



La théorie des extrêmes et la gestion des risques de marché

Antoine BEZAT & Ashkan NIKEGHBALI

Groupe de Travail ENSAE
sous la direction de Gaël Riboulet et Thierry Roncalli

Mai 2000

Remerciements

Nous tenons tout d'abord à remercier Nicole El Karoui qui nous a mis en contact avec le Groupe de Recherche Opérationnelle (GRO) du Crédit Lyonnais et qui a ainsi été à l'origine de cette passionnante et enrichissante expérience au sein du GRO.

Nous remercions également Antoine Frachot qui nous a particulièrement bien accueilli au sein de son équipe et qui a mis à notre disposition les outils informatiques adéquats à cette étude.

Enfin, nous remercions très chaleureusement Gaël Riboulet et Thierry Roncalli pour leur disponibilité de tous les instants. Cette étude leur doit beaucoup et nous avons beaucoup appris à leur contact.

Table des matières

Remerciements	iii
Introduction	1
1 La mesure du risque	3
1.1 Définition d'une mesure des risques	3
1.2 L'approche classique du risque en finance	5
1.3 La valeur en risque (<i>Value-at-Risk</i>)	6
2 La théorie univariée des extrêmes appliquée aux risques de marché	9
2.1 Présentation de la théorie classique des extrêmes	9
2.1.1 Le cadre d'analyse	9
2.1.2 Le théorème limite de Fisher-Tippet	10
2.1.3 Estimation des paramètres de la distribution GEV	13
2.2 Une mesure de risque cohérente (ou ' <i>expected shortfall</i> ')	14
2.2.1 Mesures de risques extrêmes et distribution de Pareto généralisée	14
2.2.2 Estimation de la fonction de distribution des excès	15
2.2.3 Estimation des paramètres de la distribution GPD	16
2.2.4 Estimateurs de la <i>VaR</i> et de l' <i>expected shortfall</i>	17
2.3 Applications financières	17
2.3.1 Description de la base de données LME	17
2.3.2 Le calcul de la valeur en risque d'un portefeuille	18
2.3.3 La construction des scénarios de crises	22
3 La théorie des copules	25
3.1 Présentation générale	25
3.1.1 Introduction mathématique	25
3.1.2 Théorème de Sklar et premières propriétés des copules	27
3.1.2.1 Le théorème de Sklar	27
3.1.2.2 La relation d'ordre sur les copules	28
3.1.2.3 Les copules et les variables aléatoires	28
3.1.2.4 Les relations de dépendance	29
3.2 Méthodes d'estimation	33
3.2.1 Les copules empiriques	33
3.2.2 Estimation par la méthode du maximum de vraisemblance	34
3.2.3 Estimation par la méthode IFM	35

3.3	Construction des familles de copules	35
3.3.1	Familles de distributions multivariées générées par des mélanges	35
3.3.2	Les copules archimédiennes	39
3.3.2.1	Quelques généralités	39
3.3.2.2	Propriétés des copules archimédiennes	40
3.3.2.3	Identification d'une copule archimédienne	42
3.3.3	Les copules extrêmes	43
3.3.3.1	Résultats généraux	43
3.3.3.2	Méthode d'estimation paramétrique	46
3.3.3.3	Estimation non paramétrique	46
3.3.4	Les copules archimax	47
3.4	Applications financières	49
3.4.1	Modélisation bivariable des scénarios de crise	49
3.4.2	Quantification d'un scénario de crise multidimensionnel	50
3.4.3	Un retour sur la valeur en risque	55
	Conclusion	59
	Références	61
	Table des figures	65

Introduction

Jusqu'à la fin des années 80, le risque bancaire prenait essentiellement deux formes : le risque de crédit (ou de contrepartie) qui correspond à la défaillance d'un emprunteur et le risque de liquidité qui correspond à la difficulté potentielle de réaliser rapidement une transaction. Le développement rapide des instruments dérivés sur les marchés financiers durant les années 90, et la libéralisation des mouvements internationaux de capitaux ont été accompagnés par une augmentation considérable de l'activité de trading dans les banques. Cette évolution les expose de façon importante au risque de marché, c'est-à-dire à des pertes dues aux variations des prix des actifs. Le risque de marché s'est brutalement manifesté durant la crise financière de 1987 et les années 90 avec des conséquences parfois catastrophiques comme la faillite de la Banque Barings en février 1995.

En réponse à ces "accidents", les autorités réglementaires ont décidé d'imposer un certain nombre de normes en matière de contrôle des risques bancaires et en particulier du risque de marché. Ces règlements sont inspirés des recommandations du Comité de Bâle, groupe de travail institué en 1975 par les gouverneurs des banques centrales, qui tente d'établir des normes minimales de contrôle bancaire communes à tous les pays.

Les établissements financiers ne sont pas en mesure de se couvrir parfaitement contre le risque de marché. Les normes de contrôle des risques ne visent donc pas à faire disparaître les risques, mais à les maîtriser, à les mesurer, et à couvrir ceux jugés les plus importants.

En pratique, ces normes s'expriment essentiellement comme une exigence pour la banque d'un montant minimal de fonds propres réglementaires destiné à la couverture des risques de marché. C'est l'accord de Bâle en 1988 qui a développé le cadre réglementaire. Cet accord a depuis été amendé et la Directive Européenne d'Adéquation des fonds propres (CAD) a établi en 1993 le cadre dans lequel les banques peuvent développer leurs propres modèles de contrôle des risques de marché — voir les rapports [1] et [2] du comité de Bâle.

Les établissements financiers ont le choix du modèle d'évaluation des risques, mais ils doivent respecter un certain nombre de règles. Ils doivent calculer la perte potentielle quotidiennement pour une période de détention de dix jours et pour un niveau de confiance de 99% (la baisse de valeur du portefeuille sur une période de dix jours ne dépasse donc théoriquement ce niveau qu'une fois sur 100). Notons cette perte $P(t)$. A chaque date t , l'établissement calcule l'exigence de fonds propres $FP(t)$ de la façon suivante :

$$FP(t) = \max \left(P(t-1), (3 + \xi) \times \frac{1}{60} \sum_{i=1}^{60} P(t-i) \right)$$

où ξ , le complément éventuel compris entre 0 et 1, est fixé par la Commission Bancaire en fonction des performances du modèle évaluées a posteriori. La perte potentielle $P(t)$ est aussi appelée *Value-at-Risk* (VaR).

De nombreuses études ont été consacrées à l'estimation de la VaR . Cette démarche est certainement utile puisque les banques sont tenues d'estimer leur *Value-at-Risk*. Cependant, nous adopterons ici un point de vue différent. Dans notre première partie, nous allons définir les critères que doit vérifier une bonne mesure de risque et nous étudierons le cas particulier de deux mesures de risque : la volatilité et la VaR . Ces mesures apparaîtront inadaptées à la prise en compte des événements extrêmes, qui sont justement ceux qui font peser les risques les plus grands sur les établissements financiers, car elles se placent explicitement ou implicitement dans le cadre de l'approximation par la loi normale. Dans la seconde partie, nous introduirons la théorie des valeurs extrêmes univariées et l'appliquerons à des séries financières. Nous construirons dans cette seconde partie une mesure de risque cohérente. Nous déduirons de la théorie des extrêmes des informations telles que les quantiles extrêmes ou les temps de retour d'événements extrêmes en élaborant des scénarios de crise. Nous montrerons ainsi que la théorie des valeurs extrêmes est un outil indispensable pour construire des indicateurs ou des mesures de risque qui viennent compléter la VaR . Nous montrerons aussi que la théorie des extrêmes permet de calculer la VaR au delà du seuil de confiance de 99%. Dans la pratique, elle est très facile à mettre en place si la VaR de l'établissement financier est de type historique¹. Pourquoi vouloir calculer des quantiles aussi extrêmes ? Tout d'abord, la réglementation peut évoluer et certains établissements financiers publient même des mesures de risque à 99,96% ! Mais la réponse la plus évidente concerne la mise en place des projets d'allocation de fonds propres sur base économique — c'est-à-dire des fonds propres calculés de façon plus rigoureuse que les méthodes forfaitaires par des modèles internes. Que l'allocation soit de type *bottom-up* (SCHMOCK et STRAUMANN [1999] et TASCHE [1999]) ou *top-down* (BAUD, FRACHOT, IGIGABEL, MARTINEU et RONCALLI [1999]), les seuils de confiance utilisés sont de toute façon très largement supérieurs à 99%. Par exemple, on considère qu'une banque notée AAA doit mesurer ces risques avec un seuil de 99,97%. Dans ce cas, les méthodes analytique, historique et Monte Carlo paraissent inadaptées. La théorie des extrêmes apportent alors une réponse pour les risques de marché².

Enfin, nous tenterons dans le troisième chapitre de construire des outils pour la gestion du risque dans le cas multivarié. Cette généralisation s'avère être souvent très délicate et n'est pas toujours traitée. Nous utiliserons pour cela la théorie des copules. Bien que cette branche des probabilités existe depuis les années 1960, elle n'a été découverte que tardivement par les milieux de la recherche en finance. Depuis deux ou trois ans, certains chercheurs du milieu académique proches des préoccupations des banques ou des assurances ont publié des articles exposant des idées d'utilisation des copules pour la gestion des risques. Après avoir exposé en détail la théorie générale des copules et les techniques d'estimation, nous soulignerons l'intérêt de tels outils pour la gestion du risque. Enfin, nous illustrerons notre propos par une application inédite : l'élaboration de scénarios de crise multivariés à l'aide des copules. Nous mettrons notamment en relief l'importance de la structure de dépendance à travers le calcul de temps des retour implicites.

¹ Contrairement à l'idée reçue, la VaR historique est de loin la méthode la plus utilisée par les établissements financiers.

² Cette réponse est d'autant plus fiable que la VaR historique existe depuis longtemps dans l'établissement financier.

1

La mesure du risque

1.1 Définition d'une mesure des risques

Le risque est inhérent à la plupart des activités humaines. Le transport aérien, l'énergie nucléaire ou le jeu par exemple comportent des risques facilement identifiables. Il semble cependant délicat de mesurer ces risques. Un exemple fameux fournissant une bonne illustration de la difficulté de mesurer le risque consiste à demander à une personne :

1. si vous savez que vous avez une probabilité de 5% de gagner à la prochaine loterie nationale, prenez vous le risque de jouer ?
2. si vous êtes directeur d'une centrale nucléaire et que vous savez qu'il y a une probabilité de 5% qu'il se produise un incident, prenez vous le risque de faire fonctionner cette centrale nucléaire ?

De par leur nature, les marchés financiers fournissent d'innombrables séries temporelles de chiffres. Une mesure objective du risque peut donc y sembler plus aisée que dans d'autres domaines où la qualification du risque est plus floue et où les données sont moins nombreuses. Mais mesurer le risque reste une activité délicate et avant d'étudier les problèmes d'estimation, il est important de choisir une mesure pertinente du risque.

Dans notre cas, le risque est lié à la valeur future des actifs et on cherche à déterminer le montant de fonds propres (actifs sûrs) nécessaires pour la couverture d'un portefeuille risqué. Comme il est impossible d'accéder à l'ensemble de la loi de distribution de cette valeur future, on va chercher un indicateur qui nous permettra de décider du montant de fonds propres à allouer et qui nous donnera donc une idée du niveau d'exposition au risque.

Dans cet chapitre, nous reprenons la démarche de ARTZNER, DELBAEN, EBER et HEATH [1999]. Soit Ω l'espace des états possibles de la nature. On suppose Ω fini, la généralisation au cas infini a été traitée dans DELBAEN [1998]. La valeur future du portefeuille dans les différents états de la nature est donc un vecteur noté $\mathbf{X}$. L'actif sûr est noté $\mathbf{r}$, ce vecteur a toutes ses coordonnées positives (c'est en ce sens qu'il est qualifié de sûr).

Dans un premier temps, il est indispensable de définir une classe $\mathcal{A}$ de portefeuilles admissibles constitués d'actifs sûrs et d'actifs risqués : ce sont les portefeuilles dont on estime le risque acceptable. La couverture devra donc permettre de ramener un portefeuille jugé trop risqué dans cette classe, en ajoutant des actifs sans risque.

ARTZNER, DELBAEN, EBER et HEATH [1999] donnent quatre axiomes nécessaires pour que la région de portefeuilles acceptables soit définie de manière cohérente. Ces quatre axiomes portant sur $\mathcal{A}$ sont les suivants :

1. $\mathcal{A}$ contient le cône des vecteurs ayant toutes leurs coordonnées positives. Un portefeuille assurant un revenu positif dans tous les états possibles est bien sûr acceptable.
2. $\mathcal{A}$ ne doit pas avoir d'intersection avec l'ensemble des vecteurs ayant toutes leur coordonnées strictement négatives. Un portefeuille qui donne une perte dans tous les états ne peut être accepté.
3. $\mathcal{A}$ est convexe. Cet axiome est, lui aussi, naturel. Si deux positions sont acceptables, une position intermédiaire entre les deux sera aussi acceptable.
4. $\mathcal{A}$ est un cône positivement homogène. Cette condition sera discutée plus loin après avoir défini la mesure de risque.

Une fois qu'on s'est donné une région de portefeuilles acceptables, c'est à dire des limites sur le niveau de risque du portefeuille, on obtient de manière duale la mesure de risque comme quantité d'actif sûr à ajouter pour ramener le portefeuille dans cette classe :

$$\rho(X) = \inf\{m, mr + X \in \mathcal{A}\} \quad (1.1)$$

Réciproquement, à partir de la mesure de risque, on retrouve la région de portefeuilles acceptables

$$\mathcal{A}_\rho = \{X, \rho(X) \leq 0\} \quad (1.2)$$

car les portefeuilles admissibles sont ceux pour lesquels il n'est pas nécessaire d'ajouter des actifs sûrs supplémentaires.

ARTZNER, DELBAEN, EBER et HEATH [1999] donnent ainsi cinq axiomes caractérisant une mesure de risque cohérente. Ils montrent que ces cinq axiomes sont équivalents aux précédents au sens où si ρ est une mesure cohérente, $\mathcal{A}_\rho$ est une région admissible cohérente et réciproquement. Ces cinq axiomes sont les suivants :

1. **Invariance par translation :**

$$\forall X, \forall \alpha \quad \rho(X + \alpha r) = \rho(X) - \alpha \quad (1.3)$$

Cet axiome assure que $\rho(X + \rho(X)r) = 0$ et s'interprète facilement avec la définition de $\mathcal{A}_\rho$.

2. **Sous-additivité :**

$$\rho(X_1 + X_2) \leq \rho(X_1) + \rho(X_2) \quad (1.4)$$

Le fait de diviser un portefeuille ne peut réduire le risque total.

3. **Homogénéité positive :**

$$\forall \lambda \geq 0, \forall X \quad \rho(\lambda X) = \lambda \rho(X) \quad (1.5)$$

La taille du portefeuille ne doit pas avoir de conséquence sur la mesure de risque, c'est à dire sur le fait d'être acceptable ou pas.

4. **Monotonie :** si $X \leq Y$ alors $\rho(Y) \leq \rho(X)$. Si X offre un rendement inférieur à celui de Y dans tous les états de la nature, alors il est plus risqué que Y .

5. **Pertinence :**

$$\forall X, X \leq 0, X \neq 0, \quad \rho(X) > 0 \quad (1.6)$$

Cet axiome est nécessaire pour éviter que certains risques ne soient oubliés.

La condition sur la région d'admission que nous n'avons pas interprétée stipulant que $\mathcal{A}$ doit être un cône homogène découle de la sous-additivité et de l'homogénéité positive.

Cette démarche ne définit pas de manière pratique une mesure du risque, mais constitue une première approche fort utile. La caractérisation équivalente pour la région de portefeuilles admissibles et la mesure de risque fournit cinq critères indispensables à vérifier lorsqu'on va chercher à mesurer le risque en pratique.

1.2 L'approche classique du risque en finance

Traditionnellement, on associe au risque financier l'incertitude statistique sur la variation du prix des actifs. Sa mesure canonique est la volatilité. Cette mesure est couramment utilisée par exemple dans la théorie du portefeuille de Markowitz. Elle n'est pourtant pas une mesure cohérente du risque. Les gains et les pertes sont pris en compte de manière analogue. Il est évident que cela ne correspond pas à la réalité. Si l'incertitude pèse essentiellement sur le gain, l'exposition au risque est moins grande que dans le cas symétrique où l'incertitude pèse sur les pertes. Cette mesure ne vérifie donc pas la condition de monotonie, ni celle d'invariance par translation puisque nous avons $\sigma(X + \alpha r) = \sigma(X)$.

L'argument principal en faveur de la volatilité repose sur le théorème central limite. Le théorème dit que pour une suite (X_n) de variables aléatoires indépendantes et identiquement distribuées, on a :

$$\sum_{i=1}^n \frac{X_i - \mathbb{E}[X_i]}{\sigma(X_i) \sqrt{n}} \longrightarrow \mathcal{N}(0, 1) \quad (1.7)$$

Une autre raison en faveur de la volatilité est que les nombreuses données que fournissent les marchés permettent de la calculer assez facilement.

BOUCHAUD et POTTERS [1997] critiquent cette vision gaussienne. La démarche qui permet d'utiliser le théorème central limite est la suivante. On découpe la période de temps T en N intervalles, ce qui permet d'écrire :

$$X(T) = X_0 + \sum_{k=1}^{N-1} \delta X_k \quad (1.8)$$

où $\delta X_k = X_k \eta_k$, $X_k = X\left(k \frac{T}{N}\right)$ et où η_k est par définition le rendement instantané :

$$\ln \left(\frac{X(T)}{X(0)} \right) = \sum_{k=1}^{N-1} \ln(1 + \eta_k) \quad (1.9)$$

En supposant que les rendements η_k sont des variables indépendantes identiquement distribuées, on peut appliquer le théorème central limite pour N tendant vers l'infini, c'est à dire pour des intervalles de temps très courts. On obtient alors que $\eta(t)$ suit une loi gaussienne $\mathcal{N}(mT, \sigma\sqrt{T})$ et donc

$$x(T) = x_0 \exp(mT + \sigma\sqrt{T}\xi) \quad \text{où} \quad \xi \sim \mathcal{N}(0, 1)$$

Dans ce cadre, il est facile de voir comment σ va caractériser le risque. Cependant, cette interprétation n'est vraie que dans le cas où les accroissements des rendements sont indépendants, ce qui a été contredit par de nombreuses études économétriques¹. En outre, dans la réalité, N est fini ; on ne peut donc pas appliquer le théorème central limite. BOUCHAUD et POTTERS [1997] soulignent par exemple qu'un mois boursier correspond à $N \simeq 320$ intervalles de 30 minutes. La convergence utilisée dans le théorème central limite est une convergence en loi (qui n'est donc pas uniforme). Elle n'est donc valable que dans une région centrée autour de la moyenne. FELLER [1968] a montré que même dans le cas favorable où la loi est "fine" (tous les moments existent), l'approximation n'est valable qu'autour de $n^{3/4}$ écarts-types de la moyenne.

Cet argument est aussi retenu par BOULIER, BRABANT, DALAUD et DIEU [1997] dans le cadre d'une approche APT où les rentabilités sont décrites par la combinaison linéaire d'une trentaine de facteurs économiques qu'ils supposent indépendants et de même loi. L'approximation ne sera alors valable qu'autour de $30^{3/4} \simeq 12$ écarts-types de la moyenne. Or, l'écart-type de la loi normale résultante est $\sqrt{30}$ fois plus grand que l'écart-type initial. Cela implique que l'approximation n'est plus valable au-delà de la probabilité de 99%.

Cette approximation n'est donc pas valable pour les queues de distribution, là où justement se situent les risques les plus importants. D'ailleurs, dans le calcul de la variance, les valeurs extrêmes sont souvent considérées comme aberrantes et exclues. C'est pourquoi, la réglementation bancaire s'attache à une autre mesure du risque, plus apte à contrôler les événements extrêmes qui sont ceux qui font peser un réel danger sur les établissements financiers.

