

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

UNIVERSITE MOHAMED KHIDER BISKRA

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



## MEMOIRE

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

Par

Imane BENELMIR

Thème

---

---

*LES L-MOMENTS:  
APPLICATION EN HYDROLOGIE*

---

---

Option

Probabilités & Statistiques

Soutenu publiquement le :

Devant le jury :

Président: Brahim MEZERDI  
Rapporteur: Abdelhakim NECIR  
Examineur: Khaled MELKEMI  
Examineur: Djamel MERAGHNI

Pr.	U.M.K BISKRA
Pr.	U.M.K BISKRA
Pr.	U.M.K BISKRA
M.C. (A)	U.M.K BISKRA

# REMERCIEMENTS

*Au nom du Dieu Clément et Miséricordieux,*

*Louange à Allah, maître de l'univers, que sa grâce, sa paix, sa miséricorde et sa bénédiction soient accordées au meilleur de ses messagers, à notre prophète Mohamed ainsi qu'aux membres de sa famille et à tous ses compagnons.*

*Pour le humble effort durant ces années, je prie Allah le tout puissant de m'avoir avec courage et volonté permis de mener à bien mon travail.*

*Je tiens tout d'abord à remercier le Professeur **Necir Abdelhakim** pour l'intérêt qu'il porte à mon travail et aux statistiques en générale. Il a su sagement m'orienter durant ces années. Ses compétences scientifiques m'ont permis de découvrir le domaine de la recherche sur des bases solides. Nos échanges (scientifiquement parfois vigoureux) ont toujours débouché sur des perspectives et des études passionnantes.*

*Je remercie évidemment le professeur **Mezerdi Brahim** d'avoir bien voulu participer à ce jury et d'en avoir été le président.*

*Je souhaite remercier vivement les professeurs **Melkmi Khaled** et **Meraghni Djamel** d'avoir accepté d'être examinateurs, malgré leurs nombreuses charges respectives.*

*Un grand hommage à mon professeur **Bahlali Seid** qui a laissé un immense vide dans nos cœurs et à notre université Mohamed Khider.*

*Merci à Monsieur **Brahimi Brahim** pour ses conseils scientifiques qui ont souvent éclairci mes perspectives.*

*Je ne remercierai jamais assez mes parents **Omar** et **Akila** pour leurs encouragements, leurs patiences, leurs soutiens et surtout leur amour. Ils ont su mener une part essentielle à l'aboutissement de mon travail.*

*Merci à **Benelmir Samir** (mon très cher frère) et ma très chère sœur **Samia** d'être là chaque fois que j'ai besoin d'eux...*

**IMANE**

## TABLE DES MATIÈRES

TABLE DES FIGURES . . . . .	<b>4</b>
LISTE DES TABLEAUX . . . . .	<b>5</b>
Introduction générale . . . . .	6
Analyse fréquentielle régionale . . . . .	8
Contexte et perspective . . . . .	10
CHAPITRE 1. <i>L</i> -MOMENTS UNIVARIÉS PAR J. R. M. HOSKING . . . . .	<b>12</b>
1.1. <i>L</i> -moments univariés . . . . .	13
1.1.1. <i>L</i> -moments pour des échantillons de données . . . . .	14
1.1.2. <i>L</i> -moments pour les distributions de probabilités . . . . .	15
1.1.3. <i>L</i> -moments en tant que fcts linéaires des statistiques d'ordres . . . . .	15
1.1.4. Exemples . . . . .	17
1.2. Techniques de base . . . . .	18
1.2.1. Représentations des <i>L</i> -moments . . . . .	18
1.2.2. Estimation . . . . .	20
1.2.3. Théorie de la distribution asymptotique pour $\hat{\lambda}_k$ et $\hat{\tau}_k$ . . . . .	22
1.3. Conclusion . . . . .	23
CHAPITRE 2. CONTRIBUTION AUX <i>L</i> -MOMENTS MULTIVARIÉS . . . . .	<b>24</b>
2.1. <i>L</i> -Comoments et <i>L</i> -Moments multivariés . . . . .	24
2.1.1. <i>L</i> -moments multivariés . . . . .	25
2.1.2. <i>L</i> -moments bivariés . . . . .	25
2.2. Représentations de $\lambda_{k[12]}$ en termes de concomitants . . . . .	29
2.3. Estimation des <i>L</i> -Comoments . . . . .	31
2.3.1. Exemple : ( <i>Estimation du 2<sup>ème</sup> L-Comoment</i> ) . . . . .	32
2.4. Convergence asymptotique . . . . .	32
2.5. Conclusion . . . . .	33
CHAPITRE 3. THÉORIE DES COPULES . . . . .	<b>34</b>
3.1. Introduction aux copules . . . . .	35
3.2. Généralités sur les copules bivariées . . . . .	35
3.2.1. Définition d'une copule bivariée (dimension 2) . . . . .	37
3.2.2. Propriétés immédiates des copules bivariées . . . . .	38
3.3. Applications des copules . . . . .	43
3.3.1. Application à la construction des copules . . . . .	43
3.3.2. Application à la modélisation multivariée des distributions . . . . .	45
3.4. Propriétés fondamentales des copules . . . . .	47

3.5. "Copules" associées à une copule . . . . .	49
3.5.1. Copule de survie . . . . .	49
3.5.2. Copule duale . . . . .	51
3.5.3. Co-copule . . . . .	51
3.5.4. Copule mixte . . . . .	51
3.6. Quelques copules classiques . . . . .	54
3.6.1. Copules usuelles . . . . .	54
3.6.2. Copules elliptiques . . . . .	55
3.6.3. Copules Archimédiennes et Archimax . . . . .	57
3.6.4. Copules des valeurs extrêmes . . . . .	60
3.6.5. Copule empirique . . . . .	61
3.7. Mesures de concordance . . . . .	61
3.7.1. Fonction de concordance . . . . .	61
3.7.2. Propriétés de la fonction de concordance . . . . .	62
3.7.3. Mesures de concordance . . . . .	69
3.8. Représentation des $L$ -comoments en termes de Copules . . . . .	75
3.8.1. Représentation de $\lambda_{2[12]}$ en termes de copule . . . . .	76
3.8.2. Représentation de $\lambda_{3[12]}$ en termes de Copule . . . . .	77
3.8.3. Représentation de $\lambda_{4[12]}$ en termes de Copule . . . . .	79
3.9. Conclusion . . . . .	83
<b>CHAPITRE 4. APPLICATION : MODÉLISATION PAR UNE COPULE DE GUMBEL</b>	<b>84</b>
4.1. Copule de Gumbel . . . . .	84
4.2. Caractéristiques . . . . .	85
4.2.1. Copule archimédienne . . . . .	85
4.2.2. Fonction de répartition . . . . .	85
4.2.3. Densité . . . . .	86
4.2.4. Copule singulière . . . . .	87
4.3. Mesures de dépendance . . . . .	88
4.3.1. Coefficient de corrélation de Kendall . . . . .	88
4.3.2. Coefficient de corrélation de Spearman . . . . .	89
4.4. Copule multivariée . . . . .	91
4.5. Mise en œuvre des copules . . . . .	92
4.5.1. Techniques de simulation . . . . .	92
4.5.2. Estimation et implémentation . . . . .	94
4.5.3. Outils graphiques d'ajustement de la copule . . . . .	97
4.5.4. Package R gumbel . . . . .	98
4.6. Conclusion . . . . .	99
<b>CONCLUSION GÉNÉRALE . . . . .</b>	<b>100</b>
<b>ANNEXE A. . . . .</b>	<b>105</b>

ANNEXE B. . . . . 107

ANNEXE C. . . . . 108

## TABLE DES FIGURES

FIGURE 1.	Portrait de Wallis John . . . . .	11
FIGURE 3.1.	Notion de copule . . . . .	35
FIGURE 4.1.	Distribution de la copule de Gumbel pour $(\alpha = 2)$ et $(\alpha = 4)$ . . . . .	86
FIGURE 4.2.	Densité de la copule de Gumbel pour $(\alpha = 2)$ et $(\alpha = 4)$ . . . . .	87
FIGURE 4.3.	Simulation des paires $(u, v)$ ayant une distribution de la copule de Gumbel . . . . .	94
FIGURE C.1.	Simulation de la copule de Gumbel pour $(\alpha = 1)$ et $(\alpha = 2)$ . . . . .	115

## LISTE DES TABLEAUX

TABLE 3.1. Familles classiques de copules archimédiennes . . . . .	58
--	----

## Introduction générale

Les applications de la modélisation et de l'analyse des modèles pour les phénomènes extrêmes sont très variées : modélisation de l'ampleur des vagues (construction de digues ou de plateformes maritimes), pannes des systèmes complexes, météorologies (force du vent, hydrologie,...), stratégies du management, études biomédicales, analyse des risques alimentaires, analyse des risques sismiques, assurance, finance, etc. Citons quelques exemples :

1. **Crues des rivières** : Pour savoir la hauteur à laquelle il faut construire une digue, il faut être capable de prédire des hauteurs de grandes crues, plus grandes que celles qui ont été observées dans les 50 dernières années si on veut que la digue ait une durée de vie de plus de 50 ans.
2. **Vent** : Prédire la probabilité que la vitesse du vent dépasse un certain seuil (Alerte météorologique, 2008 – 2009, France).
3. **Assurance** : On s'est rendu compte que 20% des sinistres représentent 80% du montant des sinistres. Ces sinistres très coûteux correspondent à des phénomènes extrêmes qu'il faut pouvoir prédire.
4. **Chutes de neige** : Sur l'aéroport de Caroline du Nord (US) enregistrées entre (1948 – 1998). Comment utiliser ces données pour prédire la probabilité d'observer des chutes de neiges bien plus importantes telles que celles observées en 2000.

Les événements hydrologiques sont complexes et souvent caractérisés par le comportement commun de plusieurs variables aléatoires (sécheresses, inondations, tempêtes de pluie, etc), qui ne sont généralement pas indépendantes.

Des exemples de représentations multivariées des phénomènes hydrologiques inclus :

- Durée de la tempête et l'intensité (*Yue, 2001 ; Salvadori & Michele, 2004*).
- Pics des inondations, le volume et la durée (*Ashkar, 1980 ; Yue & al., 1999 ; Ouarda & al., 2000 ; Yue, 2001 ; Yue & Rasmussen, 2002 ; Shiau, 2003 ; Michele & al., 2005 ; Zhang & Singh, 2006*).
- Volume de la sécheresse, la durée et la grandeur (*Kim & al., 2003 ; Ashkar & al., 1998*).

Cette compréhension des inondations multivariées, les tempêtes, ou les événements caractéristiques de sécheresse sont essentiels pour plusieurs planifications d'ingénierie, la conception et les activités de gestion. Les approches multivariées représentent ces événements mieux que les outils classiques univariés. *Snyder (1962)* et *Wong (1963)* ont réalisé les premières applications des outils d'analyse multivarié en hydrologie.



Une compréhension approfondie des événements hydrologiques multivariés exige l'étude du comportement probabiliste commun de plusieurs variables aléatoires corrélées qui caractérisent les événements (Yue *et al.*, 2001). Les  $L$ -moments multivariés sont utilisées comme un outil pour définir des statistiques et des techniques de copules. Pour examiner le comportement statistiques des variables dépendantes, les copules sont récemment montrées pour représenter un outil mathématique pour des applications hydrologiques (El Adlouni *et al.*, 2004 ; Salvadori *et De Michele*, 2004).

Dans ce travail, les  $L$ -moments ont été introduites. Ils sont devenus des outils populaires pour résoudre divers problèmes liés à l'estimation des paramètres, l'identification de la distribution et la régionalisation.

En effectuant une analyse fréquentielle régionale multivariée pour ces événements, il est important de conjointement considérer toutes ces caractéristiques.

- Le premier chapitre de cette thèse présente les  $L$ -moments univariés et leurs caractéristiques principales ainsi que certains résultats de base pour les  $L$ -moments univariés.
- Le deuxième chapitre s'appuie sur les  $L$ -moments multivariés. Nous développons les propriétés de base, représentations, estimations et convergence asymptotique des  $L$ -comoments.
- Dans le chapitre trois nous mettons en oeuvre la théorie des copules dans ses grandes lignes, notamment le théorème de Sklar qui clarifiera la notion de dépendance. Enfin nous présentons des mesures de dépendance qui peuvent être exprimées à partir des copules.
- Le dernier chapitre met en avant les copules utilisées au cours de notre étude, soit la copule de Gumbel et donne toutes les formules nécessaires, les techniques de simulation, l'implémentation des divers outils présentés jusqu'ici, ainsi que les méthodes de calibration.
- En conclusion nous récapitulons les apports des méthodes proposées dans ce travail.
- L'outil utilisé pour mener à bien cette étude est le logiciel *R*.

## Analyse fréquentielle régionale

Les événements environnementaux extrêmes, tels que les inondations, sécheresses, tempêtes de pluie et forces des vents, ont de graves conséquences pour la société humaine. Comment la fréquence d'un événement d'une grandeur (ampleur ou magnitude) donnée peut être prévu à se produire est d'une grande importance. Projetant pour des urgences dues aux conditions atmosphériques, la conception des structures de génie civil, la gestion de réservoir, la lutte contre la pollution et les calculs des risques d'assurance, tous se fondent sur la connaissance de la fréquence de ces événements extrêmes. L'estimation de ces fréquences est difficile parce que les événements extrêmes sont par définition rares et les échantillons sont souvent courts.

L'analyse de fréquence régionale résout ce problème par des données de plusieurs sites sont utilisées pour estimer les fréquences des événements de n'importe quel site. Les  $L$ -moments sont un développement récent dans les statistiques. Ils forment la base d'une théorie mathématique et peuvent être utilisés pour faciliter le processus d'estimation dans l'analyse de fréquence régionale. Les méthodes des  $L$ -moments sont manifestement supérieures à ceux qui ont été utilisés précédemment et sont maintenant adoptées par des organismes principaux dans le monde entier. Pour plus de détail voir *J. R.M. Hosking & J. R. Wallis, (1997). [Analyse de fréquence régionale : Une approche basée sur les  $L$ -moments. Université De Cambridge]*.

Plusieurs applications concernent l'environnement, pour chaque variable d'intérêt, par exemple l'inondation, des séries distinctes d'observations prises à différents sites de mesure dans un réseau. Il en résulte pour un échantillon de données de plusieurs variables de nature similaire, avec une dépendance possibles à l'intérieur et entre les échantillons. Un objectif est d'estimer la fonction quantile marginal, en particulier à fins de déterminer la partie du quantile supérieure correspondant à la présence d'un événement extrême.

Dans de nombreuses applications la taille de l'échantillon d'un site est trop petite pour permettre une estimation efficace des quantiles supérieures et des données pour tous les sites dans une région où sont combinés par "l'analyse de fréquence régional" sous des hypothèses efficaces et plificatrices.

Dans ce contexte, les méthodes des  $L$ -moments se sont avérées très efficaces dans la prestation stable et les estimations fiables moins sensibles à l'influence des observations extrême que les méthodes classiques des moments. Voir *Hosking et Wallis* pour compléter l'exposition de l'analyse de fréquence régionale par la méthode des  $L$ -moments qui a été développée dans des domaines tels que l'hydrologie, la climatologie et la science environnementale.

En analyse de fréquence régionale, une étape importante à la partition d'un réseau

de sites en sites de régions approximativement homogènes avec des distributions de fréquences similaires pour la variable d'intérêt. Pour chaque site, le vecteur des quatre premiers  $L$ -moments ou des coefficients  $L$ -moments est obtenu et les sites "insolite" sont identifiés via une mesure de discordance.

On a pu observer que les séries de données issues de stations très proches géographiquement peuvent être fortement corrélées. L'hypothèse d'indépendance des séries n'est alors pas valable. Une approche intéressante consiste alors à utiliser des copules, qui sont des fonctions permettant de modéliser la structure de dépendance indépendamment des distributions marginales.

Autrement dit, elles permettent de décrire le comportement individuel de chaque station et couplent les lois marginales pour obtenir la loi jointe. Grâce à elles, on peut donc former des distributions multidimensionnelles avec différentes marginales et la structure de dépendance sera donnée par la copule.

Toutefois, dans de nombreux cas, il existe plusieurs variables d'intérêt (inondation, température, précipitation, vent, etc) toutes sont mesurées à chaque site dans le réseau. Au lieu d'effectuer séparément l'analyse de fréquence régionale avec les approches des  $L$ -moments multivariés pour chaque variable, à chaque génération d'une partition différente en régions homogènes, on peut maintenant mettre au point une longue analyse de fréquence régionale conduisant à une seule partition en groupes homogènes de régions basées sur toutes les variables conjointement considéré

## Contexte et perspective

Pour mesurer les caractéristiques descriptives d'une distribution univariée, les moments centraux sont très populaires, mais leur utilisation est suffisamment limitée aux distributions à queues ordinaires. Une alternative attrayante est fournie par la série des  $L$ -moments, qui ont la forme de fonction linéaire d'espérance des statistiques d'ordres stratégiquement choisies.

Tandis que le premier  $L$ -moment est juste la moyenne ordinaire, l'ordre supérieur enferme non seulement la mesure de diffusion (spread), l'asymétrie (skewness), l'aplatissement (kurtosis), etc.

Par exemple, le  $L$ -moment de n'importe quel ordre  $k$  existe sous une seule hypothèse de premier moment fini, ce qui rend l'ensemble de la série des  $L$ -moments disponibles pour la plus part des distributions à queues lourdes.

Avec des antécédents en *Sillitto & Downton*, un traitement formel des  $L$ -moments a été mis au point par *Hosking*, qui a établi des résultats fondamentaux soutenant une nouvelle méthode d'analyse des données et à l'inférence statistique basée sur les  $L$ -moments.

L'importance du potentiel offert par l'approche des  $L$ -moments est l'utilisation des distributions à queues lourdes en modélisation statistique. Dans certains contextes, la modélisation de la fréquence des événements extrêmes est particulièrement préoccupante. À cet égard, une vaste méthode des  $L$ -moments a été développée à l'appui de l'analyse fréquentielle régionale en science environnementale.

Cette dernière traite les distributions des quantiles telles que les précipitations maximales annuelles, cours d'eau, ou vitesse du vent observée à chaque emplacement dans un réseau donné (*Hosking & Wallis*).

L'approche des  $L$ -moments est particulièrement plus utile pour les applications et les estimations descriptives tout comme les moments centraux habituels. De tels soucis surgissent par exemple, dans l'estimation de la volatilité dans la gestion des risques financiers impliquant des variables du marché tels que les indices boursiers, les taux d'intérêt, etc (*Hosking, Bonti & Siegel*).

Les données sont en général de nature multidimensionnelle et demande l'analyse statistique en tenant compte de la structure de la dépendance fondamentale appropriée et de la géométrie. Cependant jusqu'ici, les  $L$ -moments ont été définis seulement dans le cas univarié. L'utilisation de ces marginales ne permet pas de saisir les caractéristiques des composantes des variables conjointement distribuées.

Ce qu'il faut, c'est une extension de la notion des  $L$ -moments au cas multivarié. Excepté pour les simples extensions de la moyenne univariée à la moyenne du vecteur multivarié. Ceci est resté un problème ouvert, en partie par manque d'extension

immédiate de la notion de la fonction linéaire des statistiques d'ordres à espace de dimensions supérieures. En effet, *Hosking* a écrit "aucune extension des  $L$ -moments de distributions multivariées n'est immédiatement apparente".

D'autre part, il mentionne aussi, mais sans donner de précisions, que "l'approche le plus prometteur en apparence serait d'utiliser la notion des statistiques d'ordres pour mesurer l'association entre deux variables aléatoires". Dans cette thèse, nous développons la perspicacité de Hosking en solution efficace.

Pour la perspective, nous constatons que l'analyse classique multivariée tourne autour de deux principales entités paramétriques : la valeur moyenne du vecteur et la matrice de covariance, qui sont respectivement l'extension de la moyenne et de la variance univariée. En outre, le troisième et le quatrième moment central univarié, sont respectivement la mesure de l'asymétrie et de l'aplatissement.

Puisque les  $L$ -comoments sont similaires à la structure et au comportement aux  $L$ -moments, les matrices des  $L$ -moments multivariées sont de nouveaux outils descriptifs efficaces ayant une utilité semblable à la matrice classique utilisée, ou plus précisément analogues mais moins connues que les matrices  $L$ -comoments d'ordres élevés.



FIG. 1. Portait de Wallis John

## Chapitre 1

### *L*-MOMENTS UNIVARIÉS PAR J. R. M. HOSKING

L'analyse statistique multivariée s'appuie fortement sur les hypothèses du moment du second ordre et plus. Cependant avec l'augmentation de l'attention au problème de la modélisation avec des distributions à queues lourdes, dans par exemple les sciences environnementales et l'analyse des risques financiers. Il est important de remédier à cette lacune.

Nous pouvons être en mesure de décrire les caractéristiques des distributions multivariées, non seulement par la dispersion mais également par l'asymétrie (skewness) et l'aplatissement (kurtosis) simplement sous les hypothèses du premier moment d'ordre. Ici nous présentons une nouvelle méthode d'analyse multivariée contribuant vers ce but.

Notre approche consiste à développer les extensions appropriées des *L*-moments univariés par *Hosking* (1990), qui sont analogues aux moments centraux univariés définis pour tout ordre simplement sous les hypothèses du premier moment.

Nous présentons dans ce chapitre les *L*-moments univariés ainsi leurs caractéristiques principales. Ensuite nous fournissons certain nécessaire résultat de base, pour les *L*-moments univariés (représentation, estimation et comportement asymptotique).

## 1.1 $L$ -moments univariés

Les  $L$ -moments sont des statistiques pour les distributions de probabilités et des échantillons de données. Ils sont semblables à des moments ordinaires. Ils fournissent des mesures de l'emplacement, de la dispersion, de la dissymétrie "skewness", de l'aplatissement "kurtosis" et d'autres aspects de la forme de la distribution de probabilité ou des échantillons de données. Mais ils sont calculés à partir des combinaisons linéaires des valeurs de données ordonnées.

Les  $L$ -moments ont les avantages théoriques suivants :

- Pour que les distributions de probabilité des  $L$ -moments soient significatives, nous exigeons seulement que la distribution à une moyenne finie. Aucun moments d'ordre élevé ne doit être fini (*J. R. M. Hosking, 1990*).
- Pour que les erreurs standards des  $L$ -moments soient finies, nous exigeons seulement que la distribution ait une variance finie.
- Bien que les proportions des moments peuvent être arbitrairement grandes, l'échantillon des proportions des moments ont des limites algébriques (*J. Dalen, 1987. Statistics and Probability Letters*).

De plus, les propriétés suivantes se tiennent dans une vaste gamme de situations pratiques :

- Les approximations asymptotiques à l'échantillonnage des distributions sont meilleures pour les  $L$ -moments que pour les moments (*Hosking, 1990*).
- Ils sont moins sensibles à la périphérie des valeurs (*P. Royston, 1992. Statistics in Medicine. R. M. Vogel & N. M. Fennessey, 1993. Water Resources*).
- Ils fournissent une meilleure identification de la distribution qui génère un échantillon de données particulier (*Hosking, 1990*).

### 1.1.1 $L$ -moments pour des échantillons de données

Les moments de probabilités pondérées, définis par *J. A. Greenwood & al. (1979)*, sont des précurseurs des  $L$ -moments. L'échantillon des moments de probabilités pondérées est calculé à partir des valeurs de données  $X_1, X_2, \dots, X_n$ , arrangées dans l'ordre croissant. Ces moments de probabilités pondérées sont donnés par :

$$\begin{cases} b_0 := n^{-1} \sum_{j=1}^n X_j, \\ b_r := n^{-1} \sum_{j=r+1}^n \frac{(j-1)(j-2)\dots(j-r)}{(n-1)(n-2)\dots(n-r)} X_j. \end{cases}$$

Les  $L$ -moments sont des combinaisons linéaires des moments de probabilités pondérées qui ont des interprétations simples comme les mesures de l'emplacement, la dispersion et la forme de l'échantillon de données.

Les quelques premiers  $L$ -moments sont définis par :

$$\begin{cases} l_1 = b_0, \\ l_2 = 2b_1 - b_0, \\ l_3 = 6b_2 - 6b_1 + b_0, \\ l_4 = 20b_3 - 30b_2 + 12b_1 - b_0. \end{cases}$$

Les coefficients sont ceux "des polynômes de Legendre déplacé".

- Le premier  $L$ -moment est la moyenne de l'échantillon, appelé : mesure de l'emplacement.
- Le deuxième  $L$ -moment est la différence moyenne de *Gini*, appelé : mesure de la dispersion des moyennes des valeurs de données au tour de la moyenne.

En divisant l'ordre supérieur des  $L$ -moments par la mesure de dispersion, nous obtenons les proportions des  $L$ -moments :

$$t_r := l_r/l_2; \quad r \geq 3.$$

- $t_3$  : mesure de l'asymétrie "skewness".
- $t_4$  : mesure de l'aplatissement "kurtosis".

Celles-ci sont respectivement le  $L$ -skewness et le  $L$ -kurtosis. Ils prennent des valeurs entre  $-1$  et  $+1$ .



Le  $L$ -moment analogue au coefficient de variation (écart type divisé par la moyenne), est le  $L$ -CV qui est défini par :

$$t := l_2/l_1.$$

Il prend des valeurs entre 0 et 1.

### 1.1.2 $L$ -moments pour les distributions de probabilités

Pour une distribution de probabilité avec fonction de distribution (fd)  $F(x)$ , les moments de probabilités pondérées sont définis par :

$$B_r := \int x \{F(x)\}^r dF(x); \quad r \geq 0.$$

Les  $L$ -moments sont définis en termes des moments de probabilités pondérées. Par analogie à l'échantillon des  $L$ -moments nous obtenons :

$$\begin{cases} \lambda_1 = B_0, \\ \lambda_2 = 2B_1 - B_0, \\ \lambda_3 = 6B_2 - 6B_1 + B_0, \\ \lambda_4 = 20B_3 - 30B_2 + 12B_1 - B_0. \end{cases}$$

Les proportions des  $L$ -moments sont :

$$\tau_r := \lambda_r/\lambda_2; \quad r \geq 3.$$

### 1.1.3 $L$ -moments en tant que fcts linéaires des statistiques d'ordres

Soient  $X_1, X_2, \dots$  une suite de variables aléatoires (va) indépendantes et identiquement distribuées (iid) avec fonction de distribution  $F(x) := \mathbb{P}(X \leq x); x \in \mathbb{R}$ . Pour chaque entier  $k \geq 1$ ,  $X_{1,k} \leq X_{2,k} \leq \dots \leq X_{k,k}$  sont des statistiques d'ordres (so) fondées sur l'échantillon de taille  $k$ , les  $k^{\text{ème}}$   $L$ -moments univariés relatifs à cet échantillon sont définis par :

$$\lambda_k := k^{-1} \sum_{j=0}^{k-1} (-1)^j \binom{k-1}{j} E(X_{k-j,k}); \quad k \geq 1. \quad (1)$$

La notation  $\binom{k}{j} := k! / (k-j)!j!$ .

