leq k_p\}$$

et la copule paramétrique C , avec $R_{j,i}, j = 1, \dots, p$ et $i = 1, \dots, n$ dénote le rang de $X_{j,i}$ parmi les observations $X_{j,1}, \dots, X_{j,n}$. La dépendance retenue est donc celle qui assure que C soit la

plus proche possible de \hat{C} .

2.2.2 Sélection d'une copule à partir du test d'adéquation du chi-deux

La sélection de la meilleure copule doit s'appuyer sur un test statistique. Le test du chi-deux utilisé dans le cadre de l'ajustement d'une distribution paramétrique à une distribution empirique constitue à cet égard un outil intéressant pour le choix d'une copule. Habituellement mis en oeuvre pour ajuster une distribution univariée, il offre l'avantage de facilement s'adapter au cas des distributions multivariées. Il est vrai qu'il nécessite de procéder à un découpage en classes qui réduit sa puissance et que la manière de choisir ces classes peut être critiquée. Néanmoins, sa facilité d'utilisation en fait un outil de décision très intéressant pour sélectionner la meilleure copule.

Nous proposons ci-après une méthode statistique de sélection de copule basée sur un ajustement bivarié entre la copule paramétrique et la copule empirique effectué à partir de la statistique du chi-deux. Notons que l'extension de cette méthodologie au cas trivarié est immédiate en veillant à bien adapter le calcul des probabilités trivariées. Notre approche s'inspire d'un article de Hurlimann [2001a] dans lequel le test du chi-deux est utilisé dans une version modifiée de la méthode IFM. Le processus de sélection d'une copule est décrit ci-après.

- **Estimation du paramètre de la copule :**

On estime le paramètre de la copule bivariée à partir de la méthode CML afin de s'affranchir d'une éventuelle erreur de spécification des lois marginales. On note $C_{\hat{\theta}}(u_1, u_2)$ la copule paramétrique obtenue.

- **Calcul de la copule empirique :**

On définit un treillis d'ordre T adapté au nombre d'observations et aux contraintes de temps de calcul puis on calcule la copule empirique bivariée sur ce treillis

$$\hat{C}\left(\frac{t_1}{T}, \frac{t_2}{T}\right) = \frac{1}{T} \sum_{i=1}^T \prod_{k=1}^2 1_{\{r_k^i \leq t_k\}}$$

- **Choix d'un nombre d'intervalles pour découper les uniformes univariées :**

On se donne un nombre k afin de découper chacune des uniformes en k intervalles et de construire un tableau de contingence bivarié des effectifs. k et T doivent être choisis de concert afin d'obtenir aisément les effectifs empiriques à partir de la copule empirique. Sans a priori sur la distribution empirique bivariée, on découpe chaque uniforme sur des intervalles identiques. On note $(b_0, b_1], (b_1, b_2], \dots, (b_{k-1}, b_k]$ et $(c_0, c_1], (c_1, c_2], \dots, (c_{k-1}, c_k]$ les k intervalles pour découper respectivement u_1 et u_2 .

- **Calcul des effectifs et construction du tableau de contingence bivarié :**

L'objet de cette étape est de calculer les effectifs empiriques et théoriques sur chacun des k^2 intervalles bidimensionnels $(b_{i-1}, b_i] \times (c_{j-1}, c_j]$, $i, j = 1, \dots, k$.

Pour ce faire, il convient de calculer les probabilités bivariées $p_{i,j}$ d'appartenance à un intervalle. La formule de calcul est la suivante

$$p_{i,j} = C(b_i, c_j) - C(b_{i-1}, c_j) - C(b_i, c_{j-1}) + C(b_{i-1}, c_{j-1})$$

Où

$$i, j = 1, \dots, k \text{ et } C(x, y) = \hat{C}(x, y) \text{ ou } C_{\hat{\theta}}(x, y)$$

Les effectifs $f_{i,j}$ sont calculés en multipliant les probabilités bivariées par le nombre total d'observations. On obtient alors un tableau de contingence bivarié donnant les effectifs empiriques issus de la copule empirique et les effectifs théoriques issus de la copule paramétrique. Il a la forme suivante

$u_1 \backslash u_2$	$(c_0, c_1]$	$(c_1, c_2]$...	$(c_{k-1}, c_k]$
$(b_0, b_1]$	$f_{1,1}$	$f_{1,2}$...	$f_{1,k}$
$(b_1, b_2]$	$f_{2,1}$	$f_{2,2}$...	$f_{2,k}$
...
$(b_{k-1}, b_k]$	$f_{k,1}$	$f_{k,2}$...	$f_{k,k}$

- **Application du critère de Cochran et regroupement en classes :**

Dans la pratique, les effectifs théoriques de certaines cases peuvent être très faibles voir proches de zéro. Il faut alors procéder à un regroupement des k^2 intervalles initiaux en n classes permettant de respecter le critère de Cochran qui recommande d'avoir des effectifs théoriques au moins supérieurs à 1 du nombre total d'observations dans chaque classe, et supérieurs à 5 du nombre total d'observations dans au moins 80 des classes. Ce regroupement est aussi à effectuer sur le tableau bivarié des effectifs empiriques. Selon les zones de dépendance recherchées, on pourra notamment isoler les cellules des queues de distribution ou retenir des directions particulières comme la diagonale principale par exemple.