¹ Les marchés financiers ont de la mémoire.

1.3 La valeur en risque (*Value-at-Risk*)

L'outil recommandé par la Commission Bancaire est la *Value-at-Risk* (VaR). La VaR désigne la perte potentielle maximale sur un certain horizon T et pour un niveau de probabilité α . La perte $P(t)$ à laquelle nous faisons référence en introduction est en fait la VaR pour $T = 10$ jours de trading et un niveau α de 99%. Si on note $\mathcal{P}_T(\partial X)$ la loi de distribution des variations de valeur des actifs durant la période T , la VaR_α est telle que

$$\mathcal{P}_T(\partial X \geq -VaR_\alpha) = 99\% \quad (1.10)$$

L'interprétation de ce seuil de confiance est que la probabilité que la perte soit supérieure à la perte potentielle (VaR_α) est de 1%. Cela doit donc se produire théoriquement une fois toutes les 100 périodes. Comme la période de détention est de 10 jours, on a théoriquement un retour tous les quatre ans (car $100 \text{ périodes} \times 10 \text{ jours} = 1000 \text{ jours de trading} \simeq 4 \text{ années calendaires}$). Notons cependant que cette VaR_α ne prend pas en compte les pertes cumulées, la perte totale pouvant dépasser la VaR , ni les pertes maximales à l'intérieur d'une période. On ne considère ici que la variation du prix durant la période $[kT, (k+1)T]$.

Trois méthodes de calcul sont développées pour estimer la VaR :

1. **La VaR paramétrique (ou analytique)** : cette approche consiste à représenter le gain algébrique d'un portefeuille comme combinaison linéaire de facteurs gaussiens. Si on note $V(t)$ la valeur du portefeuille à l'instant t , on a

$$V(t) = a^\top F(t) \quad (1.11)$$

où $F(t)$ est le vecteur gaussien des facteurs de loi $\mathcal{N}(\mu, \Sigma)$ et a le vecteur de sensibilité à ces facteurs. En t , $V(t+1)$ est donc une variable aléatoire gaussienne de loi $\mathcal{N}(a^\top \mu, a^\top \Sigma a)$. La valeur de la VaR pour un seuil de confiance α correspond alors à :

$$Pr(V(t+1) - V(t) \geq -VaR_\alpha) = \alpha \quad (1.12)$$

On obtient alors

$$VaR_\alpha = \Phi^{-1}(\alpha) \sqrt{a^\top \Sigma a} \quad (1.13)$$

où Φ^{-1} est la fonction de répartition inverse d'une loi gaussienne. Notons les hypothèses très fortes qui ont été faites pour réaliser cette estimation :

- (a) l'indépendance temporelle des variations de la valeur du portefeuille ;
 - (b) la normalité des facteurs $F(t)$;
 - (c) la relation linéaire entre les facteurs et la valeur du portefeuille.
2. **La VaR historique** : elle consiste à obtenir sur la base d'un historique des variations de facteurs de risque à un horizon de temps donné, une distribution des variations de valeur du portefeuille. De cette distribution, on peut extraire un quantile qui permet de lire la VaR pour un seuil de confiance donné. Cette méthode est relativement simple mais présente un risque de mesure lié au choix de l'échantillon. Si celui-ci est trop court, on s'expose à un risque lié au fait qu'on n'aura pas suffisamment de données pour estimer correctement le quantile à 99% (la variance de l'estimateur sera alors très grande). Si, au contraire, on le choisit trop long, on court le risque que la distribution des facteurs change, ce qui induit un risque sur l'estimation du quantile.
 3. **La VaR Monte Carlo** : cette méthode est une variation de la VaR historique dans le cas où seule la matrice de variance-covariance est disponible. On génère alors par une simulation de Monte Carlo des scénarios compatibles avec cette matrice.

Il est important de s'interroger maintenant sur la pertinence de la VaR comme mesure du risque. ARTZNER, DELBAEN, EBER et HEATH [1997] ont montré que la VaR n'est pas une mesure cohérente du risque car elle ne satisfait pas la condition de sous-additivité. Pour cela, ils proposent le cas de deux options digitales. La première A , dont le prix initial est u , paie 1000 si la valeur de l'action à la date T est supérieure

à un certain U , rien sinon. La seconde B , dont le prix initial est l , paie 1000 si la valeur de l'action en T est inférieure à L ($L < U$) et rien sinon. On choisit L et U de sorte que $\Pr(S_T < L) = \Pr(S_T > U) = 0.008$. Observons la *Value-at-Risk* à 99% de 2 traders vendant respectivement une option A et une option B . Elles sont évidemment nulles. La *VaR* d'un trader vendant une action A et une action B sera, elle, de 1000. La *VaR* n'est donc pas **sous-additive**. C'est une contradiction majeure de l'utilisation de la *VaR* comme mesure de risque qui contredit le principe de diversification. On montrerait bien sûr de la même manière que la région de portefeuilles admissibles n'est pas convexe.

Dans le cas particulier où la variation de prix d'un portefeuille X suit une loi normale, on a

$$VaR_\alpha(X) = \Phi^{-1}(\alpha) \sigma(X) \quad (1.14)$$

Or on a

$$\sigma(X + Y) \leq \sigma(X) + \sigma(Y) \quad (1.15)$$

Dans le cadre normal, la *VaR* est bien une mesure cohérente de risque. La *VaR* ne peut donc être utilisée de manière pertinente que lorsque cette approximation peut être effectuée ou que le portefeuille est linéaire (ce qui n'est en particulier pas le cas des portefeuilles comportant des options).

En résumé, nous noterons que l'utilisation de la *VaR* suppose à double titre que l'on puisse approximer la distribution des rendements par une loi normale :

1. en tant que mesure de risque (la *VaR* n'est pas une mesure cohérente dans le cas général) ;
2. et par les techniques d'estimation (*VaR* paramétrique et Monte Carlo).

Si la *VaR* présente un intérêt certain pour le régulateur — elle permet de fixer une norme pour tous les établissements bancaires — elle semble mal adaptée à **une mesure d'allocation de fonds propres économiques** par les établissements financiers. Pour ceux-ci, l'introduction de la théorie des valeurs extrêmes va permettre de s'intéresser effectivement aux régions de la distribution de la valeur future des actifs qui constituent un risque pour l'établissement, d'élaborer des scénarios de crise et de fournir des informations telles que le temps de retour ou l'*expected shortfall*.

2

La théorie univariée des extrêmes appliquée aux risques de marché

La gestion du risque est une gestion des événements inattendus, des événements rares. Or avec l'utilisation de la distribution gaussienne, ces événements ont un caractère “prédictable” et rien ne sert pour cela d'étudier leur fréquence (en effet la distribution gaussienne est entièrement caractérisée par ses deux premiers moments et la fréquence de ces événements rares s'en déduit). Nous allons ainsi dans ce chapitre tenter d'appréhender le risque grâce à des mesures qui donnent des informations sur l'occurrence des événements extrêmes : nous utiliserons pour cela la théorie des valeurs extrêmes (**EVT**).

L'approche mathématique classique pour modéliser le risque se fait grâce à la théorie des probabilités. Le risque est une variable aléatoire qui va des états futurs du monde (et imprédictables) dans $\mathbb{R}$ et qui représente les pertes ou les profits. Cette variable aléatoire suit une certaine loi de probabilité que l'on ne connaît pas de manière exacte (bien que l'on dispose de données historiques concernant les pertes passées et qui peuvent partiellement nous renseigner sur cette distribution). On peut dire qu'un événement extrême survient lorsque le risque prend ses valeurs au niveau de la queue de distribution.

Ainsi la théorie des valeurs extrêmes qui consiste à analyser les occurrences présentant des fréquences très faibles apparaît comme un outil particulièrement bien adapté pour notre étude. Nous exposerons dans la section suivante la théorie classique des extrêmes pour en déduire des informations sur les pertes maximales.

De plus nous utiliserons la théorie des valeurs extrêmes pour compléter la *VaR* et obtenir une nouvelle mesure de risque qui soit cohérente : l'*expected shortfall*. En effet, comme nous l'avons vu précédemment, la *VaR* peut être critiqué sur deux points. Tout d'abord, la *VaR* n'est pas sous-additive. Enfin elle ne donne aucune information sur l'ampleur des pertes lorsque celles-ci dépassent la *VaR*. Des auteurs proposent l'utilisation de l'*expected shortfall* au lieu de la *VaR*. L'*expected shortfall* représente en fait la perte moyenne au delà de la *VaR*. La troisième section sera précisément consacrée à l'élaboration de cette mesure de risque en utilisant un groupe de modèles spécifiques appelés “*peaks-over-threshold models*” ou “*POT models*”. Ces modèles tiennent compte de toutes les observations qui sont plus grandes qu'un certain seuil et font intervenir la distribution de Pareto généralisée (**GPD**).

2.1 Présentation de la théorie classique des extrêmes

2.1.1 Le cadre d'analyse

Supposons données n variables aléatoires (va) $X_1, \dots, X_n$ iid de fonction de répartition $\mathbf{F}$ définie par

$$\mathbf{F}(x) = \Pr(X \leq x) \quad (2.1)$$

Une manière simple d'étudier le "comportement" des événements extrêmes est de considérer la variable aléatoire

$$M_n = \max \{X_1, \dots, X_n\} \quad (2.2)$$

Pour notre étude, M_n représente la plus grande perte observée sur les n pertes observées $X_1, \dots, X_n$ (nous adopterons la convention que la perte est un nombre positif). Comme les va sont *iid* on obtient facilement

$$\begin{aligned} \Pr \{M_n \leq x\} &= \Pr \{X_1 \leq x, \dots, X_n \leq x\} \\ &= [\mathbf{F}(x)]^n \end{aligned} \quad (2.3)$$

La difficulté provient du fait que l'on ne connaît pas en général la fonction de répartition $\mathbf{F}$. C'est la raison pour laquelle on s'intéresse au comportement asymptotique de la *va* M_n . Ainsi en exhibant la famille de loi vers laquelle M_n va converger, on pourra remplacer $\mathbf{F}$ par cette dernière pour des grandes valeurs de n .

2.1.2 Le théorème limite de Fisher-Tippett

Avant d'énoncer le principal théorème de cette section, définissons des classes d'équivalences sur l'ensemble des fonctions de répartition (sur les distributions de probabilité).

Définition 1 Les distributions $\mathbf{F}$ et $\mathbf{F}^*$ sont dites de même type si

$$\exists a, b \in \mathbb{R}, \forall x \in \mathbb{R}, \quad \mathbf{F}^*(ax + b) = \mathbf{F}(x) \quad (2.4)$$

Nous pouvons à présent énoncer le théorème de Fisher-Tippett qui permet de caractériser la loi de distribution des extrêmes :

Théorème 1 S'il existe des constantes $a_n > 0$ et b_n telles que

$$\lim_{n \rightarrow \infty} \Pr \left\{ \frac{M_n - b_n}{a_n} \leq x \right\} = \mathbf{G}(x) \quad (2.5)$$

pour une fonction de distribution $\mathbf{G}$ non dégénérée, alors $\mathbf{G}$ appartient à l'un des trois types suivants

$$\begin{aligned} &\text{Type I : Gumbel} \\ \mathbf{G}(x) &= \exp(-e^{-x}) \quad x \in \mathbb{R} \end{aligned}$$

$$\begin{aligned} &\text{Type II : Fréchet} \\ \mathbf{G}(x) &= 0 \quad \text{si } x \leq 0 \\ \mathbf{G}(x) &= \exp(-x^{-\alpha}) \quad \text{si } x > 0, \alpha > 0 \end{aligned}$$

$$\begin{aligned} &\text{Type III : Weibull} \\ \mathbf{G}(x) &= \exp(-(-x)^{-\alpha}) \quad \text{si } x \leq 0, \alpha > 0 \\ \mathbf{G}(x) &= 1 \quad \text{si } x > 0 \end{aligned}$$

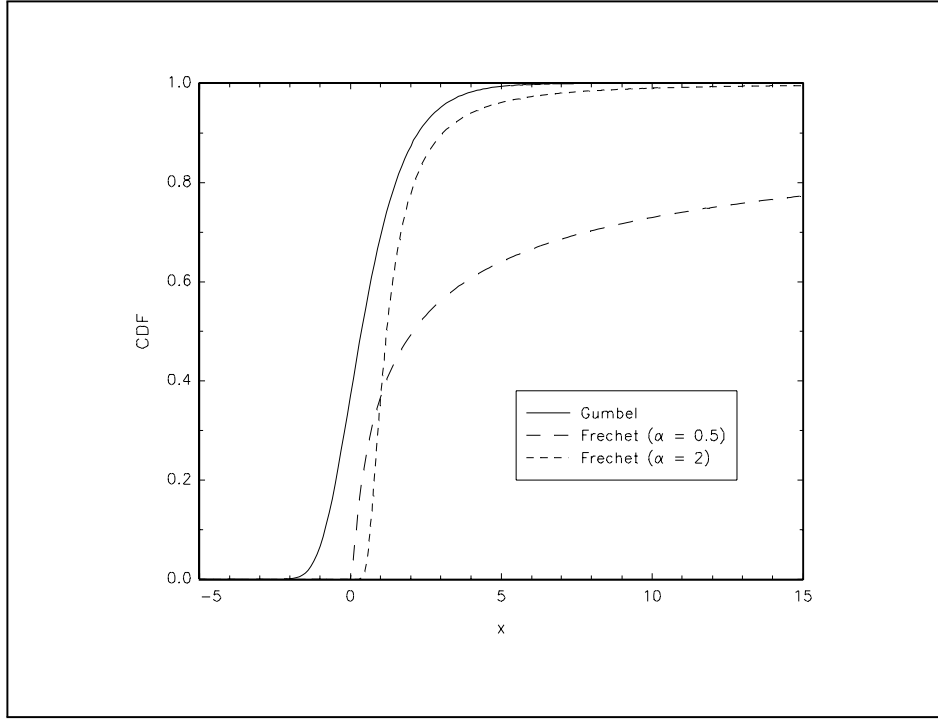
Nous avons représenté les graphiques des fonctions $\mathbf{G}$ dans le cas d'une Gumbel et d'une Fréchet sur la figure (2.1). Nous remarquons que le graphique de la fonction de densité de Fréchet est toujours en dessous de celui de la Gumbel. De plus on remarque que plus le facteur α baisse (dans le cas de la fonction de densité de Fréchet), plus la distribution est à queue épaisse : on retrouve là le fait que le facteur $\frac{1}{\alpha}$ est un indice de queue de distribution.

Exemple 1 Supposons que $\mathbf{F}$ soit une distribution exponentielle de paramètre 1, c'est-à-dire

$$\mathbf{F}(x) = 1 - \exp(-x) \quad (2.6)$$

Alors

$$\mathbf{F}(x)^n = (1 - e^{-x})^n \quad (2.7)$$



Graphique 2.1. Distributions de Gumbel et Fréchet

Ainsi

$$\begin{aligned} (\mathbf{F}(x + \ln(n)))^n &= (1 - e^{-x - \ln(n)})^n \\ &\longrightarrow_{n \rightarrow \infty} \exp(-e^{-x}) \end{aligned} \quad (2.8)$$

Prenons $a_n = 1$ et $b_n = \ln(n)$. $(M_n - b_n)/a_n$ tend asymptotiquement vers une distribution de Gumbel. On a illustré cette convergence sur la figure (2.2). On remarque que pour $n = 50$, la distribution de $(M_n - b_n)/a_n$ est pratiquement confondue avec la distribution de Gumbel. Sur cet exemple, nous avons donc une convergence très rapide. Toutefois cela n'est pas toujours le cas.

En général, on applique le théorème précédent comme suit : on considère n suffisamment grand et fixé ; on fait alors l'approximation suivante

$$\Pr \left\{ \frac{M_n - b}{a} \leq x \right\} \approx \mathbf{G}(x) \quad (2.9)$$

pour un certain $a > 0$ et un certain b . De manière équivalente, cela s'écrit

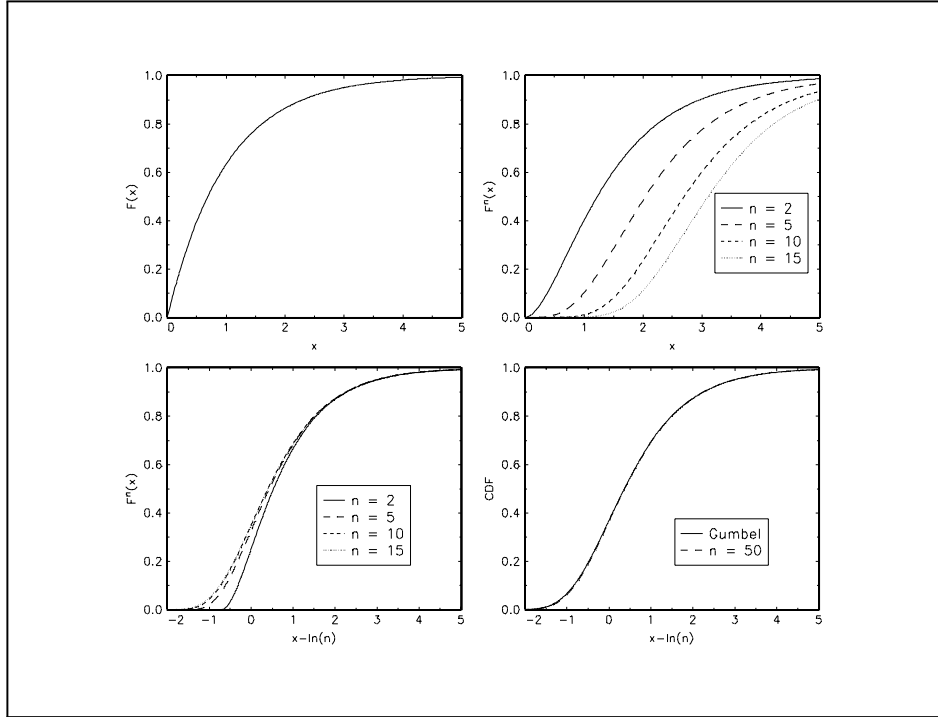
$$\Pr \{M_n \leq x\} \approx \mathbf{G}\left(\frac{x - b}{a}\right) = \mathbf{G}^*(x) \quad (2.10)$$

En fait nous pouvons caractériser les trois types de distribution précédentes par une distribution unique

$$\mathbf{G}(x) = \exp \left\{ - \left[1 + \xi \left(\frac{x - \mu}{\sigma} \right) \right]^{\frac{-1}{\xi}} \right\} \quad (2.11)$$

défini sur le support $\Delta = \{x : 1 + \xi \left(\frac{x - \mu}{\sigma} \right) > 0\}$. Cette fonction de distribution correspond à la loi de probabilité de Von Mises, mais elle est plus connue sous le nom de **“Generalized Extreme Value distribution” (GEV)**. Nous avons alors les correspondances suivantes :

$$\begin{array}{lll} \text{Fréchet} & \xi & = \alpha^{-1} > 0 \\ \text{Weibull} & \xi & = -\alpha^{-1} < 0 \\ \text{Gumbel} & \xi & \rightarrow 0 \end{array}$$



Graphique 2.2. Convergence de la distribution exponentielle vers la Gumbel

Remarquons que les paramètres μ et σ sont en fait les limites de b_N et a_N .

A titre d'illustration, voyons quelques exemples de distribution **GEV**. Pour cela nous pouvons montrer facilement que la fonction de densité correspondante est

$$g(x) = \frac{1}{\sigma} \left[1 + \xi \left(\frac{x - \mu}{\sigma} \right) \right]^{-\frac{1+\xi}{\xi}} \exp \left\{ - \left[1 + \xi \left(\frac{x - \mu}{\sigma} \right) \right]^{\frac{-1}{\xi}} \right\} \quad (2.12)$$

On remarque une translation des fonctions de densité lorsque μ varie. μ est donc un paramètre de **localisation**. σ joue le rôle d'une variance, c'est pourquoi nous le considérons comme un paramètre de **dispersion**. Enfin le paramètre ξ est lié au caractère leptokurtique de la fonction de distribution **F**. C'est pourquoi on lui donne généralement le nom d'indice de queue ou d'indice de valeur extrême (voir HOEUNG, RIBOULET et RONCALLI [1999]).