Les coefficients dans (1) sont aussi ceux de l'opérateur de différence d'ordres  $k = 1$  et  $k \geq 2$  :

---

"L-Moments univariés"

---

- Pour  $k = 1$  :  $\lambda_k$  est la  $(k - 1)^{\text{ème}}$  itération de l'opérateur de différence d'ordre appliqué à la séquence  $\{E(X_{j,k}) : j = 1, \dots, k\}$ .

Ainsi, le premier  $L$ -moment est la moyenne de la fd  $F$ .

$$\lambda_1 = E(X_{1,1}); k = 1.$$

- Tandis que, pour  $k \geq 2$  :  $\lambda_k$  est la  $(k - 2)^{\text{ème}}$  itération de l'opérateur de différence d'ordre appliqué à la séquence  $\{E(X_{j+1,k}) - E(X_{j,k}) : j = 1, \dots, k - 1\}$ .

Les  $L$ -moments peuvent être caractérisés en termes d'espacements prévus :

$$\chi_{j,k} := E(X_{j+1,k} - X_{j,k}); 1 \leq j \leq k - 1.$$

Pour un échantillon de taille  $k$  on a :

$$\lambda_k = k^{-1} \sum_{j=0}^{k-2} (-1)^j \binom{k-2}{j} \chi_{k-1-j,k}; k \geq 2. \quad (2)$$

Cette représentation facilite beaucoup l'interprétation et la compréhension des  $L$ -moments. L'idée de lien avec les espacements prévus est due à *Sillitto*.

Voici les quatre premiers  $L$ -moments pour une distribution univariée :

1. Le deuxième  $L$ -moment  $\lambda_2$  mesure clairement la dispersion (spread) :

$$\lambda_2 = \frac{1}{2} E(X_{2,2} - X_{1,2}) = \chi_{1,2}/2.$$

- En fait, il est bien connu que cette mesure est la moitié de la différence moyenne classique de Gini.

- Outre intérêt indispensable,  $\lambda_2$  est utilisé pour obtenir des mesures descriptives appelées les proportions des  $L$ -moments :

$$\tau_k := \lambda_k / \lambda_2; k \geq 3.$$

Pour plus de détails, on se réfère à *Hosking & Wallis (1997)*.

- Hosking (1989) a démontré que les proportions des  $L$ -moments satisfaisaient :

$$-1 \leq \tau_k \leq 1; k \geq 3.$$

2. Le troisième  $L$ -moment  $\lambda_3$  :

$$\lambda_3 = \frac{1}{3} E(X_{3,3} - 2X_{2,3} + X_{1,3}) = (\chi_{2,3} - \chi_{1,3})/3.$$

3. Le quatrième  $L$ -moment  $\lambda_4$  :

$$\begin{aligned}\lambda_4 &= \frac{1}{4}E(X_{4:4} - 3X_{3:4} + 3X_{2:4} - X_{1:4}) \\ &= (\chi_{3:4} - 2\chi_{2:4} + \chi_{1:4})/4 \\ &= [(\chi_{3:4} + \chi_{1:4})/2 - \chi_{2:4}]/2.\end{aligned}$$

#### 1.1.4 Exemples

1. Distribution uniforme sur  $(0, 1)$  :

$$\begin{cases} \lambda_1 = 1/2, \\ \lambda_2 = 1/6, \\ \tau_3 = 0, \\ \tau_4 = 0. \end{cases}$$

2. Distribution normale avec moyenne 0 et variance 1 :

$$\begin{cases} \lambda_1 = 0, \\ \lambda_2 = 1/\sqrt{\pi}, \\ \tau_3 = 0, \\ \tau_4 \approx 0.123. \end{cases}$$

Dans le cas d'une distribution symétrique tous les  $L$ -moments impairs d'ordre supérieur sont égales à zéro. Par conséquent elles peuvent être considérées comme des mesures généralisées de l'asymétrie (*Hosking*).

3. Distribution exponentielle avec moyenne 1 :

$$\begin{cases} \lambda_1 = 1, \\ \lambda_2 = 1/2, \\ \tau_3 = 1/3, \\ \tau_4 = 1/6. \end{cases}$$

## 1.2 Techniques de base

Nous fournissons ici certains nouveaux résultats de base pour les  $L$ -moments univariés, couvrant les représentations et l'estimation. Ces résultats sont indispensables aux développements des extensions multivariées.

Tout au long, nous considérons une fd univariée  $F$  dont la fonction quantile  $F^{-1}$  et la séquence  $L$ -moment  $\{\lambda_k\}$ .

### 1.2.1 Représentations des $L$ -moments

Différents types de représentation des  $L$ -moments s'avèrent utiles : en tant qu'une  $L$ -fonctionnelle, une covariance et une fonction linéaire des so.

#### Représentation de $\lambda_k$ en tant que $L$ -fonctionnelle

Soit la fonction quantile relative à la fd  $F$ .

$$F^{-1}(u) := \inf(t : F(t) \geq u); \quad 0 < u < 1.$$

La fonction quantile  $F^{-1}$  donne la représentation  $L$ -fonctionnelle classique.

Le  $k^{\text{ème}}$   $L$ -moment est donné par :

$$\lambda_k := \int_0^1 F^{-1}(u) P_{k-1}^*(u) du; \quad k \geq 1. \quad (3)$$

Où :

$$\left\{ \begin{array}{l} P_k^*(u) := \sum_{j=0}^k p_{k,j}^* u^j; \quad 0 < u < 1; \quad P_0^*(u) = 1, \\ p_{k,j}^* := (-1)^{k-j} \binom{k}{j} \binom{k+j}{j}. \end{array} \right.$$

$P_{k,j}^*(u)_{j=0,1,2,\dots}$  : Système de Legendre déplacé des polynômes orthogonaux.

Les deux premiers  $L$ -moments sont :

$$\begin{aligned} \lambda_1 &= \int_0^1 F^{-1}(u) du \quad (\text{la moyenne de } F). \\ \lambda_2 &= \int_0^1 F^{-1}(u)(2u - 1) du. \end{aligned}$$

### Représentation de $\lambda_K$ en tant que Covariance

$\lambda_k$  peut être réécrite :

$$\lambda_k := \sum_{j=0}^{k-1} p_{k-1,j}^* \beta_j; \quad k \geq 1. \quad (4)$$

Avec :

$$\beta_j := \int_0^1 F^{-1}(u) u^j du = E(X F^j(X)); \quad j = 0, 1, 2, \dots, k.$$

Comme indiqué par *Hosking & Wallis* (1997), en utilisant l'orthogonalité des fonctions  $P_k^*$  nous obtenons facilement :

$$\lambda_k := \begin{cases} E(X), & k = 1. \\ Cov(X, P_{k-1}^*(F(X))); & k \geq 2. \end{cases} \quad (5)$$

Les trois premiers  $L$ -moments sont :

$$\begin{aligned} \lambda_1 &= E(x). \\ \lambda_2 &= 2Cov(X, F(X)) = Cov(X, (2F(X) - 1)). \\ \lambda_3 &= -6Cov(X, F(X)(1 - F(X))). \end{aligned}$$

### Représentation de $\lambda_k$ en tant que fonctions linéaires so

On a :

$$\begin{aligned} \forall r \leq n : \quad E(X_{r,n}) &:= r \binom{n}{r} \int_0^1 F^{-1}(u) u^{r-1} (1-u)^{n-r} du \\ &= n \binom{n-1}{r-1} \int_{-\infty}^{\infty} x [F(x)]^{r-1} [1-F(x)]^{n-r} dF(x) \end{aligned} \quad (6)$$

En utilisant la formule (6) et la définition de  $\beta_j$  de *David & Nagaraja* (2003) on obtient :

$$\forall r \leq n : \quad E(X_{r,n}) = n \binom{n-1}{r-1} \sum_{j=0}^{n-r} \binom{n-r}{j} (-1)^{n-r-j} \beta_{n-1-j}.$$

Par conséquent, on a pour  $k = 1, 2, \dots, n$  :

$$\beta_k = n^{-1} \binom{n-1}{k}^{-1} \sum_{j=k+1}^n \binom{j-1}{k} E(X_{j,n}). \quad (7)$$

---

"Techniques de base"

---

En utilisant (7) dans (4) on obtient la représentation suivante :

$$\begin{aligned}
\lambda_k &= \sum_{j=0}^{k-1} p_{k-1,j}^* \beta_j \\
&= \sum_{j=0}^{k-1} p_{k-1,j}^* n^{-1} \binom{n-1}{j}^{-1} \sum_{r=j+1}^n \binom{r-1}{j} E(X_{r,n}) \\
&= n^{-1} \sum_{r=1}^n \omega_{r,n}^{(k)} E(X_{r,n}).
\end{aligned} \tag{8}$$

Où

$$\begin{aligned}
\omega_{r,n}^{(k)} &:= \sum_{j=0}^{\min\{r-1, k-1\}} (-1)^{k-1-j} \binom{k-1}{j} \binom{k-1+j}{j} \binom{n-1}{j}^{-1} \binom{r-1}{j}. \\
\lambda_{k[ij]} &:= n^{-1} \sum_{r=1}^n \omega_{r,n}^{(k)} E(X_{r,n}).
\end{aligned}$$

En particulier :

$$\begin{cases} \omega_{r,n}^{(1)} \equiv 1. \\ \omega_{r,n}^{(2)} \equiv (2r - n - 1) / (n - 1); \quad r = 1, \dots, n. \\ \omega_{r,n}^{(k)} \equiv (-1)^{k-1} \omega_{n-r+1, n}^{(k)}. \end{cases}$$

Par l'intermédiaire de (4), les séquences  $\{\lambda_k\}$  et  $\{\beta_k\}$  sont équivalentes et alternativement, elles sont équivalentes à  $\{E(X_{k,k})\}$ , puisque par l'intermédiaire de (7), nous avons :

$$\beta_j = E(XF^j(X)) = (j+1)^{-1} E(X_{j+1, j+1}).$$

Ainsi d'après (4) on obtient une expression de  $\lambda_k$  en termes de valeurs extrêmes attendues :

$$\lambda_k = \sum_{j=0}^{k-1} p_{k-1,j}^* (j+1)^{-1} E(X_{j+1, j+1}).$$

### 1.2.2 Estimation

L'essentiel au développement des  $L$ -comoments est la compréhension des  $L$ -moments univariés.

Toute une série d'aspects favorables proviennent des caractéristiques comme :

- Espérances définissant les  $L$ -statistiques (fonctions linéaires des so).
- Noyaux définissant les  $U$ -statistiques.

Nous allons maintenant considérer l'estimation de  $\lambda_k$  basée sur un échantillon de taille  $n$ . Les estimateurs des proportions  $\tau_k$  suivent automatiquement.

### Représentation de $\hat{\lambda}_k$ en tant que $L$ -Statistique

Une application immédiate de (8) est de proposer un estimateur sans biais de  $\lambda_k$  (*Hosking & Wallis*), qui a la forme suivante :

$$\hat{\lambda}_k := n^{-1} \sum_{r=1}^n \omega_{r,n}^{(k)} X_{r,n}. \quad (9)$$

### Représentation de $\hat{\lambda}_k$ en tant que $U$ -Statistique

D'après la formule (9) on a :

Pour  $k = 1$  :  $\hat{\lambda}_1 = \bar{X}$ , moyenne de l'échantillon.

Pour  $k = 2$  :  $\hat{\lambda}_2 = \frac{1}{2}G$ .

La différence moyenne de *Gini* :

$$G = \binom{n}{2}^{-1} \sum_{1 \leq i < j \leq n} |X_i - X_j|.$$

Non seulement :  $\hat{\lambda}_1$  et  $\hat{\lambda}_2$  sont des  $U$ -statistiques, mais également  $\hat{\lambda}_k$  est une  $U$ -statistique.

D'abord notons que  $\lambda_k = E(h(X_1, \dots, X_k))$  et à partir de la formule (1) on obtient :

$$h(x_1, \dots, x_k) = k^{-1} \sum_{j=0}^{k-1} (-1)^j \binom{k-1}{j} x_{k-j,k}. \quad (10)$$

En général, pour tout noyau  $h(x_1, \dots, x_k)$  qui est une combinaison linéaire des so, il n'est pas difficile de montrer que les  $U$ -statistiques basées sur un échantillon de taille  $n$  peuvent être exprimées aussi comme une combinaison linéaire des so d'un échantillon.

D'une simple dérivation, ou en utilisant une technique de *Blom*, on obtient les coefficients appropriés à cette combinaison linéaire d'un polynôme associé à un noyau donné.

Par conséquent, les  $U$ -statistiques basées sur le noyau (10) s'avèrent conforme à la représentations des  $L$ -statistiques données par (9).

**Estimateur de  $\lambda_k$** 

Soit  $\hat{F}_n$  désignant l'estimateur de la fd basé sur l'échantillon  $X_1, \dots, X_n$  avec  $\hat{F}_n^{-1}$  la fonction de quantile correspondante, le remplacement de  $\hat{F}_n^{-1}$  dans (3) donne un autre estimateur des  $L$ -statistiques :

$$\hat{\lambda}_k^* = \int_0^1 \hat{F}_n^{-1}(u) P_{k-1}^*(u) du = \sum_{i=1}^n c_{n,i} X_{i:n}. \quad (11)$$

Où :  $c_{n,i} = \int_{(i-1)/n}^{i/n} P_{k-1}^*(u) du.$

- Pour  $k = 1$  :  $\hat{\lambda}_1^* = \hat{\lambda}_1 = \bar{X} \longrightarrow$  les estimateurs conviennent.

- Pour  $k = 2$  :  $\hat{\lambda}_2^* = n^{-1}(n-1)\hat{\lambda}_2 \longrightarrow$  les estimateurs sont proches.

Par une dérivation il suit qu'en général le  $\hat{\lambda}_k^* - \hat{\lambda}_k = o_p(n^{-1})$ ,  $n \longrightarrow \infty$ , ce qui donne l'équivalence de ces estimateurs à l'égard de la théorie de distribution asymptotique.

**1.2.3 Théorie de la distribution asymptotique pour  $\hat{\lambda}_k$  et  $\hat{\tau}_k$** 

Des résultats antérieurs, sous les conditions du second moment sur  $F$ , la théorie standard pour les  $U$ -statistiques et les  $L$ -statistiques montrent que le vecteur des  $k$  premiers  $L$ -moments est asymptotiquement normal de  $k$ -va, avec un résultat similaire pour le vecteur des proportions des  $L$ -moments.



### 1.3 Conclusion

Plusieurs applications de la théorie des  $L$ -moments ont été faites, tant pour l'estimation des paramètres, le choix de la meilleure distribution que dans les études de régionalisation. De nombreuses études rappellent un des avantages des  $L$ -moments, soit leur robustesse face aux valeurs extrêmes. Ce point a été critiqué par *Bernier* (1993), dans le sens où le diagramme des  $L$ -moments peut éliminer l'information issue d'une observation extrême qui est considérée comme étant aberrante, alors qu'il s'agit d'une information très précieuse si elle est réelle.

*Bobée & Rasmussen* (1995) expliquent, bien que cette faible sensibilité aux valeurs extrêmes puisse sembler souhaitable, il se peut que les estimateurs des  $L$ -moments soient trop robustes. Si on accepte l'hypothèse que les valeurs des échantillons donnant de l'information sont proches sur la queue de distribution, il pourrait alors être préférable de choisir une méthode d'estimation qui soit suffisamment sensible pour refléter cette information.

## Chapitre 2

# CONTRIBUTION AUX $L$ -MOMENTS MULTIVARIÉS

Pour obtenir une extension multivariée, on définit la notion des  $L$ -comoments qui maintient les mêmes caractéristiques des  $L$ -moments et inclut des notions analogues aux moments centraux classiques de la covariance, du coskewness et du cokurtosis. En conséquence, nos extensions multivariées des  $L$ -moments pour tout ordre supérieur à deux se composent des matrices  $L$ -covariance,  $L$ -coskewness,  $L$ -cokurtosis, qui ont des rôles analogues à la matrice classique de la covariance.

Une conclusion importante dans notre traitement est que si des va  $X$  et  $Y$  sont conjointement distribuées avec des distributions marginales  $F_X$  et  $F_Y$  respectivement. Alors sous les hypothèses du second moment l'échantillon de la  $L$ -corrélacion estime le même paramètre que l'échantillon de la corrélacion classique de Pearson et demeure également valable et significatif seulement sous l'hypothèse du premier moment.

C'est-à-dire, sous ces hypothèses, la  $L$ -corrélacion fournit une extension cohérente de la corrélacion classique de Pearson. Manifestement, ce résultat peut être clairement appliqué non seulement dans le modèle paramétrique, mais également dans le semi et non paramétrique.

### 2.1 $L$ -Comoments et $L$ -Moments multivariés

L'approche des  $L$ -moments offre de forts avantages pour la modélisation des distributions à queues-lourdes comme certaines distributions utilisées en hydrologie. Les  $L$ -moments multivariés sont principalement développés par *Serfling & Xiao* (2006).

Nous présentons dans ce chapitre les  $L$ -comoments, les  $L$ -moment multivariés, leur représentations en termes de concomitants, les estimateurs appropriés et la convergence asymptotique.

### 2.1.1 $L$ -moments multivariés

Pour un vecteur aléatoire de dimension  $d$ ,  $X = (X^{(1)}, \dots, X^{(d)})'$  ayant une fd  $F$  :

- Le premier  $L$ -moment multivarié est simplement la moyenne du vecteur  $X$  supposé fini :

$$\lambda_k := E(X); \quad k = 1.$$

- Pour  $k \geq 2$ , le  $k^{\text{ème}}$   $L$ -moment multivarié est la matrice du  $k^{\text{ème}}$   $L$ -comoment pour tous les couples  $(X^{(i)}, X^{(j)})_{1 \leq i, j \leq d}$  :

$$\Lambda_k := (\lambda_{k[ij]})_{d \times d}.$$

Avec  $\Lambda_2, \Lambda_3$  et  $\Lambda_4$  les matrices  $L$ -covariance,  $L$ -coskewness et  $L$ -cokurtosis respectivement.

Dans la suite les  $L$ -moments bivariés sont brièvement présentés.

### 2.1.2 $L$ -moments bivariés

Pour un vecteur aléatoire de dimension 2,  $X = (X^{(1)}, X^{(2)})'$  ayant la fd  $F$  avec une moyenne finie, des distributions marginales  $F_1, F_2$  et des séquences  $L$ -moments  $\{\lambda_k^{(1)}\}$  et  $\{\lambda_k^{(2)}\}$ .

Soit  $X^{(j)}$  une va de distribution  $F_j$ , pour  $j = 1, 2$ . Les  $L$ -moments multivariés  $\Lambda_k$  sont des matrices formées des éléments  $L$ -comoments, définis par :

$$\lambda_{k[ij]} := Cov(X^{(i)}, P_{k-1}^*(F_j(X^{(j)}))) ; \quad i, j = 1, 2 \text{ et } k \geq 2. \quad (12)$$

Où  $P_k^*$  est le polynôme de Legendre déplacé.

$$P_k^*(u) := \sum_{j=0}^k (-1)^{k-j} C_k^j C_{k+1}^j u^j ; \quad 0 < u < 1 ; \quad P_0^*(u) = 1.$$

Notons que les éléments  $\lambda_{k[ij]}$  et  $\lambda_{k[ji]}$  ne sont pas nécessairement égaux.

Par analogie avec la représentation de la covariance (5) pour les  $L$ -moments univariés et également par analogie avec les comoments centraux, nous définissons :

- Le  $k^{\text{ème}}$   $L$ -comoment de  $X^{(1)}$  par rapport à  $X^{(2)}$  d'ordre  $k \geq 2$  par :

$$\lambda_{k[12]} := Cov(X^{(1)}, P_{k-1}^*(F_2(X^{(2)}))).$$

- De même, le  $k^{\text{ème}}$   $L$ -comoment de  $X^{(2)}$  par rapport à  $X^{(1)}$  d'ordre  $k \geq 2$  par :

$$\lambda_{k[21]} := Cov(X^{(2)}, P_{k-1}^*(F_1(X^{(1)}))).$$

- La matrice des  $L$ -comoments dans le cas bivarié est :

$$\begin{aligned}\Lambda_k &= (\lambda_{k[ij]})_{1 \leq i, j \leq 2} \\ &= \begin{pmatrix} \lambda_{k[11]} & \lambda_{k[12]} \\ \lambda_{k[21]} & \lambda_{k[22]} \end{pmatrix} \\ &= \begin{pmatrix} \lambda_k^{(1)} & \lambda_{k[12]} \\ \lambda_{k[21]} & \lambda_k^{(2)} \end{pmatrix}.\end{aligned}$$

Proportions  $L$ -comoments :

$$\tau_{k[ij]} := \lambda_{k[ij]} / \lambda_2^{(i)}; \quad 1 \leq i, j \leq 2; \quad k \geq 3.$$

Des versions de mesures appropriées sont ainsi données par les proportions  $L$ -comoments :

$$\tau_{k[12]} := \lambda_{k[12]} / \lambda_2^{(1)}; \quad k \geq 3.$$

$$\tau_{k[21]} := \lambda_{k[21]} / \lambda_2^{(2)}; \quad k \geq 3.$$

Notons que :

- $\tau_{k[12]}$  : la  $L$ -correlation de  $X^{(1)}$  par rapport à  $X^{(2)}$ .
- $\tau_{k[21]}$  : la  $L$ -correlation de  $X^{(2)}$  par rapport à  $X^{(1)}$ .
- Pour  $k = 2$  :  $\lambda_{2[12]} = \lambda_{2[21]}$ .
- Pour  $k = 2$  :  $\tau_{2[12]} = \tau_{2[21]}$ .

Lorsque  $X^{(1)} = X^{(2)}$ ,  $\lambda_{k[12]} = \lambda_{k[21]} = \lambda_{k[11]} = \lambda_{k[22]}$ , que nous notons également  $\lambda_k^{(1)}$  et  $\lambda_k^{(2)}$ .

D'autre part lorsque  $X^{(1)}$  et  $X^{(2)}$  sont indépendants, les  $L$ -comoments correspondant à tout ordre  $k \geq 2$  sont égales à zéro.

En particulier, les premiers éléments  $L$ -comoments sont les suivants :

$$\begin{aligned}
\lambda_1 &:= E(X), \quad X = (X^{(1)}, X^{(2)})' \\
\lambda_{2[12]} &:= 2Cov(X^{(1)}, F_2(X^{(2)})) : L\text{-covariance.} \\
\lambda_{3[12]} &:= Cov(X^{(1)}, P_2^*(F_2(X^{(2)}))) \\
&= 6Cov\left(X^{(1)}, (F_2(X^{(2)}) - 1/2)^2\right) \\
&= -6Cov(X^{(1)}, F_2(X^{(2)})(1 - F_2(X^{(2)}))) \\
&: L\text{-coskewness.} \\
\lambda_{4[12]} &:= Cov(X^{(1)}, P_3^*(F_2(X^{(2)}))) \\
&= Cov\left(X^{(1)}, \begin{pmatrix} 20(F_2(X^{(2)}) - 1/2)^3 \\ -3(F_2(X^{(2)}) - 1/2) + 1 \end{pmatrix}\right) \\
&= Cov\left(X^{(1)}, \begin{pmatrix} 20F_2^3(X^{(2)}) - 30F_2^2(X^{(2)}) \\ +12F_2(X^{(2)}) - 1 \end{pmatrix}\right) \\
&: L\text{-cokurtosis.}
\end{aligned}$$

En particulier, nous appelons :

- $\lambda_{3[12]}$  et  $\lambda_{4[12]}$  le coskewness et le cokurtosis respectivement.
- $\tau_{3[12]}$  et  $\tau_{4[12]}$  les proportions du coskewness et du cokurtosis de  $X^{(1)}$  par rapport à  $X^{(2)}$  respectivement.

Les  $L$ -comoments fournissent ainsi une hiérarchie de la covariance classique et des comoments qui sont intuitivement analogues. Leurs interprétations et comparaisons sont facilitées par le fait qu'elles soient définies en termes d'opérateur classique de covariance.

La matrice des proportions  $L$ -comoments est donnée par :

$$\Lambda_k^* = (\tau_{k[ij]})_{i,j=1,2} := \begin{pmatrix} \tau_{k[11]} & \tau_{k[12]} \\ \tau_{k[21]} & \tau_{k[22]} \end{pmatrix}.$$

En particulier, pour  $k = 2$ , la matrice des proportions des  $L$ -covariances est :

$$\Lambda_2^* := \begin{pmatrix} \tau_{2[11]} & \tau_{2[12]} \\ \tau_{2[21]} & \tau_{2[22]} \end{pmatrix}.$$

**Théorème 2.2.** (*2<sup>ème</sup> L-comoment de la somme des va*)

*Avec les comoments nous pouvons construire des versions symétriques dans le cas des L-comoments de 2<sup>ème</sup> ordre.*

---

"L-Comoments et L-Moments multivariés"

---

En effet, les comoments de 2<sup>ème</sup> ordre et les corrélations relatives servent à la décomposition du 2<sup>ème</sup> L-comoment " $\lambda_2(S)$ " de la somme des va  $Y_1, \dots, Y_n$  tel que :

$$S = \sum_{i=1}^n Y_i = Y_1 + \dots + Y_n.$$

**Démonstration.** □

Le 2<sup>ème</sup> L-comoment de la somme des va est la somme des L-comoments d'ordres 2 et des L-corrélations d'ordres 2.

$$\begin{aligned} \lambda_2(S) &= 2Cov(S, F_S(S)) \\ &= 2Cov\left(\sum_{i=1}^n Y_i, F_S(S)\right) \\ &= 2 \sum_{i=1}^n Cov(Y_i, F_S(S)) \\ &= \sum_{i=1}^n 2Cov(Y_i, F_S(S)) \\ &= \sum_{i=1}^n \lambda_{2[Y_i, S]}. \end{aligned}$$

On a:  $\tau_{2[Y_i, S]} = \lambda_{2[Y_i, S]} / \lambda_2^{(Y_i)} \implies \lambda_{2[Y_i, S]} = \tau_{2[Y_i, S]} \lambda_2^{(Y_i)}$ .

Alors :

$$\lambda_2(S) = \sum_{i=1}^n \tau_{2[Y_i, S]} \lambda_2^{(Y_i)}.$$

□

### 2.3 Représentations de $\lambda_{k[12]}$ en termes de concomitants

De telles représentations servent de manière équivalente pour définir les  $L$ -comoments.

Considérons maintenant un échantillon bivarié  $\{(X_l^{(1)}, X_l^{(2)})\}_{1 \leq l \leq n}$  ayant distribution  $F(x^{(1)}, x^{(2)})$  avec les distributions marginales  $F_1$  et  $F_2$ . Correspondant aux valeurs  $X^{(2)}$  ordonnées  $X_{1:n}^{(2)} \leq X_{2:n}^{(2)} \leq \dots \leq X_{n:n}^{(2)}$ .