- **Calcul de la statistique bivariée du chi-deux et sélection de la copule :**

Après avoir défini le regroupement en n classes vérifiant le critère de Cochran, on calcule la statistique bivariée du chi-deux observée comme suit

$$\chi_{obs}^2 = \sum_{j=1}^n \frac{(O_j - E_j)^2}{E_j}$$

Où O_j et E_j représentent respectivement les effectifs observés et attendus dans chaque classe j , $j = 1, \dots, n$. Cette statistique suit une loi du chi-deux à $(n - r - 1)$ degrés de liberté avec r le nombre de paramètres estimés de la copule. On se donne ensuite un seuil critique α pour la zone de rejet et on calcule la p-valeur $\hat{\alpha} = \mathbb{P}(\chi_{n-r-1}^2 > \chi_{obs}^2)$.

On rejette l'adéquation de la copule paramétrique si $\hat{\alpha} < \alpha$. Il faut préciser pour les

praticiens que la p-valeur exacte $\hat{\alpha}$ se situe en fait dans l'intervalle suivant

$$\mathbb{P}(\chi_{n-1}^2 > \chi_{obs}^2) \leq \hat{\alpha} \leq \mathbb{P}(\chi_{n-r-1}^2 > \chi_{obs}^2).$$

Cet encadrement étant de faible amplitude lorsque r est petit vis à vis de n , le seuil critique α sera le plus souvent en dehors de cet encadrement et on conclura sans ambiguïté .

Parmi les copules acceptées, on sélectionne alors celle qui maximise la p-valeur.

2.2.3 Test de Cramér-Von Mises

On suppose qu'on cherche à tester $C \in \mathcal{C}$ (l'hypothèse nulle), où C est une famille de copules (l'hypothèse alternative étant $C \notin \mathcal{C}$). Si la famille \mathcal{C} est une famille théorique, Fermanian ou Genest et Remillard ont suggéré d'utiliser le test de Cramér-von Mises, avec comme statistique

$$T = n \int \int_{I^2} \left(\hat{C}(u, v) - C_{\hat{\theta}}(u, v) \right)^2 d\hat{C}(u, v)$$

Une approche un peu plus simple est d'utiliser la fonction de Kendall définie par : $K(t) = \mathbb{P}(C(u, v) \leq t)$. Si on suppose que cette fonction appartient à une famille paramétrique, alors la statistique précédente peut s'écrire

$$T = n \int \int_{I^2} \left(\hat{K}(t) - K_{\hat{\theta}}(t) \right)^2 dK_{\hat{\theta}}(t)$$

Comme suggéré par Genest et Rivest. On rejette l'hypothèse nulle si la valeur de ce indicateur T est supérieure au seuil critique α .

Applications

3.1 Exemple d'utilisation du modèle copule en physique

Cette application est inspiré d'un travail de recherche (2016) de Mr Biljana Popović, intitulé "AN EXAMPLE OF THE APPLICATION OF COPULA MODELS IN PHYSICS".

Les modèles de copule ne sont pas encore courants en physique. Le but principal de cet article est de montrer comment les copules peuvent être utiles dans la modélisation des distributions concernant la ventilation des gaz.

Dans de nombreuses situations réelles, nous nous intéressons à la valeur d'une quantité physique difficile ou impossible à mesurer directement.

La décomposition électrique des gaz revêt une grande importance en raison des nombreuses applications des gaz en électronique et en technologie. Le temps qui s'écoule à partir du moment de l'application de la tension supérieure à la tension de claquage statique t_s à la claquage est le délai de claquage t_d . Il se compose du délai statistique et du délai formatif. Ainsi, le délai de décomposition peut être considéré comme la somme du délai statistique t_s et du délai de formation t_f , c'est-à-dire

$$t_d = t_s + t_f$$

Le délai de décomposition, ainsi que le délai statistique et formatif sont des variables stochastiques et une approche statistique peut être très utile. Pendant ce temps, les mesures simultanées directes du temps statistique et du temps de formation ne sont pas possibles. Les scientifiques ont recours à des méthodes statistiques pour les séparer théoriquement. Une suggestion est de choisir la valeur minimale mesurée du délai dans chaque série de mesures afin de jouer le rôle de délai de formation.

Certains auteurs développent la théorie de la décomposition en considérant le délai statistique et le délai formatif comme deux variables aléatoires indépendantes. Mais ils sont dépendants positivement sauf lorsque le taux de production d'électrons est très faible (temps de relaxation élevé).

Nous devrions modéliser la distribution générale de t_d à partir des distributions marginales de t_s et de t_f et de t_s .

une solution possible du problème

Soient F_{t_d}, F_{t_s} et F_{t_f} fonctions des répartitions et, f_{t_s} et f_{t_f} fonctions de densités des variables aléatoires t_d, t_s et t_f , concernant la ventilation, respectivement. Alors que la variable aléatoire t_f a fondamentalement, une distribution normale (gaussienne) pour un mécanisme déclencheur, la distribution de la variable t_s et peut être exprimée sous la forme d'un mélange convexe de deux distributions, normale et exponentielle. La distribution de t_s se change en fonction du temps de relaxation κ

$$f_{t_s}(X) = a_G \frac{1}{\sigma\sqrt{2\pi}} \exp \left[-\frac{x-\mu}{2\sigma^2} \right] + a_E \frac{1}{\lambda} \exp \left[-\frac{x}{\lambda} \right], \quad x \in \mathbb{R} \quad (3.1)$$

où, $a_E \geq 0$, et $a_G + a_E = 1$. Remarquons que lorsque le problème traité est réel, nous ne considérons que la région du temps réel et positif, c'est-à-dire, $x > 0$. Bien sûr, entre autres, le délai de claquage dépend du type et de la pression du gaz, du matériau et de la géométrie de l'électrode, de l'éclairage, de la présence d'impuretés, de la tension de travail et du temps

de relaxation, mais entre autres conditions identiques, il peut être traité comme fonction du temps de relaxation κ (seulement). Pour de petites valeurs de temps de relaxation κ , il s'agit de Gauss, $a_E = 0$, puis de mélange de distributions Gauss – exponentiel, $a_G > 0$ et $a_E > 0$ et enfin, pour un nombre de fois suffisamment élevé κ , il devient exponentiel, $a_G = 0$.