Un exemple d'application classique de ce modèle est l'estimation des quantiles d'extrême, c'est-à-dire l'estimation de x_p défini par

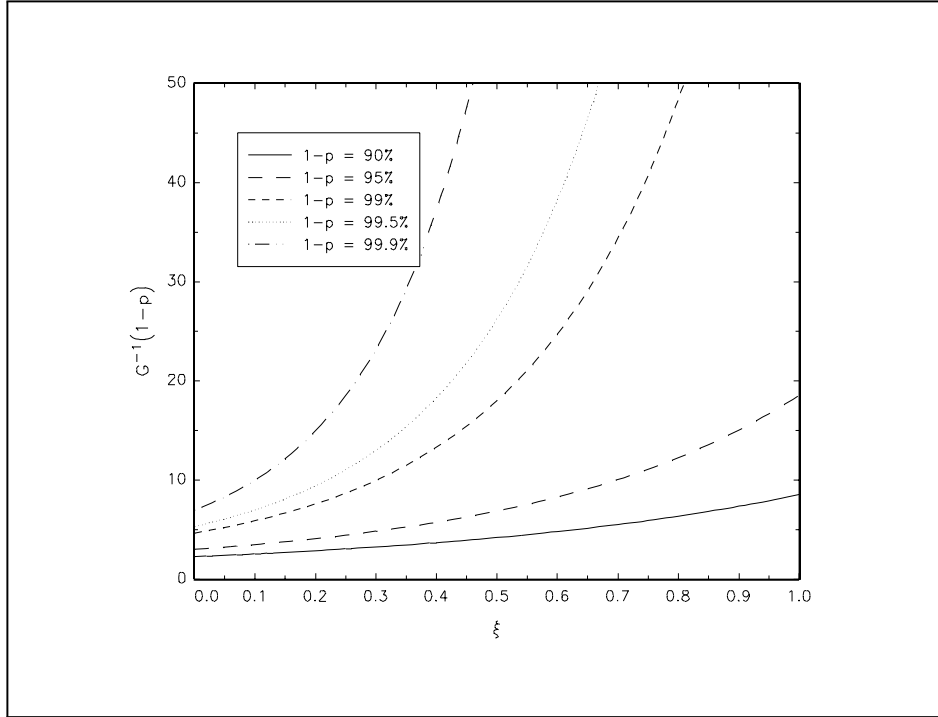
$$\mathbf{G}(x_p) = 1 - p \quad (2.13)$$

avec p un nombre réel proche de 1 (cela signifie que sur une période correspondant aux n observations, la perte sera plus petite que x_p avec une probabilité de $1 - p$). Il suffit pour cela d'inverser la fonction G pour obtenir

$$x_p = \mu - \frac{\sigma}{\xi} \left[1 - (-\ln(1 - p))^{-\xi} \right] \quad (2.14)$$

Ce quantile est donc influencé par les deux paramètres σ et ξ . A titre d'illustration, la figure (2.3) présente différentes courbes de quantiles en fonction de ξ . Intuitivement, nous comprenons que plus la distribution **F** possède de la variance et plus elle est leptokurtique, plus ce quantile sera élevé. Si ξ est négatif, les quantiles sont supérieurement limités.

Nous pouvons également calculer à partir des résultats précédents l'espérance du pire, c'est-à-dire $\mathbb{E}[M_n]$ par des techniques d'intégration numérique de type Gauss-Legendre ou Gauss-Hermite. Nous



Graphique 2.3. Quantiles de la distribution GEV

pouvons également calculer des grandeurs telles que le mode du pire ou les temps de retour. Toutes ces grandeurs sont autant d'éléments utiles pour la gestion du risque. Mais notons toutefois qu'en pratique nous ne connaissons pas les paramètres μ , σ et ξ ; il faut donc les estimer à partir des données et les remplacer par leurs estimations.

2.1.3 Estimation des paramètres de la distribution **GEV**

Supposons que nous ayons déjà observé une suite de k maxima $M_{n,1}, \dots, M_{n,k}$ et on suppose que le nombre d'observations n est suffisamment grand pour que l'approximation par une **GEV** soit valable. Le problème est alors d'estimer les paramètres (μ, σ, ξ) . Il existe plusieurs méthodes pour estimer les paramètres de la distribution **GEV**. Nous pouvons par exemple citer les méthodes d'estimation de l'indice de queue, la méthode des moments ou encore les méthodes de seuil (basée par exemple sur la distribution de Pareto généralisée). Mais celle qui reste la plus populaire et qui sous certaines conditions¹ est la plus efficace est la méthode du maximum de vraisemblance.

Soit θ le vecteur de paramètres. Nous avons

$$\theta = \begin{bmatrix} \mu \\ \sigma \\ \xi \end{bmatrix} \quad (2.15)$$

Nous considérons un échantillon de données $\{X_t\}$ de dimension $T = kn$ avec $k \in \mathbb{N}$. Nous divisons cet échantillon en n blocs et nous définissons χ_n^+ de la façon suivante

$$\chi_q^+ = \max(\{X_{1+q(t-1)}, t = 1, \dots, k\}) \quad (2.16)$$

L'expression de la vraisemblance de l'observation est donc

$$L(\chi_q^+; \theta) = \frac{1}{\sigma} \left[1 + \xi \left(\frac{\chi_q^+ - \mu}{\sigma} \right) \right]^{-\frac{1+\xi}{\xi}} \exp \left\{ - \left[1 + \xi \left(\frac{\chi_q^+ - \mu}{\sigma} \right) \right]^{\frac{-1}{\xi}} \right\} \quad (2.17)$$

¹La méthode de maximum de vraisemblance est valable pour $\xi > -0.5$. De plus les données empiriques en finance ou assurance montrent que l'on a généralement $\xi > 0$.

Nous en déduisons l'expression suivante pour la log-vraisemblance

$$\ell(\chi_q^+; \theta) = -\ln \sigma - \left(\frac{1+\xi}{\xi} \right) \ln \left(1 + \xi \left(\frac{\chi_q^+ - \mu}{\sigma} \right) \right) - \left[1 + \xi \left(\frac{\chi_q^+ - \mu}{\sigma} \right) \right]^{\frac{-1}{\xi}} \quad (2.18)$$

L'estimateur du maximum de vraisemblance correspond alors à

$$\hat{\theta}_{ML} = \arg \max_{\theta \in \Theta} \sum_{q=1}^N \ell(\chi_q^+; \theta) \quad (2.19)$$

avec Θ l'espace des paramètres.

HOEUNG, RIBOULET et RONCALLI [1999] ont simulé plusieurs séries de données de loi $\mathbf{G}(3, 2, \frac{1}{2})$ de longueur 1000. Pour chacune des séries, ils estiment les paramètres de la distribution **GEV** en prenant respectivement les cinquante, cent et cinq cents premières observations, ainsi que l'échantillon entier. Ils remarquent que l'adéquation des fonctions de densité estimées à la vraie loi est meilleure lorsque l'on dispose de plus de données. Cela peut poser des problèmes importants pour l'application de la théorie des extrêmes à la mesure du risque, puisqu'elle nécessite un échantillon de longueur assez importante pour obtenir des bonnes estimations. En effet, les pertes potentielles vont être calculées comme pour les méthodes *VaR* historique et Monte Carlo. La différence provient du fait qu'on ne calcule pas directement le quantile, mais qu'on estime la distribution à partir des maxima. Par exemple, si nous avons un historique de 250 jours et que nous considérons des blocs de 25 jours de trading pour définir χ_q^+ , l'estimation portera sur un échantillon de 10 extrema. **Nous ne pouvons donc appliquer la théorie des extrêmes que dans le cadre d'un système performant de gestion de bases de données qui a été mis en place relativement tôt afin de disposer d'historiques assez longs.**

C'est pourquoi nous allons développer dans la section suivante une autre approche fondée également sur la théorie des extrêmes et qui utilise les “*peaks-over-threshold models*” ou *méthodes de seuil*. Cette méthode aura l'avantage d'utiliser plus de données (précisément celles qui dépassent un certain seuil suffisamment élevé).

2.2 Une mesure de risque cohérente (ou ‘*expected shortfall*’)

2.2.1 Mesures de risques extrêmes et distribution de Pareto généralisée

Rappelons que mathématiquement les mesures de risques extrêmes sont données grâce à la fonction de distribution de perte $\mathbf{F}$. Si $q \in \mathbb{R}$, la valeur en risque est le q -ième quantile de la distribution $\mathbf{F}$

$$VaR_q = \mathbf{F}^{-1}(q) \quad (2.20)$$

où $\mathbf{F}^{-1}$ est la fonction réciproque de $\mathbf{F}$ ². On définit également l'expected shortfall ou ES qui va mesurer la taille moyenne des pertes sachant que celles-ci dépassent la VaR

$$\text{ES}_q = \mathbb{E}[X \mid X > VaR_q] \quad (2.22)$$

Cette mesure de risque est une mesure cohérente. De plus, nous mentionons de nouveau la *VaR* pour deux raisons essentiellement : tout d'abord parce que l'on va obtenir une nouvelle méthode pour calculer la *VaR* à partir de la théorie des extrêmes (grâce à la distribution de Pareto généralisée ou **GPD**), et puis enfin parce qu'elle va nous être utile pour le calcul de l'ES.

Comme la distribution $\mathbf{F}$ elle-même, ces deux dernières mesures représentent des quantités théoriques que l'on ne connaît pas et qu'il nous faudra estimer. Nous donnerons dans les sections suivantes des méthodes “simples” pour estimer ces quantités. Mais avant cela, nous présenterons la distribution de Pareto

²Dans le cas où $\mathbf{F}$ n'est pas inversible, on définit l'inverse généralisée comme suit

$$\mathbf{F}^{\leftarrow}(q) = \inf \{x \in \mathbb{R}, F(x) \geq p\} \quad (2.21)$$

et on fait les mêmes raisonnements avec $\mathbf{F}^{\leftarrow}$.

généralisée ou **GPD** (*Generalized Pareto Distribution*) qui aura pour notre étude la même importance que la distribution gaussienne pour le calcul de la VaR par la méthode analytique.

La **GPD** (Distribution de Pareto Généralisée) est une distribution à deux paramètres de fonction de distribution

$$\mathbf{G}_{\xi,\beta}(x) = 1 - \left(1 + \frac{\xi x}{\beta}\right)^{-\frac{1}{\xi}} \quad \text{si } \xi \neq 0 \quad (2.23)$$

et

$$\mathbf{G}_{\xi,\beta}(x) = 1 - \exp\left(\frac{-x}{\beta}\right) \quad \text{si } \xi = 0 \quad (2.24)$$

avec $\beta > 0, x \geq 0$ quand $\xi \geq 0$ et $0 \leq x \leq -\beta/\xi$ quand $\xi < 0$. Le paramètre ξ est lié au caractère leptokurtique de la fonction de distribution et β est un paramètre d'échelle. Le cas $\xi > 0$ est le plus intéressant dans les modèles de gestion de risque car la **GPD** est à queue épaisse. Alors que la loi normale a des moments de tout ordre, une distribution à queue épaisse possède des moments d'ordre infini. Par exemple, dans le cas de la **GPD** avec $\xi > 0$ on a :

$$\mathbb{E}[X^k] = +\infty \quad \text{pour } k \geq \frac{1}{\xi} \quad (2.25)$$

Quand $\xi = 1/2$, la **GPD** n'a pas de variance ; quand $\xi = 1/4$, la **GPD** a un moment d'ordre 4 infini.

2.2.2 Estimation de la fonction de distribution des excès

La fonction de distribution des excès de perte par rapport à un *seuil* élevé u est définie par

$$\begin{aligned} \mathbf{F}_u(y) &= \Pr\{X - u \leq y \mid X > u\} \\ \text{pour } 0 &\leq y < x_0 - u \\ \text{avec } x_0 &= \sup\{x \in \mathbb{R} : \mathbf{F}(x) < 1\} \end{aligned} \quad (2.26)$$

Remarque 1 On a généralement $x_0 = +\infty$.

La fonction de distribution des excès représente la probabilité qu'une certaine perte dépasse le seuil u d'au plus une quantité y , sachant qu'elle dépasse u . Remarquons que l'on peut écrire cette fonction de distribution sous la forme

$$\mathbf{F}_u(y) = \frac{\mathbf{F}(y+u) - \mathbf{F}(u)}{1 - \mathbf{F}(u)} \quad (2.27)$$

On supposera, comme le suggère la remarque précédent, que $x_0 = +\infty$ pour **F** dans ce qui suit.

On va à présent énoncer un théorème limite qui va être le résultat théorique central de notre étude. On va volontairement l'énoncer sous deux formes : l'une imprécise mais qui donne une idée intuitive du contenu du théorème ainsi que de l'utilisation qui en sera faite ; l'autre rigoureuse avec toutes les hypothèses.

Théorème 2 Pour une "grande classe" de fonctions de distribution on peut trouver $\beta(u)$ positif tel que

$$\lim_{u \rightarrow x_0} \sup_{0 \leq y < x_0 - u} |\mathbf{F}_u(y) - \mathbf{G}_{\xi,\beta(u)}(y)| = 0 \quad (2.28)$$

Ce théorème signifie que pour une "grande classe" de fonctions de distributions **F**, quand le seuil u augmente, la fonction de distribution des excès converge vers une **GPD**. On peut dire que la "grande classe" de fonctions de distributions contient toutes les fonctions de distributions continues classique des statistiques et de l'actuariat (normale, lognormale, χ^2 , Student, Fisher, gamma, exponentielle, uniforme, beta, etc.).

D'après le théorème ci-dessus, la manière la plus naturelle de modéliser la fonction de distribution des excès au delà de seuils suffisamment élevés est l'utilisation de la **GPD**. Cela signifie que pour un certain u suffisamment grand, on peut poser

$$\mathbf{F}_u(y) = \mathbf{G}_{\xi,\beta}(y) \quad (2.29)$$

pour un certain ξ et β . Il suffit alors d'estimer les paramètres ξ et β par la méthode de maximum de vraisemblance.

Avant de détailler la démarche à suivre pour estimer les différents paramètres du modèle, donnons un énoncé plus précis du théorème précédent. Pour cela, nous aurons besoin de nouvelles définitions et notations. On note

$$\mathbf{H}_{\xi,\mu,\beta}(x) = \exp \left\{ - \left[1 + \xi \left(\frac{x-\mu}{\sigma} \right) \right]^{\frac{-1}{\xi}} \right\} \quad (2.30)$$

sur $\Delta = \{x : 1 + \xi \left(\frac{x-\mu}{\sigma} \right) > 0\}$ et

$$\mathbf{H}_{\xi,0,1} = \mathbf{H}_{\xi} \quad (2.31)$$

Définition 2 Si $\mathbf{F}$ vérifie les hypothèses du théorème (1) pour une certaine fonction de distribution $\mathbf{G}$ que l'on peut représenter à l'aide de $\mathbf{H}_{\xi,\mu,\beta}$, alors on dit que $\mathbf{F} \in MDA(\mathbf{H}_{\xi,\mu,\beta})$ (MDA pour "maximum domain of attraction").

Théorème 3 Soit $\mathbf{F}$ une fonction de distribution et $\mathbf{F}_u$ sa fonction de distribution des excès associée. Alors, pour $\xi \in \mathbb{R}$, $\mathbf{F} \in MDA(H_{\xi})$ si et seulement si il existe une fonction positive mesurable $\beta(u)$ telle que

$$\lim_{u \rightarrow x_0} \sup_{0 \leq y < x_0 - u} |\mathbf{F}_u(y) - \mathbf{G}_{\xi,\beta(u)}(y)| = 0 \quad (2.32)$$

2.2.3 Estimation des paramètres de la distribution **GPD**

La détermination de u est assez délicate. u doit être choisi suffisamment grand pour que l'on puisse appliquer le théorème précédent avec une bonne approximation mais u ne doit pas être trop grand car il faut suffisamment de données pour avoir des estimateurs de bonne qualité. Il existe des techniques efficaces pour évaluer u : nous utiliserons celle qui est la plus utilisée et qui est en fait une technique graphique. Pour cela nous aurons besoin d'une nouvelle définition.

Définition 3 On appelle fonction d'excès en moyenne ("mean excess function") la fonction $e(u)$ définie par

$$e(u) = \mathbb{E}[X - u \mid x > u] \quad \text{pour } u \geq 0 \quad (2.33)$$

Un estimateur empirique de cette fonction en statistique est donné par

$$e_n(u) = \frac{\sum_{i=1}^n (X_i - u)^+}{\sum_{i=1}^n I_{\{X_i > u\}}} \quad (2.34)$$

En outre, on montre facilement que pour la **GPD** on a

$$e(u) = \mathbb{E}[X - u \mid x > u] = \frac{\beta + \xi u}{1 - \xi} \quad (2.35)$$

pourvu que $\xi < 1$ (ce qui est souvent le cas avec les données dont on dispose en finance ou en assurance). Ainsi $e(u)$ est une fonction linéaire en u . Donc étant données les observations de perte $X_1, \dots, X_n$ on trace graphiquement $e_n(u)$ en fonction de u . On repère alors les valeurs de u à partir desquelles le graphique est presque linéaire. Lorsque u est déterminé, on compte le nombre de pertes N_u qui dépassent u sur les n pertes observées. On utilise alors une méthode de maximum de vraisemblance pour déterminer les estimateurs de ξ et β .