Nous appelons l'élément de  $\{X_1^{(1)}, \dots, X_n^{(1)}\} = \{X_l^{(1)}\}_{1 \leq l \leq n}$  qui est lié à  $X_{r,n}^{(2)}$ , le concomitant de  $X_{r,n}^{(2)}$  qui est noté par  $X_{[r,n]}^{(12)}$  (*David & Nagaraja, 2003*).

Il est clair que  $E(X_{[r,n]}^{(12)}) = nE(X_1^{(1)} | X_1^{(2)} = X_{r,n}^{(2)})$ , car :

$$\begin{aligned} E(X_{[r,n]}^{(12)}) &= \sum_{l=1}^n E(X_l^{(1)} | X_l^{(2)} = X_{r,n}^{(2)}) \\ &= nE(X_1^{(1)} | X_1^{(2)} = X_{r,n}^{(2)}). \end{aligned}$$

Cela conduit immédiatement à la formule suivante analogue à (6) :

$$\begin{aligned} E(X_{[r,n]}^{(12)}) &= n \binom{n-1}{r-1} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^{(1)} [F_2(x^{(2)})]^{r-1} [1 - F_2(x^{(2)})]^{n-r} dF(x^{(1)}, x^{(2)}) \\ &= n \binom{n-1}{r-1} E(X^{(1)} [F_2(X^{(2)})]^{r-1} [1 - F_2(X^{(2)})]^{n-r}); \quad r \leq n. \end{aligned} \tag{12*}$$

Nous utilisons (12) pour établir la représentation suivante, qui indique que les  $L$ -comoments peuvent être définis en termes de valeurs des concomitants exactement de la même manière que les  $L$ -moments sont définis en termes des valeurs des statistiques d'ordres.

**Definition 1.** Le  $k^{\text{ème}}$   $L$ -comoment de  $X^{(1)}$  par rapport à  $X^{(2)}$  peut être représenté comme suit :

$$\lambda_{k[12]} = k^{-1} \sum_{j=0}^{k-1} (-1)^j \binom{k-1}{j} E(X_{[k-j,k]}^{(12)}). \tag{13}$$

On pose  $L_k = \lambda_{k[12]}$ . En insérant (12\*) en (13) qui sont semblables à l'insertion de (6) en (1), cela donne les  $L$ -comoments analogues à (5) :

$$L_k = Cov(X^{(1)}, P_{k-1}^*(F_2(X^{(2)}))); \quad k \geq 2.$$



---

"Estimation des L-Comoments"

---

Cependant le rôle principal de la proposition précédente est de rendre simple l'obtention des principaux résultats des  $L$ -comoments analogues à ceux des  $L$ -moments, avec des concomitants à la place des so.

Ainsi, par la définition  $\chi_{[j,k]}^{(12)} = E(X_{[j+1,k]}^{(12)}) - E(X_{[j,k]}^{(12)})$  pour  $k \geq 2$ , d'après la formule (2) nous obtenons :

$$\lambda_{k[12]} = k^{-1} \sum_{j=0}^{k-2} (-1)^j \binom{k-2}{j} \chi_{[k-1-j,k]}^{(12)}.$$

**Remarque :**

Pour un échantillon de taille  $n$  et à la base de l'estimation sans biais des comoments, pour  $k \geq 2$ , on a la formule suivante analogue à (8) :

$$\lambda_{k[12]} = n^{-1} \sum_{r=1}^n \omega_{r,n}^{(k)} E(X_{[r,n]}^{(12)}). \quad (14)$$

## 2.4 Estimation des $L$ -Comoments

En utilisant la proposition précédente on obtient un estimateur sans biais du  $k^{\text{ème}}$   $L$ -comoment :

$$\hat{\lambda}_{k[12]} = n^{-1} \sum_{r=1}^n \omega_{r,n}^{(k)} X_{[r,n]}^{(12)}, \quad k \geq 2. \quad (15)$$

Qui est une  $L$ -statistique des concomitants. En outre, chaque  $\hat{\lambda}_{k[12]}$  est une  $U$ -statistique.

D'après la proposition précédente :

$$\begin{aligned} \lambda_{k[12]} &= k^{-1} \sum_{j=0}^{k-1} (-1)^j \binom{k-1}{j} E(X_{[k-j,k]}^{(12)}). \\ \lambda_{k[12]} &= E(h^{(k)}((X_1^{(1)}, X_1^{(2)}), \dots, (X_k^{(1)}, X_k^{(2)}))). \end{aligned}$$

Donc :

$$h^{(k)}((x_1^{(1)}, x_1^{(2)}), \dots, (x_k^{(1)}, x_k^{(2)})) = k^{-1} \sum_{j=0}^{k-1} (-1)^j \binom{k-1}{j} x_{[k-j,k]}^{(12)}. \quad (16)$$

Pour un noyau  $h$  avec des arguments bivariés qui est une combinaison linéaire des concomitants, les  $U$ -statistiques correspondantes basées sur un échantillon de taille  $n$

peuvent être exprimés comme une combinaison linéaire des concomitants. Ainsi, la  $U$ -statistique basée sur le noyau (16) s'avère conforme à la représentation de la  $L$ -statistique donnée par (15).

#### 2.4.1 Exemple : (*Estimation du 2<sup>ème</sup> L-Comoment*)

Pour  $k = 2$ , d'après (15) et (16), on obtient les représentations des  $L$ -statistiques et des  $U$ -statistiques :

$$\begin{aligned}\hat{\lambda}_{2[12]} &= n^{-1} \sum_{r=1}^n \omega_{r,n}^{(2)} X_{[r,n]}^{(12)} \\ &= n^{-1} \sum_{r=1}^n \frac{2r - n - 1}{n - 1} X_{[r,n]}^{(12)}, \text{ car } \omega_{r,n}^{(2)} = \frac{2r - n - 1}{n - 1} \\ &= \binom{n}{2}^{-1} \sum_{1 \leq i < j \leq n} (X_{[j,n]}^{(12)} - X_{[i,n]}^{(12)})/2.\end{aligned}$$

Cependant la représentation de  $U$ -statistique ne peut pas être réexprimée en tant que la moitié de la différence moyenne de Gini des concomitants, car le noyau des concomitants  $(x_{[2,2]}^{(12)} - x_{[1,2]}^{(12)})/2$ , n'est pas le même que le noyau  $|x_{[2,2]}^{(12)} - x_{[1,2]}^{(12)}|/2$  de la différence moyenne de Gini.

## 2.5 Convergence asymptotique

La distribution asymptotique d'un vecteur des estimateurs des  $L$ -comoments résulte de la théorie standard des  $U$ -statistiques (*Serfling*).

Sous les hypothèses du seconds moment, pour  $k \geq 2$ , le vecteur des estimateurs des  $L$ -comoments  $(\hat{\lambda}_{2[12]}, \dots, \hat{\lambda}_{k[12]})'$  est asymptotiquement normal.

Alternativement, le vecteur correspondant aux estimateurs des proportions  $\hat{\tau}_{i[12]}$ ,  $2 \leq i \leq k$ , est asymptotiquement normal.

"Conclusion"

---

## 2.6 Conclusion

Les  $L$ -moments multivariés définis précédemment sont fondés sur une distribution de population théorique. Cependant leurs versions d'échantillon fini sont utiles pour définir des tests statistiques et aussi pour l'estimation des paramètres des distributions multivariées.

### Chapitre 3

## THÉORIE DES COPULES

On a pu observer que les séries de données issues de  $va$  très proches géographiquement peuvent être fortement corrélées. L'hypothèse d'indépendance des séries n'est alors pas valable. Une approche intéressante consiste alors à utiliser des copules, qui sont des fonctions permettant de modéliser la structure de dépendance indépendamment des distributions marginales. Autrement dit, elles permettent de décrire le comportement individuel de chaque  $va$  et couplent les lois marginales pour obtenir la loi jointe. Grâce à elles, on peut donc former des distributions multidimensionnelles avec différentes marginales, la structure de dépendance sera donnée par la copule.

Jusqu'à présent la plupart des techniques utilisées présentaient certains inconvénients. Nous pouvons, à titre d'exemple, citer l'utilisation de distributions bidimensionnelles comme la Pareto bivariée. Mais dans ce cas précis, les lois marginales seront nécessairement des lois de Pareto. L'idéal serait donc de pouvoir créer des distributions multivariées et à fortiori bivariées en choisissant nous mêmes nos lois marginales. De manière générale les copules sont des fonctions qui vont constituer un véritable lien entre la distribution multivariée et les lois marginales.

Dans la littérature, en dehors du domaine de la statistique, c'est surtout en finance et en actuariat que l'on retrouve l'approche par copules. Les livres de *Joe* (1997) et de *Nelsen* (1999) présentent une introduction générale aux copules. Les copules sont utiles pour la compréhension et la modélisation dans différents domaines. On peut se référer au livre de *Salvadori & al.* (2007) pour des applications au domaine des événements extrêmes météorologiques. Elle a été aussi largement appliquée en biologie (*Frees & Valdez*, 1998), en assurance et réassurance (*Klugman & Parsa*, 1999 ; *Belguise*, 2001 ; *Venter*, 2002, 2003 ; *Charpentier*, 2003).

Dans le domaine spécifique de l'hydrologie, on peut citer *Favre & al.* (2004), *Genest & Favre* (2007) et *Genest & al.* (2007) présentant notamment deux applications. La première concerne la modélisation des débits d'un cours d'eau avec tributaire et la seconde concerne la modélisation conjointe des débits maximaux annuels en fonction des volumes écoulés. *De Michele & al.* (2005) utilisent la copule de Gumbel pour modéliser la dépendance positive entre le débit de pointe et le volume de crue. *Zhang & Singh* (2006) utilisent les copules bivariées pour l'analyse fréquentielle des variables : débit de pointe et volume écoulé, volume de crue et sa durée.

Dans la section 1 nous présentons d'abord les copules bivariées ainsi que leurs propriétés élémentaires puis on établit le lien entre ces copules et les  $va$  via le théorème de *Sklar* qui clarifiera la notion de dépendance.

## "Introduction aux copules"

---

Nous discuterons dans la section 2 des familles paramétriques usuelles des copules bivariées. Cette partie mettra également en avant les copules utilisées au cours de notre étude, et donnera toutes les formules nécessaires.

La section 3 traite les notions de concordance à travers des mesures de concordance telles que le tau de Kendall et le rho de Spearman.

Enfin nous présentons dans la section 4 les  $L$ -comoments en termes de copules.

### 3.1 Introduction aux copules

L'étude des copules (copulas en anglais) et de leurs applications en statistiques est un phénomène récent qui date de la fin des années 50. Il y a encore 20 ans, il était malgré tout difficile de trouver le mot copule dans la littérature statistique. Que sont les copules ?

Les copules sont des fonctions qui couplent les fonctions de répartition multivariées avec les fonctions de répartition univariées (lois marginales). Alternativement, les couples sont des fonctions de distribution multivariées (théorème de *Sklar*) dont les marginales unidimensionnelles sont uniformes sur l'intervalle  $[0, 1]$ .

Comme le dit *Fisher* (1997), les copules intéressent les statisticiens pour deux raisons. Premièrement comme manière de mesurer la dépendance et deuxièmement comme point de départ pour construire des familles de lois bivariées.

Le mot latin "copula" signifie "un lien". Il a été utilisé en mathématique par *Abe Sklar* (1959) dans le théorème qui porte son nom.

Pour amener progressivement le lecteur vers les copules multivariées, nous présentons en premier lieu les copules bivariées et leurs propriétés élémentaires avant d'établir, dans la deuxième section, le lien entre ces copules et les variables aléatoires via le théorème de *Sklar*. Les familles usuelles des copules bivariées sont données dans la troisième section.

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.

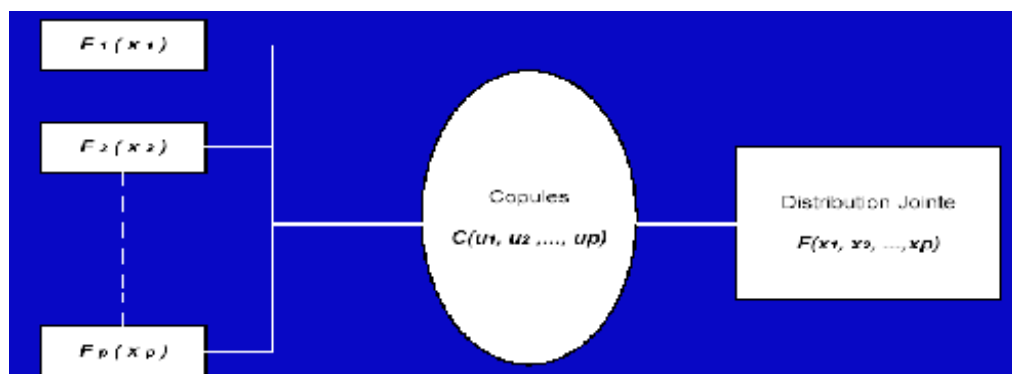


FIG. 3.1. Notion de copule

### 3.2 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.

Les fonctions copules permettent une représentation de la dépendance statistique fondée sur la mesure traditionnelle de corrélation qui présente les limites dans l'étude de la dépendance entre deux variables (*Embrechts & al., 1999*).

Elles permettent la construction de distributions multidimensionnelles assez générales et ce, indépendamment des lois marginales qui peuvent avoir des lois différentes et quelconques.

Par ailleurs, l'approche par les copules a beaucoup contribué 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 possible des extensions naturelles de certains résultats obtenus dans le cas univarié au cas multivarié.

Les distributions multidimensionnelles ainsi obtenues sont davantage en adéquation avec la réalité surtout dans l'utilisation financière des statistiques.

L'idée de base des copules est de travailler sur le couple  $(F_X(x), F_Y(y))$  et non pas sur le couple  $(X, Y)$ . On travaille donc avec un vecteur aléatoire  $(U = F_X(x), V = F_Y(y))$  dont les lois marginales suivent une loi uniforme sur  $[0, 1]$ .

Les copules serviront pour la modélisation de la loi conjointe entre les va  $X$  et  $Y$ .

### 3.2.1 Définition d'une copule bivariée (dimension 2)

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

Comme dans la théorie univariée le terme "distribution bivariée" sera associé à la fd d'un couple de va continues  $(U, V)$ .

**Definition 2.** On appelle copule bivariée toute fonction  $C$  définie de  $I^2 \rightarrow I$  vérifiant les propriétés suivantes :

$$i/ \quad \forall u, v \in I : \begin{cases} C(u, 0) = C(0, v) = 0. \\ C(u, 1) = u \text{ et } C(1, v) = v. \end{cases}$$

ii/  $C$  est 2-croissante (2-increasing) i.e.

$$\forall u, u', v, v' \in I : C(u, v) - C(u, v') - C(u', v) + C(u', v') \geq 0; \quad u \leq u' \text{ et } v \leq v'.$$

L'interprétation de cette notion de croissance se fait en notant que, soient  $U$  et  $V$  deux va uniformes sur  $I$  et  $(U, V)$  admet une fd  $C$ , alors on a :

$$\mathbb{P}(u \leq U \leq u', v \leq V \leq v') = C(u, v) - C(u, v') - C(u', v) + C(u', v') \geq 0.$$

La mesure  $\mathbb{P}$  étant nécessairement positive.

On établira par la suite le lien entre les propriétés  $i/$  et  $ii/$  :

- Propriété  $i/$  : signifie que toute copule est une distribution dont les distributions marginales sont de loi uniforme définie sur  $I$ .

En effet en considérant le vecteur aléatoire  $(U, V)$  où  $U$  et  $V$  sont deux va uniformes sur  $I$  alors on a :

$$C(u, v) := P(U \leq u, V \leq v).$$

- Propriété  $ii/$  : est la 2-croissance ou inégalité du rectangle d'une copule de la distribution  $C$ . Elle se résume à démontrer la positivité de sa densité, lorsque celle-ci existe, i.e. elle traduit le fait que si  $C$  admet une densité  $c(u, v)$  alors elle est positive :

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

1

<sup>1</sup>Rappelons qu'une loi uniforme  $U$  sur  $[0, 1]$  a pour fonction de répartition :

$$P(U \leq u) = \begin{cases} 0 & \text{si } u \leq 0 \\ u & \text{si } 0 < u < 1 \\ 1 & \text{si } u \geq 1 \end{cases}$$

**Exemple 1.** (*Exemples de copules*)

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

En effet :  $\forall u, v \in I$  :

$$\bullet \left\{ \begin{array}{l} M(u, 0) = M(0, v) = 0 \\ M(u, 1) = u \text{ et } M(1, v) = v \end{array} \right. \implies \left\{ \begin{array}{l} \min(u, 0) = \min(0, v) = 0. \\ \min(u, 1) = u \text{ et } \min(1, v) = v. \end{array} \right.$$

$$\implies M \text{ vérifie } i/.$$

- De même  $C$  est 2-croissante (2-increasing) :

$$\forall u, u', v, v' \in I : C(u, v) - C(u, v') - C(u', v) + C(u', v') \geq 0; u \leq u' \text{ et } v \leq v'.$$

$$\forall u, u', v, v' \in I : \min(u, v') \geq \min(u, v) \text{ et } \min(u', v') \geq \min(u', v);, u \leq u' \text{ et } v \leq v'.$$

$$\implies M \text{ vérifie } ii/.$$

Par conséquent  $M$  est une copule.

2. De la même façon on établit que les fonctions :  $W(u, v) = \max(u + v - 1, 0)$  et  $\Pi(u, v) = uv$  définissent aussi des copules.

$M, W$  et  $\Pi$  sont des copules usuelles.

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

**3.2.2 Propriétés immédiates des copules bivariées**

La dépendance entre les  $v$  est parfaitement décrite par leur distribution jointe. Nous pouvons cependant distinguer les comportements des distributions marginales de la structure de dépendance. La copule est l'outil permettant d'extraire la structure de dépendance d'une distribution jointe et ainsi de séparer dépendance et comportement marginal.

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. En 1959 *Abel Sklar* apporte une solution partielle à ce problème pour les distributions marginales univariées.

Nous donnons d'abord le théorème de *Sklar* dans sa version probabiliste.

**Théorème de Sklar**

Par un théorème qui porte son nom, *Sklar* établit le lien défini par la copule  $C$  entre les df marginales univariées  $F_1$  et  $F_2$  et la distribution jointe bivariée  $F$ .



"Généralités sur les copules bivariées"

---

**Théorème 3.3.** (*Théorème de Sklar*)

- Soit  $F$  une fd jointe de dimension  $d$  dont les marges (univariées) sont précisément  $F_1, \dots, F_d$ . La copule  $C$  associée à  $F$  s'écrit :

$$\begin{aligned} C(u_1, \dots, u_d) &= C(F_1(x_1), \dots, F_d(x_d)) \\ &= F(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d)). \end{aligned}$$

Ce théorème permet d'associer à chaque distribution multidimensionnelle une copule.

- Réciproquement, si  $F$  est une fd jointe de dimension  $d$ , il existe une copule  $C$  telle que :

$$F(x_1, \dots, x_d) := C(F_1(x_1), \dots, F_d(x_d)).$$

Où les  $F_i$  sont les lois marginales de  $F$ .

**Remarques :**

1. Si ces lois marginales sont toutes continues, la copule  $C$  du vecteur  $(X_1, \dots, X_d)$  est alors unique et donnée par la relation :

$$C(u_1, \dots, u_d) := F(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d)).$$

2. La copule d'un vecteur aléatoire  $(X_1, \dots, X_d)$  est alors la fd du vecteur aléatoire  $(F_1(X_1), \dots, F_d(X_d))$  que l'on note  $(U_1, \dots, U_d)$ .
3. La densité  $f$  d'une loi bivariée peut s'écrire aussi en fonction de la densité  $c$  de la copule associée et des densités des marginales  $f_X$  et  $f_Y$ .

Si la distribution bivariée  $F$  est absolument continue, alors elle admet une densité  $f$  :

$$f(x, y) := \frac{\partial^2 C(F_X(x), F_Y(y))}{\partial x \partial y} = c(F_X(x), F_Y(y)) \times f_X(x) \times f_Y(y).$$

Précisons que :

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

La copule  $C$  est dite copule des va  $U$  et  $V$  ou copule de la distribution  $F$ . On la note  $C$ .

**Exemple 2.** (*Copule produit*)

Dans le cas de la copule produit nous avons :

$$C^\perp(u, v) := uv \implies c^\perp(u, v) = \frac{\partial^2}{\partial u \partial v} C^\perp(u, v) = 1.$$

## "Généralités sur les copules bivariées"

---

Cette copule caractérise l'indépendance entre deux va puisque la densité bivariée d'un vecteur aléatoire  $(U, V)$  dont la structure de dépendance est déterminée par la copule produit s'écrit :

$$f(x, y) = c(F_X(x), F_Y(y)) \times f_X(x) \times f_Y(y) = f_X(x) \times f_Y(y).$$

### Bornes de Fréchet

Définissons tout d'abord deux copules importantes :

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

Les copules minimum  $C^-$  et maximum  $C^+$  sont essentielles puisqu'elles représentent les extrêmes de toute copule, que l'on appelle les bornes de *Fréchet*.

### Théorème 3.4. (Bornes de Fréchet)

Pour toute copule  $C$  on a :

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

En effet :

$$\text{a) } \forall u, v \in I : \begin{cases} C(u, v) \leq C(u, 1) = u \\ C(u, v) \leq C(1, v) = v. \end{cases}$$

Donc :

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

b) La propriété ii/  $\implies$

$$\forall u, v \in I : \begin{cases} C(u, v) \geq C(u, 1) + C(1, v) - C(1, 1). \\ C(u, v) \geq u + v - 1. \\ \text{Or : } C(u, v) \geq 0. \end{cases}$$

Alors :

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

De (a) et (b)  $\implies \max(u + v - 1, 0) \leq C(u, v) \leq M(u, v)$ .

$$\forall u, v \in I^2 : W(u, v) \leq C(u, v) \leq M(u, v).$$

Tel que :

- Borne inférieure de *Fréchet-Hoeffding* (copule minimale) :  $W(u, v) := \max(u + v - 1, 0)$ .

---

"Généralités sur les copules bivariées"

---

- Borne supérieure de *Fréchet-Hoeffding* (copule maximale) :  $M(u, v) := \min(u, v)$ .

### Théorème d'invariance

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

#### Théorème 3.5. (*Théorème d'invariance*)

Soient deux va continues  $U$  et  $V$  de marges  $F_X$  et  $F_Y$  et de copule  $C_{XY}$  associée à la distribution  $F$  du vecteur aléatoire  $U = (U, V)$ . Si  $h_X$  et  $h_Y$  sont deux fonctions strictement croissantes sur  $\text{Im}(X)$  et  $\text{Im}(Y)$  respectivement, alors :

$$C_{h_X h_Y}(u, v) = C_{XY}(u, v).$$

La copule est invariante par transformations strictement croissantes des va.

### Dépendance de queue

Définissons 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 :

$$\bullet \mathbb{P}[U \leq u | V \leq v] = \frac{\mathbb{P}[U \leq u, V \leq v]}{\mathbb{P}[V \leq v]} = \frac{C(u, v)}{v}.$$

$$\mathbb{P}[U \leq u | V \leq v] = \frac{C(u, v)}{v}.$$

"Généralités sur les copules bivariées"

---

$$\begin{aligned}
 \bullet \mathbb{P}[U > u|V > v] &= \frac{\mathbb{P}[U > u, V > v]}{\mathbb{P}[V > v]} \\
 &= \frac{\mathbb{P}[U \leq 1, V > v] - \mathbb{P}[U \leq u, V > v]}{1 - v}, \text{ car } \mathbb{P}[V \leq v] = v \\
 &= \frac{\mathbb{P}[U \leq 1, V \leq 1] - \mathbb{P}[U \leq 1, V \leq v] - \mathbb{P}[U \leq u, V > v]}{1 - v} \\
 &= \frac{1 - v - \mathbb{P}[U \leq u, V \leq 1] + \mathbb{P}[U \leq u, V \leq v]}{1 - v} \\
 &= \frac{1 - v - u + C(u, v)}{1 - v}.
 \end{aligned}$$

$$\mathbb{P}[U > u|V > v] = \frac{1 - v - u + C(u, v)}{1 - v}.$$

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

**Definition 3.** (*Lower tail dependence*)

Une copule  $C$  a une dépendance de queue à gauche si :

$$\lambda_L := \lim_{u \rightarrow 0^+} \frac{C(u, u)}{u} \text{ existe et } \lambda_L \in I.$$

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

**Definition 4.** (*Upper tail dependence*)

Une copule  $C$  a une dépendance de queue à droite si :

$$\lambda_U := \lim_{u \rightarrow 1^-} \frac{1 - 2u + C(u, u)}{1 - u} \text{ existe et } \lambda_U \in I.$$

Si  $\lambda_U = 0$  alors elle n'a pas de dépendance de queue à droite.

## 3.6 Applications des copules

Le nom copule a été choisi pour exprimer la manière avec laquelle une copule couple une jointe et les marginales univariées. Le théorème de *Sklar* est le résultat le plus important concernant les copules et est utilisé dans toutes les applications des copules pour la modélisation statistique.

### 3.6.1 Application à la construction des copules

Soient  $X_1$  et  $X_2$  deux va de distributions respectives  $F_1$  et  $F_2$  supposées strictement monotones. On peut, à partir de leur distribution conjointe  $F$ , construire la copule associée aux deux variables  $X_1$  et  $X_2$ . Introduisons pour ce fait, la définition suivante pour généraliser cette construction aux distributions.

**Definition 5.** (*Pseudo-inverse ou inverse généralisé*)

Soit  $F$  une distribution. On définit la pseudo-inverse  $F^{(-1)}$  de  $F$  telle que :

i/ Si  $t \in \text{Im } F$ , alors :  $F^{(-1)}(t) := \{x \in \mathbb{R} / F(x) = t\}$ .

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

Si la fonction  $F$  est strictement croissante alors la notion de pseudo-inverse coïncide avec la notion d'inverse i.e.  $F^{(-1)} = F^{-1}$ .

En utilisant la notion de pseudo-inverse dans le théorème de *Sklar* on obtient le résultat suivant :

Soit  $F$  une distribution bivariée dont les distributions marginales univariées sont  $F_1$  et  $F_2$  de pseudo-inverses respectives  $F_1^{-1}$  et  $F_2^{-1}$ . Alors il existe une copule  $C$  telle que :

$$\forall u, v \in I : C(u, v) := F(F_1^{-1}(u), F_2^{-1}(v)).$$

Le corolaire donne ainsi une procédure de construction d'une copule associée à un couple de va connaissant la distribution conjointe et les distributions marginales continues.

"Applications des copules"

---

**Exemple 3.** (*Construction de copules*)

1. Copule de Galambos :

Pour tout réel  $\theta \geq 1$ ,  $x$  et  $y$  sont définies sur  $[0, +\infty[ \times [0, +\infty[$ .