Cela veut dire que

$$t_s : \begin{cases} \mathcal{N}(\mu_{t_s}, \sigma_{t_s}^2) & \kappa \text{ petit} \\ a_E \varepsilon(\lambda) + a_G \mathcal{N}(\mu, \sigma^2) & \kappa \text{ moderate} \\ \varepsilon(\lambda_{t_s}) & \kappa \text{ grand} \end{cases}$$

et

$$t_f : \mathcal{N}(\mu_{t_f}, \sigma_{t_f}^2)$$

Alors, comment déterminer la distribution de la somme de t_s et de t_f ? La question de la détermination du pdf d'une somme de deux variables aléatoires dépendantes a été largement discutée dans la littérature et a été soulevée à l'origine par A.N. Kolmogorov. Le problème du calcul df d'une combinaison linéaire d'un certain nombre n de df-s connues. La conclusion était que le problème pose des difficultés de calcul. il est connu que pour, la somme de seulement deux variables aléatoires dépendantes, les bornes exactes et optimales pour F_{t_d} sont :

$$\underline{F}_{t_d} = \max_{x,y,x+y=z} \max(F_{t_s}(x) + F_{t_f}(y) - 1, 0)$$

$$\overline{F}_{t_d}(z) = \min_{x,y,x+y=z} \min(F_{t_s}(x) + F_{t_f}(y), 1)$$

où \underline{F}_{t_d} et \overline{F}_{t_d} sont la limite inférieure et supérieure pour F_{t_d} respectivement. Par ailleurs, dans de nombreuses applications, il est nécessaire de connaître et d'utiliser une estimation

"ponctuelle" de F_{t_d} c'est-à-dire de f_{t_d} .

Les copules fournissent un cadre naturel pour l'étude de ce problème. En particulier, dans le modèle Copula, nous pouvons intégrer différentes familles de distributions de probabilité, ce qui est important dans notre cas. Quand les marges sont fixes les copules sont alors impliqués :

$$F_{t_d}(z) = \int \int_{x+y \leq z} dC(F_{t_s}(x), F_{t_f}(y))$$

où C est la copule, étant donné que F_{t_f} et F_{t_s} . Ce travail résulte d'un intérêt de longue date pour le problème de la détermination théorique de la distribution du temps de retard de décomposition dans les gaz. L'article traite des copules comme la seule solution possible à ce problème à l'heure actuelle. Pendant ce temps, une des difficultés est, en général, le choix de la copule. Il n'existe pas de méthode systématique rigoureuse pour le choix de la copule. Rien ne peut nous dire que la copule choisie, ou même la famille de copules, convergera vers la dépendance réelle à la structure des données soulignées. Notre problème sera toujours de choisir la copule qui correspond le mieux aux données.

3.2 Application en finance

Dans l'analyse des risques des marchés financiers, l'un des problèmes majeurs est la quantification de ce risque, les méthodes probabilistes et statistiques sont actuellement au cœur de cet axe de recherche, en mettant en place des indicateurs d'exposition au risque.

3.2.1 Notions de base de risque

Définition du risque

Dans la littérature et au sein de l'industrie, le risque est la valeur potentielle d'une conséquence négative non désirée d'un événement ou d'une activité. Cependant, le risque est associé à une situation de faible probabilité ; si celle-ci était élevée, on utiliserait, plus justement le terme "danger". Ainsi, le risque est considéré comme une contingence négative et peu probable.

Facteur de risque

Un facteur de risque est un élément qui déclenche une perte en constituant l'origine d'un risque ou d'un ensemble de risques. On retrouve parmi les exemples classiques de facteurs de risque, la complexité des applications, la taille d'un projet donné, les nouveautés techniques, le manque d'expérience etc.

3.2.2 Mesure de risque

Les mesures de risques sont des outils de quantification de risque. Elle permettent d'évaluer un niveau de dangerosité d'un risque mais également de comparer les différents risques et de les classer selon le niveau de dangerosité.

Définition 3.2.1. *Soit un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$*

Une mesure de risque est une application

$$\rho : \mathcal{F} \longrightarrow \mathbb{R}$$

$$X \longmapsto \rho(X)$$

censée à quantifier le risque porté par la variable aléatoire X définie sur $(\Omega, \mathcal{F}, \mathbb{P})$.

Définition 3.2.2. On dit que ρ est une mesure de risque monétaire si :

- ρ est croissante : $X \geq Y \implies \rho(X) \geq \rho(Y)$
- invariance par translation : $\rho(X + \alpha) = \rho(X) + \alpha \quad \forall X, \forall \alpha$

D'après cette définition, nous pouvons dire que l'espérance, la variance ou l'écart-type, lorsqu'ils existent, sont des mesures de risque. Si beaucoup d'applications répondent à cette définition, pour être jugée "satisfaisante", une mesure de risque doit pouvoir vérifier un certain nombre de propriétés élémentaires.