Le but est ici d'obtenir une estimation de $\mathbf{F}(x)$ pour $x = u + y$ avec u un seuil suffisamment élevé et $y \geq 0$. En combinant (2.27) et (2.29) on peut écrire

$$\mathbf{F}(x) = (1 - \mathbf{F}(u))\mathbf{G}_{\xi,\beta}(x - u) + \mathbf{F}(u) \quad (2.36)$$

$$\text{avec } x = u + y \quad (2.37)$$

On veut utiliser cette dernière formule pour donner une estimation de la queue de distribution. On a donc besoin d'un estimateur de $\mathbf{F}(u)$. Pour estimer $\mathbf{F}(u)$, on utilise l'estimateur historique $n^{-1}(n - N_u)$. On n'utilise pas cet estimateur pour $\mathbf{F}(x)$ car il est mauvais au niveau des queues de distribution où les données sont trop dispersées. En revanche, on choisit u pour qu'il y ait suffisamment de données au-dessus de u , de sorte que cet estimateur est satisfaisant. En remplaçant dans (2.36) les différents paramètres par leurs estimateurs on obtient une estimation de la queue de distribution

$$\widehat{F}(x) = 1 - \frac{N_u}{n} \left(1 + \widehat{\xi} \frac{x - u}{\widehat{\beta}} \right)^{\frac{-1}{\widehat{\xi}}} \quad (2.38)$$

Remarquons que cet estimateur est valable pour $x > u$. **Cet estimateur peut être vu comme un estimateur historique amélioré par la théorie des extrêmes.**

2.2.4 Estimateurs de la VaR et de l'expected shortfall

Par définition de la VaR, et pour une probabilité $q > \mathbf{F}(u)$, un estimateur de la VaR est obtenu en inversant l'estimateur de la queue de distribution (2.38) :

$$\widehat{VaR}_q = u + \frac{\widehat{\beta}}{\widehat{\xi}} \left(\left(\frac{n}{N_u} (1 - q) \right)^{-\widehat{\xi}} - 1 \right) \quad (2.39)$$

L'expected shortfall est relié à la VaR par la formule

$$\mathbb{ES}_q = VaR_q + \mathbb{E}[X - VaR_q \mid X > VaR_q] \quad (2.40)$$

Notre modèle pour la fonction de distribution des excès au delà du seuil u possède une jolie propriété de stabilité. En effet, pour n'importe quel seuil plus grand que u , tel que VaR_q pour $q > \mathbf{F}(u)$, alors la fonction de distribution des excès au delà de ce seuil plus élevé est encore une **GPD** avec le même paramètre ξ mais de paramètre d'échelle différent. On montre facilement que

$$\mathbf{F}_{VaR_q}(y) = \mathbf{G}_{\xi, \beta + \xi(VaR_q - u)}(y) \quad (2.41)$$

L'intérêt de la formule (2.41) provient du fait qu'on a une expression simple et explicite de la distribution des excès au delà de la VaR. En remarquant que la moyenne de la distribution donnée dans (2.41) est

$$\frac{\beta + \xi(VaR_q - u)}{1 - \xi} \quad (2.42)$$

on peut calculer l'expected shortfall. On trouve

$$\frac{\mathbb{ES}_q}{VaR_q} = \frac{1}{1 - \xi} + \frac{\beta - \xi u}{(1 - \xi) VaR_q} \quad (2.43)$$

Considérons ce ratio dans le cas où $x_0 = \infty$. Dans ce cas ce ratio est en grande partie déterminé par le premier terme $(1 - \xi)^{-1}$. En effet, le second terme devient négligeable quand q est de plus en plus proche de 1. Cette observation sur le comportement asymptotique met en évidence l'importance du paramètre ξ . Il détermine comment nos deux mesures de risque diffèrent dans les régions extrêmes de la distribution des pertes. En remplaçant dans (2.43) les différents paramètres par leurs estimateurs on obtient comme estimateur de l'expected shortfall

$$\widehat{\mathbb{ES}}_q = \frac{\widehat{VaR}_q + \widehat{\beta} - \widehat{\xi}u}{1 - \widehat{\xi}} \quad (2.44)$$

2.3 Applications financières

2.3.1 Description de la base de données LME

Nous utilisons dans cette section les bases de données du *London Metal Exchange*³. Nous nous intéressons plus particulièrement au cours (spot) de l'Aluminium (AL), du Cuivre (CU) ainsi qu'au cours à

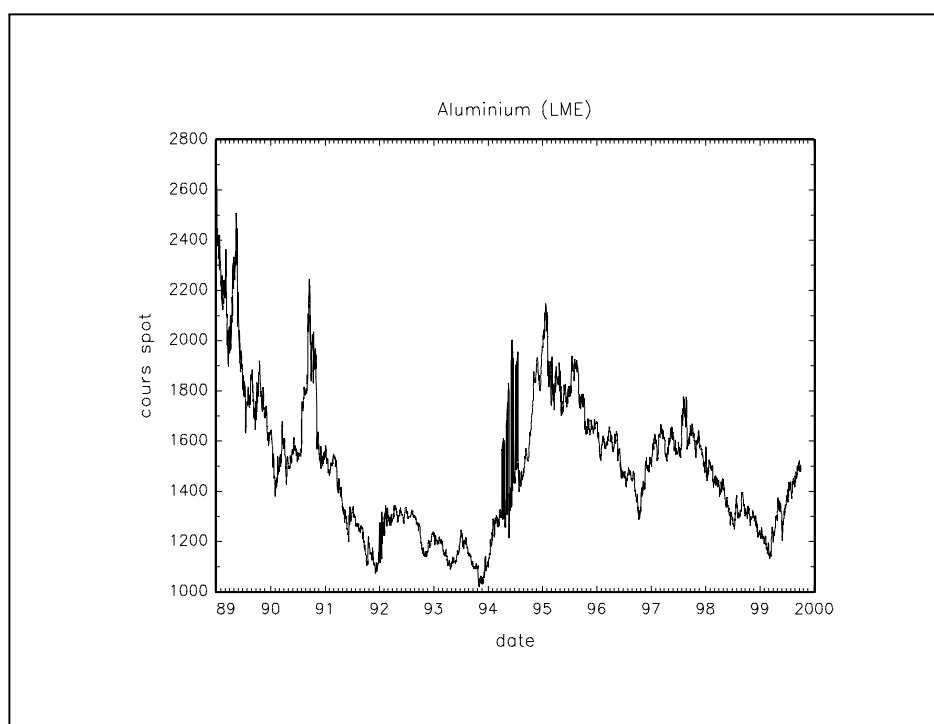
³Nous invitons le lecteur à consulter le site web du LME (<http://www.lme.co.uk>) pour obtenir de plus ample informations sur la base de données ou pour refaire les calculs.

terme ou forward (à 15 mois) de l'Aluminium (AL-15). Nous avons représenté sur les graphiques (2.4), (2.5) et (2.6) la variation des cours pour AL, CU et AL-15.

Nous constituons à partir de ces trois produits neuf portefeuilles que nous notons $P_1, \dots, P_9$ sous forme d'un vecteur ligne à trois colonnes et qui sont constitués comme suit :

1. $P_1 = [1, 0, 0]$
2. $P_2 = [-1, 0, 0]$
3. $P_3 = [0, 1, 0]$
4. $P_4 = [0, -1, 0]$
5. $P_5 = [0, 0, 1]$
6. $P_6 = [0, 0, -1]$
7. $P_7 = [1, 1, 1]$
8. $P_8 = [1, -1, -1]$
9. $P_9 = [-3, 5, 2]$

Le premier élément est relatif à AL, le second à CU et le troisième à AL-15. Les quantités positives indiquent une position longue et les quantités négatives une position courte.



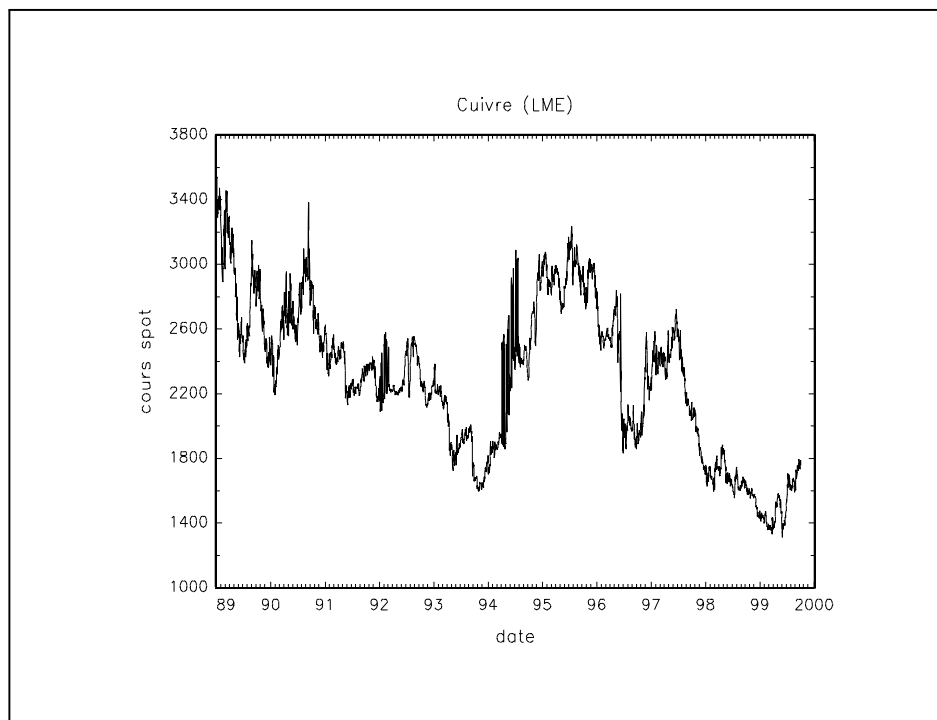
Graphique 2.4. Cours de l'aluminium

2.3.2 Le calcul de la valeur en risque d'un portefeuille

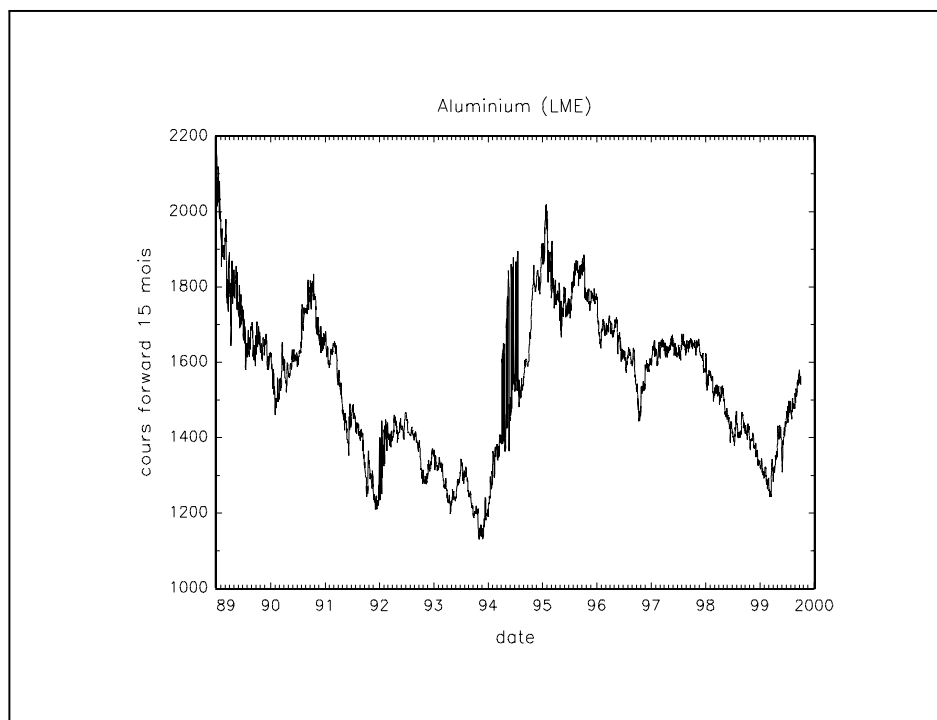
Nous allons à présent appliquer les méthodes développées en début de chapitre pour calculer la *VaR* analytique, la *VaR* historique et les quantiles à $\alpha\%$ avec la méthode **GEV**.

Nous avons représenté sur la figure (2.7) la fonction de densité des variations normalisées du cours des portefeuilles. Ces fonctions de densité ont été obtenues par la méthode des noyaux. Pour chaque portefeuille, nous avons également représenté sa densité sous l'hypothèse de normalité.

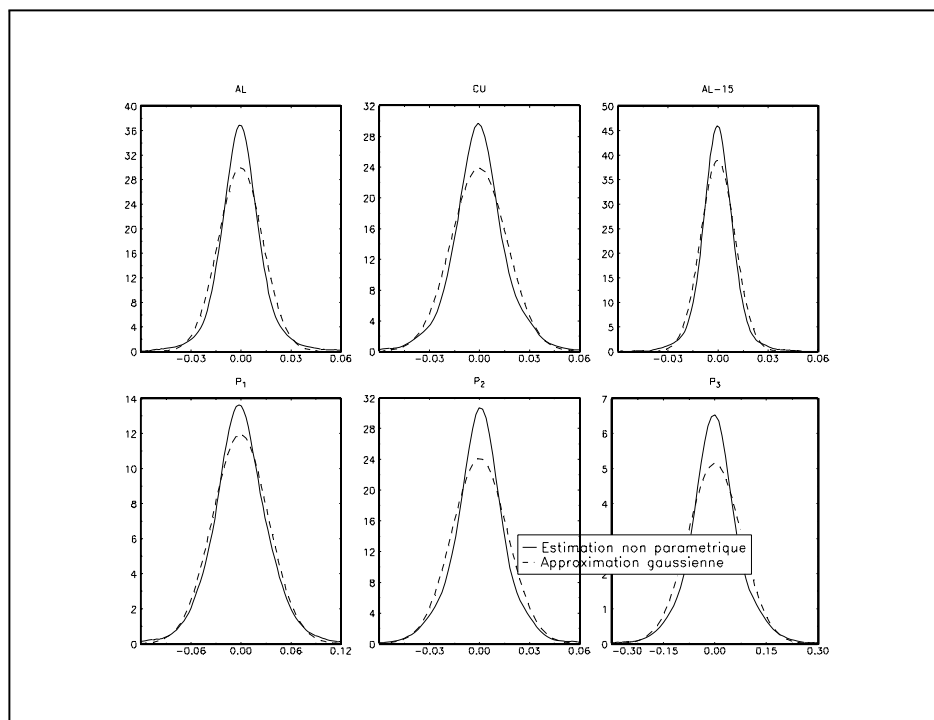
Nous remarquons que les courbes des fonctions de densité estimées de manière non paramétrique se trouvent "à l'intérieur" des courbes des fonctions de densité obtenues sous l'hypothèse de normalité,



Graphique 2.5. Cours du cuivre



Graphique 2.6. Cours à terme de l'aluminium



Graphique 2.7. Estimations paramétrique et non paramétrique des rendements

autour de la moyenne. En revanche les fonctions de densité estimées de manière non paramétrique sont à queues plus épaisses. L'analyse gaussienne consiste donc à se focaliser sur les événements qui sont concentrés autour du mode de la fonction de densité. Cela revient à s'éloigner suffisamment pour ne distinguer que la forme générale sans considérer les parties accidentées de cette fonction de densité.

Les résultats observés dans les tableaux (2.1) et (2.2) appellent de nombreux commentaires. Les différentes *VaR* correspondent à l'exigence en fond propre (en dollars) pour se couvrir contre des risques à 90%, 95%, 99%, 99.6% et 99.9%. Tout d'abord, pour l'horizon à 1 jour, nous remarquons que la *VaR* analytique, comparée à la *VaR* historique, sous-estime les risques pour des quantiles supérieurs à 99% alors qu'elle le sur-estime pour les quantiles à 90%. Ceci est en fait dû à l'approximation gaussienne qui a été faite pour le calcul de la *VaR* analytique ; en effet comme nous l'avons expliqué plus haut, l'approximation gaussienne conduit à accorder moins de poids aux événements de la queue de distribution, c'est à dire aux événements extrêmes. De plus les portefeuilles P_i et $-P_i$ présentent la même *VaR* analytique en raison du caractère symétrique de la densité gaussienne. A 99% et 99.6%, la *VaR* historique est toujours supérieur à la *VaR* **GEV** : ceci provient du fait que lorsque l'on prend pour référence des quantiles supérieurs à 99%, les événements extrêmes sont de moins en moins nombreux et du fait des méthodes d'estimation, les éventuels points aberrants joueront un rôle plus important dans l'estimation par la méthode *VaR* historique. Ainsi, pour la mesure de risques extrêmes, la *VaR* **GEV** semble être un indicateur pertinent. A 99.9%, les *VaR* historique et **GEV** sont sensiblement les mêmes. Cependant on peut dire que la mesure *VaR* **GEV** est plus pertinente car elle fait intervenir un plus grand nombre de points dans les calculs. En effet, pour des quantiles 99.6% ou plus, et sur une période de 250 jours, la *VaR* historique ne tient compte que d'une seule donnée, à savoir la plus grande perte, alors que la théorie des extrêmes prend en compte 10 valeurs si on découpe les 250 jours en période de 25 jours. On peut donc penser que la *VaR* **GEV** est un indicateur plus robuste. En ce qui concerne le seuil à 99%, on peut difficilement se prononcer. En effet la *VaR* analytique est, sur nos données, toujours supérieur à la *VaR* **GEV** mais inférieur à la *VaR* historique. On peut difficilement donner la préférence à une de ces trois mesures : elles ont chacune leurs faiblesses. En revanche, si on se place dans une optique dynamique, c'est à dire si on prend en compte la possibilité que la réglementation change, et que le seuil dépasse les 99%, alors la *VaR* **GEV** devient un outil incontournable. Cette hypothèse n'est pas à écarter. De plus, en raison de leurs objectifs de rating, les banques ciblent sur des seuils supérieurs à 99% dans leurs

Portefeuille	Méthode VaR	90%	95%	99.0%	99.6%	99.9%
P_1	VaR analytique	29	37	53	60	70
	VaR historique	23	31	75	92	151
	VaR GEV	✓	✓	46	71	129
P_2	VaR analytique	29	37	53	60	70
	VaR historique	22	32	70	90	120
	VaR GEV	✓	✓	52	76	125
P_3	VaR analytique	52	67	95	108	126
	VaR historique	43	65	120	152	239
	VaR GEV	✓	✓	94	139	231
P_4	VaR analytique	52	67	95	108	126
	VaR historique	45	65	107	150	198
	VaR GEV	✓	✓	93	131	200
P_5	VaR analytique	21	27	39	44	51
	VaR historique	18	25	50	63	93
	VaR GEV	✓	✓	37	53	83
P_6	VaR analytique	21	27	39	44	51
	VaR historique	18	25	45	65	89
	VaR GEV	✓	✓	40	56	85
P_7	VaR analytique	85	109	154	175	204
	VaR historique	73	104	190	258	331
	VaR GEV	✓	✓	146	209	340
P_8	VaR analytique	51	65	92	105	123
	VaR historique	43	63	103	138	207
	VaR GEV	✓	✓	93	134	212
P_9	VaR analytique	246	316	447	509	593
	VaR historique	194	306	513	759	1158
	VaR GEV	✓	✓	✓	✓	✓

TABLE 2.1. *Value-at-Risk* 1 jour

programmes d'allocation de fonds propres. **Une banque notée AAA se doit par exemple d'utiliser un seuil de 99.97%.**

La lecture du second tableau soulève le problème du scaling. En effet, le comité Bâle impose une période de détention de 10 jours pour mesurer les risques d'un portefeuille. Il conseille alors de convertir la *VaR* 1 jour à la *VaR* 10 jours par scaling. Le scaling consiste à multiplier la volatilité à 1 jour par $\sqrt{T}$, T étant l'horizon considéré. Cette méthode soulève une certaine polémique quant à son application et il n'y a toujours pas un consensus clair sur le sujet. Elle repose essentiellement sur une hypothèse de normalité et d'indépendance des accroissements des prix des actifs. Ces hypothèses sont souvent discutées. Toutefois, cette méthode est appréciée pour sa simplicité et sa très facile mise en œuvre. L'alternative à cette méthode repose sur la théorie des extrêmes. Il faut estimer le coefficient α de la loi extrême correspondant aux données et multiplier ensuite par $T^{\frac{1}{\alpha}}$. Une autre méthode consiste à déterminer la matrice de variance-covariance des rendements à T jours. Cette approche nécessite un historique assez long des données et ne peut être appliquée à tous les horizons. Nous avons choisi de multiplier la *VaR* analytique et la *VaR* historique par $\sqrt{T}$ et la *VaR* **GEV** par $T^{\frac{1}{\alpha}}$. On remarque ainsi que la *VaR* analytique devient presque toujours supérieur à la *VaR* **GEV** pour un horizon de 10 jours. Le scaling conduit ainsi à une surestimation de la *VaR*. En conclusion nous pouvons dire que nos résultats sont cohérents avec les idées rencontrées dans la littérature. En effet les articles publiés sur le sujet s'accordent à dire que l'approche *VaR* analytique conduit à une sur-estimation de la mesure de risque à 10 jours par rapport à un calcul basé sur la théorie des extrêmes. Si nous assimilons la mesure basée sur les extrêmes au véritable risque encouru, nous pouvons donc affirmer que la méthode du scaling est conservatrice.