La distribution bivariable est donnée par :

$$\forall x, y \in \mathbb{R} : F_\theta(x, y) = \exp \left\{ - \left[ (x + y) - (x^{-\theta} + y^{-\theta})^{-1/\theta} \right] \right\}.$$

$\forall x \in [0, +\infty[ : F_1(x) = F_2(x) = \exp(-x)$ , les marges  $F_1$  et  $F_2$  respectivement.

$\forall u \in [0, 1[ : F_1^{-1}(u) = F_2^{-1}(u) = -\ln u$ , les fonctions quantiles de  $F_1$  et  $F_2$  respectivement.

La copule associée à  $F_\theta$  est donnée par :

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

$C_\theta(u, v)$  est la copule de Galambos.

2. Copule normale :

Soit  $F_\rho$  une distribution normale bivariable de coefficient de corrélation  $\rho$  telle que :

$$\begin{aligned} \forall x, y \in \mathbb{R} : F_\rho(x, y) &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \int_{-\infty}^x \int_{-\infty}^y \exp \left\{ \frac{-1}{2(1-\rho^2)} \left( \left( \frac{t_1 - m_1}{\sigma_1} \right)^2 \right. \right. \\ &\quad \left. \left. - \frac{2\rho(t_1 - m_1)(t_2 - m_2)}{\sigma_1\sigma_2} + \left( \frac{t_2 - m_2}{\sigma_2} \right)^2 \right) \right\} dt_1 dt_2. \end{aligned}$$

Supposons que  $F_\rho$  est la loi conjointe des va normales  $X_1 \sim N(m_1, \sigma_1)$  et  $X_2 \sim N(m_2, \sigma_2)$ .

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  $F_\rho$  est alors donnée par :

$$\begin{aligned} \forall u, v \in I^2 : C_\rho(u, v) &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \int_{-\infty}^{F_1^{-1}(u)} \int_{-\infty}^{F_2^{-1}(v)} \exp \left\{ \frac{-1}{2(1-\rho^2)} \left( \left( \frac{x - m_1}{\sigma_1} \right)^2 \right. \right. \\ &\quad \left. \left. - \frac{2\rho(x - m_1)(y - m_2)}{\sigma_1\sigma_2} + \left( \frac{y - m_2}{\sigma_2} \right)^2 \right) \right\} dx dy. \end{aligned}$$

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

## "Applications des copules"

---

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

$$\forall u, v \in I : C_\rho(u, v) = \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{F_1^{-1}(u)} \int_{-\infty}^{F_2^{-1}(v)} \exp\left\{-\frac{x^2 - 2\rho xy + y^2}{2(1-\rho^2)}\right\} dx dy.$$

Est la copule de Gauss de corrélation  $\rho$ .

Comme sa distribution, la copule de Gauss est souvent caractérisée par sa densité  $c_\rho$ .

Soit  $f_\rho$  la densité associée à la distribution  $F_\rho$ , on déduit que :

$$\forall x, y \in \mathbb{R} : f_\rho(x, y) = f_1(x)f_2(y)c_\rho(F_1(x), F_2(y)).$$

D'où :

$$\begin{aligned} f_\rho(x, y) &= \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{-\frac{x^2 - 2\rho xy + y^2}{2(1-\rho^2)}\right\} \\ &= \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{x^2}{2}\right\} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{y^2}{2}\right\} c_\rho(F_1(x), F_2(y)). \end{aligned}$$

Par suite :

$$\forall x, y \in \mathbb{R}^2 : c_\rho(F_1(x), F_2(y)) = \frac{1}{\sqrt{1-\rho^2}} \exp\left\{-\frac{x^2 - 2\rho xy + y^2}{2(1-\rho^2)} + \frac{x^2 + y^2}{2}\right\}.$$

### 3.6.2 Application à la modélisation multivariée des distributions

Le corolaire précédent signifie d'autre part que si  $F$  est une distribution bivariée continue de distributions marginales  $F_1$  et  $F_2$  dont les inverses sont  $F_1^{-1}$  et  $F_2^{-1}$ , alors la fonction  $C$  définie telle que :  $\forall u, v \in I : C(u, v) = F(F_1^{-1}(u), F_2^{-1}(v))$  est l'unique copule vérifiant la représentation canonique de  $F$ .

En effet l'unicité résulte des propriétés suivantes :

#### Propriétés :

- i/ Si  $X$  est une va continue de distribution  $F$  alors la variable  $U = F(X)$  suit une loi uniforme  $U_{[0,1]}$ .
- ii/ Si  $F$  est une distribution continue d'inverse  $F^{-1}$  alors :
  - Si  $(X, Y) \sim F$  de marges  $F_1$  et  $F_2$  alors :  $C(F_1(x), F_2(y)) \sim C$ .
  - Réciproquement si  $(U, V) \sim C$  alors :  $F(F_1^{-1}(u), F_2^{-1}(v)) \sim F$ .

#### Conclusion :

- i/  $\forall u, v \in I : C(u, 1) = C(1, u) = u$  signifie que les distributions marginales de cette distribution sont de lois uniformes.

---

"Applications des copules"

---

ii/ De plus, toute copule est une distribution indépendante de ses distributions marginales.

Ainsi, le théorème de *Sklar* fait des copules un outil puissant à l'analyse multivariée car elles permettent de construire des modèles de distributions multivariées compatibles avec les modèles marginaux unidimensionnels (puisqu'on part de ces marges).

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 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.

**Exemple 4.** (*Construction de distributions*)

Partant de la copule de Galambos :

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

Construisons la distribution bivariable  $F_\theta$  ayant une marge normale  $N(0, 1)$  et une marge uniforme  $U_{[0,1]}$ .

$$x \sim N(0, 1) : F_1(x) = \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{x^2}{2} \right\} = \Phi(x); x \geq 0.$$

Où  $\Phi$  est la fonction de répartition de la loi normale standard.

$$y \sim U_{[0,1]} : F_2(y) = y; 0 \leq y \leq 1.$$

$\forall \theta \geq 1; \forall x, y \in \mathbb{R} :$

$$\begin{aligned} F_\theta(x, y) &= C_\theta(F_1(x), F_2(y)) \\ &= C_\theta(\Phi(x), y) \\ &= \Phi(x)y \exp \left\{ - \left[ (-\ln \Phi(x))^{-\theta} + (-\ln y)^{-\theta} \right]^{-\frac{1}{\theta}} \right\} \\ &= \frac{1}{\sqrt{2\pi}} \exp \left\{ -\frac{x^2}{2} \right\} y \exp \left\{ - \left[ (-\ln \Phi(x))^{-\theta} + (-\ln y)^{-\theta} \right]^{-\frac{1}{\theta}} \right\}. \\ F_\theta(x, y) &= \frac{y}{\sqrt{2\pi}} \exp \left\{ -\frac{x^2}{2} - \left[ (-\ln \Phi(x))^{-\theta} + (-\ln y)^{-\theta} \right]^{-\frac{1}{\theta}} \right\}. \end{aligned}$$



### 3.7 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*).
- D'autre part elles constituent un outil de mesure de dépendance entre les distributions univariées tout en restant invariantes sous des transformations strictement monotones de celles-ci.

Les résultats suivants résument les propriétés fondamentales des copules.

**Théorème 3.8.** (*Caractérisation de la copule indépendante*)

Soient  $X, Y$  deux va continues et  $C_{XY}$  la copule associée.  $X$  et  $Y$  sont indépendantes si et seulement si  $C_{XY} = \Pi$ .

**Démonstration.** □

En effet :

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

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

$$C_{XY}(u, v) = \Pi(u, v).$$

$\impliedby$  Si  $C_{XY} = \Pi$  alors :  $\forall u, v \in I : C_{XY}(u, v) = \Pi(u, v) = uv$ .

$$\begin{aligned} \forall x, y \in \mathbb{R} : F(x, y) &= C_{XY}(F(x), F(y)) \\ &= \Pi(F(x), F(y)) \\ &= F(x)F(y). \end{aligned}$$

$$F(x, y) = F(x)F(y). \quad \square$$

Les variables  $x$  et  $y$  sont indépendantes.

---

"Propriétés fondamentales des copules"

---

**Théorème 3.9.** *Soient deux va continues  $X$  et  $Y$  de copule associée  $C_{XY}$ . Si  $\alpha$  et  $\beta$  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.

**Démonstration.** □

Notons d'une part,  $F$  la distribution conjointe des va continues  $X$  et  $Y$  de lois respectives  $F_1$  et  $F_2$  et d'autre part,  $\hat{F}$  la distribution de la transformée  $(\alpha(x), \beta(y))$  de marges respectives  $\hat{F}_1$  et  $\hat{F}_2$ , on a :

$$\begin{aligned} \forall x \in \mathbb{R} : \hat{F}_1(x) &= \mathbb{P}(\alpha(X) \leq x) \\ &= \mathbb{P}(X \leq \alpha^{-1}(x)) \\ &= F_1(\alpha^{-1}(x)) \quad (\alpha^{-1} \text{ existe car } \alpha \text{ est strictement } \nearrow) \\ &= u. \end{aligned}$$

$$\forall x \in \mathbb{R} : x = \hat{F}_1^{-1}(u).$$

$$\hat{F}_1(x) = u = F_1(\alpha^{-1}(x)) \implies \alpha^{-1}(x) = F_1^{-1}(u).$$

$$\forall x \in \mathbb{R} : x = \alpha(F_1^{-1}(u)).$$

De même on a :  $\forall y \in \mathbb{R} : \hat{F}_2(y) = F_2(\beta^{-1}(y))$ .

Et donc :

$$\hat{F}_1^{-1}(u) = \alpha(F_1^{-1}(u)) \text{ et } \hat{F}_2^{-1}(v) = \beta(F_2^{-1}(v)). \quad (*)$$

Alors :

$$\begin{aligned} \forall u, v \in I : C_{\alpha(X)\beta(Y)}(u, v) &= F(\hat{F}_1^{-1}(u), \hat{F}_2^{-1}(v)) \\ &= \mathbb{P} \left[ \alpha(X) \leq \hat{F}_1^{-1}(u), \beta(Y) \leq \hat{F}_2^{-1}(v) \right] \\ \text{De } (*) \implies C_{\alpha(X)\beta(Y)}(u, v) &= \mathbb{P} \left[ X \leq \alpha^{-1}(\hat{F}_1^{-1}(u)), Y \leq \beta^{-1}(\hat{F}_2^{-1}(v)) \right] \\ &= \mathbb{P} \left[ X \leq F_1^{-1}(u), Y \leq F_2^{-1}(v) \right] \\ &= F(F_1^{-1}(u), F_2^{-1}(v)) = C_{XY}(u, v). \end{aligned}$$

Donc :

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

---

"Copules associées à une copule"

---

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). Dans le cas de transformations strictement décroissantes des distributions marginales d'une copule  $C$  on a le résultat suivant.

**Théorème 3.10.** *Soient  $X$  et  $Y$  deux vz continues de copule  $C_{XY}$ . Si  $\alpha$  et  $\beta$  sont des fonctions strictement monotones sur  $\text{Im}(X)$  et  $\text{Im}(Y)$  respectivement :*

**i/** *Si  $\alpha$  est strictement croissante ( $\nearrow$ ) et  $\beta$  strictement décroissante ( $\searrow$ ) alors :*

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

**ii/** *Si  $\alpha$  est strictement décroissante ( $\searrow$ ) et  $\beta$  strictement croissante ( $\nearrow$ ) alors :*

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

**iii/** *Si  $\alpha$  et  $\beta$  sont strictement décroissantes ( $\searrow$ ) alors :*

$$\forall u, v \in I : C_{\alpha(X)\beta(Y)}(u, v) = u + v - 1 - C_{XY}(1 - u, 1 - v).$$

### 3.11 "Copules" associées à une copule

#### 3.11.1 Copule de survie

Soit  $X$  une va de distribution  $F_X$ . On note  $\bar{F}_X$  la fonction de survie associée à  $F_X$  i.e.  $\forall x \in \mathbb{R} : \bar{F}_X(x) = \mathbb{P}(X > x) = 1 - F_X(x)$ . Par exemple si la va continue  $X$  modélise la durée de vie d'un individu au sein d'une population alors  $\bar{F}_X(x)$  est la probabilité que l'individu vive ou survive au delà du temps  $x$ .

De même, on peut associer à la distribution conjointe  $F$  d'un couple de va  $(X, Y)$  une distribution de survie  $\bar{F}$  par :  $\bar{F}(x, y) = \mathbb{P}(X > x, Y > y)$  dont les distributions marginales  $\bar{F}_1(x) = \lim_{y \rightarrow +\infty} \bar{F}(x, y) = \int_{-\infty}^x \bar{f}(x, y) dy$  et  $\bar{F}_2(y) = \lim_{x \rightarrow +\infty} \bar{F}(x, y) = \int_{-\infty}^y \bar{f}(x, y) dx$ .

**Problème :** Comment  $\bar{F}$  s'exprime en fonction de ses marges ?

$$\begin{aligned} \forall (x, y) \in \mathbb{R} : \bar{F}(x, y) &= \mathbb{P}(X > x, Y > y) \\ &= 1 - \mathbb{P}(X \leq x, Y \leq y) \\ &= \mathbb{P}(\overline{X \leq x, Y \leq y}) \\ &= 1 - F_1(x) - F_2(y) + F(x, y). \end{aligned}$$

---

"Copules associées à une copule"

---

En effet, probabilité élémentaire :

$$\begin{aligned}
 \mathbb{P}(\bar{A} \cap \bar{B}) &= \mathbb{P}(\overline{A \cup B}) \\
 &= 1 - \mathbb{P}(A \cup B) \\
 &= 1 - [\mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)] \\
 &= 1 - \mathbb{P}(A) - \mathbb{P}(B) + \mathbb{P}(A \cap B).
 \end{aligned}$$

$$\begin{aligned}
 \forall (x, y) \in \mathbb{R} : \bar{F}(x, y) &= 1 - F_1(x) - F_2(y) + F(x, y) \\
 &= [1 - F_1(x)] + [\bar{F}_2(y) - 1] + F(x, y) \\
 &= \bar{F}_1(x) + \bar{F}_2(y) - 1 + F(x, y) \\
 &= \bar{F}_1(x) + \bar{F}_2(y) - 1 + C(F_1(x), F_2(y)).
 \end{aligned}$$

$$\bar{F}(x, y) = \bar{F}_1(x) + \bar{F}_2(y) - 1 + C(1 - \bar{F}_1(x), 1 - \bar{F}_2(y)).$$

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

$\forall u, v \in I^2 : \hat{C}(u, v) = u + v - 1 + C(1 - u, 1 - v)$  alors on obtient :

$$\forall (x, y) \in \mathbb{R} : \bar{F}(x, y) = \hat{C}(\bar{F}_1(x), \bar{F}_2(y)).$$

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

Il ne faut pas confondre la copule de survie  $\hat{C}$  et la distribution de survie  $\bar{C}$  associée à  $C$ .

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 liés par la relation :

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

Ou simplement  $\forall u, v \in I :$

$$\hat{C}(u, v) = \bar{C}(\bar{u}, \bar{v}) \text{ avec } \bar{u} = 1 - u \text{ et } \bar{v} = 1 - v.$$

Dans un raisonnement similaire on peut exprimer les probabilités  $\mathbb{P}(X \leq x \text{ ou } Y \leq y)$ ,  $\mathbb{P}(X > x \text{ ou } Y > y)$ ,  $\mathbb{P}(X > x, Y \leq y)$  et  $\mathbb{P}(X \leq x, Y > y)$  en terme de couple.

### 3.11.2 Copule duale

$$\begin{aligned}
\mathbb{P}(X \leq x \text{ ou } Y \leq y) &= \mathbb{P}(X \leq x) + \mathbb{P}(Y \leq y) - \mathbb{P}(X \leq x, Y \leq y) \\
&= F_1(x) + F_2(y) - F(x, y) \\
&= F_1(x) + F_2(y) - C(F_1(x), F_2(y)).
\end{aligned}$$

En posant :  $\tilde{C}(u, v) = \tilde{C}(F_1(x), F_2(y)) = F_1(x) + F_2(y) - C(F_1(x), F_2(y))$  on vérifie que  $\tilde{C}$  est dite copule duale de  $C$  :

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

### 3.11.3 Co-copule

$$\begin{aligned}
\mathbb{P}(X > x \text{ ou } Y > y) &= \mathbb{P}(X > x) + \mathbb{P}(Y > y) - \mathbb{P}(X > x, Y > y) \\
&= \bar{F}_1(x) + \bar{F}_2(y) - \bar{F}(x, y) \\
&= 1 - C(1 - \bar{F}_1(x), 1 - \bar{F}_2(y)).
\end{aligned}$$

$$\text{car } \bar{F}(x, y) = \bar{F}_1(x) + \bar{F}_2(y) - 1 + C(1 - \bar{F}_1(x), 1 - \bar{F}_2(y)).$$

En posant :  $C^*(u, v) = C^*(F_1(x), F_2(y)) = 1 - C(1 - \bar{F}_1(x), 1 - \bar{F}_2(y))$  on vérifie que  $C^*$  est dite co-copule de  $C$  :

$$\forall u, v \in I : C^*(u, v) = 1 - C(1 - \bar{u}, 1 - \bar{v}).$$

### 3.11.4 Copule mixte

1. La probabilité élémentaire :

$$\begin{aligned}
\mathbb{P}(X > x, Y \leq y) &= \mathbb{P}(\bar{A} \cap B) \\
&= \mathbb{P}(B) - \mathbb{P}(A \cap B) \\
&= F_2(y) - F(x, y) \\
&= F_2(y) - C(F_1(x), F_2(y)).
\end{aligned}$$

En posant :  $\hat{C}(u, v) = \hat{C}(F_1(x), F_2(y)) = F_2(y) - C(F_1(x), F_2(y))$  on vérifie que  $\hat{C}$  est dite copule mixte de  $C$  :

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

$$\begin{aligned}
\mathbb{P}(X > x, Y \leq y) &= \mathbb{P}(\bar{A} \cap B) \\
&= \mathbb{P}(B) - \mathbb{P}(A \cap B) \\
&= F_2(y) - F(x, y) \\
&= F_2(y) - C(F_1(x), F_2(y)) \\
&= F_2(y) - [\bar{C}(F_1(x), F_2(y)) - 1 + F_1(x) + F_2(y)] \\
&= F_2(y) - \bar{C}(F_1(x), F_2(y)) + 1 - F_1(x) - F_2(y) \\
&= 1 - F_1(x) - \bar{C}(F_1(x), F_2(y)) \\
&= 1 - (1 - F_1(x)) - \bar{C}(1 - F_1(x), 1 - F_2(y)) \\
&= F_1(x) - \bar{C}(1 - F_1(x), 1 - F_2(y)). \\
\mathbb{P}(X > x, Y \leq y) &= -\bar{C}(1 - F_1(x), 1 - F_2(y)) + F_1(x) \\
&= 1 - F_2(y) - \bar{C}(1 - F_1(x), 1 - F_2(y)) + F_1(x) + F_2(y) - 1 \\
&= \bar{F}_2(y) - [\bar{C}(1 - F_1(x), 1 - F_2(y)) - F_1(x) - F_2(y) + 1] \\
&= \bar{F}_2(y) - [\hat{C}(F_1(x), F_2(y)) - F_1(x) - F_2(y) + 1] \\
&= \bar{F}_2(y) - [C(\bar{F}_1(x), \bar{F}_2(y))] \\
&= \bar{F}_2(y) - C(1 - F_1(x), 1 - F_2(y)) \\
&= 1 - F_2(y) - C(1 - F_1(x), 1 - F_2(y)).
\end{aligned}$$

En posant :  $\hat{C}(u, v) = \hat{C}(F_1(x), F_2(y)) = 1 - F_2(y) - C(1 - F_1(x), 1 - F_2(y))$  on vérifie que  $\hat{C}$  est dite copule mixte de  $C$  :

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

"Quelques copules classiques"

---

2. De même on établit que :

$$\begin{aligned}
\mathbb{P}(X \leq x, Y > y) &= \mathbb{P}(A \cap \bar{B}) \\
&= \mathbb{P}(A) - \mathbb{P}(A \cap B) \\
&= F_1(x) - F(x, y) \\
&= F_1(x) - C(F_1(x), F_2(y)) \\
&= F_1(x) - [\bar{C}(F_1(x), F_2(y)) - 1 + F_1(x) + F_2(y)] \\
&= F_1(x) - \bar{C}(F_1(x), F_2(y)) + 1 - F_1(x) - F_2(y) \\
&= 1 - F_2(y) - \bar{C}(F_1(x), F_2(y)) \\
&= 1 - (1 - F_2(y)) - \bar{C}(1 - F_1(x), 1 - F_2(y)) \\
&= F_2(y) - \bar{C}(1 - F_1(x), 1 - F_2(y)). \\
\mathbb{P}(X \leq x, Y > y) &= -\bar{C}(1 - F_1(x), 1 - F_2(y)) + F_2(y) \\
&= 1 - F_1(x) - \bar{C}(1 - F_1(x), 1 - F_2(y)) + F_2(y) + F_1(x) - 1 \\
&= \bar{F}_1(x) - [\bar{C}(1 - F_1(x), 1 - F_2(y)) - F_2(y) - F_1(x) + 1] \\
&= \bar{F}_1(x) - [\hat{C}(F_1(x), F_2(y)) - F_2(y) - F_1(x) + 1] \\
&= \bar{F}_1(x) - [C(\bar{F}_1(x), \bar{F}_2(y))] \\
&= -[F_1(x) - 1 + C(1 - F_1(x), 1 - F_2(y))].
\end{aligned}$$

En posant :  $(u, v) = \hat{C}(F_1(x), F_2(y)) = -[F_1(x) - 1 + C(1 - F_1(x), 1 - F_2(y))]$ ,  
on vérifie que  $\hat{C}$  est dite copule mixte de  $C$  :

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

2. De même on établit que :

$$\begin{aligned}
 \mathbb{P}(X \leq x, Y > y) &= \mathbb{P}(A \cap \bar{B}) \\
 &= \mathbb{P}(A) - \mathbb{P}(A \cap B) \\
 &= F_1(x) - F(x, y) \\
 &= F_1(x) - C(F_1(x), F_2(y)).
 \end{aligned}$$

En posant :  $\hat{C}(u, v) = \hat{C}(F_1(x), F_2(y)) = F_1(x) - C(F_1(x), F_2(y))$ , on vérifie que  $\hat{C}$  est dite copule mixte de  $C$  :

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

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

- $\hat{C}(u, v) = u + v - 1 + C(1-u, 1-v) \longrightarrow$  La copule de survie.
- $\tilde{C}(u, v) = u + v - C(u, v) \longrightarrow$  La copule duale.
- $C^*(u, v) = 1 - C(u, v) \longrightarrow$  La co-copule.
- $\begin{cases} \hat{C}(u, v) = v - C(u, v) \\ \hat{C}(u, v) = u - C(u, v) \end{cases} \longrightarrow$  Les copules mixtes.

Ces copules associées sont très importantes dans la modélisation.

### 3.12 Quelques copules classiques

Les principales propriétés des copules ayant été définies, nous pouvons désormais en présenter quelques exemples.

Nous exposons dans la présente section les principales copules paramétriques. Ces copules présentent un intérêt particulier en hydrologie, de par le fait qu'elles autorisent la construction de modèles paramétriques.

#### 3.12.1 Copules usuelles

Parmi les copules usuelles :

- Copule produit (copule indépendante) :  $\forall u, v \in I : \Pi(u, v) := uv$ ,  $u$  et  $v$  sont des va indépendantes.
- Copule du maximum (borne inférieure de *Fréchet*) :  $\forall u, v \in I : W(u, v) := \max(u + v - 1, 0)$ .



---

"Quelques copules classiques"

---

- Copule du minimum (borne supérieure de *Fréchet*) :  $\forall u, v \in I : M(u, v) := \min(u, v)$ .

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 de  $M, \Pi$  et  $W$ .

Par exemple la copule de *Mardia* :  $\forall \theta \in [-1, 1], \forall u, v \in I :$

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

### 3.12.2 Copules elliptiques

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

Or les copules qui ont été retenus dans le cadre de nos travaux sont les copules archimédiennes (Gumbel, Clayton, Frank) dont les fonctions de répartition sont mentionnées ci-dessous.

**Definition 6.** *On appelle copule elliptique toute copule de la forme :*

$$\begin{aligned} \forall u, v \in I : C_\rho(u, v) &= \frac{1}{\sqrt{1 - \rho^2}} \int_{-\infty}^{F_{g,1}^{-1}(u)} \int_{-\infty}^{F_{g,2}^{-1}(v)} g\left(\frac{x^2 - 2\rho xy + y^2}{1 - \rho^2}\right) dx dy \\ &= F_\rho(F_{g,1}^{-1}(u), F_{g,2}^{-1}(v)). \end{aligned}$$

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

Parmi les copules elliptiques :

---

"Quelques copules classiques"

---

**- Copule normale bivariée :**

Cette copule ne permet pas de mesurer la dépendance entre les queues des distributions marginales. Il s'agit d'une propriété contraignante lorsqu'on veut évaluer la dépendance entre les événements rares.

Un des types de copules les plus utilisées dans la modélisation est la copule normale bivariée.

Si  $X$  et  $Y$  sont deux variables normales standard de corrélation  $\rho$  leur copule est donnée par :

$$C_\rho(u, v) = \frac{1}{2\pi\sqrt{1-\rho^2}} \int_{-\infty}^{F_1^{-1}(u)} \int_{-\infty}^{F_2^{-1}(v)} \exp\left\{-\frac{x^2 - 2\rho xy + y^2}{2(1-\rho^2)}\right\} dx dy.$$

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

**- Copule de Student :**

Comparativement à la copule gaussienne, la copule Student permet grâce à son degré de liberté, de mieux tenir compte des queues de distribution épaisses. En outre, lorsque le degré de liberté tend vers l'infini, la copule de Student est égale à la copule Gaussienne.

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

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

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

Où  $t_v^{-1}$  est la fonction inverse de la distribution standard de Student  $t_{\rho, v}$  à  $v$  degrés de liberté.

Les copules normale et de Student 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.

**- 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  $C_{\rho_1}$  et  $C_{\rho_2}$  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 :

$$\forall u, v \in I : C_p(u, v) = pC_{\rho_1}(u, v) + (1 - p)C_{\rho_2}(u, v).$$

### 3.12.3 Copules Archimédiennes et Archimax

Parmi les copules paramétriques, les copules archimédiennes (*Schweizer & Sklar, 1983*) sont plus faciles de construction et ont des propriétés intéressantes.

La théorie des copules archimédienne a été introduite et développée par le professeur *Christian Genest* de l'université de Laval au Québec.

#### Copule Archimédienne

Une classe particulièrement importante des copules est celle des copules Archimédiennes.

**Definition 7.** (*Copule archimédienne*)

*Une copule archimédienne est une copule de la forme :*

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

$$\begin{cases} C(u, v) > 0 & \text{si } \lim_{t \rightarrow 0^+} \varphi(t) = +\infty. \\ C(u, v) = 0 & \text{si } u \text{ ou } v \text{ sont égales à } 0. \end{cases}$$

La fonction  $\varphi$  est dite générateur archimédien de la copule  $C$ ,  $\varphi^{-1}$  la pseudo inverse de  $\varphi$ .