Définition 3.2.3. Une mesure de risque ρ est dite cohérente si, pour deux variables aléatoires X et Y , elle satisfait les propriétés suivantes :

1. **Invariance par translation :**

$$\forall X, \forall \alpha, \rho(X + \alpha) = \rho(X) + \alpha$$

2. **La sous additivité**

$$\rho(X + Y) \leq \rho(X) + \rho(Y)$$

3. **Homogénéité positive**

$$\rho(\lambda X) = \lambda \rho(X)$$

4. **Monotonie**

$$X \geq Y \implies \rho(X) \geq \rho(Y)$$

3.2.3 Valeur-à-Risque

Définition de la VaR

La VaR est un outil très répandu dans les marchés financiers dû à sa promesse implicite d'améliorer la gestion des risques en offrant une mesure complète des risques. La valeur-à-risque a comme fonction de donner une information de synthèse sur le risque d'un titre

financier ou d'un portefeuille de titres financiers. Elle se définit comme la perte maximale espérée à l'intérieur d'un horizon temporel étant donné un niveau de confiance.

Formulation mathématique de la VaR

soit X une variable aléatoire dont la fonction de répartition F_X d'une position financière risquée, sur un horizon de temps fixé τ . Pour un niveau $\alpha \in [0, 1]$,

$$VaR_\alpha(X) = \inf\{x, F_X \geq \alpha\}$$

ce n'est rien d'autre que le quantile d'ordre α de la perte potentielle pour un horizon donnée.

Si F est strictement croissante, alors $F^{-1} = F^{(-)}$, c'est-à-dire le quasi-inverse est égale à l'inverse usuel,

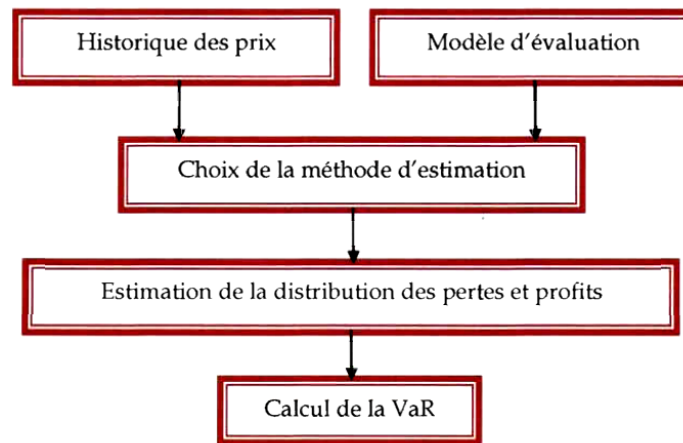
$$VaR_\alpha(X) = F_X^{-1}(\alpha)$$

L'avantage de la VaR est d'être un concept simple et facile à calculer. Elle vérifie les propriétés d'invariance par translation, d'homogénéité positive, de monotonie.

En revanche, elle n'est pas sous-additive et ne tient pas compte de la sévérité de la ruine, ce qui constitue la critique la plus souvent formulée à son égard. Cependant cette mesure de risque peut faire sens dans un objectif de solvabilité.

Méthodes d'estimation de la VaR

L'étape critique dans le calcul de la Valeur-à-Risque d'un portefeuille, est l'estimation de la distribution des pertes et profits de ce dernier. La problématique générale de l'estimation de la distribution des changements de la valeur du portefeuille et de la VaR peut se résumer par le schéma de la figure suivante :



Étapes de calcul de la VaR :

* Historique des prix :

L'historique des prix représente les observations des différents actifs pour un certain nombre de périodes passées. Cette étape est fondamentale pour estimer les variations potentielles de la valeur du portefeuille dans le futur proche.

* Modèle d'évaluation :

Le risque de marché est le risque auquel est confronté l'investisseur en raison de sa méconnaissance de l'évolution future des variables fondamentales du marché que sont les cours des titres, les taux d'intérêts, les taux de change,... Ces variables, également appelées facteurs de risque, déterminent le prix des titres, des actifs conditionnels, des portefeuilles,... Si on désigne par P_p le prix d'un portefeuille et par X_1, X_2, \dots, X_n les facteurs de risque qui expliquent le prix, on a la relation générale suivante

$$P_p = f(X_1, X_2, \dots, X_n) + \varepsilon$$

où,

- f est une fonction qui décrit le modèle d'évaluation,
- ε est le résidu représentant l'écart existant entre le prix effectif P_p et la fonction d'évaluation f .

Ainsi, la première étape consiste à choisir le modèle d'évaluation adéquat f , qui pourrait être linéaire comme des actions et non-linéaire comme des obligations et des options.

* **Choix de la méthode d'estimation :**

Dans la littérature, il existe plusieurs méthodes pour calculer la VaR. Parmi ces méthodes la méthode de la simulation historique

* **Estimation de la distribution des pertes et profits et de la VaR :**

Une fois la série historique des prix est obtenue, le modèle d'évaluation est spécifié et le choix de la méthode d'estimation est déterminé, nous passons à l'étape qui consiste à estimer la distribution des pertes et profits laquelle nous servira à la déduction de la valeur de la VaR d'un portefeuille.

3.2.4 La gestion de risque en utilisant la copule

Dans cette section on va présenter les principales étapes pour estimer la VaR pour un portefeuille linéaire en utilisant la copule.