Portefeuille	Méthode VaR	90%	95%	99.0%	99.6%	99.9%
P_1	VaR analytique	92	118	166	189	221
	VaR historique	71	98	237	291	477
	VaR GEV	✓	✓	129	190	335
P_2	VaR analytique	92	118	166	189	221
	VaR historique	68	101	223	286	380
	VaR GEV	✓	✓	124	166	251
P_3	VaR analytique	166	212	300	343	399
	VaR historique	136	208	378	479	755
	VaR GEV	✓	✓	231	313	483
P_4	VaR analytique	166	212	300	343	399
	VaR historique	142	206	338	475	626
	VaR GEV	✓	✓	201	258	365
P_5	VaR analytique	67	86	122	139	162
	VaR historique	57	79	158	200	294
	VaR GEV	✓	✓	83	110	167
P_6	VaR analytique	67	86	122	139	162
	VaR historique	56	79	142	206	280
	VaR GEV	✓	✓	86	109	154
P_7	VaR analytique	268	344	487	555	646
	VaR historique	231	330	600	816	1047
	VaR GEV	✓	✓	340	461	721
P_8	VaR analytique	161	206	292	333	388
	VaR historique	135	198	325	437	656
	VaR GEV	✓	✓	212	279	412
P_9	VaR analytique	778	999	1413	1610	1876
	VaR historique	614	968	1622	2399	3661
	VaR GEV	✓	✓	✓	✓	✓

TABLE 2.2. *Value-at-Risk* 10 jours

2.3.3 La construction des scénarios de crises

La théorie des extrêmes peut également être un outil performant dans l'élaboration des scénarios de crise. Comme le préconisent les autorités réglementaires, les établissements financiers doivent effectuer régulièrement des simulations de crise afin de connaître le montant des pertes potentielles en cas de fluctuations dangereuses et importantes du marché. Il faut tout de suite préciser que malgré le caractère obligatoire, ces simulations de crise ne sont en aucun cas utilisés pour calculer l'exigence en fonds propres. Cependant, l'utilisation des simulations de crise est l'un des quatre principes de base définis par la Commission bancaire pour qu'un modèle interne soit validé. Le but de ces simulations est alors de compléter le dispositif de mesure des risques, et notamment les mesures de VaR⁴.

Ces simulations de crise, plus connu sous le terme anglais *stress testing*, doivent être des outils pour appréhender l'exposition de la banque à une crise grave. Contrairement à la *VaR*, elles doivent répondre à la question : quel est le montant de perte auquel la banque doit faire face lors de la prochaine crise si le portefeuille de négociation ne change pas ? On adopte donc là une vision du risque extrême.

Nous allons, pour illustrer une méthodologie⁵ utilisant la théorie des extrêmes, établir des scénarios de crise sur la variation des cours de AL, AL-15 et CU. Pour cela, introduisons une notion essentielle qui est le **temps de retour**. Lorsque nous nous fixons un seuil de 99% par exemple, la probabilité qu'un événement extrême (une perte) survienne est de 0.01. Lorsque l'on travaille sur la *VaR* à un jour, cela

⁴HOEUNG, RIBOULET et RONCALLI [1999] citent une étude menée par Tucker et Lawrence en 1998 sur le risk management auprès de 679 institutions de 25 pays différents et qui montre que 93% des risk managers considèrent qu'une mesure *VaR* est insuffisante comme mesure de risque.

⁵Il existe deux grandes familles de stress testing : les méthodes *objectives* et les méthodes dites *subjectives*. Le premier type utilise les faits historiques pour construire des scénarios, alors que la seconde famille est basée sur des hypothèses de travail.

signifie que le temps de retour de cet événement est de $\frac{1}{0.01} = 100$ jours (on utilise juste le fait que la probabilité d'apparition d'une crise est de 0.01 et que la variable aléatoire Z qui est égale à 1 si une crise survient et 0 sinon suit une loi géométrique de probabilité 0.01 a une espérance de $1/0.01$). De manière plus générale, lorsque l'on cherche à se protéger contre un risque donnée à un seuil de $\alpha = 1 - p$, le temps de retour τ est de

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

Cependant, lorsque nous avons utilisé la *VaR* **GEV** (avec des blocs de 20 jours), **pour pouvoir faire des comparaisons avec les autres *VaR* à un jour**, nous avons du choisir un autre quantile que le quantile à $\alpha\%$. Si nous appelons α_{GEV} le seuil correspondant à un temps de retour $\tau = (1 - \alpha)^{-1}$, on a la relation

$$\tau = 20 \times \frac{1}{1 - \alpha_{\text{GEV}}} \quad (2.45)$$

Ainsi pour un horizon de 1 jour, on compare la *VaR* analytique avec un seuil de α à la *VaR* **GEV** avec un quantile α_{GEV} .

Nous élaborons des scénarios de crise de la manière suivante. On se donne des temps de retour élevés de l'ordre de 5 ans, 10 ans, 25 ans, 50 ans, 75 ans et 100 ans et on utilise la formule (2.45) pour calculer la variation maximale négative des cours sur une journée. Les résultats obtenus sont⁶ :

Echelles de Risque (en variation journalière) pour des positions longues (en %)			
Temps de Retour	AL	AL-15	CU
5	-6.74	-4.82	-8.03
10	-8.57	-5.90	-9.35
25	-11.75	-7.68	-11.23
50	-14.91	-9.36	-12.76
75	-17.14	-10.49	-13.70
100	-18.92	-11.38	-14.39

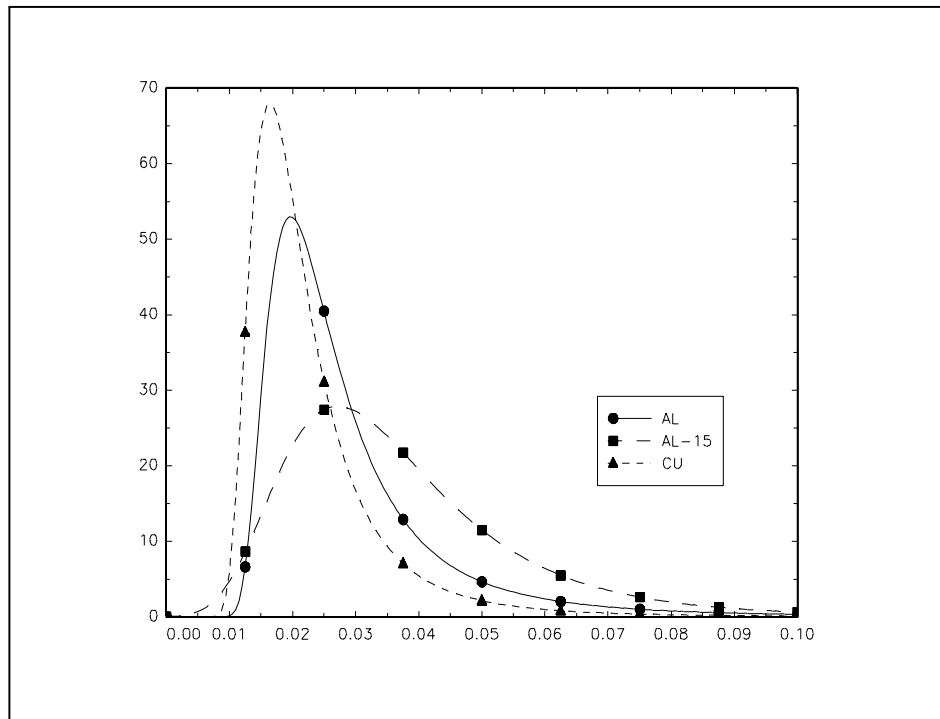
Echelles de Risque (en variation journalière) pour des positions courtes (en %)			
Temps de Retour	AL	AL-15	CU
5	6.96	5.79	7.53
10	8.35	7.10	8.82
25	10.46	9.20	10.73
50	12.32	11.12	12.34
75	13.52	12.39	13.36
100	14.43	13.38	14.12

Nous avons reporté en outre sur les figures (2.8) et (2.9) la densité des lois d'extrêmes estimées ainsi que la relation entre l'échelle de risque et le temps de retour dans le cas des positions longues.

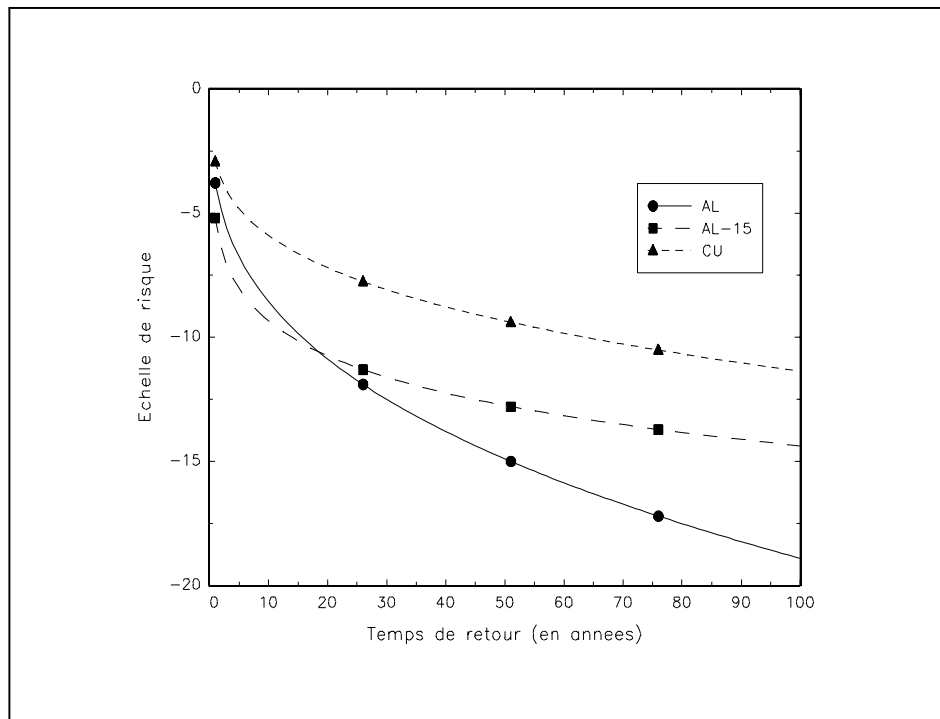
Notons qu'il existe d'autres alternatives à l'utilisation de la **GEV**. COSTINOT, RIBOULET et RONCALLI [2000] utilisent une approche dite semi-paramétrique en s'appuyant sur le fait que les séries financières appartiennent au domaine d'attraction de la distribution de Fréchet. Les quantiles extrêmes se déduisent des quantiles empiriques à partir d'une simple homothétie paramétrée par l'indice de queue (voir aussi les travaux de STRAETMANS [1999,2000]). Dans cette approche, l'indice de queue est alors donné par l'estimateur de Hill.

⁶ Pour information, la valeurs des paramètres estimés sont

Paramètres	AL	AL-15	CU
μ	0.022	0.018	0.028
σ	0.007	0.006	0.013
ξ	0.344	0.275	0.095



Graphique 2.8. Densité des lois d'extrêmes estimées



Graphique 2.9. Relation entre l'échelle de risque et le temps de retour

3

La théorie des copules

La dépendance entre les variables aléatoires est parfaitement décrite par leur distribution jointe. On peut toutefois séparer le comportement des distributions marginales de la structure de dépendance. On réalise ceci grâce à l'utilisation des copules qui sont un outil statistique permettant de relier la densité jointe aux densités marginales. C'est la fonction copule qui va contenir toute l'information sur la structure de dépendance du modèle.

Les copules ont été introduites en 1959 dans des études sur les espaces de probabilité métriques. Ce n'est que depuis 1999 que cette théorie commence à être appliquée à la finance (voir par exemple EMBRECHTS, MCNEIL et STRAUMANN [1999]). Les perspectives d'application sont nombreuses du fait de la simplification apportée à l'étude du cas multivarié : on calcule les densités marginales et la structure de dépendance des modèles séparément. En outre, les copules présentent des propriétés intéressantes telles que l'invariance par des transformations croissantes.

Cette partie constitue une introduction à la théorie des copules et à quelques une de ses applications à la finance, comme l'élaboration de scénarios de crises dans le cas multivarié. La présentation qui est faite ici est très largement inspirée des exposés de JOE [1997], NELSEN [1999] et BOUYÉ, NIKEGHBALI, RIBOULET et RONCALLI [2000].

3.1 Présentation générale

3.1.1 Introduction mathématique

Définition 4 Soient $S_1, \dots, S_n$ des sous-ensembles non vides de $\overline{\mathbb{R}}$. Soit $\mathbf{H}$ une fonction réelle de n variables dont le domaine de définition est $\text{Dom}(\mathbf{H}) = S_1 \times \dots \times S_n$ et soit $[a, b]$, avec $a = (a_1, \dots, a_n)$ et $b = (b_1, \dots, b_n)$ un hypercube inclus dans $\text{Dom}(\mathbf{H})$. On appelle $\mathbf{H}$ -volume de $[a, b]$ le réel défini par

$$V_{\mathbf{H}}(B) = \sum_{c \in S_{a,b}} \text{sgn}(c) \mathbf{H}(c) \quad (3.1)$$

où $S_{a,b}$ est l'ensemble des sommets de $[a, b]$ (il y en a 2^n) et où $\text{sgn}(c)$ est défini par

$$\text{sgn}(c) = \begin{cases} 1 & \text{si } c_k = a_k \text{ pour un nombre pair de } k \\ -1 & \text{si } c_k = a_k \text{ pour un nombre impair de } k \end{cases}$$

Définition 5 Une fonction $\mathbf{H}$ de n variables à valeurs réelles est dite n -croissante si $V_{\mathbf{H}}(B) \geq 0$ pour tout hypercube inclus dans $\text{Dom}(\mathbf{H}) = S_1 \times \dots \times S_n$. Si de plus chaque S_k a un plus petit élément a_k , et si $\mathbf{H}(t) = 0$ pour tout $t = (t_1, \dots, t_n)$ tel qu'il existe au moins un k pour lequel $t_k = a_k$, on dit alors que $\mathbf{H}$ est "grounded".

Si chaque S_k est non vide et possède un plus grand élément b_k , alors on dit que H possède des marginales, et les marginales uni-dimensionnelles (ou marginales) de $\mathbf{H}$ sont les fonctions $\mathbf{H}_k$ définies sur $\text{Dom}(\mathbf{H}_k) = S_k$ par

$$\mathbf{H}_k(x) = \mathbf{H}(b_1, \dots, b_{k-1}, x, b_{k+1}, \dots, b_n) \quad \text{pour tout } x \in S_k \quad (3.2)$$

On peut également définir les marginales k -dimensionnelles pour $k \geq 2$ en fixant $n - k$ variables. On peut à présent donner la définition d'une copule (NELSEN [1999]).

Définition 6 Une copule n -dimensionnelle ou copule est une fonction ayant les propriétés suivantes :

1. $\text{Dom}(\mathbf{C}) = \mathbf{I}^n$ où $\mathbf{I} = [0, 1]$;
2. $\mathbf{C}$ est grounded et n -croissante ;
3. $\mathbf{C}$ possède des marginales $\mathbf{C}_k$, $k = 1, 2, \dots, n$ qui vérifient $\mathbf{C}_k(u) = u$ pour tout $u \in S_k$.

Remarque 2 De manière équivalente, on peut dire qu'une copule est une fonction de $\mathbf{I}^n$ dans $\mathbf{I}$ avec les propriétés suivantes :

1. Pour tout $u \in \mathbf{I}^n$, $\mathbf{C}(u) = 0$ si au moins l'une des coordonnées de u est nulle. De plus si toutes les coordonnées de u sont égales à 1 sauf u_k , alors $\mathbf{C}(u) = u_k$;
2. Pour tout a et b dans $\mathbf{I}^n$, tels que $a \leq b^1$ on a $V_{\mathbf{C}}([a, b]) \geq 0$.

Remarque 3 Pour toute copule n -dimensionnelle, chaque marginale k -dimensionnelle est une k -copule.

Il existe un résultat d'uniforme continuité sur les copules :

Théorème 4 Soit $\mathbf{C}$ une copule. Alors pour tout $(u, v) \in \mathbf{I}^2$

$$|\mathbf{C}(v) - \mathbf{C}(u)| \leq \sum_{k=1}^n |v_k - u_k| \quad (3.3)$$

Ainsi toute copule est uniformément continu.

On parlera souvent dans la suite des dérivées partielles des copules. Nous allons à présent justifier la dérivabilité des copules. Nous allons citer les résultats sans en donner de démonstrations².

Théorème 5 Soit C une copule. Pour $v \in \mathbf{I}$, la dérivée partielle $\frac{\partial \mathbf{C}}{\partial u}$ existe pour presque tout u , et quand cette dérivée existe, on a

$$0 \leq \frac{\partial \mathbf{C}}{\partial u}(u, v) \leq 1 \quad (3.4)$$

De même, pour tout $u \in \mathbf{I}$, la dérivée partielle $\frac{\partial \mathbf{C}}{\partial v}$ existe pour presque tout v , et quand cette dérivée existe, on a

$$0 \leq \frac{\partial \mathbf{C}}{\partial v}(u, v) \leq 1 \quad (3.5)$$

De plus, les fonctions $u \mapsto \frac{\partial \mathbf{C}}{\partial v}(u, v)$ et $v \mapsto \frac{\partial \mathbf{C}}{\partial u}(u, v)$ sont définies et croissantes presque partout sur $\mathbf{I}$.

¹ $a \leq b$ signifie que $\forall k \in \{1, \dots, n\}$, nous vérifions

$$a_k \leq b_k$$

² Cependant, celles-ci ne sont pas difficiles ; elles utilisent essentiellement le fait que toute fonction monotone d'une variable réelle est presque partout dérivable.

3.1.2 Théorème de Sklar et premières propriétés des copules

3.1.2.1 Le théorème de Sklar

Le théorème de Sklar est le théorème fondamental dans la théorie des copules : c'est lui qui va permettre la liaison entre la densité jointe d'un vecteur aléatoire et ses densités marginales.

Définition 7 En dimension n , une fonction de distribution $\mathbf{H}$ est une fonction dont le domaine de définition est $\overline{\mathbb{R}}^n$ et qui vérifient les propriétés suivantes :

- (a) $\mathbf{H}$ est n -croissante ;
- (b) $\mathbf{H}(t) = 0$ pour tout $t \in \overline{\mathbb{R}}^n$ tel que $t_k = -\infty$ pour au moins un indice k et $\mathbf{H}(\infty, \dots, \infty) = 1$.

$\mathbf{H}$ est donc grounded et on note ses marginales $\mathbf{F}_1, \dots, \mathbf{F}_n$.

Théorème 6 (Théorème de Sklar) Soit $\mathbf{H}$ une fonction de distribution avec des marginales $\mathbf{F}_1, \dots, \mathbf{F}_n$. Alors il existe un copula $\mathbf{C}$ tel que pour tout $x = (x_1, \dots, x_n) \in \overline{\mathbb{R}}^n$

$$\mathbf{H}(x_1, \dots, x_n) = \mathbf{C}(\mathbf{F}_1(x_1), \dots, \mathbf{F}_n(x_n)) \quad (3.6)$$

Si les $\mathbf{F}_1, \dots, \mathbf{F}_n$ sont toutes continues, alors $\mathbf{C}$ est unique ; sinon le copula $\mathbf{C}$ est déterminé de manière unique sur $\mathbf{F}_1(\mathbf{I}) \times \dots \times \mathbf{F}_n(\mathbf{I})$. Réciproquement, si $\mathbf{C}$ est un copula et $\mathbf{F}_1, \dots, \mathbf{F}_n$ des fonctions de distribution, alors la fonction $\mathbf{H}$ définie par (3.6) est une fonction de distribution de marginales $\mathbf{F}_1, \dots, \mathbf{F}_n$.

Remarque 4 Une preuve de ce théorème peut être trouvée dans SCHWEIZER et SKLAR [1983].

On voit ainsi que pour des fonctions de distribution continues, on peut séparer la structure de dépendance de la densité jointe et des densités marginales ; cette structure de dépendance est représentée dans la copule.