$\varphi$  est au moins deux fois continument dérivable, est une fonction strictement décroissante telle que :  $\varphi'(u) < 0$  et  $\varphi''(u) > 0, \forall 0 \leq u \leq 1$ .

$$\begin{cases} \varphi : [0, 1] \longrightarrow [0, +\infty[ \\ \varphi(0) = +\infty & \text{si } \lim_{t \rightarrow +0} \varphi(t) = +\infty. \\ \varphi^{-1}(t) = 0 & \text{si } t \geq \varphi(0). \end{cases}$$

**Example 5.** (*Familles classiques de copules archimédiennes bivariées de paramètre  $\theta$* )

#### Copule Gumbel :

Le générateur archimédien :  $\theta \geq 1; \forall t > 0 : \varphi_\theta(t) := (-\ln t)^\theta$ .

La pseudo inverse de  $\varphi : \varphi_\theta^{-1}(t) := \exp(-t^{\frac{1}{\theta}})$ .

"Quelques copules classiques"

---

La copule  $C_\theta$  associée est telle que :

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

Pour  $\theta = 1$  correspond à la copule produit  $\Pi(u, v) = uv$ .

Nous présentons ici quelques familles classiques de copules archimédiennes, à savoir la copule indépendante, la copule de Frank, la copule de Gumbel et celle de Clayton.

- Copule produit obtenue lorsque le générateur  $\varphi_\theta(t) := -\ln t$ .  
 $\implies C_\theta(u, v) = \Pi(u, v) = uv$ .
- Copule de Clayton obtenue lorsque le générateur  $\varphi_\theta(t) := (t^{-\theta} - 1)$ ;  $\theta \geq 0$ .  
 $\implies C_\theta(u, v) = (u^{-\theta} + v^{-\theta} - 1)^{-\frac{1}{\theta}}$ ,  $\theta \geq 0$ .
- Copule de Gumbel obtenue lorsque le générateur  $\varphi_\theta(t) := (-\ln t)^\theta$ ;  $\theta \geq 1$ .  
 $\implies C_\theta(u, v) = \exp \left\{ - \left[ (-\ln u)^\theta + (-\ln v)^\theta \right]^{1/\theta} \right\}$ .
- Copule de Frank obtenue lorsque le générateur  $\varphi_\theta(t) := -\log \left( \frac{\ell^{-\theta t} - 1}{\ell^{-\theta} - 1} \right)$ .  
 $\implies C_\theta(u, v) = \frac{-1}{\theta} \log \left( 1 + \frac{(\ell^{-\theta u} - 1)(\ell^{-\theta v} - 1)}{\ell^{-\theta} - 1} \right)$ ;  $\theta \neq 0$ .

Cette copule est la seule qui soit symétrique dans la queue inférieure et supérieure.

**Résumé :**

Nom	Générateur $\varphi_\theta(t)$	Copule bivariable $C_\theta(u, v)$
Produit	$-\ln t$	$uv$
Clayton ( $\theta \geq 0$ )	$t^\theta - 1$	$(u^{-\theta} + v^{-\theta} - 1)^{-\frac{1}{\theta}}$
Gumbel ( $\theta \geq 1$ )	$(-\ln t)^\theta$	$\exp \left\{ - \left[ (-\ln u)^\theta + (-\ln v)^\theta \right]^{1/\theta} \right\}$
Frank ( $\theta \neq 0$ )	$-\log \left( \frac{\ell^{-\theta t} - 1}{\ell^{-\theta} - 1} \right)$	$-\frac{1}{\theta} \log \left( 1 + \frac{(\ell^{-\theta u} - 1)(\ell^{-\theta v} - 1)}{\ell^{-\theta} - 1} \right)$

TAB. 3.1. Familles classiques de copules archimédiennes

### Copules Archimax

**Definition 8.** Une fonction bivariable est une copule Archimax si et seulement si elle est de la forme :

$$\forall 0 \leq u, v \leq 1 : C_{\varphi, A}(u, v) := \varphi^{-1} \left[ (\varphi(u) + \varphi(v)) A \left( \frac{\varphi(u)}{\varphi(u) + \varphi(v)} \right) \right].$$

$A : [0, 1] \rightarrow [1/2, 1]$  est une fonction convexe tel que  $\max(t, 1 - t) \leq A(t) \leq 1$ ;  $\forall 0 \leq t \leq 1$ .

**Remarque :** (Cas particuliers)

- Pour  $\varphi(t) = -\ln t$  on obtient la forme générale des copules des valeurs extrêmes :

$$\forall u, v \in I : C_{\varphi, A}(u, v) = C_A(u, v) = \exp \left[ -(\ln u + \ln v) A \left( \frac{\ln u}{\ln u + \ln v} \right) \right].$$

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

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

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

**Example 6.** (Fonction de dépendance associée à la copule de Galambos)

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

On a :

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

---

"Quelques copules classiques"

---

De (\*\*):

$$\begin{aligned}
 (\ln u + \ln v)A\left(\frac{\ln u}{\ln u + \ln v}\right) &= (\ln u + \ln v) + [(-\ln u)^{-\theta} + (-\ln v)^{-\theta}]^{-\frac{1}{\theta}} \\
 &= (\ln u + \ln v)\left[1 - \left[\left(\frac{-\ln u}{\ln u + \ln v}\right)^{-\theta} + \left(\frac{-\ln v}{\ln u + \ln v}\right)^{-\theta}\right]^{-\frac{1}{\theta}}\right].
 \end{aligned}$$

En posant :  $t = \frac{\ln u}{\ln u + \ln v}$  et  $\frac{\ln v}{\ln u + \ln v} = 1 - t$ , on retrouve la fonction de dépendance :

$$A_\theta(t) = 1 - [t^{-\theta} + (1 - t)^{-\theta}]^{-\frac{1}{\theta}}.$$

### 3.12.4 Copules des valeurs extrêmes

Pour construire une distribution des valeurs extrêmes bivariée, il suffit ainsi de coupler des marges issues de loi de la théorie des valeurs extrêmes avec une copule des valeurs extrêmes.

**Definition 9.** On appelle copule des valeurs extrêmes toute copule  $C$  vérifiant la propriété suivante :

$$\left\{ \begin{array}{l} \forall u, v \in I; \forall t > 0 : C(u^t, v^t) := C^t(u, v). \\ \text{Où} \\ \forall t > 0 : C^{\frac{1}{t}}(u^t, v^t) := C(u, v). \end{array} \right.$$

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.

**Exemple 7.** Considérons la copule de Galambos :  $\forall \theta \geq 1 : C_\theta(u, v) = uv \exp \left\{ - [(-\ln u)^{-\theta} + (-\ln v)^{-\theta}]^{-\frac{1}{\theta}} \right\}$

$$\begin{aligned}
 \forall u, v \in I; \forall t > 0 : C_\theta(u^t, v^t) &= u^t v^t \exp \left\{ - [(-\ln u^t)^{-\theta} + (-\ln v^t)^{-\theta}]^{-\frac{1}{\theta}} \right\} \\
 &= u^t v^t \exp \left\{ - [(-t \ln u)^{-\theta} + (-t \ln v)^{-\theta}]^{-\frac{1}{\theta}} \right\} \\
 &= u^t v^t \exp \left\{ -t [(-\ln u)^{-\theta} + (-\ln v)^{-\theta}]^{-\frac{1}{\theta}} \right\} \\
 &= \left( uv \exp \left\{ - [(-\ln u)^{-\theta} + (-\ln v)^{-\theta}]^{-\frac{1}{\theta}} \right\} \right)^t \\
 &= C_\theta^t(u, v).
 \end{aligned}$$

Donc  $C_\theta$  est une copule bivariée des valeurs extrêmes notée BEV.

### 3.12.5 Copule empirique

Soit  $(X, Y)$  un couple de va continues. Considérons un échantillon de taille  $n$  de ce couple  $E_n = \{(x_k, y_k); k = 1, \dots, n\}$ . La copule empirique associée à cet échantillon est la fonction  $C_n$  donnée par :

$$C_n(u, v) := \sum_{i=1}^n \sum_{j=1}^n f_n\left(\frac{i}{n}, \frac{j}{n}\right).$$

Où la fonction fréquence est définie par :  $f_n\left(\frac{i}{n}, \frac{j}{n}\right) := \begin{cases} \frac{1}{n} & \text{si } (x_i, y_j) \in E_n \\ 0 & \text{si non} \end{cases}$

## 3.13 Mesures de concordance

La copule  $C$  d'un vecteur aléatoire continue  $(X, Y)$  est une normalisation de la distribution conjointe  $H$  après avoir éliminé les effets des marges. C'est donc une structure de dépendance entre les deux va  $X$  et  $Y$  connaissant leurs distributions respectives.

Cette structure permet des estimations pour étudier cette dépendance à travers les mesures de concordance (corrélation linéaire, concordance, tau de Kendall et rho de Spearman).

### 3.13.1 Fonction de concordance

**Definition 10.** (*Notion de concordance*)

Deux observations  $(x_1, y_1)$  et  $(x_2, y_2)$  d'un couple de va continues  $(X, Y)$  de lois conjointes  $H_1$  et  $H_2$  mais de marges communes  $F$  et  $G$  (i.e.  $F$  pour  $X_1$  et  $X_2$ ,  $G$  pour  $Y_1$  et  $Y_2$ ). C'est deux observations sont dites :

- *Concordantes* si :  $\begin{cases} (x_1 < x_2 \text{ et } y_1 < y_2) \text{ ou } (x_1 > x_2 \text{ et } y_1 > y_2) \\ \text{ie } (x_1 - x_2)(y_1 - y_2) > 0. \end{cases}$
- *Discordantes* si :  $\begin{cases} (x_1 < x_2 \text{ et } y_1 > y_2) \text{ ou } (x_1 > x_2 \text{ et } y_1 < y_2) \\ \text{ie } (x_1 - x_2)(y_1 - y_2) < 0. \end{cases}$

Plus généralement, soit  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$  un échantillon de  $n$  observations d'un couple  $(X, Y)$ . Il existe  $C_n^2$  paires distinctes de couples  $(x_i, y_i)$  et  $(x_j, y_j)$  qui sont soit concordantes, soit discordantes.

**Definition 11.** (*Fonction de concordance*)

La fonction de concordance entre ces deux couples  $(X_1, Y_1)$  et  $(X_2, Y_2)$  est :

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

C'est la différence entre la probabilité de concordance et celle de discordance.

### 3.13.2 Propriétés de la fonction de concordance

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.

**Definition 12.** Soient  $(X_1, Y_1), (X_2, Y_2)$  deux couples de va indépendantes 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 :

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

**Démonstration.**

□

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

Avec :

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



Par ailleurs :

$$\begin{aligned}
\mathbb{P}[X_1 > X_2, Y_1 > Y_2] &= \mathbb{P}[X_2 < X_1, Y_2 < Y_1] \\
&= \iint_{\mathbb{R}^2} \mathbb{P}[X_2 < x, Y_2 < y] dH_1[x, y] \\
&= \iint_{\mathbb{R}^2} \mathbb{P}[X_2 < x, Y_2 < y] dC_1[F(x), G(y)] \\
&= \iint_{\mathbb{R}^2} C_2[F(x), G(y)] dC_1[F(x), G(y)] \\
&= \iint_{I^2} C_2(u, v) dC_1(u, v).
\end{aligned}$$

De façon similaire :

$$\begin{aligned}
\mathbb{P}[X_1 < X_2, Y_1 < Y_2] &= \mathbb{P}[X_2 > X_1, Y_2 > Y_1] \\
&= \iint_{\mathbb{R}^2} \mathbb{P}[X_2 > x, Y_2 > y] dH_1[x, y] \\
&= \iint_{\mathbb{R}^2} (1 - \mathbb{P}[X_2 \leq x, Y_2 \leq y]) dH_1[x, y] \\
&= \iint_{\mathbb{R}^2} (1 - H_2[x, y]) dC_1[F(x), G(y)] \\
&= \iint_{\mathbb{R}^2} \bar{H}_2(x, y) dH_1(x, y) \quad (\bar{H}_2 \text{ fct de survie de } H_2) \\
&= \iint_{I^2} [1 - u - v + C_2(u, v)] dC_1[u, v] \\
&= 1 - \frac{1}{2} - \frac{1}{2} + \iint_{I^2} C_2(u, v) dC_1[u, v].
\end{aligned}$$

$$\mathbb{P}[X_1 < X_2, Y_1 < Y_2] = \iint_{I^2} C_2(u, v) dC_1[u, v].$$

Par suite :

$$\begin{aligned}
Q &= 2\mathbb{P}[(X_1 - X_2)(Y_1 - Y_2) > 0] - 1 \\
&= 2(\mathbb{P}[X_1 < X_2, Y_1 < Y_2] + \mathbb{P}[X_1 > X_2, Y_1 > Y_2]) - 1 \\
&= 2\left(\iint_{I^2} C_2(u, v) dC_1(u, v) + \iint_{I^2} C_2(u, v) dC_1(u, v)\right) - 1 \\
&= 4\iint_{I^2} C_2(u, v) dC_1(u, v) - 1.
\end{aligned}$$

$$Q := 4\iint_{I^2} C_2(u, v) dC_1(u, v) - 1. \quad \square$$

Le résultat suivant résume les propositions essentielles de la fonction  $Q$ .

**Proposition**

Soit  $C_1, C_2$  et  $Q$  donnés par le théorème précédent Alors :

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

**Démonstration. ii)**

$$Q(\acute{C}_1, \acute{C}_2) - Q(C_1, C_2) = 4\left[\iint_{I^2} \acute{C}_2(u, v) d\acute{C}_1(u, v) - \iint_{I^2} C_2(u, v) dC_1(u, v)\right] : \quad \square$$

Or  $\frac{\partial^2 C_1}{\partial u \partial v}$  et  $\frac{\partial^2 \acute{C}_1}{\partial u \partial v}$  sont croissantes par rapport à chaque variable (2-croissante).

Puisque  $C_1 < \acute{C}_1 \implies \frac{\partial^2 C_1}{\partial u \partial v} < \frac{\partial^2 \acute{C}_1}{\partial u \partial v}$ .

Donc :  $dC_1(u, v) < d\acute{C}_1(u, v)$ .

$$\text{Puisque : } \iint_{I^2} \acute{C}_2(u, v) d\acute{C}_1(u, v) > \iint_{I^2} \acute{C}_2(u, v) dC_1(u, v).$$

$$\text{Et : } \iint_{I^2} \acute{C}_2(u, v) dC_1(u, v) > \iint_{I^2} C_2(u, v) dC_1(u, v).$$

$$\text{Alors : } \iint_{I^2} \acute{C}_2(u, v) d\acute{C}_1(u, v) - \iint_{I^2} C_2(u, v) dC_1(u, v) > 0.$$

$$\implies 4 \iint_{I^2} \acute{C}_2(u, v) d\acute{C}_1(u, v) - 4 \iint_{I^2} C_2(u, v) dC_1(u, v) > 0.$$

$$\implies 4 \iint_{I^2} \acute{C}_2(u, v) d\acute{C}_1(u, v) - 1 - 4 \iint_{I^2} C_2(u, v) dC_1(u, v) + 1 > 0.$$

$$Q(\acute{C}_1, \acute{C}_2) - Q(C_1, C_2) > 0. \quad \square$$

iii) Pour toute copule  $C$ , on a : 
$$\begin{cases} dC(u, v) = \partial^2 C(u, v) / \partial u \partial v. \\ \text{et } \hat{C}(u, v) = 1 - u - v + C(u, v). \end{cases}$$

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

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

$$\text{Or : } 4 \iint_{I^2} (1 - u - v) dC_1(u, v) = 4 \left[ 1 - \frac{1}{2} - \frac{1}{2} \right] = 0.$$

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

$$Q(C_1, C_2) = Q(\hat{C}_1, \hat{C}_2). \quad \square$$

**Definition 13.** (*Fonctions de dépendance des copules usuelles*)

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

$$Q(M, M) = 1.$$

$$Q(\Pi, \Pi) = 0.$$

$$Q(W, W) = -1.$$

$$Q(M, \Pi) = 1/3.$$

$$Q(W, \Pi) = 1/3.$$

$$Q(M, W) = -1.$$

Pour toute copule  $C$  on a :

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

$$-1 \leq Q(C, W) \leq 1/3.$$

$$0 \leq Q(C, \Pi) \leq 1/3.$$

**Démonstration.**

□

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

$$\text{Donc : } \begin{cases} M(u, v) &= \min(u, v) &= \min(u, u) &= u. \\ W(u, v) &= \max(u + v - 1, 0) &= \max(2u - 1, 0) &= 2u - 1. \\ \Pi(u, v) &= uv &= u \times u &= u^2. \end{cases}$$

$$\begin{aligned} \bullet Q(M, M) &= 4 \iint_{I^2} M(u, v) dM(u, v) - 1 \\ &= 4 \iint_{I^2} M(u, u) dM(u, u) - 1 \\ &= 4 \int_0^1 u du - 1 = 4 \left[ \frac{u^2}{2} \right]_0^1 - 1 = 1. \end{aligned}$$

$$\begin{aligned} \bullet Q(M, \Pi) &= 4 \iint_{I^2} \Pi(u, v) dM(u, v) - 1 \\ &= 4 \int_0^1 u^2 du - 1 = 1/3. \end{aligned}$$

$$\begin{aligned} \bullet Q(M, W) &= 4 \iint_{I^2} W(u, v) dM(u, v) - 1 \\ &= 4 \int_0^1 (2u - 1) du - 1 = -1. \end{aligned}$$

2. Le support de la copule usuelle  $W$  est l'ensemble  $D_W = \{u, v \in I / u = 1 - v\}$ .

$$\begin{aligned} \bullet Q(W, \Pi) &= 4 \iint_{I^2} W(u, v) d\Pi(u, v) - 1 \\ &= 4 \iint_{I^2} (2u - 1) du^2 - 1 \end{aligned}$$

$$\begin{aligned}\bullet Q(W, \Pi) &= 4 \iint_{I^2} W(u, v) d\Pi(u, v) - 1 \\ &= 4 \iint_{I^2} (2u - 1) du^2 - 1 \\ &= 4 \int_0^1 (2u - 1) 2u du - 1 \\ &= 4 \int_0^1 (4u^2 - 2u) du - 1 = 1/3.\end{aligned}$$

$$\begin{aligned}
&= 4 \int_0^1 (2u - 1)2udu - 1 \\
&= 4 \int_0^1 (4u^2 - 2u)du - 1 = 1/3.
\end{aligned}$$

$$\begin{aligned}
\bullet Q(W, W) &= 4 \iint_{I^2} W(u, v)dW(u, v) - 1 \\
&= 0 - 1 = -1.
\end{aligned}$$

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

3. Pour toute copule quelconque  $C$  on a :

$$\begin{aligned}
\bullet W \leq C \leq M &\implies Q(W, M) \leq Q(C, M) \leq Q(M, M) \\
&\implies -1 \leq Q(C, M) \leq 1.
\end{aligned}$$

De la même manière, on établit que :  $\begin{cases} W \leq C \leq \Pi \implies -1 \leq Q(C, W) \leq 1/3, \\ \Pi \leq C \leq M \implies 0 \leq Q(C, \Pi) \leq 1/3. \end{cases}$

### 3.13.3 Mesures de concordance

Plus classiquement, les mesures d'association couramment utilisées (*Agostino & Stephens, 1986 ; Joe, 1997 ; Genest & Favre, 2007*) pour la caractérisation de la dépendance sont : le coefficient de corrélation linéaire de Pearson, le coefficient de corrélation de rang Spearman  $\rho$  et le coefficient de corrélation de rang Kendall  $\tau$ .

Le coefficient de corrélation de Pearson mesure la dépendance linéaire entre deux ra. Pour l'étude d'une dépendance non linéaire, les deux autres coefficients sont plus adaptés puisqu'ils ne dépendent que du rang de chaque observation par opposition au coefficient de corrélation de Pearson qui utilise les réalisations des variables. En outre,  $\rho$  et  $\tau$  sont invariants par transformation strictement croissante des va,  $X$  et  $Y$ . Ces deux derniers sont utilisés dans cette partie.

**Definition 14.** (Caractérisation d'une mesure de concordance)

Une mesure d'association  $K$  entre deux va continues  $X$  et  $Y$  de copule  $C$  est une mesure de concordance si elle vérifie propriétés suivantes :

1.  $K$  est définie pour tout couple  $(X, Y)$  de va continues.
2.  $-1 \leq K_{X,-X} \leq K_{X,Y} \leq 1 = K_{X,X}$ .
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 :  $K_{C_1} \leq K_{C_2}$ .
7. Si  $\{(X_n, Y_n)\}_{n \geq 1}$  est une suite de couples aléatoires dont la copule est  $C_n$ , si  $\{C_n\}_{n \geq 1}$  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.

**Definition 15.** (Tau de Kendall)**Tau de Kendall (version échantillon) :**

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 par :

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

**Tau de Kendall (version population) :**

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

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

Si  $C$  est la copule associée au couple  $(X, Y)$ , le résultat suivant donne la conséquence immédiate du théorème de la fonction de concordance.

Si les couples sont identiquement distribués i.e.  $H_1 = H_2 = H$  alors  $\tau_{XY} = Q_{XY}$ .



---

"Mesures de concordance"

---

Soient  $X$  et  $Y$  deux va continues de copule  $C$ . Le tau de Kendall de  $X$  et  $Y$  est donné par :

$$\tau_{XY} = \tau_C = Q_C = Q(C, C) := 4 \int_0^1 \int_0^1 C(u, v) dC(u, v) - 1 = Q_{XY}.$$

Si les couples sont uniformément distribués sur  $I$ , donc :

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

Ainsi, on peut écrire :

$$\tau_{XY} = \tau_C = 4E [C(u, v)] - 1.$$

Où  $F_c$  est la distribution de la va  $C(u, v)$ .

Pour toute copule de la famille de Fréchet :  $C_{\alpha, \beta} = \alpha M + (1 - \alpha - \beta)\Pi + \beta W$ .

Par suite on a :

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

"Mesures de concordance"

---

$$\begin{aligned}
\iint_{I^2} C_{\alpha,\beta}(u,v)dC_{\alpha,\beta}(u,v) &= \alpha^2 \iint_{I^2} MdM + \alpha(1-\alpha-\beta) \iint_{I^2} Md\Pi \\
&+ \alpha\beta \iint_{I^2} MdW + \alpha(1-\alpha-\beta) \iint_{I^2} \Pi dM \\
&+ (1-\alpha-\beta)^2 \iint_{I^2} \Pi d\Pi + \beta(1-\alpha-\beta) \iint_{I^2} \Pi dW \\
&+ \alpha\beta \iint_{I^2} WdM + \beta(1-\alpha-\beta) \iint_{I^2} Wd\Pi \\
&+ \beta^2 \iint_{I^2} WdW.
\end{aligned}$$

$$\begin{aligned}
4 \iint_{I^2} C_{\alpha,\beta}(u,v)dC_{\alpha,\beta}(u,v) &= \alpha^2 \left( 4 \iint_{I^2} MdM \right) + \alpha(1-\alpha-\beta) \left( 4 \iint_{I^2} Md\Pi \right) \\
&+ \alpha\beta \left( 4 \iint_{I^2} MdW \right) + \alpha(1-\alpha-\beta) \left( 4 \iint_{I^2} \Pi dM \right) \\
&+ (1-\alpha-\beta)^2 \left( 4 \iint_{I^2} \Pi d\Pi \right) + \beta(1-\alpha-\beta) \\
&\left( 4 \iint_{I^2} \Pi dW \right) + \alpha\beta \left( 4 \iint_{I^2} WdM \right) \\
&+ \beta(1-\alpha-\beta) \left( 4 \iint_{I^2} Wd\Pi \right) + \beta^2 \left( 4 \iint_{I^2} WdW \right).
\end{aligned}$$

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

Alors :

$$\begin{aligned}\tau_{\alpha\beta} &= \alpha^2 \times 1 + (1 - \alpha - \beta)^2 \times 0 + \beta^2 \times (-1) \\ &+ 2 \left[ \alpha(1 - \alpha - \beta) \frac{1}{3} + \alpha\beta \times 0 - \beta(1 - \alpha - \beta) \times \frac{1}{3} \right] \\ &= \alpha^2 - \beta^2 + 2 \left[ \frac{1}{3}\alpha(1 - \alpha - \beta) - \frac{1}{3}\beta(1 - \alpha - \beta) \right] \\ &= \alpha^2 + \frac{2}{3} [\alpha(1 - \alpha - \beta) - \beta(1 - \alpha - \beta)] - \beta^2 \\ &= \frac{(\alpha - \beta)(\alpha + \beta + 2)}{3}.\end{aligned}$$

**Definition 16.** (*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  $C$ .

**Rho de Spearman (version population) :**

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  $(X_1, Y_1)$ ,  $(X_2, Y_3)$  :

$$\rho_{XY} = \rho_C := 3([\mathbb{P}(X_1 - X_2)(Y_1 - Y_3) > 0] - [\mathbb{P}(X_1 - X_2)(Y_1 - Y_3) < 0]).$$

La distribution de  $(X_1, Y_1)$  étant  $H$  et celle de  $(X_2, Y_3)$  étant  $\Pi$  (car les variables  $X_2$  et  $Y_3$  sont indépendantes) alors d'après ce qui précède on a le théorème suivant.

**Théorème 3.14.** *Soit  $(X, Y)$  un couple de variables continues de copule  $C$ . La version population du rho de Spearman pour  $X$  et  $Y$  est donnée par :*

$$\begin{aligned} \rho_{XY} &= \rho_C \\ &= 3Q(C, \Pi) \\ &= 3 \times \left[ 4 \iint_{I^2} C(u, v) d\Pi(u, v) - 1 \right]. \end{aligned}$$

$$\rho_{XY} = 12 \iint_{I^2} C(u, v) dudv - 3.$$

### 3.15 Représentation des $L$ -comoments en termes de Copules

Soient  $X^{(1)}$  et  $X^{(2)}$  deux va ayant une fd jointe  $H$  et des marges  $F_1$  et  $F_2$  respectivement. Il existe une copule  $C$  qui représente la fd jointe en termes des marges, telles que :

$$H(x^{(1)}, x^{(2)}) = \mathbb{P}(X^{(1)} \leq x^{(1)}, X^{(2)} \leq x^{(2)}) := C(F_1(x^{(1)}), F_2(x^{(2)})).$$

Supposons que  $F$  est absolument continue,  $\partial^2 H(x, y) / \partial x \partial y := h(x, y)$ .