Calcul de la VaR par la copule

Soit $W = (w_1, \dots, w_n) \in \mathbb{R}^n$ un portefeuille de position qui se compose de n actifs (l'attribution des actifs) et $S_t = (S_{1,t}, \dots, S_{n,t})^T$ le vecteur aléatoire non négatif représentant les prix des actifs à l'instant t . La valeur V_t du portefeuille est donné par

$$V_t = \sum_{i=1}^n w_i S_{i,t}$$

Les prix des actifs dans l'économétrie financière sont considérés comme des variables aléatoires. La valeur d'un portefeuille V_t au temps t est également une variable aléatoire.

Soit une période du temps τ , le profit dans l'intervalle $[t, t + \tau]$ est donnée par $(V_{t+\tau} - V_t)$. Cette différence est dite la distribution de profit perte qui exprime le changement dans la valeur du portefeuille entre t et $t + \tau$. Soit $X_{t+\tau}$ la fonction log rendement (log retour) entre t et $t + \tau$

$$X_{t+\tau} = \ln \frac{S_{t+\tau}}{S_t}$$

La fonction de profit perte $K_{t+\tau} = V_{t+\tau} - V_t$ peut s'écrire en fonction des log retours comme suit

$$\begin{aligned} V_{t+\tau} - V_t &= \sum_{i=1}^n w_i S_{i,t+\tau} - \sum_{i=1}^n w_i S_{i,t} \\ &= \sum_{i=1}^n w_i (S_{i,t+\tau} - S_{i,t}) \\ &= \sum_{i=1}^n w_i S_{i,t} \left(\frac{S_{i,t+\tau}}{S_{i,t}} - 1 \right) \\ &= \sum_{i=1}^n w_i S_{i,t} (\exp\{X_{i,t+\tau}\} - 1), \\ K_{t+\tau} &= \sum_{i=1}^n w_i S_{i,t} (\exp\{X_{i,t+\tau}\} - 1) \end{aligned} \tag{3.2}$$

De plus, si la fonction de distribution de K est donnée par $F_K(x) = \mathbb{P}[K \leq x]$, la VaR du portefeuille au niveau de probabilité α est définie comme le quantile d'ordre de la fonction F_K et sa fonction est donnée par

$$VaR_\alpha(K) = F_K^{-1}(\alpha)$$

On remarque que la fonction de distribution unidimensionnelle F_K dépend de la fonction de distribution n-dimensionnelle des log retours F_X . En utilisant la copule, les fonctions de distributions marginales F_{X_j} peuvent être modélisées séparément de leur structure de la dépendance et être alors couplées ensemble pour former la distribution multivariée F_X .

Remarque 3.2.4. *Le rendement logarithmique est une approximation du rendement simple, en effet*

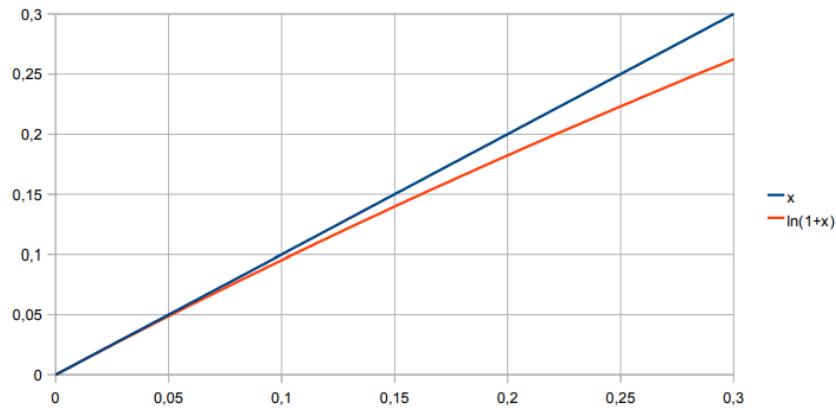


FIGURE 3.1 – Courbes des fonctions $x \mapsto \ln(x + 1)$ et $x \mapsto x$.

Le taux de rentabilité "simple" ou "arithmétique" est défini par

$$R_\alpha = \frac{S_{t+\tau} - S_t}{S_t}$$

Le taux de rentabilité "logarithmique" est défini par

$$R_l = \ln \frac{S_{t+\tau}}{S_t}$$

R_l est une approximation de R_α

$$R_l = \dots = \ln(1 + R_\alpha) \approx R_\alpha$$

mais comme $x \geq \ln(1 + x)$, $R_l \leq R_\alpha$.

Estimation de la VaR par la copule

Les fonctions de distribution marginales des log retours F_{X_j} ont une fonction de distribution conjointe F_X et qui peut être écrite comme fonction copule

$$F_X(x_1, \dots, x_n) = C_\theta(F_{X_1}(x_1), \dots, F_{X_n}(x_n))$$

où C_θ est une copule qui appartient à la famille paramétrique $C = C_\theta, \theta \in \Theta$ Pour un porte-

feuille qui se compose de n actifs et un échantillon $\{x_{j,t}\}_{t=1}^T$, avec $j = 1, \dots, n$ des log retours, la VaR est estimée en suivant les étapes :

1. **Spécification et estimation des distributions marginales** F_{X_j} ;
2. **Spécification de la copule paramétrique** $C(u_1, \dots, u_d; \theta)$; Dans nos essais numériques, nous emploierons quelques copules étudiées dans le chapitre 1 pour le cas bidimensionnel (Nelsen (2006)).
3. **Estimation de la copule paramétrique** C_θ **pour obtenir le paramètre de la dépendance** $\hat{\theta}$;
4. **Génération d'un échantillon de la perte** K ;
5. **Estimation de la** $\widehat{Var}_\alpha(K)$ **par le quantile d'ordre** α **de la fonction** K ;