Corollaire 1 Soit $\mathbf{H}, \mathbf{C}, \mathbf{F}_1, \dots, \mathbf{F}_n$ les fonctions définies dans le théorème précédent et soient $\mathbf{F}_1^{(-1)}, \dots, \mathbf{F}_n^{(-1)}$ les inverses généralisées de $\mathbf{F}_1, \dots, \mathbf{F}_n$ respectivement. Alors pour tout $\mathbf{u} \in \mathbf{I}^n$

$$\mathbf{C}(u_1, \dots, u_n) = \mathbf{H}\left(\mathbf{F}_1^{(-1)}(u_1), \dots, \mathbf{F}_n^{(-1)}(u_n)\right) \quad (3.7)$$

Exemple 2 (la distribution bivariable de Gumbel) Soit $\mathbf{H}_\theta$ la fonction de distribution donnée par

$$\mathbf{H}_\theta(x, y) = \begin{cases} 1 - e^{-x} - e^{-y} + e^{-(x+y+\theta xy)}, & x \geq 0, y \geq 0 \\ 0 & \text{si non} \end{cases} \quad (3.8)$$

où θ est un paramètre dans $[0, 1]$. Les marginales sont des exponentielles, les fonctions inverses sont donc respectivement $\mathbf{F}^{(-1)}(u) = -\ln(1-u)$ et $\mathbf{G}^{(-1)}(v) = -\ln(1-v)$ pour $u, v \in \mathbf{I}$. On en déduit que le copula correspondant est

$$\mathbf{C}_\theta(u, v) = u + v - 1 + (1-u)(1-v)e^{-\theta \ln(1-u) \ln(1-v)} \quad (3.9)$$

Exemple 3 (La distribution bivariable gaussienne) Soit ρ le coefficient de corrélation. La copule normale bivariable est donnée par

$$\mathbf{C}_\rho(u, v) = \mathcal{N}_\rho(\Phi^{-1}(u), \Phi^{-1}(v)) \quad (3.10)$$

On a de manière explicite

$$\mathbf{C}_\rho(u, v) = \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi(1-\rho^2)^{1/2}} \exp\left(-\frac{1}{2(1-\rho^2)}(s^2 - 2\rho st + t^2)\right) ds dt \quad (3.11)$$

On voit de la même manière qu'en choisissant n'importe quelle copule différente de $\mathbf{C}_\rho$, on peut obtenir une loi bivariable qui n'est pas normale mais dont les marginales sont gaussiennes.

3.1.2.2 La relation d'ordre sur les copules

Il est intéressant de voir qu'on peut *dans un certain sens* mettre une relation d'ordre **partielle** sur les copules, et obtenir des bornes inférieures et supérieures valables pour n'importe quelle copule. On considère les copules particulières suivantes :

$$\begin{aligned}\mathbf{C}^+(\mathbf{u}) &= \min(u_1, \dots, u_n) \\ \mathbf{C}^\perp(\mathbf{u}) &= u_1 \times u_2 \times \dots \times u_n \\ \mathbf{C}^-(\mathbf{u}) &= \max(u_1 + \dots + u_n - n + 1, 0)\end{aligned}$$

On peut facilement vérifier que $\mathbf{C}^+$ et $\mathbf{C}^\perp$ sont des copules alors que $\mathbf{C}^-$ n'est pas une copule pour $n > 2$. En effet, le $V_{\mathbf{C}^-}$ volume de l'hypercube $[1/2, 1]^n$ est négatif. Par symétrie, on a

$$\begin{aligned}V_{\mathbf{C}^-}([1/2, 1]^n) &= \max(1 + \dots + 1 - n + 1, 0) \\ &\quad - n \max(1/2 + 1 + \dots + 1 - n + 1, 0) \\ &\quad + \binom{n}{2} \max(1/2 + 1/2 + 1 + \dots + 1 - n + 1, 0) \\ &\quad \dots \\ &\quad + \max(1/2 + 1/2 + \dots + 1/2 - n + 1, 0) \\ &= 1 - n \times \frac{1}{2} + 0 + \dots + 0 \\ &= 1 - \frac{n}{2}\end{aligned}\tag{3.12}$$

ce qui montre que pour $n \geq 3$, $\mathbf{C}^-$ n'est pas une copule.

Théorème 7 Soit $\mathbf{C}$ une copule. Alors pour tout $\mathbf{u} \in \mathbf{I}$,

$$\mathbf{C}^-(\mathbf{u}) \leq \mathbf{C}(\mathbf{u}) \leq \mathbf{C}^+(\mathbf{u})\tag{3.13}$$

(Pour la preuve, le lecteur peut consulter NELSEN [1999]). $\mathbf{C}^-$ et $\mathbf{C}^+$ sont appelées respectivement les bornes inférieures et supérieures de Fréchet-Hoeffding.

Définition 8 Soient $\mathbf{C}_1$ et $\mathbf{C}_2$ deux copules. On dit que $\mathbf{C}_1$ est plus petite que $\mathbf{C}_2$ et on note $\mathbf{C}_1 \prec \mathbf{C}_2$ si $\mathbf{C}_1(\mathbf{u}) \leq \mathbf{C}_2(\mathbf{u})$ pour tout $\mathbf{u} \in \mathbf{I}$.

Cette relation d'ordre est **partielle** car on ne peut pas comparer toutes les copules entre elles. Néanmoins, le théorème précédent montre que l'ensemble des copules muni de la relation d'ordre précédemment définie est inférieurement et supérieurement borné.

3.1.2.3 Les copules et les variables aléatoires

On se donne dans ce qui suit des variables aléatoires $X_1, \dots, X_n$ avec des fonctions de distribution $\mathbf{F}_1, \dots, \mathbf{F}_n$ respectivement, de densité jointe $\mathbf{H}$ et de copule $\mathbf{C}$ où $\mathbf{C}$ est donnée par la relation (3.6).

Théorème 8 Soient $X_1, \dots, X_n$ des variables aléatoires continues de copule $\mathbf{C}$. Les variables $X_1, \dots, X_n$ sont indépendantes si et seulement si $\mathbf{C} = \mathbf{C}^\perp$.

On obtient ainsi une caractérisation de l'indépendance par les copules. Une autre propriété intéressante des copules est leur invariance par des transformations strictement croissantes.

Théorème 9 Soient $X_1, \dots, X_n$ des variables aléatoires continues de copule $\mathbf{C}$. Si $\alpha_1, \dots, \alpha_n$ sont des fonctions strictement croissantes sur $\text{Im}(X_1), \dots, \text{Im}(X_n)$ respectivement³, alors $\alpha_1(X_1), \dots, \alpha_n(X_n)$ ont la même copule $\mathbf{C}$. Ainsi $\mathbf{C}$ est invariante par toute transformation strictement croissante des variables aléatoires $X_1, \dots, X_n$.

³On note $\text{Im}(X)$ l'ensemble des valeurs que prend la variable aléatoire X .

Démonstration 1 Notons $\mathbf{G}_1, \dots, \mathbf{G}_n$ les distributions de $\alpha_1(X_1), \dots, \alpha_n(X_n)$ respectivement et $\mathbf{C}_\alpha$ la copule associée. Comme pour tout k , α_k est strictement croissante et on a

$$\mathbf{G}_k(x) = \Pr[\alpha_k(X_k) \leq x] = \Pr[X_k \leq \alpha_k^{-1}(x)] = \mathbf{F}_k(\alpha_k^{-1}(x)) \quad (3.14)$$

pour tout $x \in \overline{\mathbb{R}}$. D'où

$$\begin{aligned} \mathbf{C}_\alpha(\mathbf{G}_1(x_1), \dots, \mathbf{G}_n(x_n)) &= \Pr[\alpha_1(X_1) \leq x_1, \dots, \alpha_n(X_n) \leq x_n] \\ &= \Pr[X_1 \leq \alpha_1^{-1}(x_1), \dots, X_n \leq \alpha_n^{-1}(x_n)] \\ &= \mathbf{C}(\mathbf{F}_1(\alpha_1^{-1}(x_1)), \dots, \mathbf{F}_n(\alpha_n^{-1}(x_n))) \\ &= \mathbf{C}(\mathbf{G}_1(x_1), \dots, \mathbf{G}_n(x_n)) \end{aligned} \quad (3.15)$$

Ainsi $\mathbf{C}_\alpha = \mathbf{C}$.

On peut également s'intéresser aux transformations décroissantes. Il existe un théorème général, dont la démonstration est identique à celle du théorème précédent.

Théorème 10 Soient $X_1, \dots, X_n$ des variables aléatoires continues de copule $\mathbf{C}_{X_1, \dots, X_n}$. Soient $\alpha_1, \dots, \alpha_n$ des fonctions strictement monotones sur $\text{Im}(X_1), \dots, \text{Im}(X_n)$ respectivement, et notons $\mathbf{C}_{\alpha_1(X_1), \dots, \alpha_n(X_n)}$ la copule de $\alpha_1(X_1), \dots, \alpha_n(X_n)$. Supposons par exemple que α_k est strictement décroissante pour un certain k , par exemple $k = 1$. Alors on a

$$\mathbf{C}_{\alpha_1(X_1), \dots, \alpha_n(X_n)}(u_1, \dots, u_n) = \mathbf{C}_{\alpha_2(X_2), \dots, \alpha_n(X_n)}(u_2, \dots, u_n) - \mathbf{C}_{X_1, \alpha_2(X_2), \dots, \alpha_n(X_n)}(1 - u_1, \dots, u_n) \quad (3.16)$$

Remarque 5 On voit aisément par récurrence que le copula $\mathbf{C}_{\alpha_1(X_1), \dots, \alpha_n(X_n)}$ s'exprime à l'aide du copula $\mathbf{C}_{X_1, \dots, X_n}$.

Exemple 4 Plaçons-nous en dimension 2 et considérons deux transformations α_1 et α_2 strictement décroissantes. Par application du théorème précédent, on obtient

$$\mathbf{C}_{\alpha_1(X_1), \alpha_2(X_2)}(u_1, u_2) = u_1 + u_2 - 1 + \mathbf{C}_{X_1, X_2}(1 - u_1, 1 - u_2) \quad (3.17)$$

On peut grâce à cette formule obtenir la copule de survie $\hat{\mathbf{C}}$ de (X_1, X_2)

$$\overline{\mathbf{H}}(x_1, x_2) = \Pr[X_1 > x_1, X_2 > x_2] = \hat{\mathbf{C}}(\overline{\mathbf{F}}_1(x_1), \overline{\mathbf{F}}_2(x_2)) \quad (3.18)$$

3.1.2.4 Les relations de dépendance

Nous avons vu que la structure de dépendance entre variables aléatoires était contenue dans la fonction copule. Ainsi, celle-ci va constituer un bon cadre pour étudier et mesurer la dépendance entre les variables aléatoires. KIMELDORF et SAMPSON [1989] présente les conditions que doit satisfaire une mesure d'association pour être une mesure de concordance.

Définition 9 Une mesure d'association χ entre deux variables aléatoires continues X_1 et X_2 de copule $\mathbf{C}$ est une mesure de concordance si elle satisfait les propriétés qui suivent :

1. χ est définie pour toute paire de variables aléatoires continues ;
2. $-1 \leq \chi_{X,Y} \leq 1$, $\chi_{X,X} = 1$ et $\chi_{X,-X} = -1$;
3. $\chi_{X,Y} = \chi_{Y,X}$;
4. si X et Y sont indépendantes, alors $\chi_{X,Y} = \chi_{\mathbf{C}^\perp} = 0$;
5. $\chi_{-X,Y} = \chi_{X,-Y} = -\chi_{X,Y}$;
6. si $\mathbf{C}_1$ et $\mathbf{C}_2$ sont deux copules telles que $\mathbf{C}_1 \prec \mathbf{C}_2$ alors $\chi_{\mathbf{C}_1} \leq \chi_{\mathbf{C}_2}$;
7. si $\{(X_n, Y_n)\}$ est une suite continue de variables aléatoires avec des copules $\{\mathbf{C}_n\}$, et si $\{\mathbf{C}_n\}$ converge point par point vers $\mathbf{C}$, alors $\lim_{n \rightarrow \infty} \chi_{\mathbf{C}_n} = \chi_{\mathbf{C}}$.

Du fait des résultats d'invariance des copules par transformation strictement croissante, on a

Théorème 11 Soit χ une mesure de concordance pour des variables aléatoires continues X et Y .

1. si Y est presque sûrement une fonction croissante de X , alors $\chi_{X,Y} = \chi_{\mathbf{C}^+} = 1$;
2. si Y est presque sûrement une fonction décroissante de X , alors $\chi_{X,Y} = \chi_{\mathbf{C}^-} = -1$;
3. si α et β sont des fonctions strictement croissantes, alors $\chi_{\alpha(X),\beta(Y)} = \chi_{X,Y}$.

Nous allons à présent étudier deux mesures de concordance qui sont très utilisées : le tau de Kendall et le rho de Spearman (voir JOE [1997] ou NELSEN [1999]).

Définition 10 Soient $\mathbf{F}$ une fonction de distribution bivariée et (X_1, X_2) et (X'_1, X'_2) deux vecteurs aléatoires indépendants de distribution $\mathbf{F}$. On définit alors le **tau de Kendall** par

$$\begin{aligned}\tau &= \Pr((X_1 - X'_1)(X_2 - X'_2) > 0) - \Pr((X_1 - X'_1)(X_2 - X'_2) < 0) \\ &= 2\Pr((X_1 - X'_1)(X_2 - X'_2) > 0) - 1 \\ &= 4 \int \mathbf{F} d\mathbf{F} - 1\end{aligned}\tag{3.19}$$

En utilisant les copules, cette relation devient

$$\tau = 4 \int \mathbf{C} d\mathbf{C} - 1\tag{3.20}$$

Pour le calcul du tau de Kendall, la condition $(X_1 - X'_1)(X_2 - X'_2) > 0$ correspond à deux paires de variables concordantes au sens où un des vecteurs a ses deux composantes supérieures à celles de l'autre. La condition $(X_1 - X'_1)(X_2 - X'_2) < 0$ correspond à deux paires de variables aléatoires discordantes dans le sens où pour chaque vecteur une des composantes est plus grande que celle de l'autre vecteur alors que la seconde est plus petite.

Définition 11 Soit $\mathbf{F}$ une fonction de distribution bivariée continue de marginales $\mathbf{F}_1$ et $\mathbf{F}_2$. Soit (X_1, X_2) un vecteur aléatoire de distribution $\mathbf{F}$. Le **rho de Spearman** est alors défini comme le coefficient de corrélation linéaire entre les variables $\mathbf{F}_1(X_1)$ et $\mathbf{F}_2(X_2)$. Comme $\mathbf{F}_1(X_1)$ et $\mathbf{F}_2(X_2)$ suivent une loi uniforme sur $[0, 1]$, leur espérance vaut $1/2$ et leur variance $1/12$. Ainsi, on a

$$\varrho_S = 12 \int \int \mathbf{F}_1(x_1)\mathbf{F}_2(x_2) d\mathbf{F}(x_1, x_2) - 3\tag{3.21}$$

ou en utilisant les copules

$$\varrho_S = 12 \int \int uv d\mathbf{C}(u, v) - 3\tag{3.22}$$

Exemple 5 Notons Q la mesure d'association ϱ_S ou τ . On a alors les égalités suivantes :

$$\begin{aligned}Q(\mathbf{C}^+, \mathbf{C}^+) &= 1 \\ Q(\mathbf{C}^+, \mathbf{C}^\perp) &= \frac{1}{3} \\ Q(\mathbf{C}^+, \mathbf{C}^-) &= 0 \\ Q(\mathbf{C}^-, \mathbf{C}^-) &= -1 \\ Q(\mathbf{C}^\perp, \mathbf{C}^\perp) &= 0\end{aligned}$$

Exemple 6 Soient X et Y deux variables aléatoires suivant chacune une loi exponentielle et dont la copule est celle de Farlie-Gumbel-Morgenstern, c'est à dire

$$\mathbf{C}(u, v) = \mathbf{C}_\theta(u, v) = uv + \theta uv(1-u)(1-v) \quad \text{avec } \theta \in [-1, 1]\tag{3.23}$$

Cette copule peut se voir comme une perturbation du copula indépendant. La loi de distribution jointe de (X, Y) est donnée par

$$\mathbf{H}(x, y) = \mathbf{C}(1 - e^{-x}, 1 - e^{-y}) \quad (3.24)$$

On a

$$\begin{aligned} \varrho_S &= 12 \int \int \mathbf{C}(u, v) \, dudv - 3 \\ &= 12 \int \int (uv + \theta uv(1-u)(1-v)) \, dudv - 3 \\ &= 12 \left(\frac{1}{4} + \frac{\theta}{36} \right) - 3 \\ &= \frac{\theta}{3} \end{aligned} \quad (3.25)$$

De même, pour le tau de Kendall, on obtient

$$\tau = \frac{2\theta}{9} \quad (3.26)$$

On peut à la vue de ces valeurs penser que cette famille sera peu souvent utilisée du fait qu'elle présente un spectre de dépendance trop limité.

On peut également citer d'autres mesures de dépendance qui peuvent s'avérer utiles pour obtenir des informations sur le phénomène modélisé. On ne donne que les définitions pour le cas bivarié — le lecteur peut consulter JOE [1997] pour la généralisation au cas multivarié.

1. **La dépendance positive (PQD).** Soit (X_1, X_2) un vecteur aléatoire de fonction de distribution $\mathbf{F}$ et de marginales $\mathbf{F}_1$ et $\mathbf{F}_2$. On dit que $\mathbf{F}$ est **PQD** si

$$\Pr(X_1 \leq a_1, X_2 \leq a_2) \geq \Pr(X_1 \leq a_1) \Pr(X_2 \leq a_2) \quad \forall a_1, a_2 \in \mathbb{R} \quad (3.27)$$

ou encore

$$\mathbf{C}(\mathbf{F}_1(a_1), \mathbf{F}_2(a_2)) \geq \mathbf{F}_1(a_1) \mathbf{F}_2(a_2) \quad \forall a_1, a_2 \in \mathbb{R} \quad (3.28)$$

On peut remarquer que cette expression est équivalente à l'inégalité

$$\Pr(X_1 > a_1, X_2 > a_2) \geq \Pr(X_1 > a_1) \Pr(X_2 > a_2) \quad \forall a_1, a_2 \in \mathbb{R}$$

Cela signifie que les variables aléatoires X_1 et X_2 ont plus de chances de prendre des valeurs élevées ou faibles ensembles par rapport au cas où elles auraient été indépendantes.

2. **La dépendance positive stochastique croissante (DPSC).** Soit (X_1, X_2) un vecteur aléatoire de fonction de distribution $\mathbf{F}$ et de marginales $\mathbf{F}_1$ et $\mathbf{F}_2$. On dit que X_2 est stochastiquement croissante en X_1 si

$$\Pr(X_2 > x_2 \mid X_1 = x_1) = 1 - \mathbf{F}_{2|1}(x_2 \mid x_1) \quad (3.29)$$

est croissante en x_1 pour tout x_2 . De même en inversant le rôle des variables on peut définir la croissance stochastique de X_1 en X_2 . La croissance stochastique signifie que l'une des variables a plus de chance de croître lorsque l'autre croît.

3. **La dépendance de queue.** Le concept de dépendance de queue renseigne sur la "quantité" de dépendance au niveau des queues de distribution. C'est un outil très pertinent pour l'étude de la dépendance des valeurs extrêmes. De plus, comme c'est une quantité qui sera définie par l'intermédiaire des copules, elle sera invariante par transformation strictement croissante.

Définition 12 Définissons le survivor copula $\overline{\mathbf{C}}$ par

$$\overline{\mathbf{C}}(u, v) = 1 - u - v + \mathbf{C}(u, v) \quad (3.30)$$

Soient X et Y deux variables aléatoires continues de fonction de distribution $\mathbf{F}_1$ et $\mathbf{F}_2$ respectivement. Le coefficient de **dépendance de queue supérieure** de X et Y est défini par

$$\lambda_U = \lim_{u \rightarrow 1^-} \Pr[Y > \mathbf{F}_2^{-1}(u) \mid X > \mathbf{F}_1^{-1}(u)] \quad (3.31)$$

pourvu que la limite $\lambda_U \in [0, 1]$ existe. Si $\lambda_U \in (0, 1]$, X et Y sont dites asymptotiquement dépendantes au niveau supérieur de la queue de distribution. Si $\lambda_U = 0$, X et Y sont dites asymptotiquement indépendantes au niveau supérieur de la queue de distribution.