Rappelons que pour toutes va  $U$  et  $V$  avec  $E|U|$ ,  $E|V|$  et  $E|UV|$  finies, nous avons :

$$Cov(U, V) := \iint_{\mathbb{R}^2} [F_{U,V}(u, v) - F_U(u)F_V(v)] dudv.$$

La fd jointe de  $(U, V)$  :

$$F_{U,V}(s, t) := \int_{-\infty}^s \int_{-\infty}^t f_{U,V}(u, v) dudv; \quad s \in \mathbb{R}, \quad t \in I.$$

La densité jointe de  $(U, V)$  :

$$f_{U,V}(u, v) := J(u, v) h(x, y).$$

Le Jacobien de la transformation inverse :

$$J(u, v) := \begin{vmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{vmatrix}.$$

La densité jointe de  $(X, Y)$  :

$$h(x, y) := f_1(x) f_2(y) c(F_1(x), F_2(y)). \quad (*)$$

### 3.15.1 Représentation de $\lambda_{2[12]}$ en termes de copule

Soient :  $U = X^{(1)}$  et  $V = F_2(X^{(2)}) \implies X^{(2)} = F_2^{-1}(V)$ .

L'inverse de la matrice jacobienne de la transformation  $X^{(1)} = U$  et  $X^{(2)} = F_2^{-1}(V)$  est :

$$J(u, v) = 1/f_2(F_2^{-1}(v)).$$

La densité conjointe du vecteur bivarié  $(U, V)$  est :

$$f_{U,V}(u, v) = J(u, v) h(u, F_2^{-1}(v)).$$

De (\*) on a :

$$h(u, F_2^{-1}(v)) = f_1(u) f_2(F_2^{-1}(v)) c(F_1(u), F_2(F_2^{-1}(v))).$$

On a :

$$\begin{aligned} \frac{\partial^2 H(u, F_2^{-1}(v))}{\partial u \partial v} &= h(u, F_2^{-1}(v)) / f_2(F_2^{-1}(v)). \\ &= f_1(u) c(F_1(u), v). \end{aligned}$$

Alors :

$$h(u, F_2^{-1}(v)) = f_1(u) f_2(F_2^{-1}(v)) c(F_1(u), v).$$

La densité jointe de  $(U, V)$  devient :

$$f_{U,V}(u, v) = f_1(u) c(F_1(u), v).$$

Par conséquent :

$$\begin{aligned} F_{U,V}(s, t) &= \int_{-\infty}^s \int_0^t f_1(u) c(F_1(u), v) dudv \\ &= \int_0^{F_1(s)} \int_0^t c(u, v) dudv \\ &= C(F_1(s), t); \quad s \in \mathbb{R}; \quad t \in I. \end{aligned}$$

---

"Représentation des L-comoments en termes de Copules"

---

Ainsi, la fd jointe de  $(X^{(1)}, F_2(X^{(2)}))$  est :

$$F_{(X^{(1)}, F_2(X^{(2)}))}(s, t) = C(F_1(s), t); \quad s \in \mathbb{R}, \quad t \in I.$$

Puis :

$$\begin{aligned} Cov(X^{(1)}, F_2(X^{(2)})) &= \int_{-\infty}^{+\infty} \int_0^1 [F_{(X^{(1)}, F_2(X^{(2)}))}(s, t) - F_1(s)t] ds dt \\ &= \int_{-\infty}^{+\infty} \int_0^1 [C(F_1(s), t) - F_1(s)t] ds dt \\ &= \int_0^1 \int_0^1 [C(s, t) - st] dF_1^{-1}(s) dt, \quad \text{car } F_1^{-1}(s) = s \\ &= \int_0^1 \int_0^1 C(s, t) - st / f_1(F_1^{-1}(s)) ds dt. \end{aligned}$$

Enfin :

$$\begin{aligned} \lambda_{2[12]} &= 2Cov(X^{(1)}, F_2(X^{(2)})) \\ &= \int_0^1 \int_0^1 2(C(s, t) - st) / f_1(F_1^{-1}(s)) ds dt. \end{aligned}$$

### 3.15.2 Représentation de $\lambda_{3[12]}$ en termes de Copule

Soient :  $U = X^{(1)}$  et  $V = F_2^2(X^{(2)}) \implies X^{(2)} = F_2^{-1}(\sqrt{V})$ .

L'inverse de la matrice jacobienne de la transformation  $X^{(1)} = U$  et  $X^{(2)} = F_2^{-1}(\sqrt{V})$  est :

$$J(u, v) = 1/2\sqrt{v}f_2(F_2^{-1}(\sqrt{v})).$$

La densité conjointe du vecteur bivarié  $(U, V)$  est :

$$f_{U,V}(u, v) = J(u, v) h(u, F_2^{-1}(\sqrt{v})).$$

"Représentation des L-comoments en termes de Copules"

---

De (\*) on a :

$$\begin{aligned} h(u, F_2^{-1}(\sqrt{v})) &= f_1(u) f_2(F_2^{-1}(\sqrt{v})) c(F_1(u), F_2(F_2^{-1}(\sqrt{v}))) \\ &= f_1(u) f_2(F_2^{-1}(\sqrt{v})) c(F_1(u), \sqrt{v}). \end{aligned}$$

On a :

$$\frac{\partial^2 H(u, F_2^{-1}(\sqrt{v}))}{\partial u \partial v} = h(u, F_2^{-1}(\sqrt{v})) / 2\sqrt{v} f_2(F_2^{-1}(\sqrt{v})).$$

Alors :

$$h(u, F_2^{-1}(\sqrt{v})) = f_1(u) c(F_1(u), \sqrt{v}) / 2\sqrt{v}.$$

La densité conjointe du vecteur bivarié  $(U, V)$  est :

$$f_{U,V}(u, v) = f_1(u) c(F_1(u), \sqrt{v}) / 2\sqrt{v}.$$

Par conséquent :

$$\begin{aligned} F_{U,V}(s, t) &= \int_{-\infty}^s \int_0^t f_1(u) c(F_1(u), \sqrt{v}) / 2\sqrt{v} dudv \\ &= \int_0^{F_1(s)} \int_0^t c(u, \sqrt{v}) / 2\sqrt{v} dudv. \end{aligned}$$

On pose :  $\sqrt{v} = v' \implies dv = 2v' dv'$ .

$$\begin{aligned} F_{U,V}(s, t) &= \int_0^{F_1(s)} \int_0^{\sqrt{t}} c(u, v') dudv' \\ &= C(F_1(s), \sqrt{t}); \quad s \in \mathbb{R}, t \in I. \end{aligned}$$

Ainsi, la fd jointe de  $(X^{(1)}, F_2^2(X^{(2)}))$  est :

$$F_{(X^{(1)}, F_2^2(X^{(2)}))}(s, t) = C(F_1(s), \sqrt{t}); \quad s \in \mathbb{R}, t \in I.$$



Puis :

$$\begin{aligned}
Cov(X^{(1)}, F_2^2(X^{(2)})) &= \int_{-\infty}^{+\infty} \int_0^1 [F_{(X^{(1)}, F_2^2(X^{(2)}))}(s, t) - F_1(s)\sqrt{t}] ds dt \\
&= \int_{-\infty}^{+\infty} \int_0^1 [C(F_1(s), \sqrt{t}) - F_1(s)\sqrt{t}] ds dt \\
&= \int_0^1 \int_0^1 [C(s, \sqrt{t}) - s\sqrt{t}] dF_1^{-1}(s) dt \\
&= \int_0^1 \int_0^1 \frac{C(s, \sqrt{t}) - s\sqrt{t}}{f_1(F_1^{-1}(s))} ds dt.
\end{aligned}$$

On pose :  $\sqrt{t} = t' \implies dt = 2t' dt'$ .

$$Cov(X^{(1)}, F_2^2(X^{(2)})) = \int_0^1 \int_0^1 \frac{2t(C(s, t) - st)}{f_1(F_1^{-1}(s))} ds dt.$$

Enfin :

$$\begin{aligned}
\lambda_{3[12]} &= 6Cov(X^{(1)}, F_2^2(X^{(2)})) - 6Cov(X^{(1)}, F_2(X^{(2)})) \\
&= \int_0^1 \int_0^1 \frac{12(t-1)(C(s, t) - st)}{f_1(F_1^{-1}(s))} ds dt.
\end{aligned}$$

### 3.15.3 Représentation de $\lambda_{4[12]}$ en termes de Copule

Soient :  $U = X^{(1)}$  et  $V = F_2^3(X^{(2)}) \implies X^{(2)} = F_2^{-1}(\sqrt[3]{v})$ .

L'inverse de la matrice jacobienne de la transformation  $X^{(1)} = U$  et  $X^{(2)} = F_2^{-1}(\sqrt[3]{v})$  est :

$$J(u, v) = \sqrt[3]{v}/3v f_2(F_2^{-1}(\sqrt[3]{v})).$$

La densité conjointe du vecteur bivarié  $(U, V)$  est :

$$f_{U,V}(u, v) = J(u, v) h(u, F_2^{-1}(\sqrt[3]{v})).$$

"Représentation des L-comoments en termes de Copules"

---

De (\*) on a :

$$\begin{aligned} h(u, F_2^{-1}(\sqrt[3]{v})) &= f_1(u) \sqrt[3]{v} c(F_1(u), F_2(F_2^{-1}(\sqrt[3]{v}))) / 3v \\ &= \sqrt[3]{v} f_1(u) c(F_1(u), \sqrt[3]{v}) / 3v. \end{aligned}$$

On a :

$$\frac{\partial^2 H(u, F_2^{-1}(\sqrt[3]{v}))}{\partial u \partial v} = \frac{\sqrt[3]{v} h(u, F_2^{-1}(\sqrt[3]{v}))}{3v f_2(F_2^{-1}(\sqrt[3]{v}))}.$$

Alors :

$$h(u, F_2^{-1}(\sqrt[3]{v})) = f_1(u) f_2(F_2^{-1}(\sqrt[3]{v})) c(F_1(u), \sqrt[3]{v}).$$

La densité conjointe du vecteur bivarié  $(U, V)$  est :

$$f_{U,V}(u, v) = \sqrt[3]{v} f_1(u) c(F_1(u), \sqrt[3]{v}) / 3v.$$

Par conséquent :

$$\begin{aligned} F_{U,V}(s, t) &= \int_{-\infty}^s \int_{-\infty}^t \sqrt[3]{v} f_1(u) c(F_1(u), \sqrt[3]{v}) / 3v \, du \, dv \\ &= \int_0^{F_1(s)} \int_0^t c(u, \sqrt[3]{v}) \sqrt[3]{v} / 3v \, du \, dv. \end{aligned}$$

On pose :  $\sqrt[3]{v} = v' \implies dv = 3v'^2 dv'$ .

$$\begin{aligned} F_{U,V}(s, t) &= \int_0^{F_1(s)} \int_0^{\sqrt[3]{t}} c(u, v') \, du \, dv' \\ &= C(F_1(s), \sqrt[3]{t}); \quad s \in \mathbb{R}, \quad t \in I. \end{aligned}$$

Ainsi, la fd jointe de  $(X^{(1)}, F_2^3(X^{(2)}))$  est :

$$F_{(X^{(1)}, F_2^3(X^{(2)}))}(s, t) = C(F_1(s), \sqrt[3]{t}); \quad s \in \mathbb{R}, \quad t \in I.$$

"Représentation des L-comoments en termes de Copules"

---

Puis :

$$\begin{aligned}
Cov(X^{(1)}, F_2^3(X^{(2)})) &= \int_{-\infty}^{+\infty} \int_0^1 [F_{X^{(1)}, F_2^3(X^{(2)})}(s, t) - F_1(s)\sqrt[3]{t}] ds dt \\
&= \int_{-\infty}^{+\infty} \int_0^1 [C(F_1(s), \sqrt[3]{t}) - F_1(s)\sqrt[3]{t}] ds dt \\
&= \int_0^1 \int_0^1 [C(s, \sqrt[3]{t}) - s\sqrt[3]{t}] d(F_1^{-1}(s)) dt \\
&= \int_0^1 \int_0^1 \frac{C(s, \sqrt[3]{t}) - s\sqrt[3]{t}}{f_1(F_1^{-1}(s))} ds dt.
\end{aligned}$$

On pose :  $\sqrt[3]{t} = t' \implies dt = 3t'^2 dt'$  :

$$Cov(X^{(1)}, F_2^3(X^{(2)})) = \int_0^1 \int_0^1 \frac{3t'^2 (C(s, t') - st')}{f_1(F_1^{-1}(s))} ds dt'.$$

Enfin :

$$\begin{aligned}
\lambda_{4[12]} &= Cov(X^{(1)}, 20F_2^3(X^{(2)}) - 30F_2^2(X^{(2)}) + 12F_2(X^{(2)}) - 1) \\
&= 20 \int_0^1 \int_0^1 \frac{3t'^2 (C(s, t') - st')}{f_1(F_1^{-1}(s))} ds dt' - 30 \int_0^1 \int_0^1 \frac{2t' (C(s, t') - st')}{f_1(F_1^{-1}(s))} ds dt' \\
&\quad + 12 \int_0^1 \int_0^1 \frac{C(s, t') - ts}{f_1(F_1^{-1}(s))} ds dt'. \\
&= \int_0^1 \int_0^1 \frac{(60t'^2 - 60t' + 12) (C(s, t') - ts)}{f_1(F_1^{-1}(s))} ds dt'. \\
&= \int_0^1 \int_0^1 (60t'^2 - 60t' + 12) (C(s, t') - ts) dF_1^{-1}(s) dt'.
\end{aligned}$$

**Recapitulation :**

$$\left\{ \begin{array}{l} \lambda_{2[12]} = \int_0^1 \int_0^1 2 (C(s, t) - ts) dF_1^{-1}(s) dt. \\ \lambda_{3[12]} = \int_0^1 \int_0^1 12 (t - 1) (C(s, t) - st) dF_1^{-1}(s) dt. \\ \lambda_{4[12]} = \int_0^1 \int_0^1 (60t^2 - 60t + 12) (C(s, t) - ts) dF_1^{-1}(s) dt. \end{array} \right.$$

### 3.16 Conclusion

Dans de nombreuses applications, on s'intéresse à la modélisation de la dépendance entre deux ou plusieurs variables. Cette modélisation doit prendre en compte la dépendance extrême entre les différents processus. En effet, l'occurrence simultanée d'évènements extrêmes peut causer de graves dommages environnementaux ou structurels voire mettre en danger des populations : grandes marées et tempêtes, hauteur de vagues extrêmes, ect. Il est donc important de disposer d'outils pour construire des modèles qui prennent en compte cette dépendance extrême.

La dépendance entre les différents processus est rarement symétrique. Souvent il existe une relation d'ordre entre les variables. Par exemple on ne peut observer des vagues très hautes ayant une courte période, en revanche loin d'une tempête on observe des vagues de faibles hauteurs avec une période très longue.

Néanmoins, les copules, même si elles demeurent un outil très puissant n'ont pour l'instant pas souvent été utilisées. Les exemples exposés dans cette thèse traitent des problèmes bivariés.

En théorie, la méthode doit pouvoir être étendue à plus de deux dimensions. Mais, la plupart des articles parus à ce jour se contentent d'utiliser les copules pour des travaux en deux dimensions. Lorsque le nombre de dimensions est supérieur à deux, les auteurs se ramènent toujours à une copule elliptique (Normale ou Student) pour la simple et bonne raison que les densités de ces copules sont aisées à calculer et que ces copules sont simulables de manière simple, même lorsque le nombre de dimensions est très élevé. Or, ce type de copules n'est pas forcément le plus adapté en hydrologie.

Même s'il existe des extensions et encore sous certaines conditions, les copules Archimédiennes à plus de deux dimensions, nécessiteraient des calculs tout à fait fastidieux.

## Chapitre 4

# APPLICATION : MODÉLISATION PAR UNE COPULE DE GUMBEL

Nous nous intéressons dans cette partie à une copule particulière, soit la copule de Gumbel. Dans la littérature, elle est parfois appelée copule de Gumbel-Hougaard. Pour plus de clarté, nous traitons ici le cas bivarié. La plupart des informations sont tirées de *Nelsen* (2006).

Ce chapitre s'articule principalement autour de quatre axes. Nous définissons tout d'abord la copule de Gumbel ainsi que ses caractéristiques. Puis nous présentons les techniques de simulation. Ensuite nous présentons l'implémentation des divers outils présentés jusqu'ici, ainsi que les méthodes de calibration. Enfin, nous terminons par une application de la copule de Gumbel en hydrologie.

### 4.1 Copule de Gumbel

La copule de Gumbel (1960) est une copule des valeurs extrêmes. Elle présente l'avantage de mesurer la dépendance des événements les plus rares, puisqu'elle est asymétrique.

Cette 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 et relativement de faible dépendance sur la queue inférieure. Si les résultats sont connus pour être fortement corrélés à des valeurs élevées, mais moins corrélés à de faibles valeurs, alors la copule de Gumbel est un choix approprié.

Elle est à ce titre particulièrement adaptée en hydrologie pour étudier l'impact de la survenance d'événements de forte intensité sur la dépendance entre les variables.

## 4.2 Caractéristiques

La copule de Gumbel est une distribution de probabilité continue. Elle est utilisée pour trouver les extrêmes d'un nombre d'échantillon de plusieurs distributions.

Elle est très utilisée en hydrologie et en climatologie pour estimer les valeurs extrêmes des phénomènes. Par exemple, elle peut servir à calculer le niveau maximal d'un fleuve en possédant le relevé de crues sur dix ans. Elle peut aussi servir à prédire la probabilité d'un événement critique, comme un tremblement de terre.

### 4.2.1 Copule archimédienne

Si on prend  $\Phi_\alpha(t) = (-\ln t)^\alpha$  avec  $\alpha \geq 1$  et  $t \in I$  comme fonction génératrice.

On vérifie facilement que :

- L'inverse de la fonction génératrice est  $\Phi_\alpha^{-1}(t) = \ell^{-t^{1/\alpha}}$ .
- $\Phi_\alpha(0) = +\infty$ .
- $\Phi_\alpha(1) = 0$ .
- $\Phi'_\alpha(t) = -\frac{\alpha}{t}(-\ln t)^{\alpha-1}$  et  $\Phi'_\alpha(t) < 0$  pour  $\alpha \geq 1$ .
- $\Phi''_\alpha(t) = \frac{\alpha}{t^2}(-\ln t)^{\alpha-2}[\alpha - 1 - \ln t]$  et  $\Phi''_\alpha(t) \geq 0$  pour  $\alpha \geq 1$ .

Ainsi  $\Phi$  est une fonction continue strictement décroissante de  $I$  dans  $[0, +\infty[$ .

Nous retrouvons la famille de Gumbel par la relation  $\Phi_\alpha^{-1}(\Phi_\alpha(u) + \Phi_\alpha(v))$ . Les copules archimédiennes forment des familles de copules qui possèdent plusieurs propriétés vues précédemment.

### 4.2.2 Fonction de répartition

La fonction copule  $C_\alpha$  est par définition la fonction de répartition d'un couple de va à marginales uniformes et de loi jointe  $F_\alpha$ .

$$\forall \alpha \geq 1; \forall u, v \in I : C_\alpha(u, v) = \exp[-((-\ln u)^\alpha + (-\ln v)^\alpha)^{\frac{1}{\alpha}}].$$

Où  $\alpha \geq 1$  est le paramètre de la copule.

La figure ci-dessous présente la distribution de la copule de Gumbel pour  $\alpha = 2$  et  $\alpha = 4$ .

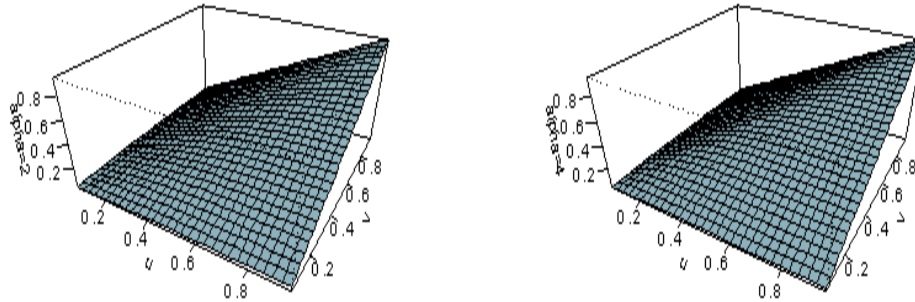


FIG. 4.1. Distribution de la copule de Gumbel pour  $(\alpha = 2)$  et  $(\alpha = 4)$ .

1

### 4.2.3 Densité

Le calcul explicite de la densité de la copule de Gumbel est complexe.

Nous savons que la densité d'une copule archimédienne, de générateur  $\Phi$  deux fois différentiable est telle que :

$$c_{\alpha}(u, v) = -\frac{\Phi'_{\alpha}(C_{\alpha}(u, v))\Phi'_{\alpha}(u)\Phi'_{\alpha}(v)}{\Phi_{\alpha}^3(C_{\alpha}(u, v))}.$$

Nous obtenons la densité suivante :

$$\forall u, v \in I : c_{\alpha}(u, v) = [C_{\alpha}(u, v)\Phi_{\alpha-1}(u)\Phi_{\alpha-1}(v)(\Phi_{\alpha}(u) + \Phi_{\alpha}(v))^{\frac{1}{\alpha}-2}(\alpha - 1 + (\Phi_{\alpha}(u) + \Phi_{\alpha}(v))^{\frac{1}{\alpha}})]/uv.$$

---

<sup>1</sup>Annexe C



## "Caractéristiques"

---

1. Sur les bords du domaine, nous avons les expressions suivantes :

$$\forall 0 < u, v < 1 : c_\alpha(u, 0) = c_\alpha(0, v) = 0; c_\alpha(u, 1) = u; c_\alpha(1, v) = v.$$

2. En sachant que la densité n'est pas définie aux points  $(0, 0)$  et  $(1, 1)$ .

2

La copule de Gumbel est caractérisée par des concentrations importantes dans les queues de distribution. Sur la figure ci-dessous, on constate que la copule de Gumbel est dissymétrique à droite.

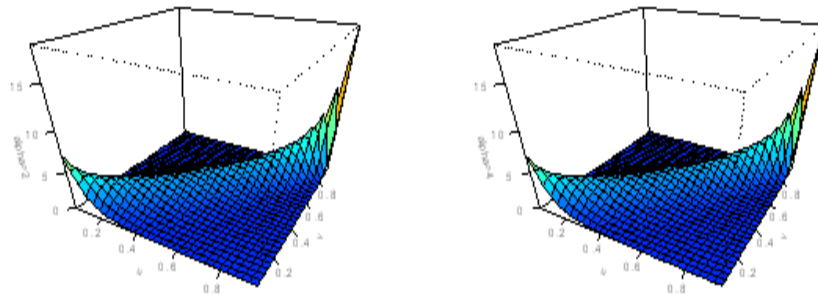


FIG. 4.2. Densité de la copule de Gumbel pour  $(\alpha = 2)$  et  $(\alpha = 4)$ .

3

### 4.2.4 Copule singulière

Comme la fonction copule  $C_\alpha$  est deux fois différentiable par rapport à  $u$  et  $v$ , alors la copule de Gumbel ne possède pas de partie singulière.

---

<sup>2</sup>On a :  $c_\alpha(0, 0) = c_\alpha(1, 1) = +\infty$

<sup>3</sup>Annexe C

### 4.3 Mesures de dépendance

Lorsque l'on évoque la dépendance, le premier mot qui vient généralement à l'esprit est celui du coefficient de corrélation linéaire. Ce coefficient de corrélation linéaire, également connu sous le nom de coefficient de corrélation de Pearson, est en effet le plus couramment utilisé.

Ce coefficient est tout à fait approprié lorsque nous étudions des distributions normales ou Student multivariées, mais celui-ci perd de son intérêt si le modèle est différent. Malheureusement, les distributions dans le domaine hydrologique suivent très rarement de telles lois.

Nous présentons ici un autre coefficient de corrélation, à savoir le coefficient de Kendall et le rho de Spearman qui présentent l'avantage de demeurer inchangé sous l'hypothèse d'une transformation strictement croissante des  $v_a$ .

#### 4.3.1 Coefficient de corrélation de Kendall

Le tau de Kendall est défini par :

$$\tau := 4 \iint_{I^2} C(u, v) dC(u, v) - 1 = 4E[C(u, v)] - 1 = 4E[X] - 1.$$

Où  $U$  et  $V$  sont deux  $v_a$  uniformes et  $X$  la  $v_a$   $C_\alpha(U, V)$ .

Pour une copule archimédienne, la fonction de répartition de  $X$  est donnée par :

$$K_C(t) := t - \frac{\Phi_\alpha(t)}{\Phi'_\alpha(t)}; \quad t \in [0, 1].$$

Soit dans le cas de la copule de Gumbel :

$$K_C(t) = t - \frac{t \ln t}{\alpha}; \quad \text{car } \Phi_\alpha(t) = (-\ln t)^\alpha.$$

Ainsi, la densité de  $X$  :

$$K'_C(t) := 1 - \frac{1}{\alpha} - \frac{\ln t}{\alpha}.$$

Par conséquent, on a :

$$\begin{aligned}
E(X) &= \int_0^1 tK'_C(t)dt = \int_0^1 t \left[ 1 - \frac{1}{\alpha} - \frac{\ln t}{\alpha} \right] dt \\
&= \left[ \int_0^1 tdt - \int_0^1 \frac{t}{\alpha} dt - \int_0^1 \frac{t \ln t}{\alpha} dt \right] \\
&= \left[ \frac{t^2}{2} \left( 1 - \frac{1}{\alpha} \right) \right]_0^1 - \int_0^1 \frac{t \ln t}{\alpha} dt \\
&= \frac{1}{2} \left( 1 - \frac{1}{\alpha} \right) - \frac{1}{\alpha} \left( \left[ \frac{t^2}{2} \ln t \right]_0^1 - \left[ \frac{t^2}{4} \right]_0^1 \right) \\
&= \frac{1}{2} \left( 1 - \frac{1}{\alpha} \right) - \frac{1}{\alpha} \left( -\frac{1}{4} \right) \\
&= \frac{1}{2} \left( 1 - \frac{1}{\alpha} \right) + \frac{1}{4\alpha} \\
&= \frac{1}{2} \left( 1 - \frac{1}{2\alpha} \right).
\end{aligned}$$

D'où le tau de Kendall s'exprime de la manière suivante :

$$\tau = 4 \left[ \frac{1}{2} \left( 1 - \frac{1}{2\alpha} \right) \right] - 1 = \frac{\alpha - 1}{\alpha}.$$

Le paramètre  $\alpha$  mesure le degré de dépendance entre les risques. Plus il est élevé, plus la dépendance est forte. On retrouve le fait que  $C_\infty = C^+$ , puisque le tau de Kendall tend vers 1, i.e si  $\alpha \rightarrow \infty$  alors :  $\tau \rightarrow 1$ .

#### 4.3.2 Coefficient de corrélation de Spearman

Le  $\rho$  de Spearman se définit par :

$$\rho := 12 \iint_{I^2} C(u, v) dudv - 3.$$

Cependant, il n'y a pas de formules explicites pour le  $\rho$  de Spearman dans le cas de la copule de Gumbel.