Application numérique

Considérons un investissement d'une période dans deux actifs nommés actif 1 et actif 2. Soit w_1 et w_2 les parts de richesse investies dans l'actif 1 et l'actif 2, respectivement, telles que $w_1 + w_2 = 1$. Soit X et Y les log retours composés d'une période, définis par :

$$X = \ln\left(\frac{S_{t+1}^1}{S_t^1}\right) \quad \text{et} \quad Y = \ln\left(\frac{S_{t+1}^2}{S_t^2}\right)$$

où S_t^1 et S_t^2 dénotent les prix des actifs 1 et 2 à la période t , respectivement. Ensuite, le log retour d'une période sur le portefeuille est donné par

$$L = \ln(w_1 e^X + w_2 e^Y)$$

Puisque L est une transformation non linéaire des variables aléatoires X et Y , la fonction de distribution de L est généralement inconnu, alors l'expression analytique de la VaR n'est pas disponible.

Donc si la fonction de distribution conjointe de X et Y est spécifié par les fonctions de

distribution marginales F_X et F_Y et une copule paramétrique C_θ , il est possible de générer des simulations aléatoires de la fonction de distribution conjointe de X et Y pour calculer des approximations numériques de la VaR d'une probabilité de perte donnée $1-\alpha$. Soit l_1, \dots, l_n indique n valeurs simulées de L basées sur n paires aléatoires simulées de X et Y . Ensuite, l'estimation basée sur la simulation est :

$$\widehat{Var}_\alpha(L) = \text{Le quantile d'ordre } \alpha \text{ de valeurs } -l_i$$

Notons que Le traitement des données sera effectué informatiquement par logiciel de statistique R, et pour l'estimation du paramètre de la copule sera effectué par La méthode de maximum de vraisemblance canonique (CML).

Exemple : Soient X et Y les log retours journalières de deux actifs d'un investisseur. Considérons un portefeuille linéaire avec $w_1 = 0.7$ et $w_2 = 0.3$. supposons $X \sim \mathcal{N}(0.002, 1.15^2)$ et $Y \sim \mathcal{N}(0.007, 1.35^2)$.

La première partie du programme nous permet de simuler et Dessiner le nuage des points des log retours X et Y , puis calculer les mesures de dépendance

```
> X <- rnorm(1000, mean = 0.002, sd = 1.15)
> Y <- rnorm(1000, mean = 0.007, sd = 1.35)
> r = cor(X, Y, method = 'spearman')
> r
-0.035
> rho = cor(X, Y, method = 'pearson')
> rho
-0.0485
> tau = cor(X, Y, method = 'kendall')
> tau
-0.032
```

La deuxième partie du programme nous permet de valider un modèle de copule adéquat et d'éliminer ceux qui ne correspondent pas à la dépendance étudiée à partir d'un test d'adéquation puis estimer le paramètre θ par la méthode CML

```
>u <- pobs(as.matrix(cbind(X,Y)))[,1]
>v <- pobs(as.matrix(cbind(X,Y)))[,2]
>selectedCopula <- BiCopSelect(u,v,familyset=NA)
>selectedCopula
```

Le traitement de ce programme donne le résultat suivant :

Bivariate copula : Frank(par = -0.29, tau = -0.032)

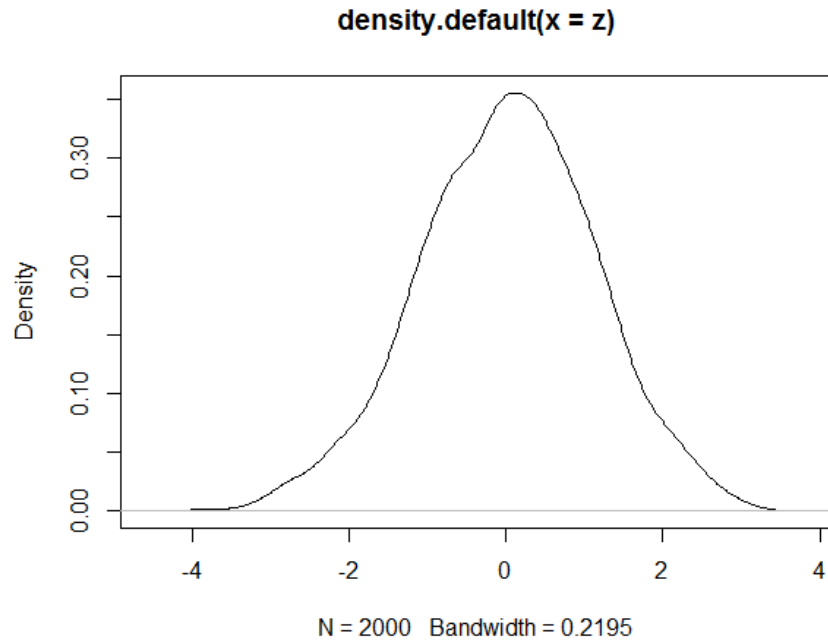
cela signifie que $\hat{\theta} = -0.29$