On peut aussi définir λ_U à l'aide des copules

$$\lambda_U = \lim_{u \rightarrow 1^-} \frac{\overline{\mathbf{C}}(u, u)}{1 - u} \quad (3.32)$$

On définit naturellement l'indice de queue de distribution inférieure par

$$\lambda_L = \lim_{u \rightarrow 0^+} \frac{\mathbf{C}(u, u)}{u} \quad (3.33)$$

Si $\lambda_U > 0$, cela signifie qu'il y a une probabilité non nulle que l'une des variables dépasse u sachant que l'autre est déjà supérieure à u . **Dit autrement, λ_U est la probabilité d'avoir un extrême sachant qu'on a déjà un extrême.** Donnons quelques exemples de calcul de λ_U .

Exemple 7 Considérons la famille de Gumbel bivariée

$$\mathbf{C}_\theta(u, v) = \exp\left(-\left[(-\ln u)^\theta + (-\ln v)^\theta\right]^{\frac{1}{\theta}}\right) \quad (3.34)$$

pour $\theta \geq 1$. On a

$$\frac{\overline{\mathbf{C}}(u, u)}{1 - u} = \frac{1 - 2u + u^{2^{1/\theta}}}{1 - u} \quad (3.35)$$

la règle de l'Hopital donne

$$\begin{aligned} \lim_{u \rightarrow 1^-} \frac{\overline{\mathbf{C}}(u, u)}{1 - u} &= 2 - \lim_{u \rightarrow 1^-} 2^{1/\theta} u^{2^{1/\theta}-1} \\ &= 2 - 2^{1/\theta} \end{aligned} \quad (3.36)$$

Ainsi, la famille de Gumbel possède pour $\theta > 1$ une dépendance au niveau supérieur de la queue de distribution.

Parfois, cette formule n'est pas directement applicable. On utilise alors le fait que pour une paire de variables aléatoires (U, V) dont chaque composante suit une loi uniforme $[0, 1]$, on a

$$\begin{aligned} \Pr[V \leq v \mid U = u] &= \frac{\partial}{\partial v} \mathbf{C}(u, v) \\ \Pr[V > v \mid U = u] &= 1 - \frac{\partial}{\partial v} \mathbf{C}(u, v) \end{aligned} \quad (3.37)$$

d'où

$$\begin{aligned} \lambda_U &= \lim_{u \rightarrow 1^-} \frac{\overline{\mathbf{C}}(u, u)}{1 - u} \\ &= - \lim_{u \rightarrow 1^-} \frac{d\overline{\mathbf{C}}(u, u)}{du} \\ &= - \lim_{u \rightarrow 1^-} \left(-2 + \frac{\partial}{\partial s} \mathbf{C}(u, u) + \frac{\partial}{\partial t} \mathbf{C}(u, u) \right) \\ &= \lim_{u \rightarrow 1^-} (\Pr[V > u \mid U = u] + \Pr[U > u \mid V = u]) \end{aligned}$$

De plus si la copule est symétrique, c'est à dire si $\mathbf{C}(u, v) = \mathbf{C}(v, u)$, alors on obtient

$$\lambda_U = 2 \lim_{u \rightarrow 1^-} \Pr[V > u \mid U = u] \quad (3.38)$$

Exemple 8 On peut alors appliquer cette formule au cas gaussien. On a

$$\begin{aligned} \lim_{u \rightarrow 1^-} \Pr[V > u \mid U = u] &= \lim_{x \rightarrow \infty} \Pr[\Phi^{-1}(V) > x \mid \Phi^{-1}(U) = x] \\ &= \lim_{x \rightarrow \infty} \Pr[X > x \mid Y = x] \end{aligned} \quad (3.39)$$

Or il est bien connu que la loi de $Y \mid X = x \sim \mathcal{N}(\rho x, 1 - \rho^2)$. On en déduit que

$$\begin{aligned} \lambda_U &= 2 \lim_{x \rightarrow \infty} \bar{\Phi}\left((x - \rho x) / \sqrt{1 - \rho^2}\right) \\ &= 2 \lim_{x \rightarrow \infty} \bar{\Phi}\left(x \sqrt{1 - \rho} / \sqrt{1 + \rho}\right) \end{aligned} \quad (3.40)$$

Donc pour $\rho < 1$, le copula gaussien n'a pas de dépendance au niveau supérieur de la queue de distribution puisque $\lambda_U = 0$. On obtient un résultat analogue dans le cas d'un copula de Student (EMBRECHTS, MCNEIL et STRAUMANN [1999]) :

$$\lambda_U = 2\bar{t}_{\nu+1}\left(\sqrt{\nu+1}\sqrt{1-\rho}/\sqrt{1+\rho}\right) \quad (3.41)$$

Dans ce cas, λ_U est égal à 0 uniquement si $\rho = -1$. On remarque aussi que cette dépendance décroît avec le degré de liberté ν de la copule. A titre d'illustration, nous avons représenté sur le graphique (3.1) la fonction $\lambda_U(u) = \frac{\bar{\mathbf{C}}(u, u)}{1-u}$ et sa limite λ_U . Pour $\rho < 1$, on remarque donc que les extrêmes ne sont pas corrélés lorsque la structure de dépendance est gaussienne, alors qu'ils le sont lorsque la dépendance est une copule de Student. Ce résultat est très important pour la modélisation financière, et pose des problèmes d'agrégation des mesures de risques lorsque nous faisons l'hypothèse de normalité.

3.2 Méthodes d'estimation

L'exposé qui suit reprend l'exposé de BOUYÉ, NIKEGHBALI, RIBOULET et RONCALLI [2000].

3.2.1 Les copules empiriques

Les copules empiriques ont été introduites par DEHEUVELS [1979]. Soit $\chi = \{(x_1^t, \dots, x_N^t)\}_{t=1}^T$ un échantillon. Le copula empirique de la distribution est donné par

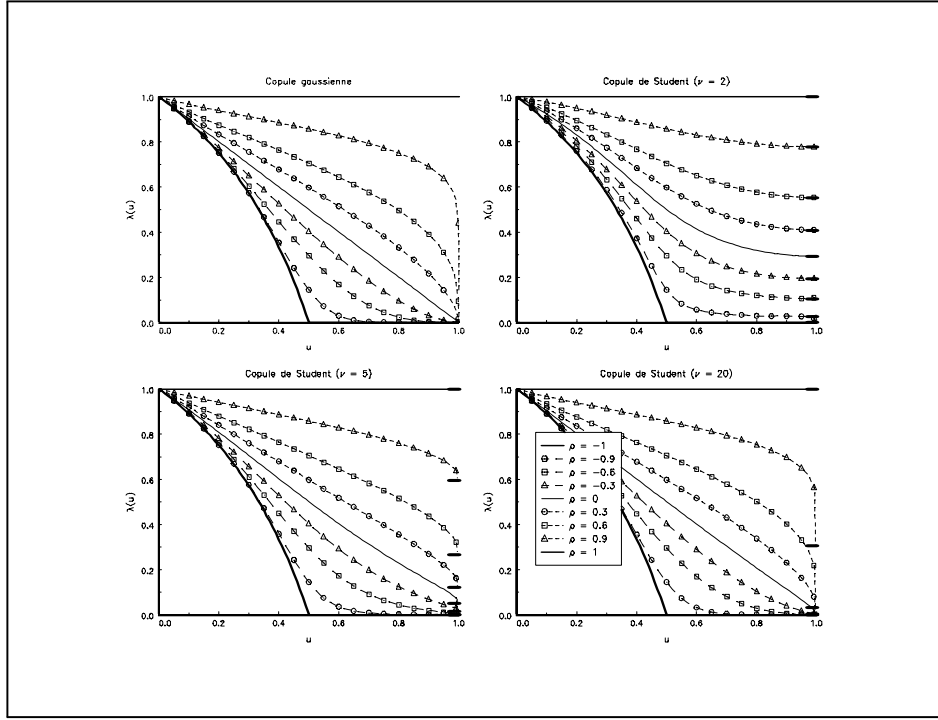
$$\hat{\mathbf{C}}\left(\frac{t_1}{T}, \dots, \frac{t_N}{T}\right) = \frac{1}{T} \sum_{t=1}^T \mathbf{1}_{[x_1^t \leq x_1^{(t_1)}, \dots, x_N^t \leq x_N^{(t_N)}]} \quad (3.42)$$

où $x_n^{(t)}$ représente la statistique d'ordre et $1 \leq t_1, \dots, t_N \leq T$. La fréquence du copula empirique est

$$\hat{\mathbf{c}}\left(\frac{t_1}{T}, \dots, \frac{t_N}{T}\right) = \begin{cases} \frac{1}{T} & \text{si } (x_1^{(t_1)}, \dots, x_N^{(t_N)}) \in \chi \\ 0 & \text{si non} \end{cases} \quad (3.43)$$

On a la relation suivante entre le copula empirique et sa fréquence

$$\hat{\mathbf{C}}\left(\frac{t_1}{T}, \dots, \frac{t_N}{T}\right) = \sum_{i_1=1}^{t_1} \dots \sum_{i_N=1}^{t_N} \hat{\mathbf{c}}\left(\frac{t_1}{T}, \dots, \frac{t_N}{T}\right) \quad (3.44)$$



Graphique 3.1. Comparaison des fonctions $\lambda_U(u)$

Les copules empiriques peuvent être utilisées pour estimer les mesures de dépendance. Par exemple, une estimation du rho de Spearman dans le cas bivarié est

$$\hat{\varrho}_S = \frac{12}{T^2 - 1} \sum_{t_1=1}^T \sum_{t_2=1}^T \left(\hat{C} \left(\frac{t_1}{T}, \frac{t_2}{T} \right) - \frac{t_1 t_2}{T^2} \right) \quad (3.45)$$

Dans le cas du copula gaussien, la relation $\rho = 2 \sin \left(\frac{\pi}{6} \hat{\varrho}_S \right)$ permet ainsi de définir une nouvelle estimation de la matrice des corrélations **sans faire aucune hypothèse sur les distributions marginales associées**.

3.2.2 Estimation par la méthode du maximum de vraisemblance

Soit θ le vecteur $K \times 1$ des paramètres que l'on doit estimer et Θ l'espace dans lequel θ prend ses valeurs. La vraisemblance pour l'observation t est notée $L_t(\theta)$. C'est une densité de probabilité. Notons $\ell_t(\theta)$ la log-vraisemblance, c'est-à-dire le logarithme de $L_t(\theta)$. Pour T observations, $\ell(\theta) = \sum_{t=1}^T \ell_t(\theta)$ représente la log-vraisemblance. $\hat{\theta}_{ML}$ est l'estimateur du maximum de vraisemblance si

$$\ell(\hat{\theta}_{ML}) \geq \ell(\theta) \quad \forall \theta \in \Theta \quad (3.46)$$

On peut montrer que $\hat{\theta}_{ML}$ est asymptotiquement normal et que

$$\sqrt{T}(\hat{\theta}_{ML} - \theta_0) \rightarrow \mathcal{N}(0, \mathcal{I}^{-1}(\theta_0)) \quad (3.47)$$

avec $\mathcal{I}(\theta_0)$ la matrice d'information de Fischer.

Or, la densité de $\mathbf{F}(x_1, \dots, x_N) = C(\mathbf{F}_1(x_1), \dots, \mathbf{F}_N(x_N))$ est

$$f(x_1, \dots, x_N) = c(\mathbf{F}_1(x_1), \dots, \mathbf{F}_N(x_N)) \prod_{n=1}^N f_n(x_n) \quad (3.48)$$

où f_n et $c(u_1, \dots, u_n) = \frac{\partial C}{\partial u_1 \dots \partial u_n}$ sont les densités respectives de $\mathbf{F}_n$ et de la copule. Ainsi, on a

$$\ell(\theta) = \sum_{t=1}^T \ln c(\mathbf{F}_1(x_1^t), \dots, \mathbf{F}_N(x_N^t)) + \sum_{t=1}^T \sum_{n=1}^N \ln f_n(x_n^t) \quad (3.49)$$

3.2.3 Estimation par la méthode IFM

La méthode précédente peut entraîner des temps de calcul très longs lorsqu'on a un grand nombre de dimension car elle nécessite le calcul de paramètres joints pour les marginales et la dépendance. Or la représentation en copule sépare les paramètres en paramètres spécifiques pour les distributions marginales et des paramètres communs pour la structure de dépendance. La log-vraisemblance peut alors s'écrire

$$\ell(\theta) = \sum_{t=1}^T \ln c(\mathbf{F}_1(x_1^t; \theta_1), \dots, \mathbf{F}_N(x_N^t; \theta_N); \alpha) + \sum_{t=1}^T \sum_{n=1}^N \ln f_n(x_n^t; \theta_n) \quad (3.50)$$

où $\theta = (\theta_1, \dots, \theta_N, \alpha)$ le vecteur des paramètres contient les paramètres θ_n de chaque marginale $\mathbf{F}_n$ et les paramètres α de la copule. JOE et XU [1996] proposent alors d'estimer θ en deux temps. Dans un premier temps, on estime les paramètres de chaque marginales

$$\hat{\theta}_n = \arg \max \sum_{t=1}^T \ln f_n(x_n^t; \theta_n) \quad (3.51)$$

Dans un second temps, on estime α en utilisant les estimations précédentes :

$$\hat{\alpha} = \arg \max \sum_{t=1}^T \ln c(\mathbf{F}_1(x_1^t; \hat{\theta}_1), \dots, \mathbf{F}_N(x_N^t; \hat{\theta}_N); \alpha) \quad (3.52)$$

Cette méthode en deux étapes est appelée *méthode des fonctions d'inférence des marginales* ou méthode IFM. En général, le paramètre $\hat{\theta}$ obtenu par cette méthode est différent de celui obtenu par maximum de vraisemblance. Dans une thèse non publiée, Xu estime que la méthode IFM est bien plus efficace que la méthode ML (JOE [1997]). En utilisant une idée proche de celle de la méthode IFM, nous remarquons que le paramètre α du copule peut être estimé sans spécifier les marginales (BOUYÉ, NIKEGBALI, RIBOULET et RONCALLI [2000]). La méthode consiste à transformer les données $(x_1^t, \dots, x_N^t)$ en variables uniformes $(\hat{u}_1^t, \dots, \hat{u}_N^t)$ et ensuite à estimer le paramètre de la manière suivante :

$$\hat{\alpha} = \arg \max \sum_{i=1}^T \ln c(\hat{u}_1^t, \dots, \hat{u}_N^t; \alpha) \quad (3.53)$$

Dans ce cas, $\hat{\alpha}$ peut être considéré comme l'estimateur du maximum de vraisemblance à partir des marginales observées — sans hypothèse sur la distribution des marginales.

3.3 Construction des familles de copules

3.3.1 Familles de distributions multivariées générées par des mélanges

Les articles sur les distributions multivariées s'intéressent particulièrement aux familles paramétriques de la forme

$$\mathbf{H}(x_1, x_2 | \theta) = \mathbf{C}(\mathbf{H}_1(x_1), \mathbf{H}_2(x_2) | \theta) \quad (3.54)$$

où θ est un paramètre ou un vecteur de paramètres. Nous allons ici exposer une méthode qui permet de construire de telles distributions. Cette méthode a été présentée pour la première fois par MARSHALL et OLKIN [1988] et très intensivement exploitée par JOE and HU [1996] pour construire des distributions d'extrêmes multidimensionnelles.

Soient $\mathbf{F}$ et $\mathbf{G}$ des fonctions de distribution univariée et supposons⁴ $\overline{\mathbf{G}}(0) = 1$. Comme $\mathbf{F}^\theta$ est encore une fonction de distribution pour tout $\theta > 0$, il s'ensuit que le **mélange**

$$\mathbf{H}(x) = \int \mathbf{F}^\theta(x) d\mathbf{G}(\theta) \quad (3.55)$$

est aussi une fonction de distribution. Ce genre de mélange est intéressant puisque si on se donne deux distributions $\mathbf{H}$ et $\mathbf{G}$ telles que $\overline{\mathbf{G}}(0) = 1$, il existe une fonction de distribution $\mathbf{F}$ telle que la relation (3.55) est vérifiée. Pour le voir, il suffit de prendre

$$\mathbf{F}(x) = \exp[-\Phi^{-1}(\mathbf{H}(x))] \quad (3.56)$$

où Φ est la transformée de Laplace⁵ de la mesure associée à la fonction de distribution $\mathbf{G}$.

Considérons maintenant une fonction de distribution bivariable $\mathbf{G}$ de marginales $\mathbf{G}_1$ et $\mathbf{G}_2$ et telle que $\overline{\mathbf{G}}(0, 0) = 1$. Les marginales $\mathbf{H}_1$ et $\mathbf{H}_2$ de la fonction de distribution $\mathbf{H}$ obtenue par le mélange

$$\mathbf{H}(x_1, x_2) = \int \int \mathbf{F}_1^{\theta_1}(x_1) \mathbf{F}_2^{\theta_2}(x_2) d\mathbf{G}(\theta_1, \theta_2) \quad (3.58)$$

sont données par

$$\mathbf{H}_i(x) = \int \mathbf{F}_i^{\theta_i}(x) d\mathbf{G}_i(\theta_i) \quad i \in \{1, 2\} \quad (3.59)$$

et sont de la forme (3.55). Ainsi, si $\mathbf{F}_i(x) = \exp[-\Phi_i^{-1}(\mathbf{H}_i(x))]$ où Φ_i est la transformée de Laplace de $\mathbf{G}_i$ pour $i \in \{1, 2\}$, alors $\mathbf{H}$ est une fonction de distribution bivariable qui a pour marginales $\mathbf{H}_1$ et $\mathbf{H}_2$. Les distributions bivariées de cette forme ont été étudiées par OAKES [1989].

La construction précédente se généralise sans difficulté pour des dimensions supérieures. De plus elle est aussi valable si on remplace $\mathbf{F}_1^{\theta_1}(x_1) \mathbf{F}_2^{\theta_2}(x_2)$ par n'importe quelle fonction de distribution ayant pour marginales $\mathbf{F}_1^{\theta_1}(x_1)$ et $\mathbf{F}_2^{\theta_2}(x_2)$ — on peut remarquer que $\mathbf{F}_1^{\theta_1}(x_1) \mathbf{F}_2^{\theta_2}(x_2)$ correspond au cas d'indépendance. Ces considérations nous conduisent à énoncer le théorème suivant :

Théorème 12 Soient $\mathbf{H}_1, \dots, \mathbf{H}_n$ des fonctions de distribution univariée et soit $\mathbf{G}$ une fonction de distribution n -dimensionnelle qui a pour marginales $(\mathbf{G}_i)_{i=1 \dots n}$ et qui est telle que $\overline{\mathbf{G}}(0, \dots, 0) = 1$. Notons Φ et Φ_i les transformées de Laplace de $\mathbf{G}$ et $\mathbf{G}_i$ respectivement. Soit $\mathbf{C}$ une n -copule. Si

$$\mathbf{F}_i(x) = \exp[-\Phi_i^{-1}(\mathbf{H}_i(x))] \quad \forall i \in \{1, \dots, n\} \quad (3.60)$$

alors

$$\mathbf{H}(x_1, \dots, x_n) = \int \dots \int \mathbf{C}(\mathbf{F}_1^{\theta_1}(x_1), \dots, \mathbf{F}_n^{\theta_n}(x_n)) d\mathbf{G}(\theta_1, \dots, \theta_n) \quad (3.61)$$

est une fonction de distribution n -variée qui a pour marginales $\mathbf{H}_1, \dots, \mathbf{H}_n$.

Le choix de $\mathbf{C}$ et $\mathbf{G}$ conduit à une grande variété de familles de distributions multivariées, ayant les marginales pour paramètres. On peut citer un corollaire intéressant de ce théorème.

Corollaire 2 Sous les hypothèses du théorème précédent, la fonction suivante

$$\mathbf{H}(x_1, \dots, x_n) = \Phi[\Phi_1^{-1}(\mathbf{H}_1(x_1)), \dots, \Phi_n^{-1}(\mathbf{H}_n(x_n))] \quad (3.62)$$

définit une distribution n -variée ayant pour marginales $\mathbf{H}_1, \dots, \mathbf{H}_n$.

⁴On rappelle que $\overline{G}(\theta) = 1 - G(\theta)$.