La dépendance des paramètres est restreinte à l'intervalle  $[1, +\infty[$ . Les valeurs de 1 et de  $\infty$  correspondent à l'indépendance et à la limite supérieure de *Fréchet*. Mais cette copule n'a pas atteint les limites inférieures de *Fréchet* pour toute valeur  $\alpha$ .

**Definition 17.** *La copule de Gumbel est une copule extrême car elle vérifie la propriété du max-stabilité :*

$$C(u^{\frac{1}{n}}, v^{\frac{1}{n}})^n := C(u, v).$$

La propriété de max-stabilité s'interprète de la manière suivante. Soient deux va  $(X_1, \dots, X_n)$  et  $(Y_1, \dots, Y_n)$  identiquement distribuées telles que les couples  $(X_i, Y_i)$  et  $(X_n, Y_n)$  ont la même copule  $C$ .

On pose  $C_{\max}$  la copule du couple  $(X_n, Y_n) = (\max_{1 \leq i \leq n} X_i, \max_{1 \leq i \leq n} Y_i)$ . D'après *Nelsen* (2005), nous obtenons la propriété suivante :

$$C_{\max} = C_{\alpha}(u^{\frac{1}{n}}, v^{\frac{1}{n}})^n := C_{\alpha}(u, v).$$

Autrement dit pour la copule de Gumbel, le couple  $(X_i, Y_i)$  et le couple  $(X_n, Y_n)$  ont la même copule.

**Démonstration.**

□

La copule de Gumbel vérifie la propriété de max-stabilité :

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

$$C_{\alpha}(u^{\frac{1}{n}}, v^{\frac{1}{n}})^n = C_{\alpha}(u, v).$$

□

#### 4.4 Copule multivariée

On sait que pour les copules archimédiennes la propriété de super modularité en dimensions  $n$  est équivalente à ce que le générateur soit complètement monotone (i.e.

$$\forall i : (-1) \frac{d^i \Phi(t)}{dt^i} \geq 0.$$

D'après *Nelson* (2006), le générateur de la copule de Gumbel est complètement monotone. Par conséquent, la copule de Gumbel multivariée se définit de la manière suivante :

$$C_\alpha(u_1, \dots, u_n) := \exp \left( - \left[ \sum_{i=1}^n (-\ln u_i)^\alpha \right]^{\frac{1}{\alpha}} \right).$$

Pour une copule archimédienne, la densité s'exprime de la manière suivante :

$$c_\alpha(u_1, \dots, u_n) = (\Phi^{-1})^{(n)} (\Phi(u_1) + \dots + \Phi(u_n)) \prod_{i=1}^n \Phi'(u_i).$$

D'après *Savu & Trede* (2006). Malheureusement pour la copule de Gumbel, la dérivée  $n^{\text{ème}}$  de l'inverse du générateur n'a pas d'expressions explicites.

Pour  $n = 3$ , nous obtenons :

$$c_\alpha(u_1, u_2, u_3) = C_\alpha(u_1, u_2, u_3) \frac{\Phi_{\alpha-1}(u_1)\Phi_{\alpha-1}(u_2)\Phi_{\alpha-1}(u_3)}{u_1 u_2 u_3} \sum^{\frac{1}{\alpha}-3} \left[ (2\alpha-1)(\alpha-1) + 3(\alpha-1) \sum^{\frac{1}{\alpha}} + \sum^{\frac{2}{\alpha}} \right].$$

Où  $\sum$  est défini par :  $\Phi(u_1) + \Phi(u_2) + \Phi(u_3)$ .

## 4.5 Mise en oeuvre des copules

Nous avons défini la copule de Gumbel, énoncé ces propriétés fondamentales. Nous allons maintenant présenter deux méthodes essentielles aux statisticiens qui utilisent les copules : les méthodes de simulation et les méthodes d'inférence statistique.

### 4.5.1 Techniques de simulation

Souhaiter simuler une copule bivariable  $C$  signifie simuler les arguments  $u_1$  et  $u_2$  de cette fonction tirés d'un vecteur aléatoire de loi uniforme  $U = (U_1, U_2)$  de distribution  $C$ . Ceci permet alors de déterminer la simulation d'un vecteur aléatoire  $X = (X_1, X_2)$  dont la structure de dépendance est définie par  $C$  et dont nous définissons des marges particulières  $F_1$  et  $F_2$ . Pour cela nous utilisons la transformation  $X = (F_1^{-1}(U_1), F_2^{-1}(U_2))$ .

Ceci est fondé sur la représentation de la copule et de la distribution  $F$  donnée par cette formule :

$$C(u_1, u_2) := F(F_1^{-1}(u_1), F_2^{-1}(u_2)).$$

Toute la difficulté est de simuler des nombres aléatoires issus de  $C$ . La simulation de la copule produit est relativement simple : on simule des réalisations de loi uniforme sur  $[0, 1]$  de manière indépendante. Mais pour la majorité des copules, cela s'avère plus complexe.

Il existe plusieurs algorithmes pour simuler des va selon une copule bien précise. Les deux algorithmes suivants sont spécifiques aux copules archimédiennes.

### Algorithmes des générateurs aléatoires

Le générateur aléatoire est effectué avec deux algorithmes :

#### Algorithme " $K_C$ "

Cette algorithme utilise le fait que pour une copule archimédienne, la fonction de distribution de la va  $C_\alpha(U, V)$  est donnée par :

$$K_C(t) := t - \frac{\Phi_\alpha(t)}{\Phi'_\alpha(t)}; \quad t \in [0, 1].$$

Soit dans le cas de la copule de Gumbel :

$$K_C(t) = t - \frac{t \ln t}{\alpha}.$$

---

 "Mise en oeuvre des copules"
 

---

La procédure à suivre pour simuler des paires  $(u, v)$  est donc :

1. Simuler indépendamment  $y$  et  $t$  suivant la loi uniforme :  $(y, t) \stackrel{iid}{\sim} \mathcal{U}(0, 1)$ .
2. Simuler  $x$  d'une distribution "K" i.e. déterminer :  $x := K_C^{-1}(t)$ .
3. Prendre : 
$$\begin{cases} u := \Phi_\alpha^{-1}(\Phi_\alpha(x)y). \\ v := \Phi_\alpha^{-1}(\Phi_\alpha(x)(1 - y)). \end{cases}$$

L'algorithme "K" ne fonctionne pas avec dimension  $> 2$  et la non existence de formule explicite pour  $K_C^{-1}$  nous ont incité à chercher un autre algorithme plus puissant.

### Représentation à facteurs communs de Marshall & Olkin (1988)

L'approche de *Marshall & Olkin* consiste à utiliser une va  $\theta$  pour un vecteur aléatoire  $(X_1, X_2)$  telles que les composantes  $X_i$  sont indépendantes conditionnellement à  $\theta$ . La loi jointe du vecteur  $(X_1, X_2)$  est donnée par :

$$F_{(X_1, X_2)}(x_1, x_2) := L_\theta \left( \sum_{i=1}^2 L_\theta^{-1}(F_{X_i}(x_i)) \right).$$

Où  $L_\theta$  désigne la transformée de Laplace de la va  $\theta$ .

De plus, l'algorithme suivant permet de simuler des va de copule archimédienne  $C_\alpha$  :

1. Simuler indépendamment  $x_1$  et  $x_2$  suivant la loi exponentielle :  $(x_1, x_2) \stackrel{iid}{\sim} \mathcal{E}(1)$ .
2. Simuler  $\theta$  suivant la loi de la transformée de Laplace  $\phi^{-1}$ .
3. Prendre :  $\forall i; u_i := \Phi_\alpha^{-1}\left(\frac{x_i}{\theta}\right)$ .

Dans le cas de la copule de Gumbel,  $\phi^{-1}(t) = \ell^{-t^{\frac{1}{\alpha}}}$ , qui est la transformée de Laplace d'une loi stable de paramètres  $(\frac{1}{\alpha}, 0, 1, 0)$ . Cet algorithme fonctionne avec n'importe quelle dimension. Pour plus de détails sur les lois stables, se reporter à *Nolan* (2005). Enfin pour simuler des va de loi stable, nous avons utilisé l'algorithme de *Chambers & al.* (1976).

### Exemple de graphiques

En traçant 10000 réalisations des composantes d'un couple  $(U, V)$  de loi jointe  $C_\alpha$  (pour  $\alpha = 1, 2, 3, 4$ ), on obtient les graphiques suivants :

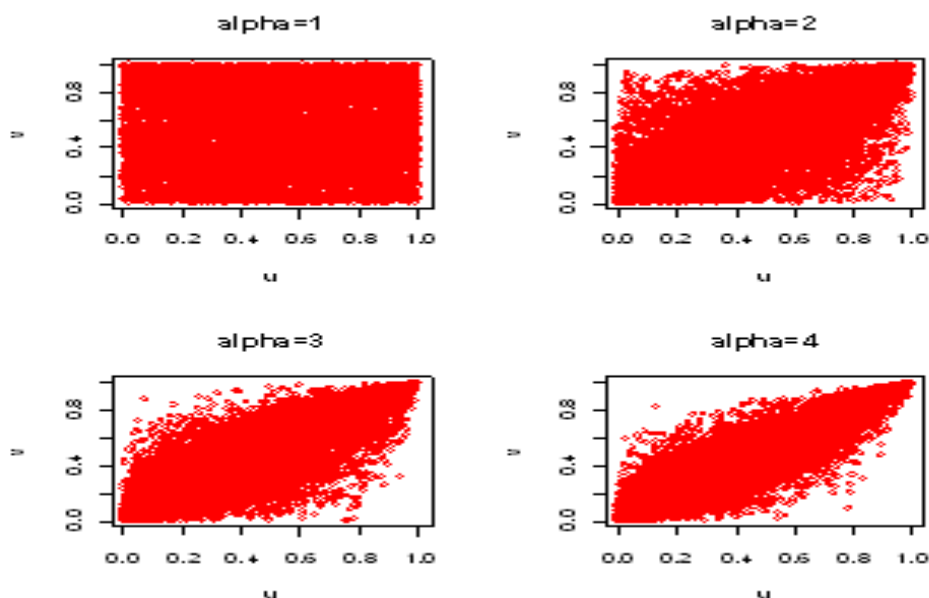


FIG. 4.3. Simulation des paires  $(u, v)$  ayant une distribution de la copule de Gumbel

4

#### 4.5.2 Estimation et implémentation

La deuxième partie sera consacrée à l'estimation des paramètres de la copule qui est une question importante lorsqu'on essaye d'obtenir une compréhension précise sur la structure de dépendance (comportement de la queue). Le comportement dans le voisinage de  $(0, 0)$  ou  $(1, 1)$  décrit plus précisément la force de la dépendance entre les événements extrêmes.

Nous rappelons d'abord les méthodes d'estimation pour calibrer une copule. Ensuite, nous présentons notre package *R* intitulé *gumbel* implémentant tous les outils nécessaires à la copule de Gumbel.

#### Méthodes de calibrage

Nous supposons avoir un échantillon de données de  $n$  couples dont nous tentons d'estimer la fonction de dépendance. Nous présentons les méthodes d'estimation classiques.

---

<sup>4</sup>Annexe C



---

 "Mise en oeuvre des copules"
 

---

Puis, nous déterminons des outils graphiques d'ajustement de la copule estimée à l'échantillon des données bivariées.

Tout comme les copules archimédiennes de Clayton et de Frank, la copule de Gumbel présente l'avantage d'avoir un seul paramètre à calibrer. Nous faisons un bref rappel des méthodes d'estimation spécifiques aux copules.

### Méthode des moments

Cette méthode consiste à estimer les paramètres  $\theta$  des lois marginales et le paramètre  $\alpha$  de la copule par la méthode des moments i.e :

1. Résoudre le système a  $d$  équations et  $d$  inconnues :

$$\begin{cases} \bar{X}_n = f(\theta_1, \dots, \theta_d). \\ \mu_{2,n} = g(\theta_1, \dots, \theta_d). \\ \mu_{3,n} = h(\theta_1, \dots, \theta_d). \\ \vdots \\ \vdots \\ \vdots \end{cases}$$

Où  $d$  désigne la dimension de  $\theta$ ,  $f$ ,  $g$  et  $h$  sont les expressions des moments (ordinaires) d'ordre 1, 2 et 3 en fonction du paramètre  $\theta$ . Répéter cette étape pour toutes les marginales.

2. Inverser le tau de Kendall où le rho de Spearman pour obtenir le paramètre de la copule.

Pour des marginales exponentielles  $\mathcal{E}(\lambda)$  et une copule de Gumbel, on trouve :

$$\hat{\lambda}_n = \frac{1}{\bar{X}_n} \text{ et } \hat{\alpha}_n = \frac{1}{1 - \hat{\tau}_n}.$$

Où  $\tau_n$  désigne le tau de Kendall empirique.

La méthode des moments est fondée dans le cas des copules sur les relations existant entre le  $\tau$  de Kendall et les paramètres des copules. C'est le cas des copules Archimédiennes par exemple.

Nous présentons le cas particulier de l'estimation par la méthode des moments du paramètre  $\alpha$  de la copule Gumbel. En effet, rappelons la relation que nous avons entre  $\alpha$  et le  $\tau$  de Kendall :

$$\tau = 1 - \frac{1}{\alpha}.$$

Nous en déduisons que l'estimation  $\hat{\alpha}$  est égale à :

$$\hat{\alpha} = \frac{1}{1 - \hat{\tau}}.$$

### Maximum de vraisemblance exact

Dans le cas où la densité de la copule existe (les lois marginales sont connues), on peut utiliser les estimateurs du maximum de vraisemblance. Pour simplifier, on suppose qu'on utilise une copule bivariee  $C_\alpha$ , ayant une densité et que les lois des marginales possèdent aussi des densités.

On note  $\theta_1$  et  $\theta_2$  les paramètres des lois marginales. Le log vraisemblance s'écrit :

$$\begin{aligned} \mathcal{L}(\alpha, \theta_1, \theta_2, x_1, \dots, x_n, y_1, \dots, y_n) &= \prod_{i=1}^n (c(F_X(x_i, \theta_1), F_Y(y_i, \theta_2), \alpha)) \\ &\quad \times \prod_{i=1}^n (f_X(x_i, \theta_1)) \times \prod_{i=1}^n f_Y(y_i, \theta_2)). \\ \ln \mathcal{L}(\alpha, \theta_1, \theta_2, x_1, \dots, x_n, y_1, \dots, y_n) &= \sum_{i=1}^n \ln(c(F_X(x_i, \theta_1), F_Y(y_i, \theta_2), \alpha)) \\ &\quad + \sum_{i=1}^n \ln(f_X(x_i, \theta_1)) + \sum_{i=1}^n \ln f_Y(y_i, \theta_2)). \end{aligned}$$

Bien souvent, il n'existe pas d'expressions explicites des estimateurs maximisant  $\ln \mathcal{L}$ .

### Inférence sur les marginales

Toujours dans l'hypothèse où la copule a une densité, on peut mélanger les deux premières approches, en estimant d'abord les paramètres des lois marginales, puis en estimant le paramètre de la copule. Cela consiste à :

1. Estimer les paramètres  $\theta_1$  et  $\theta_2$  par le maximum de vraisemblance.
2. Construire les pseudo données  $\forall 1 \leq i \leq n : u_i = F_X(x_i, \hat{\theta}_1)$  et  $v_i = F_Y(y_i, \hat{\theta}_2)$ .
3. Estimer le paramètre  $\alpha$  en maximisant le log-vraisemblance :

$$\ln \mathcal{L}(\alpha, u_1, \dots, u_n, v_1, \dots, v_n) = \sum_{i=1}^n \ln c(u_i, v_i, \alpha).$$

Cette méthode présente l'avantage d'utiliser les estimateurs classiques de maximum vraisemblance des marginales.

### Maximum de vraisemblance canonique

C'est une méthode semi-paramétrique, qui se base sur la méthode précédente :

1. Calculer les fonctions de répartition empirique  $F_{X,n}$  et  $F_{Y,n}$ .
2. Construire les pseudo données  $\forall 1 \leq i \leq n : u_i = F_{X,n}(x_i)$  et  $v_i = F_{Y,n}(y_i)$ .

1. Estimer le paramètre  $\alpha$  en maximisant le log-vraisemblance :

$$\ln \mathcal{L}(\alpha, u_1, \dots, u_n, v_1, \dots, v_n) = \sum_{i=1}^n \ln c(u_i, v_i, \alpha).$$

### 4.5.3 Outils graphiques d'ajustement de la copule

L'ajustement doit être réalisé pour retenir la copule qui modélise au mieux les observations.

Pour l'évaluation de l'ajustement d'une copule à la structure de dépendance d'un échantillon, nous pouvons utiliser des graphes appelés dépendogrammes, permettant notamment d'observer la dépendance des queues.

Le dépendogramme représente la structure de dépendance sous la forme d'un nuage de points des marges uniformes  $(u, v)$  extraites de l'échantillon ou encore de simulations d'une copule théorique. Précisons que les couples  $(u, v)$  issues de l'échantillon forment ce que nous appelons la copule empirique. Cette dernière est simplement définie par les statistiques de rang issues de l'échantillon. Nous comparons le dépendogramme de cette copule empirique à celui d'autres copules théoriques estimées sur l'échantillon.

A cet effet, *Venter* (2002, 2003) & *Belguise* (2001) proposent d'utiliser des fonctions descriptives pour tester l'adéquation d'une copule ou pour sélectionner une copule ayant des caractéristiques spécifiques par exemple au niveau des queues de distribution. Il s'agit de comparer les valeurs empiriques et les valeurs estimées pour les différentes copules. Cette comparaison peut être faite par un test de Kolmogorov-Smirnov ou du  $\chi^2$  au seuil  $\alpha$ .

Cette procédure a été proposée par *Genest & Rivest* (1993) et *Genest & al.* (2006), ils utilisent les tests de Kolmogorov-Smirnov et Cramér-von Mises pour l'adéquation des copules. Pour un ajustement complémentaire des copules, *Hürlimann* (2004) propose le test du  $\chi^2$  bidimensionnel.

#### 4.5.4 Package R gumbel

Nous avons utilisé le package *R* pour la copule de Gumbel. Les package *R* présentent le gros avantage de fournir rapidement du code documenté et donc réutilisable par quelqu'un d'autres que ses concepteurs.

Notre package gumbel fournit :

- **La fonction de répartition** : pgumbel (valable en dimension  $n$ ).
- **La densité** : dgumbel (valable pour  $n \leq 3$ ).
- **Un générateur aléatoire** : rgumbel (valable en dimension  $n$  quelconque).
- **La fonction génératrice** : phigumbel.
- **L'inverse de la fonction génératrice** : invphigumbel.

De plus, nous avons implémenté les méthodes d'estimation précédentes :

- **La méthode des moments "Moment-Based Estimation"** : gumbel.MBE.
- **Le maximum de vraisemblance exacte "Exact Maximum Likelihood"** :  
gumbel.EML.
- **L'inférence sur les marginales "Inference For Margins"** : gumbel.IFM.
- **Le maximum de vraisemblance canonique "Canonical Maximum Likelihood"** : gumbel.CML.

L'aide (en anglais) pour les différentes fonctions s'obtient avec la commande `help` le nom de la fonction ou plus simplement `help(gumbel)` après avoir chargé le package a l'aide `library("gumbel")`. Dans cette aide figure les détails sur l'utilisation des différentes fonctions ainsi que des exemples.

<sup>5</sup>

## Conclusion

La copule de Gumbel a l'avantage de décrire les dépendances asymétriques, où les coefficients de queue inférieure et de queue supérieure diffèrent. Elle possède la caractéristique de pouvoir représenter des risques dont la structure de dépendance est accentuée sur la queue supérieure. De manière générale, la copule de Gumbel est à ce titre particulièrement adaptée en assurance, en finance et en hydrologie pour étudier l'impact de la survenance d'événements de forte intensité sur la dépendance entre les va.

Parallèlement au contexte de changement climatique global, la vulnérabilité face aux événements hydrologiques extrêmes est en constante augmentation. Une urbanisation mal maîtrisée a conduit la population à être de plus en plus exposée à divers risques (inondations, tempêtes, sécheresses,...). En outre, la surmédiation de tels événements ces dernières années, comme les inondations de novembre 1999 dans l'Aude en France (*Bechtold & Bazile, 2001 ; Neppel & al., 2001*) ou de septembre 2002 dans le Gard (*Neppel & al., 2003*), a conduit à la perception d'une recrudescence des événements extrêmes.

Cependant, si l'effet du changement climatique sur les températures maximales est avéré (GIEC, 2001), son impact sur le régime des pluies extrêmes n'est pas établi. De plus, un tel changement, s'il a lieu ne peut être localisé en un point et a obligatoirement eu un impact à l'échelle régional.

## CONCLUSION GÉNÉRALE

Néanmoins, les copules, même si elles demeurent un outil très puissant n'ont pour l'instant pas souvent été utilisées. Les exemples exposés dans ce travail traitent des problèmes bivariés.

En théorie, la méthode doit pouvoir être étendue à plus de deux dimensions. Mais, la plupart des articles parus à ce jour se contentent d'utiliser les copules pour des travaux en deux dimensions. Lorsque le nombre de dimensions est supérieur à deux, les auteurs se ramènent toujours à une copule elliptique (Normale ou de Student) pour la simple et bonne raison que les densités de ces copules sont aisées à calculer et que ces copules sont simulables de manière simple, même lorsque le nombre de dimensions est très élevé.

Or ce type de copules n'est pas forcément le plus adapté à l'étude des phénomènes hydrologiques. Même s'il existe des extensions et encore sous certaines conditions, des copules Archimédiennes à plus de deux dimensions, établir les dérivées nécessiterait des calculs tout à fait fastidieux.

On note que définir un modèle régional de changement permet de détecter, le cas échéant, un changement cohérent spatialement, donc cohérent avec l'hypothèse du changement climatique. Si la copule de Gumbel semble mieux modéliser la dépendance des valeurs rares (précipitations extrêmes), le choix de la copule utilisée a une forte influence sur le calcul, l'estimation, ect.

Par conséquent, il semble nécessaire, à ce stade, de tester la qualité d'ajustement des copules afin de sélectionner la plus appropriée aux données.

Nous espérons que ce travail a permis d'appréhender précisément les tenants et les aboutissants de la modélisation de la structure de dépendance d'un échantillon à travers les copules. Malgré sa relative simplicité mathématique, l'outil copule n'est pas toujours utilisé dans toutes ses potentialités. En hydrologie, de nombreux travaux sont encore à réaliser pour appliquer cet objet rigoureusement et de manière adéquate dans des domaines variés.

## BIBLIOGRAPHIE

- [1] **A. Charpentier**, (2003). *Tail distribution and dependence measures*. Available from : <http://www.actuaries.org/ASTIN/Colloquia/Berlin/Charpentier.pdf>.
- [2] **A. C. Favre, S. El Adlouni, T. L. Perreault, N. Thiemonge & B. Bobée**, (2004). *Multivariate hydrological frequency analysis using copulas*. Water Resource. Res., 40, W01101, Pp. 1-12.
- [3] **A. Sklar**, (1959). *Fonctions de répartition à  $n$  dimensions et leurs marges*. Publications de l'Institut de Statistique de L'Université de Paris, No. 8, Pp. 229-231.
- [4] **A. S. Klugman & R. Parsa**, (1999). *Fitting bivariate loss distribution with copulas*. Insurance Mathematics and Economics, Elsevier.
- [5] **A. Stuart**, (1954). *The correlation between variate values and ranks in samples from a continuous distribution*. Journal of Matimatics Research, Vol. 1, No. 1, Pp. 8-18.
- [6] **A. W. Marshall & I. Olkin**, (1988). *Families of multivariate distributions*. Journal of the American Statistical Association, Vol. 83, No. 403.
- [7] **B. Renard & M. Lang**, (2006). *Use of a Gaussian copula for multivariate extreme value analysis : some case studies in hydrology*. Available from : <http://www.cvr.b.fr/ifrance.com/article3.pdf>.
- [8] **C. Genest & A. C. Favre**, (2007). *Everything you always wanted to know about copula modeling but were afraid to ask*. Journal of Hydrologic Engineering, No.12, Pp. 347-368.
- [9] **C. Anne-lise, D. Christophe, L. Veronique & N. Triet**, (2008). *Copule de Gumbel*. Available from : <http://www.cran.r-project.org/web/packages/gumbel/vignettes/gumbel.pdf>.
- [10] **C. De Michele & G. Salvadori**, (2003). *A generalized pareto intensity-duration model of storm rainfall exploiting 2-copulas*. Journal of Geophysical Research, 108(D2), 4067, doi : 10.1029/2002JD002534.
- [11] **C. De Michele, N. T. Kottegoda & R. Rosso**, (2002). *IDAF curves of extreme storm rainfall : a scaling approach*. Water Science and Technology, Vol. 45, No. 2, Pp 83-90.
- [12] **C. Genest & J. C. Boies**, (2003). *Detecting dependance with Kendall plots*. Journal of the American Statistician, No. 57, Pp. 275-284.
- [13] **C. Genest & L. Rivest**, (1993). *Statistical inference procedures for bivariate Archimedean copulas*. Journal of American Statistical Association, Vol. 88, No. 423.

- [14] **C. Genest & R. J. Mackay**, (1986). *Copules archimédiennes et familles de lois bidimensionnelles dont les marges sont données*. The Canadian Journal of Statistics, Vol. 14, No. 2, Pp. 145-159.
- [15] **Chambers & al**, (1976). *A method for simulating stable random variables*. Journal of American Statistical Association, Vol. 71, No. 354.
- [16] **C. R. Harvey & A. Siddique**, (2000). *Conditional skewness in asset pricing tests*. Journal of Finance, Vol. 55, No. 3, Pp. 1263-1295.
- [17] **D. H. Burn**, (1990). *Evaluation of regional flood frequency analysis with a region of influence approach*. Water Resource. Res., 26(10), Pp. 2257–2265.
- [18] **D. J. G. Farlie**, (1960). *The performance of some correlation coefficients for a general bivariate distribution*. Biometrika, 47(3-4), Pp. 307-323, doi : 10.1093/biomet/47.3-4.307.
- [19] **D. Shi**, (1995). *Fisher Information for a Multivariate extreme value distribution*. Biometrika, 82(3) : Pp. 644-649, doi : 10.1093/biomet/82.3.644.
- [20] **E. A. Elamir & A. H. Seheult**, (2004). *Exact variance structure of sample L-moments*. Journal of Statistical Planning and Inference, Vol. 124, No. 2, 1, Pp. 337-359.
- [21] **E. Gargouri-Elloze & A. Chebchoub**, (2008). *Modélisation de la structure de dépendance hauteur–durée d’événements pluvieux par la copule de Gumbel*. Hydrological Sciences Journal, Vol. 53, No. 4, Pp. 802-817.
- [22] **E. Schechtman & S. Yitzhaki**, (1987). *A measure of association based on Gini’s mean difference*. Communications in statistics. Theory and methods, Vol. 16, No. 1, Pp. 207-231.
- [23] **E. Schechtman & S. Yitzhaki**, (1999). *On the proper bounds of the Gini correlation*. Economics Letters, Vol. 63, Pp. 133-138.
- [24] **F. Chebana & T. B. M. J. Ouarda**, (2007). *Test d’homogénéité avec les L-moments multivariés*. Water Resource. Res., 43, W08406, doi : 10.1029/2006WR005639.
- [25] **G. Salvadori, C. De Michele, N. T. Kottegoda & R. Rosso**, (2007). *Extremes in nature : an approach using copulas*. Springer.
- [26] **G. Salvadori & C. De Michele**, (2004). *Frequency analysis via copulas : theoretical aspects and applications to hydrological events*. Water Resource. Res., 40, W12511, doi :10.1029/2004WR003133.
- [27] **G. Venter**, (2002). *Tails of copulas*. Available from : [http://www.actuaries.asn.au/IAA/upload/public/venter\\_copulas\\_paper.pdf](http://www.actuaries.asn.au/IAA/upload/public/venter_copulas_paper.pdf).
- [28] **H. A. William**, (2006). *The lmomco package in R : L-moments, trimmed L-moments, L-comoments, and many distributions*. Available from : <http://cran.cnr.berkeley.edu/src/contrib/Descriptions/lmomco.html>.