La troisième partie du programme nous permet d'estimer la VaR avec un niveau de confiance $\alpha = 0.95$ basé sur un échantillon de taille $n = 1000$ valeurs simulées du log retour du portefeuille $L = \ln(w_1 e^X + w_2 e^Y)$.

```
Theta = -0.29 >copule.Theta <- mvds(archmCopula(family="Frank",param=Theta),
+c("norm","norm"),
+list(list(mean= 0.0002,sd=1.15),list(mean=0.007,sd=1.35)))
>set.seed(895)
>XY.sim <- rMvdc(1000,Copule.Theta)
> R.sim <- -log(0.7 * exp(XY.sim[, 1]) + 0.3 * exp(XY.sim[, 2]))
>var.95.est = -quantile(-R.sim,probs=0.95)
```

Le traitement de la troisième partie de ce programme donne les résultats suivants :

```
>VaR.95.est
95%
1.8556
```



La valeur à risque (VaR) est une quantile de la distribution de pertes et profits associée à la détention d'un actif d'un portefeuille d'actif sur une période données.

D'après ces résultats la VaR au seuil de confiance 95% à un jour noté VaR (95%, 1jour) égale à 1.856 million dinars signifie qu'il y a 95% que la profit associe de la détention de portefeuille n'exède pas 1.856 million dinars.

Cet investisseur à une grand chance pour profiter.

Conclusion générale

La modélisation par les copules est une méthode plus récente réaliste, elle permet de modéliser la dépendance entre les distributions unidimensionnelles, les indicateurs comme alternative du coefficient de corrélation linéaire (Pearson) et le tau de Kendall qui est un coefficient de corrélation non-paramétrique, ainsi que le coefficient de dépendance de queue qui est une mesure asymptotique de dépendance.

Annexe A : Notion de base de finance

Analyse de risque

L'analyse du risque consiste à expliquer le phénomène du risque pour mieux définir des avenues de solution. Elle comporte une mise en contexte, un énoncé qui met en relation le risque et les facteurs de risque, des caractéristiques du risque, etc. Généralement, on peut retrouver des exemples d'analyse du risque dans la Banque des risques.

Gestion des risques

La gestion des risques est un processus qui comprend plusieurs étapes bien définies favorisant alors une meilleure prise de décision et ce, du fait qu'elle fournit une meilleure information sur les risques et leurs impacts. La gestion des risques est centrée sur l'identification et le traitement des facteurs de risque. De plus, elle a pour objectif d'ajouter le maximum de valeur durable à chaque activité de l'organisation en mobilisant la compréhension des aléas positifs ou négatifs qui dérivent de tous les facteurs pouvant affecter l'organisation. Ladite gestion augmente la probabilité de succès et réduit donc la probabilité d'échec et l'incertitude qui s'y attache. La gestion des risques est devenue en particulier, l'une des grandes responsabilités d'un gestionnaire d'entreprise et c'est pour cette raison, une attention particulière est portée sur la modélisation des fluctuations du marché sous un angle statistique.

Soit X_1, X_2, \dots, X_n des variables aléatoires, où X_i peut représenter, à titre d'exemple un actif, une action, une valeur mobilière, un rendement d'un placement détenu par un particulier ou une institution. En finance, le vecteur (X_1, X_2, \dots, X_n) représente alors, un portefeuille.

Pour un gestionnaire d'entreprise, la volatilité est un indicateur primordial pour la fluidité du marché. En termes statistiques, la volatilité d'un portefeuille est définie comme étant l'écart type des rendements de ce portefeuille. L'écart type est l'outil le plus connu pour mesurer la

dispersion par rapport à une moyenne.

Nous verrons dans la suite que la Valeur-à-Risque, se situe au centre de nombreux systèmes de gestion des risques, du fait qu'elle représente un indice important pour mesurer le risque associé à un portefeuille.

Les types de risque

- **Risque de crédit** : incapacité d'un débiteur (entreprise, institution financière, particulier...) respecter ses engagements financiers vis-à-vis de sa banque.
- **Risque de marché** : Risque de perte suite à une évolution défavorable des paramètres de marché pouvant causer des impacts négatifs sur la position de la banque.
- **Risques opérationnels** : Pertes résultant d'une inadéquation ou d'un échec au niveau des processus, des personnes, des systèmes (erreurs humaines, pannes systèmes, fraudes, litiges,...) ou perte résultant d'événements externes (catastrophes naturelles, incendies,...)

Paramètres d'une mesure de risque

Pour chaque mesure de risque, il faut spécifier trois paramètres essentiels : l'horizon de modélisation, le seuil de confiance et la variable financière à modéliser.

1.Sélection de l'horizon :

L'horizon est la durée de temps sur laquelle le modèle doit être projeté pour produire des informations adéquates. Un horizon relativement court a le désavantage que la plupart des stratégies ne peuvent pas avoir leurs effets complets à court terme. En revanche, un long

horizon freine la modélisation en projetant la variable économique loin dans le futur, ce qui peut entraîner la dégradation de la capacité de modélisation à tirer des conclusions fiables. En effet, plus l'horizon est long, plus la sensibilité du modèle aux hypothèses sous-jacentes est élevée.

Sélection du seuil de confiance : Le choix du seuil de confiance est un paramètre capital de mesure de risque. Il s'agit de spécifier la valeur critique de la mesure qui distingue entre le niveau acceptable et le niveau inacceptable de risque. Généralement, pour une mesure de risque donnée, on fixe un niveau de confiance (par exemple 90% pour le risque de marché) qui correspond à la probabilité que le montant des pertes ne dépasse pas cette mesure de risque en valeur absolue. Du point de vue réglementaire, l'objectif de l'utilisation de ce seuil est la minimisation du nombre de faillites.