- [29] **H. N. Nagaraja & H. A. David**, (2003). *Functions of concomitants of order statistics*. Journal of the Indian Society for Probability and Statistics, Vol. 7, Pp. 16-32.
- [30] **J.A. Tawn**, (1990). *Modelling Multivariate Extreme Value Distributions*. Biometrika, 77(2), Pp. 245-253, doi : 10.1093/biomet/77.2.245.
- [31] **J. M. Chambers, C. L. Mallows & B. W. Stuck**, (1976). *A method for simulating stable random variables*. Journal of the American Statistical Association, Vol. 71, No. 354.
- [32] **J. P. Nolan**, (2005). *Stable distributions : models for heavy tailed data*. Available from : <http://academic2.american.edu/~jpnolan/stable/chap1.pdf>.
- [33] **J. R. M. Hosking**, (1990). *L-moments : analysis and estimation of distributions using linear combinations of order statistics*. Journal of the Royal Statistical Society, Vol. 52, No. 1, Pp. 105-124.
- [34] **J. R. M. Hosking & J. R. Wallis**, (1997). *Regional frequency analysis : an approach based on L-moments*. Cambridge University Press.
- [35] **K. Ghoudi, A. Khoudraji & L. P. Rivest**, (1998). *Propriétés statistiques des copules de valeurs extrêmes bidimensionnelles*. Canadian Journal of Statistics, Vol. 26, No. 1, Pp. 187-197.
- [36] **L. Zhang & V. P. Singh**, (2006). *Bivariate flood frequency analysis using the copula method*. Journal of Hydrologic Engineering, Vol. 11, No. 2, Pp. 150-164.
- [37] **Marshall & Olkin**, (1988). *Families of multivariate distributions*. Journal of the American Statistical Association, Vol. 83, No. 403, Pp. 834-841.
- [38] **O. Belguise**, (2001). *Tempêtes : étude des dépendances entre les branches auto et incendie avec la théorie des copules*. Available from : [http://www.actuaries.org/ASTIN/Colloquia/Berlin/Belguise\\_Levi.pdf](http://www.actuaries.org/ASTIN/Colloquia/Berlin/Belguise_Levi.pdf).
- [39] **P. Claus**, (2008). *Théorie des copules*. Available from : [http://www.ensai.com/userfiles/Cours\\_Copules.pdf](http://www.ensai.com/userfiles/Cours_Copules.pdf).
- [40] **P. Embrechts, A. McNeil & D. Straumann**, (1999). *Correlation and dependence in risk management : properties and pitfalls*. Available from <http://www.math.ethz.ch/~embrechts>. <http://cite-seer.ist.psu.edu/article/embrechts99correlation>.
- [41] **R. B. Nelsen**, (1999). *An Introduction to Copulas*. Springer.
- [42] **R. B. Nelsen**, (2005). *Dependence Modeling with Archimedean Copulas*. Available from : [www.lclark.edu/~mathsci/brazil2.pdf](http://www.lclark.edu/~mathsci/brazil2.pdf).
- [43] **R. Serfling & P. Xiao**, (2007). *A contribution to multivariate L-moments : L-comoment matrices*. Journal of Multivariate Analysis, Vol. 98, No. 9, Pp. 1765-1781.

- [44] **S. El Adlouni, A. C. Favre & B. Bobée**, (2004). *Multivariate frequency analysis using Copulas*. Water Resource. Res., 40, W01101, doi : 10.1029/2003WR002456.
- [45] **S. S. Yang**, (1977). *General distribution theory of the concomitants of order statistics*. Journal of Statistics, Vol. 5, No. 5, Pp. 996-1002.
- [46] **S. S. Yang**, (1981). *Linear functions of concomitants of order statistics with application to nonparametric estimation of a regression function*. Journal of the American Statistical Association, Vol. 76, No. 375.
- [47] **T. B. M. J. Ouarda, J. C. Ondo & B. Bobée**, (1997). *Literature review of homogeneity and independent tests*. Available from : <http://ete.inrs.ca/pub/rapports/R000933.pdf>.
- [48] **T. B. M. J. Ouarda, S. Yue, B. Bobée, P. Legendre & P. Bruneau**, (1999). *The Gumbel mixed model for flood frequency analysis*. Journal of Hydrology, Vol. 226, No. 1-2, Pp. 88-100.
- [49] **T. B. M. J. Ouarda, S. Yue & B. Bobée**, (2001). *A review of bivariate gamma distributions for hydrological applications*. Journal of Hydrology, Vol. 246, No. 1-4, Pp. 1-18.
- [50] **Venter, G.** (2002) *Tails of copulas*. Available from : [http://www.actuaries.asn.au/IAA/upload/public/venter\\_copulas\\_paper.pdf](http://www.actuaries.asn.au/IAA/upload/public/venter_copulas_paper.pdf).
- [51] **W. R. Schucany, W. C. Parr & J. E. Boyer**, (1978). *Correlation structure in Farlie-Gumbel-Morgenstern distributions*. Biometrika, Vol. 71, No. 3, Pp. 633-636.

## ANNEXE A

### Historique

En ce qui concerne l'approche statistique, divers auteurs l'ont adoptée : *Kurothe & al.* (1997) et *Goel & al.* (2000) ont modélisé la relation entre l'intensité moyenne et la durée de pluie par une loi jointe bi-exponentielle (modèle bivarié). De même, *De Michele & Salvadori* (2003) analysent la dépendance entre les variables intensité moyenne et durée de la pluie en modélisant chacune des lois marginales par la loi Pareto généralisée et leur loi jointe par la copule de Frank. *Salvadori & De Michele* (2004) utilisent cette modélisation pour dériver analytiquement la distribution statistique de la hauteur de pluie (produit de l'intensité moyenne de pluie et sa durée). *Salvadori & De Michele* (2006) ont étendu leur étude à trois variables et ont considéré un modèle trivarié entre l'intensité moyenne de pluie, sa durée et l'épisode sec (période séparant deux averses) pour caractériser statistiquement la structure temporelle d'épisodes pluvieux : ils proposent une procédure d'estimation de la distribution de la hauteur de pluie en tenant compte des conditions antécédentes de l'averse.

La théorie des copules permet de modéliser la structure de dépendance entre les variables indépendamment des lois marginales (*Sklar*, 1959). Elle a été largement appliquée dans divers domaines, par exemple en hydrologie, *Favre & al.* (2004) présentent notamment deux applications : la première concerne la modélisation des débits amont et aval d'un cours d'eau avec tributaire, la seconde concerne la modélisation conjointe des débits maximaux annuels en fonction des volumes écoulés. *De Michele & al.* (2005) utilisent la copule de Gumbel pour modéliser la dépendance positive entre le débit de pointe et le volume de crue. *Zhang & Singh* (2006) utilisent les copules bivariées pour l'analyse fréquentielle des variables : débit de pointe et volume écoulé et volume de crue et sa durée.

Plus spécifiquement pour l'étude de la pluie, on peut mentionner les travaux suivants : *Salvadori & De Michele* (2004) ont utilisé la théorie des copules pour critiquer le calcul univarié des périodes de retour et ont montré que l'usage des copules permet de simplifier les calculs des isolignes des périodes de retour, dans le cas inconditionnel et conditionnel. *Grimaldi & al.* (2005) ont étendu l'étude aux copules trivariées : ils ont analysé conjointement des données observées de hauteur critique, d'intensité maximale et de hauteur totale de pluie. Afin de déterminer la distribution conjointe trivariée, une analyse bivariée du couple intensité maximale–hauteur totale, conditionnée par la hauteur critique est effectuée grâce à une fonction 3-copula. *Salvadori & De Michele* (2006) ont établi la dépendance entre l'intensité moyenne, la durée

"Annexe A"

---

et l'épisode sec via une fonction 3-copula pour caractériser la structure temporelle des épisodes pluvieux.

## ANNEXE B

### Introduction à R

**R** est un système qui est communément appelé langage et logiciel, il permet de réaliser des analyses statistiques. Plus particulièrement, il comporte des moyens qui rendent possible la manipulation des données, les calculs et les représentations graphiques. **R** a aussi la possibilité d'exécuter des programmes stockés dans des fichiers textes.

En effet **R** possède :

1. Un système efficace de manipulation et de stockage des données.
2. Différents opérations pour le calcul sur "tableau".
3. Un grand nombre d'outils pour l'analyse des données et les méthodes statistiques.
4. Des moyens graphiques pour visualiser les analyses.
5. Un langage de programmation simple et performant comportant : conditions, boucles, moyens d'entrées sorties, possibilité de définir des fonctions récursives.

Le projet **R** consiste en une implémentation libre du langage S, développé depuis les années septante dans les laboratoires Bell par *John Chambers* et son équipe, distribué depuis 1993 sous licence commerciale exclusive par Insightful Corp.

Initié dans les années nonante par *Robert Gentleman* & *Ross Ihaka* (Université d'Auckland, Nouvelle-Zélande), auxquels sont venus s'ajuster un noyau de chercheurs du monde entier en 1997. Il consiste aujourd'hui un langage et un environnement de programmation intégré d'analyse statistique.

L'objectif de cette partie est de fournir un environnement interactif d'analyse de données, doté d'outils graphiques performants et permettant une adaptation aisée aux besoins des utilisateurs, depuis l'exécution de tâches routières jusqu'au développement d'applications entières.

## ANNEXE C

### Copule de Gumbel

#### Instructions :

**Densité**  $c_\alpha(u, v)$  : dgumbel (u, v, alpha, dim).

**Fonction de distribution**  $C_\alpha(u, v)$  : pgumbel (u, v, alpha, dim).

**Générateur aléatoire**  $t$  : rgumbel (n, alpha, dim = 2, method = 1).

**Fonction génératrice**  $\Phi_\alpha(t)$  : phigumbel (t, alpha = 1).

**Inverse de la fonction génératrice**  $\Phi_\alpha^{-1}(t)$  : invphigumbel (t, alpha = 1).

Méthodes d'estimation classiques des paramètres de la copule de Gumbel :

**Méthode des Moments** : gumbel.MBE (x, y, marg).

**Maximum de vraisemblance exacte** : gumbel.EML (x, y, marg).

**Inférence sur les marginales** : gumbel.IFM (x, y, marg).

**Maximum de vraisemblance canonique** : gumbel.CML (x, y).

Pour le moment, seuls deux types de marginaux sont disponibles :

**Distribution exponentielle** : (marg = "exp").

**Distribution gamma** : (marg = "gamma").

#### Arguments :

**u** : Vecteur des quantiles  $F_X^{-1}(x)$  si l'argument est fourni, où la matrice des quantiles si l'argument  $u$  n'est pas fourni.

**v** : Vecteur des quantiles  $F_Y^{-1}(y)$ , nécessaires si  $u$  n'est pas une matrice.

**n** : Taille de l'échantillon.

**alpha** : Paramètre de la copule, doit être supérieur à 1 ( $\alpha > 1$ ).

**dim** : Un entier indiquant la dimension de la copule.

**t** : Inverse du générateur, est un tableau de dimension  $n$ .

**x, y** : Vecteurs d'observations, réalisations des va  $X$  et  $Y$ .

---

"Annexe C"

---

**marg** : Une chaîne de caractères spécifiant les marginales du vecteur  $(X, Y)$ . Elle doit être soit "exp" ou "gamma" (valeur par défaut).

**method** : Code de la méthode utilisée dans la simulation :

- 1 désigne la méthode de l'algorithme commun de fragilité (*Marshall & Olkin* [1988]).
- 2 désigne la méthode de l'algorithme "K" (valable uniquement avec  $dim = 2$ ).

**Exemples :**

**Densité et fonction de distribution de la copule de Gumbel :**

1. Cas d'indépendance ( $alpha = 1$ ) :

```
library (gumbel)
alpha = 1
u <-seq (0,1, .1)           # vecteur des quantiles
v <-u
```

- a) Paramétrisation classique :

```
dgumbel (u, v, alpha)      # Densité  $c_\alpha(u, v)$ .
      NaN 1 1 1 1 1 1 1 1 NaN
pgumbel (u, v, alpha)      # Fonction de distribution  $C_\alpha(u, v)$ .
      0.00 0.01 0.04 0.09 0.16
      0.25 0.36 0.49 0.64 0.81 1.00
```

- b) Autre paramétrisation :

```
l<-cbind(u,v)              # Matrice des quantiles.
dgumbel(l, alpha)
      NaN 1 1 1 1 1 1 1 1 NaN
pgumbel(l, alpha)
      0.00 0.01 0.04 0.09 0.16 0.25
      0.36 0.49 0.64 0.81 1.00
```

---

 "Annexe C"

2. Cas de dépendance ( $\alpha > 1$ ) :

```

library (gumbel)
alpha = 2
dim = 3
x <- cbind(u, u, u)
y <- array(u, c (1,11,3))
pgumbel(x, alpha=2, dim=3)
      0.00000000 0.01853315 0.06156706 0.41280666
      0.12426461 0.20452561 0.30102374 0.53914047
      0.67943346 0.83319317 1.00000000
pgumbel (y, alpha=2, dim=3)
      0 0.01853315 0.06156706 0.1242646 0.2045256
      0.3010237 0.4128067 0.5391405 0.6794335
      0.8331932 1
dgumbel (x, alpha=2, dim=3)
      NaN 6.922376 3.646759 2.770993 2.510089
      2.586175 3.024395 4.149303 7.366620
      23.650259 NaN
dgumbel (y, alpha=2, dim=3)
      NaN 6.922376 3.646759 2.770993 2.510089
      2.586175 3.024395 4.149303 7.36662 23.65026 NaN
  
```



**Tracé de la densité de la copule de Gumbel en trois dimensions (3D) :**

1. Tracé de la densité en (3D) :

```
op<-par(mfrow=c(2,2))
u <- seq (.05, .95, length = 30)
v <- u
z <- outer (u, v, dgumbel, alpha=2)
persp(u, v, z, theta = 30, phi = 30, expand = 0.5
      , col = "lightblue",ltheta = 100, shade =0.25
      , ticktype = "detailed",xlab = "u", ylab = "v"
      , zlab = "Densité de la copule de Gumbel pour  $\alpha = 2$ ")
z <- outer (u, v, dgumbel, alpha=4)
persp(u, v, z, theta = 30, phi = 30, expand = 0.5
      , col = "lightblue",ltheta = 100, shade =0.25
      , ticktype = "detailed",xlab = "u", ylab = "v"
      , zlab = "Densité de la copule de Gumbel pour  $\alpha = 4$ ")
```

---

 "Annexe C"

2. Tracé de la densité en 3D avec différentes couleurs :

```

ncol <- 100
zlim <- c (0, max (z))
          0.000000 7.618281
nrz <- nrow (z)
          30      # nombre de ligne de z
ncz <- ncol (z)
          30      # nombre de colonne de z
jet.colors <- colorRampPalette (c ( "# 00007F", "bleu", "# 007FFF"
                                   , "cyan", "#7FFF7F", "yellow",
                                   "#FF7F00", "red", "#7F0000"))

couleurs <- tail (jet.colors (1.2 * ncol), ncol)
fcol <- couleurs [trunc (z / zlim [2] * (ncol-1)) +1]
dim (fcol) <- c (nrz, ncz)
              30 30

fcol <- fcol [-nrz,-ncz]

persp(u, v, z, col=fcol, zlim=zlim, theta=30, phi=30, ticktype =
      "detailed", box = TRUE, xlab = "u", ylab = "v", zlab =
      "Densité de la copule de Gumbel pour  $\alpha = 2$ ")

persp(u, v, z, col=fcol, zlim=zlim, theta=30, phi=30, ticktype =
      "detailed", box = TRUE, xlab = "u", ylab = "v", zlab =
      "Densité de la copule de Gumbel pour  $\alpha = 4$ ")

```

**Tracé la distribution de la copule en trois dimensions (3D) :**

```
op<-par(mfrow=c(1,2))
u <- seq (.05, .95, length = 30)
v <- u
z <- outer (u, v, pgumbel, alpha=2)
persp (u, v, z, theta = 30, phi = 30, expand = 0.5
      , col = "lightblue", ltheta = 100, box = 0,25
      , ticktype = "detailed", xlab = "u", ylab = "v"
      , zlab = "Distribution de la copule de Gumbel pour  $\alpha = 2$ ")
persp (u, v, z, theta = 30, phi = 30, expand = 0.5
      , col = "lightblue", ltheta = 100, box = 0,25
      , ticktype = "detailed", xlab = "u", ylab = "v"
      , zlab = "Distribution de la copule de Gumbel pour  $\alpha = 4$ ")
```

**Simulation de la copule de Gumbel :**

Pour générer un échantillon aléatoire de la copule de Gumbel on utilise l'instruction : **rgumbel (n, alpha)**.

**1. Générateur aléatoire et son inverse :**a) Cas d'indépendance ( $\alpha = 1$ ) :

```

op<-par(mfrow=c(1,2))
n = 200
alpha = 1
s<-rgumbel (n,alpha=1)           # Générateur aléatoire.
u <- t (s)                       # Inverse du générateur.
plot (u [1, ], u [2, ], xlab="u",ylab="v", # Nuage de points de
      col = "green", main = "Copule de   la copule de Gumbel.
      Gumbel pour alpha=1")

```

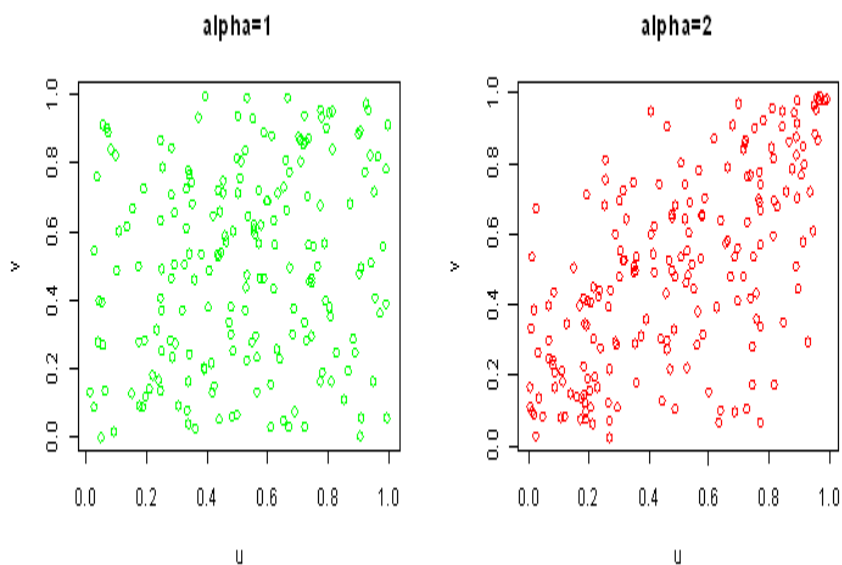
b) Cas de dépendance ( $\alpha > 1$ ) :

```

n = 200
alpha = 2
s<-rgumbel (200,alpha=2)
u <- t (s)
plot (u [1, ], u [2, ], xlab="u",ylab="v", col = "red",
      main = "Copule de Gumbel pour alpha=2")

```

## "Annexe C"

FIG. C.1. Simulation de la copule de Gumbel pour  $(\alpha = 1)$  et  $(\alpha = 2)$ **2.** Comparaison entre les deux algorithmes (générateurs aléatoires) :

```
n <- 10000
```

```
alpha = 2
```

**1<sup>ère</sup> Méthode :** Algorithme commun de fragilité (*Marshall & Olkin*)

```
system.time(rgumbel(n, alpha=2, method=1)) [3]
```

```
elapsed
```

```
0.022
```

**2<sup>ème</sup> Méthode :** Algorithme "K"

```
system.time(rgumbel(n, alpha=2, method=2)) [3]
```

```
elapsed
```

```
1.515
```

system.time : le temps chargé pour l'exécution des instructions utilisées par le système pour tout processus.

---

"Annexe C"

**3.** Fonction génératrice et son inverse :

a) Cas d'indépendance ( $alpha = 1$ ) :

```
n = 100
```

```
alpha = 1
```

```
s<-rgumbel (n,alpha=2) # Générateur aléatoire.
```

```
u <- t (s) # Inverse du générateur.
```

```
z=phigumbel(u,2) # Fonction génératrice  $\Phi_\alpha(t)$ .
```

```
l=invphigumbel(u,2) # Inverse de la fonction génératrice  $\Phi_\alpha^{-1}(t)$ .
```

b) Cas de dépendance ( $alpha > 1$ ) :

```
n = 200
```

```
alpha = 2
```

```
s<-rgumbel (200,alpha=2)
```

```
u <- t (s)
```

```
z=phigumbel(u,2)
```

```
l=invphigumbel(u,2)
```

---

"Annexe C"

4. Nuage de points de la copule de Gumbel pour différents paramètres alpha :

```
n=100
anim <-function (n, max = 50)
{
for (i in seq (1, max, length.out = n))
{
s<-rgumbel (10000, alpha=i, method = 2)
u <-t (s)
plot (u [1, ], u [2, ], col = "green", main =
      "Nuage de points de la copule de Gumbel", xlim =
      c (0,1), ylim= c (0,1), pch = "." )
cat ()
}
}
anim (20, 20) # Animation de 20 graphes.
```

**Estimation des paramètres de la copule de Gumbel :**

Pour le moment, seuls deux types de marginaux sont disponibles :

- Distribution exponentielle (marg = "exp").
- Distribution gamma (marg = "gamma").

Les méthodes d'estimations des paramètres de la copule sont :

- Méthode des Moments : gumbel.MBE (x, y, marg).
- Maximum de vraisemblance exacte : gumbel.EML (x, y, marg).
- Inférence sur les marginales : gumbel.IFM (x, y, marg).
- Maximum de vraisemblance canonique : gumbel.CML (x, y).

---

 "Annexe C"

**1.** Marginales de distributions exponentielles de paramètre  $\lambda = 1$  :

```

lambdaX = lambdaY = 1
alpha = 2
simu <- qexp(rgumbel(200, 2))           # Fonction des quantiles de la loi exponentielle.
gumbel.MBE(simu[,1], simu[,2], "exp")  # Méthode des Moments.
      1.212656 1.253483 1.633290
gumbel.IFM(simu[,1], simu[,2], "exp")  # Inférence sur les marginales.
      1.218739 1.259771 1.691185
gumbel.EML(simu[,1], simu[,2], "exp")  # Maximum de vraisemblance exacte.
      1.231117 1.275609 1.682866
gumbel.CML(simu[,1], simu[,2], "exp")  # Maximum de vraisemblance canonique.
      1.712567

```

**2.** Marginales de distributions gamma de paramètres  $\lambda = 1$   
 et  $\alpha = 2$  :

```

lambdaX = lambdaY = 1
alphaX = alphaY = 2
alpha = 3
simu <- qgamma(rgumbel(200, 3), 2, 1)
gumbel.MBE(simu[,1], simu[,2], "gamma")
      0.987068 2.111151 1.034349 2.133408 3.314457
gumbel.EML(simu[,1], simu[,2], "gamma")
      1.041724 2.199624 1.104601 2.268300 3.332762
gumbel.IFM(simu[,1], simu[,2], "gamma")
      1.018388 2.178176 1.069041 2.204767 3.375169
gumbel.CML(simu[,1], simu[,2], "gamma")
      3.354012

```



---

## Résumé

---

Plusieurs types d'événements hydrologiques sont décrits avec des caractéristiques multivariées (sécheresses, inondations, tempêtes de pluies, etc). Lorsqu'on effectue une analyse de fréquences régionales multivariée de ces événements, il est important d'étudier conjointement toutes ces caractéristiques. L'objectif de ce mémoire est de prolonger l'approche univariée de Hosking et Wallis (1993) au cas multivarié. Les L-moments multivariés sont utilisés pour définir des statistiques et des modèles de copules afin de décrire le comportement statistique des variables dépendantes. La copule la plus utilisée en hydrologie et en climatologie est celle de Gumbel. Elle permet d'estimer les valeurs extrêmes des phénomènes.

*Mots-clés:* L-moments, Univarié, Multivarié, Coskewness, Cokurtosis, Comoment, Analyse de Fréquence Régionale, Copule, Copule Archimédienne, Copule de Gumbel, Tau de Kendall, Rho de Spearman.

---

## Abstract

---

Several types of hydrological events are described with multivariate characteristics (droughts, floods, rain storms, etc). When carrying out a multivariate regional frequency analysis for these events it is important to jointly consider all these characteristics. The aim of this memory is to extend the univariate approach of Hosking and Wallis (1993) to the multivariate case. As a tool, multivariate L-moments are used to define the statistics and general copula models to describe the statistical behaviour of dependent variables. The most copulas used in hydrology and climatology is that of Gumbel. It allows to estimate the extreme value of phenomenas.

*Key-words:* L-moments, Univariate, Multivariate, Coskewness, Cokurtosis, Comoment, Regional Frequency Analysis, Copula, Archimedean Copula, Gumbel Copula, Kendall's Tau, Spearman's Rho.

---

## ملخص

---

العديد من الأحداث الهيدرولوجية لها ميزات متعددة المتغيرات (الجفاف، الفيضانات، العواصف، الأمطار، الخ). عند إجراء تحليل متعدد المتغيرات لهذه الأحداث، من المهم دراسة كل هذه الميزات. الهدف من هذه المذكرة هي تمديد النهج وحيد المتغير (Hosking و Wallis 1993) إلى حالة عدة متغيرات. اللحظات متعددة المتغيرات تستخدم لتحديد إحصائيات و نماذج الروابط لوصف السلوك الإحصائي للمتغيرات المعتمدة. الرابط الأكثر استخداما في مجال الهيدرولوجي و المناخ هو جامبل لأنه يسمح بتقدير القيم القصوى للظواهر.

مصطلحات: لحظات، أحادي المتغير، متعدد المتغيرات، Comoment، Cokurtosis، Coskewness، تحليل التردد الإقليمي، رابط ارشميدس، رابط جامبل، تو كندال، رو سبيرمان.