3. Sélection de la variable financière : Le risque est défini en termes de changement de valeur entre deux dates. Plus exactement, entre la date de modélisation (où la valeur est connue) et une date future (l'horizon de modélisation). De ce fait, la variable à modéliser est tout simplement la valeur future dans tous les états du monde, due aux changements de marché ou plus généralement à des événements incertains. Cette variable aléatoire à interprété en tant que valeur future d'une position ou d'un portefeuille actuellement détenu. De manière générale, la variable financière à modéliser peut représenter la perte d'un portefeuille donné, la perte globale d'une société,...

Choix de la méthode d'estimation

Les principales méthodes de l'approche classique de calcul de la VaR peuvent être classées en quatre catégories :

- a- Simulation Historique ;

b- Méthode de Variance-Covariance ;

c- Simulation de Monte Carlo ;

d- Méthode basée sur des tests d'ajustement. Le choix entre ces techniques d'estimation n'est pas une tâche facile pour le gestionnaire de risque. En effet, son choix se base sur les capacités de chaque méthode à répondre à ses besoins.

Pour choisir entre ces différentes méthodes de calcul de la VaR, Linsmeier et Pearson (1996) énumèrent les critères suivants :

- La capacité de la méthode à capter les risques liés aux produits dérivés comme des options par exemple ;
- La facilité dans l'implantation de la méthode choisie ;
- La facilité dans la communication des résultats de la méthode à des utilisateurs internes et externes ;
- La fiabilité des résultats obtenus ;
- La flexibilité de la méthode à incorporer des hypothèses alternatives choisies par le gestionnaire de risque (ajustement du niveau de corrélation, de volatilité, ...).

Annexe B : Abréviations et Notations

va	Variable aléatoire.
iid	identiquement Indépendantes distribuées.
H, F, G	Fonction de répartition jointe et marginales.
U, V	Variables aléatoires uniforme.
\mathbb{E}	Espérance de X.
Var (X)	Variance de X.
Cov(X, Y)	Covariance de X et Y.
W, M, Π	Copules min, max et prouit.
c	Densité de la copule.
r(X,Y)	Coefficient de corrélation linéaire.
h	Densité de la distribution conjointe.
$\tau, \tau(X, Y)$	Tau de Kendall.
$\rho, \rho(X, Y)$	Rho de Spearman.
τ_n, ρ_n	Tau de Kendall et rho de Spearman empiriques.
\bar{X}_n	La moyenne empirique.
S^2	La variance empirique.
$\hat{\theta}$	Estimateur du paramètre θ .
$\hat{\theta}_{MLE}$	Estimateur du paramètre θ par la méthode de maximum de vraisemblance.

$\hat{\theta}_{IFM}$	Estimateur du paramètre θ par la méthode des fonctions d'inférence des marginales.
$\hat{\theta}_{CML}$	Estimateur du paramètre θ par la méthode de maximum de vraisemblance canonique.
$\hat{\theta}_{TK}$	Estimateur du θ par l'inversion de tau de Kendall.
$\hat{\theta}_{RS}$	Estimateur du θ par l'inversion de rho de Spearman.
$F_n(x)$	Fonction empirique.
C_n, \hat{C}	La copule empirique.
K	La fonction de répartition de noyau.
k	La fonction de densité du noyau.
t_s	Le délai statistique.
t_f	Le délai formatif.
t_d	Le délai de décomposition.
VaR	Valeur à risque.

Bibliographie

- [1] **Abderrahmane Nour El Houda** : Modélisation De La Dépendance Par Les Copules, Université M'hamed Bougara, Boumerdes
- [2] **Anas El Khattabi** : *Quelques applications du calcul stochastique :les copules en finance*, université de Ibn Zohr, Agadir.
- [3] **Biljana Popović**, *An example of the application of copula models in physics* , article University of Niš, 2016
- [4] **C. Genest** : *The joy of copulas : Bivariate distributions with uniform marginals*, The American Statistician, 40 :280–285, 1986b.
- [5] **G. Kim** :*Comparison of semiparametric and parametric methods for estimating copulas*,Computational Statistics Data Analysis, 51 :2836–2850, 2007.
- [6] **G. Salvadori** : *Extremes in Nature, An Approach Using Copulas. Water Science and Technology Library*, volume 56. Springer, Dordrecht (NL), 2007.
- [7] **Kabot Fathia** : *Sur les copules bivariées*, université de Mohamed Khider, Biskra.
- [8] **J-F. Quessy** :*Méthodologie et application des copules : Tests d'adéquation, tests d'indépendance, et bornes pour la valeur à risque*,Thèse de Doctorat, Faculté des Sciences et de Génie, Université Laval, Québec, 2005.
- [9] **J.Hanry**, *Dependence Modeling with Copulas*, International Standard Book Number-13 : 978-1-4665-8323-8 ,2014

- [10] **Kadi Nabil** : *Estimation non-paramétrique de la distribution et densité de copule* , université de Sherbrooke, Québec, Canada, avril 2014.
- [11] **L.Michel** : *Statistique, la théorie et ses applications*, Springer-verlag, 2010
- [12] **Lounis Fadhila** : *Modélisation de la dépendance par les copules et applications*, université de Mouloude Mammeri, Tizi-Ouzou.
- [13] **R.Silvia**, *Dynamic copula methods in finance*, 2012
- [14] **Slimani** : *Estimation des paramètres d'une copule*, université de Mohamed Khider, Biskra.