⁵On rappelle que la transformée de Laplace de $\mathbf{G}$ est définie par

$$E_\theta[e^{-s\theta}] = \int e^{-s\theta} d\mathbf{G}(\theta) \quad (3.57)$$

Démonstration 2 Voir MARSHALL et OLKIN [1988].

Lorsque $\mathbf{C}$ est le copula indépendant, on obtient

$$\mathbf{H}(x_1, \dots, x_n) = \int \cdots \int \prod_{i=1}^n \mathbf{F}_i^{\theta_i}(x_i) d\mathbf{G}_1(\theta) \quad (3.63)$$

$$= \Phi[\Phi^{-1}(\mathbf{H}_1(x_1)) + \dots + \Phi^{-1}(\mathbf{H}_n(x_n))] \quad (3.64)$$

Nous verrons en fait au paragraphe suivant que l'on peut, notamment en dimension 2, considérer des distributions de la forme (3.62) pour des fonctions Φ plus générales que les transformées de Laplace.

Remarque 6 On peut refaire point par point tous les raisonnements précédents pour les fonctions de distribution de survie. On obtient alors, sous les mêmes hypothèses que le théorème précédent, que

$$\bar{\mathbf{H}}(x_1, \dots, x_n) = \int \cdots \int \bar{\mathbf{C}}(\bar{\mathbf{F}}_1^{\theta_1}(x_1), \dots, \bar{\mathbf{F}}_n^{\theta_n}(x_n)) d\mathbf{G}(\theta_1, \dots, \theta_n) \quad (3.65)$$

définit une fonction de survie (en dimension n) qui a pour marginales $\mathbf{H}_1, \dots, \mathbf{H}_n$ avec

$$\bar{\mathbf{F}}_i(x) = \exp[-\Phi_i^{-1}(\bar{\mathbf{H}}_i(x))] \quad \forall i \in \{1, \dots, n\} \quad (3.66)$$

De la même manière, on a les deux cas particuliers

$$\bar{\mathbf{H}}(x_1, \dots, x_n) = \Phi[\Phi_1^{-1}(\bar{\mathbf{H}}_1(x_1)), \dots, \Phi_n^{-1}(\bar{\mathbf{H}}_n(x_n))] \quad (3.67)$$

et

$$\bar{\mathbf{H}}(x_1, \dots, x_n) = \Phi[\Phi^{-1}(\bar{\mathbf{H}}_1(x_1)) + \dots + \Phi^{-1}(\bar{\mathbf{H}}_n(x_n))] \quad (3.68)$$

Nous allons maintenant mettre en œuvre les méthodes précédentes pour construire des distributions multivariées.

Exemple 9 On considère ici le cas simple où $\mathbf{C}$ est le copula indépendant $\mathbf{C}^+$ et le cas où $\mathbf{G} = \mathbf{C}^+$, c'est-à-dire la borne supérieure de Fréchet.

1. **La famille de Gumbel.** On va utiliser (3.63). et (3.68). GENEST et MACKAY [1986] ont montré que

$$\Phi(z) = \exp[-z^{1/m}] \quad (3.69)$$

est la transformée de Laplace d'une distribution stable. On a alors

$$\Phi^{-1}(z) = (-\ln z)^m \quad (3.70)$$

D'après la formule (3.68), on peut définir

$$[-\ln \bar{\mathbf{H}}(x_1, x_2)]^m = [-\ln \bar{\mathbf{H}}_1(x_1)]^m + [-\ln \bar{\mathbf{H}}_2(x_2)]^m \quad (3.71)$$

soit encore

$$\bar{\mathbf{H}}(x_1, x_2) = \exp\left\{-\left([- \ln \bar{\mathbf{H}}_1(x_1)]^m + [- \ln \bar{\mathbf{H}}_2(x_2)]^m\right)^{1/m}\right\} \quad (3.72)$$

On voit aisément que le copula associé à cette distribution est

$$\mathbf{C}(u, v) = \exp\left\{-[(-\ln u)^m + (-\ln v)^m]^{1/m}\right\} \quad (3.73)$$

2. **La famille de Clayton, Cook et Johnson.** $\mathbf{G}$ a pour marginales des distribution Gamma dont la transformée de Laplace est

$$\Phi_{\alpha}(z) = (1+z)^{-\alpha} \quad \forall \alpha > 0 \quad (3.74)$$

d'où

$$\Phi_{\alpha}^{-1}(z) = z^{-1/\alpha} - 1 \quad (3.75)$$

Ainsi, $\overline{\mathbf{H}}^{(\alpha)}$ définie par (3.68) est donnée par

$$\begin{aligned} \overline{\mathbf{H}}^{(\alpha)}(x_1, x_2) &= \left\{ \overline{\mathbf{H}}_1^{-1/\alpha}(x_1) + \overline{\mathbf{H}}_2^{-1/\alpha}(x_2) - 1 \right\}^{-\alpha} \\ &= \frac{\overline{\mathbf{H}}_1(x_1) \overline{\mathbf{H}}_2(x_2)}{\left\{ 1 - \left[1 - \overline{\mathbf{H}}_1^{1/\alpha}(x_1) \right] \left[1 - \overline{\mathbf{H}}_2^{1/\alpha}(x_2) \right] \right\}^{\alpha}} \end{aligned} \quad (3.76)$$

On reconnaît facilement la copule correspondant à cette distribution

$$\mathbf{C}(u, v) = \frac{uv}{\left[1 - (1 - u^{1/\alpha})(1 - v^{1/\alpha}) \right]^{\alpha}} \quad (3.77)$$

3. **La famille de Frank.** Si on choisit (on peut montrer que c'est la transformée de Laplace d'une certaine distribution — voir JOE [1997])

$$\Phi(z) = \frac{\ln(1 - \overline{\alpha}e^{-z})}{\ln \alpha} \quad 0 < \alpha < 1, \overline{\alpha} = 1 - \alpha \quad (3.78)$$

on a alors

$$\Phi^{-1}(z) = -\frac{\ln(1 - \alpha^z)}{\overline{\alpha}} \quad (3.79)$$

Dans ce cas, l'équation (3.63) nous donne

$$\mathbf{H}(x_1, x_2) = \frac{\ln \left[1 + (\alpha^{\mathbf{H}_1(x_1)} - 1)(\alpha^{\mathbf{H}_2(x_2)} - 1) \right]}{\ln \alpha} \quad 0 < \alpha < 1 \quad (3.80)$$

On en déduit la copule de Frank

$$\mathbf{C}(u, v) = \frac{\ln \left[1 + (\alpha^u - 1)(\alpha^v - 1) \right]}{\ln \alpha} \quad (3.81)$$

Exemple 10 On considère ici que $\mathbf{C}$ est la copule de Morgenstern, c'est à dire

$$\mathbf{C}(u, v) = uv [1 + \beta(1 - u)(1 - v)] \quad -1 \leq \beta \leq 1$$

En utilisant la relation (3.61), on obtient

$$\begin{aligned} \mathbf{H}(x_1, x_2) &= \int \int \mathbf{F}_1^{\theta_1}(x_1) \mathbf{F}_2^{\theta_2}(x_2) \times \left[1 + \beta \left(1 - \mathbf{F}_1^{\theta_1}(x_1) \right) \left(1 - \mathbf{F}_2^{\theta_2}(x_2) \right) \right] d\mathbf{G}(\theta_1, \theta_2) \\ &= (1 + \beta) \Phi \left[\Phi_1^{-1}(\mathbf{H}_1(x_1)), \Phi_2^{-1}(\mathbf{H}_2(x_2)) \right] - \beta \Phi \left[\Phi_1^{-1}(\mathbf{H}_1(x_1)), 2\Phi_2^{-1}(\mathbf{H}_2(x_2)) \right] \\ &\quad - \beta \Phi \left[2\Phi_1^{-1}(\mathbf{H}_1(x_1)), \Phi_2^{-1}(\mathbf{H}_2(x_2)) \right] + \beta \Phi \left[2\Phi_1^{-1}(\mathbf{H}_1(x_1)), 2\Phi_2^{-1}(\mathbf{H}_2(x_2)) \right] \end{aligned} \quad (3.82)$$

Si $\mathbf{G}$ est la borne supérieure de Fréchet $\mathbf{C}^+$, on obtient

$$\begin{aligned} \mathbf{H}(x_1, x_2) &= (1 + \beta) \Phi \left[\Phi^{-1}(\mathbf{H}_1(x_1)) + \Phi^{-1}(\mathbf{H}_2(x_2)) \right] - \beta \Phi \left[\Phi^{-1}(\mathbf{H}_1(x_1)) + 2\Phi^{-1}(\mathbf{H}_2(x_2)) \right] \\ &\quad - \beta \Phi \left[2\Phi^{-1}(\mathbf{H}_1(x_1)) + \Phi^{-1}(\mathbf{H}_2(x_2)) \right] + \beta \Phi \left[2\Phi^{-1}(\mathbf{H}_1(x_1)) + 2\Phi^{-1}(\mathbf{H}_2(x_2)) \right] \end{aligned} \quad (3.83)$$

Si $\mathbf{G}$ est la copule indépendante $\mathbf{C}^\perp$, alors la distribution devient

$$\begin{aligned}\mathbf{H}(x_1, x_2) &= (1 + \beta) \mathbf{H}_1(x_1) \mathbf{H}_2(x_2) - \beta \mathbf{H}_2(x_2) \Phi_1[2\Phi_1^{-1}(\mathbf{H}_1(x_1))] \\ &\quad - \beta \mathbf{H}_1(x_1) \Phi_2[2\Phi_2^{-1}(\mathbf{H}_2(x_2))] \\ &\quad + \beta \Phi_1[2\Phi_1^{-1}(\mathbf{H}_1(x_1))] \Phi_2[2\Phi_2^{-1}(\mathbf{H}_2(x_2))]\end{aligned}\quad (3.84)$$

On va reprendre les transformées de Laplace des exemples précédents pour construire de nouvelles familles.

1. Prenons

$$\Phi_i(z) = \exp\left[-z^{1/m}\right] \quad m \geq 1, i = 1, 2 \quad (3.85)$$

Alors l'expression (3.83) devient

$$\begin{aligned}\mathbf{H}(x_1, x_2) &= (1 + \beta) \mathbf{H}_1(x_1) \mathbf{H}_2(x_2) - \beta \mathbf{H}_1(x_1) \mathbf{H}_2^\gamma(x_2) \\ &\quad - \beta \mathbf{H}_1^\gamma(x_1) \mathbf{H}_2(x_2) + \beta \mathbf{H}_1^\gamma(x_1) \mathbf{H}_2^\gamma(x_2)\end{aligned}\quad (3.86)$$

avec $\gamma = 2^{1/m} \in [0, 2]$. Le cas $\gamma = 2$ correspond à la distribution de Morgenstern.

2. On peut également choisir de nouveau

$$\Phi(z) = (1 + z)^{-\alpha} \quad \forall \alpha > 0 \quad (3.87)$$

Nous obtenons alors pour un copula indépendant

$$\begin{aligned}\mathbf{H}(x_1, x_2) &= (1 + \beta) \left[\overline{\mathbf{H}}_1^{-1/\alpha}(x_1) + \overline{\mathbf{H}}_2^{-1/\alpha}(x_2) - 1 \right]^{-\alpha} \\ &\quad - \beta \left[\overline{\mathbf{H}}_1^{-1/\alpha}(x_1) + 2\overline{\mathbf{H}}_2^{-1/\alpha}(x_2) - 2 \right]^{-\alpha} \\ &\quad - \beta \left[2\overline{\mathbf{H}}_1^{-1/\alpha}(x_1) + \overline{\mathbf{H}}_2^{-1/\alpha}(x_2) - 2 \right]^{-\alpha} \\ &\quad + \beta \left[2\overline{\mathbf{H}}_1^{-1/\alpha}(x_1) + 2\overline{\mathbf{H}}_2^{-1/\alpha}(x_2) - 3 \right]^{-\alpha}\end{aligned}\quad (3.88)$$

avec $-1 \leq \beta \leq 1$ et $\alpha > 0$.

3.3.2 Les copules archimédiennes

Nous allons à présent donner une autre méthode pour construire des distributions multivariées, qui ressemble à la construction donnée par l'expression (3.63).

3.3.2.1 Quelques généralités

Nous exposons, pour la simplicité de l'exposé, les copules archimédiennes dans le cas bivarié.

Définition 13 Soit ϕ une fonction continue, strictement décroissante de $[0, 1]$ dans $[0, \infty]$. La pseudo-inverse de ϕ est la fonction $\phi^{[-1]}$ dont le domaine de définition est $[0, \infty]$, à valeurs dans $[0, 1]$, et qui est définie par

$$\phi^{[-1]}(t) = \begin{cases} \phi^{-1}(t) & \text{si } 0 \leq t \leq \phi(0) \\ 0 & \text{si } \phi(0) \leq t \leq \infty \end{cases} \quad (3.89)$$

On va maintenant énoncer un théorème qui nous permet de définir les copula archimédiens.

Théorème 13 Soit ϕ une fonction continue, strictement décroissante de $[0, 1]$ dans $[0, \infty]$ et telle que $\phi(1) = 0$ et soit $\phi^{[-1]}$ le pseudo-inverse de ϕ . La fonction $\mathbf{C} : [0, 1] \times [0, 1] \rightarrow [0, 1]$ définie par

$$\mathbf{C}(u, v) = \phi^{[-1]}(\phi(u) + \phi(v)) \quad (3.90)$$

est une copule si et seulement si ϕ est **convexe**.

Démonstration 3 Voir NELSEN [1999].

Les copules définies par (3.90) sont appelées **copula archimédiens**. La fonction ϕ est le générateur de la copule. On voit ainsi que l'on peut construire une grande variété de familles de copules avec cette méthode.

Exemple 11 Nous considérons différentes formes de générateur.

1. Si on prend $\phi(t) = (-\ln t)^\theta$ avec $\theta \geq 1$, on vérifie facilement que cette fonction convient et on obtient la famille de Gumbel

$$\begin{aligned} \mathbf{C}_\theta(u, v) &= \phi^{[-1]}(\phi(u) + \phi(v)) \\ &= \exp \left\{ - \left[(-\ln u)^\theta + (-\ln v)^\theta \right]^{1/\theta} \right\} \end{aligned} \quad (3.91)$$

On avait déjà trouvé cette famille par la méthode de mélange. On remarque de plus que $\mathbf{C}_1 = \mathbf{C}^\perp$ et $\mathbf{C}_\infty = \mathbf{C}^+$.

2. Soit $\phi(t) = 1 - t$ pour $t \in [0, 1]$. On a pour ce générateur

$$\phi^{[-1]}(t) = \begin{cases} 1 - t & \text{pour } t \in [0, 1] \\ 0 & \text{pour } t > 1 \end{cases} \quad (3.92)$$

c'est-à-dire $\phi^{[-1]}(t) = \max(1 - t, 0)$. On obtient ainsi

$$\begin{aligned} \mathbf{C}(u, v) &= \max(u + v - 1, 0) \\ &= \mathbf{C}^+(u, v) \end{aligned} \quad (3.93)$$

3. Considérons $\phi(t) = \theta^{-1}(t^{-\theta} - 1)$ avec $\theta \in [-1, \infty) \setminus \{0\}$. Ce générateur conduit à la famille de copules

$$\mathbf{C}_\theta(u, v) = \max \left[(u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}, 0 \right] \quad (3.94)$$

Lorsque $\theta > 0$, on obtient la famille des copules de Clayton

$$\mathbf{C}_\theta(u, v) = (u^{-\theta} + v^{-\theta} - 1)^{-1/\theta} \quad (3.95)$$

3.3.2.2 Propriétés des copules archimédiennes

On notera Ω l'ensemble des fonctions ϕ convexes strictement décroissantes de $[0, 1]$ dans $[0, 1]$ et qui vérifient $\phi(1) = 0$.

Théorème 14 Soient X et Y des variables aléatoires dont la copule $\mathbf{C}$ est archimédienne avec le générateur $\phi \in \Omega$. Alors le coefficient de corrélation de Kendall de X et Y est donné par

$$\tau = 1 + 4 \int_0^1 \frac{\phi(t)}{\phi'(t)} dt \quad (3.96)$$

Démonstration 4 Voir NELSEN [1999].

Exemple 12 Pour la copule de Gumbel, on a $\phi(t) = (-\ln t)^\theta$ avec $\theta \geq 1$. Il vient que

$$\frac{\phi(t)}{\phi'(t)} = \frac{t \ln t}{\theta} \quad (3.97)$$

d'où

$$\tau_\theta = 1 - \frac{1}{\theta} \quad (3.98)$$

Exemple 13 On considère le générateur de la famille de Clayton $\phi(t) = \theta^{-1}(t^{-\theta} - 1)$ pour $\theta \in [-1, \infty) \setminus \{0\}$. On a alors

$$\frac{\phi(t)}{\phi'(t)} = \frac{t^{\theta+1} - t}{\theta} \quad (3.99)$$

d'où

$$\tau_\theta = \frac{\theta}{\theta + 2} \quad (3.100)$$

Remarque 7 Il existe également un résultat spécifique aux copules archimédiennes qui permet de calculer l'indice de dépendance supérieure de la queue de distribution en fonction du générateur du copula.

Théorème 15 Soit $\mathbf{C}$ une copule archimédienne tel que $\phi^{[-1]}(t) = \phi^{-1}(t)$. Si $(\phi^{-1})'(0)$ est finie, alors $\mathbf{C}(u, v) = \phi^{-1}(\phi(u) + \phi(v))$ n'a pas de dépendance supérieure au niveau de la queue de distribution. Si $\mathbf{C}$ présente une dépendance supérieure au niveau de la queue de distribution, alors $(\phi^{-1})'(0) = -\infty$ et de plus on a la relation

$$\lambda_U = 2 - 2 \lim_{s \rightarrow 0} \frac{(\phi^{-1})'(2s)}{(\phi^{-1})'(s)} \quad (3.101)$$

Ce résultat va donc nous permettre de calculer assez facilement les indices de queue de distribution.

Exemple 14 Considérons de nouveau la famille de Gumbel. Rappelons que $\phi(t) = (-\ln t)^\theta$ avec $\theta \geq 1$. On a bien $\phi^{[-1]}(t) = \phi^{-1}(t)$. On peut donc appliquer le théorème précédent. On obtient

$$\begin{aligned} \lambda_U &= 2 - 2^{1/\theta} \lim_{s \rightarrow 0} \left[\frac{\exp(-(2s)^{1/\theta})}{\exp(-(s)^{1/\theta})} \right] \\ &= 2 - 2^{1/\theta} \end{aligned} \quad (3.102)$$

On peut aussi citer un résultat similaire pour les indices de queue inférieurs.

Théorème 16 Soit $\mathbf{C}$ une copule archimédienne telle que $\phi^{[-1]}(t) = \phi^{-1}(t)$. Le coefficient de dépendance inférieure de la queue de distribution de $\mathbf{C}(u, v) = \phi^{-1}(\phi(u) + \phi(v))$ est donné par

$$\lambda_L = 2 \lim_{s \rightarrow 0} \frac{(\phi^{-1})'(2s)}{(\phi^{-1})'(s)} \quad (3.103)$$

Exemple 15 Considérons la famille de Clayton $\mathbf{C}_\theta(u, v) = (u^{-\theta} + v^{-\theta} - 1)^{-1/\theta}$ avec $\theta \geq 0$. On a pour cette famille $\phi(t) = \theta^{-1}(t^{-\theta} - 1)$. Il vient donc $\phi^{-1}(t) = (1 + \theta t)^{-1/\theta}$ d'où le résultat suivant

$$\begin{aligned} \lambda_L &= 2 \lim_{s \rightarrow 0} \left[\frac{-(1 + 2\theta s)^{-1/\theta - 1}}{-(1 + \theta s)^{-1/\theta - 1}} \right] \\ &= 2^{-1/\theta} \end{aligned} \quad (3.104)$$

3.3.2.3 Identification d'une copule archimédienne

Le but de cette sous-section est de savoir comment ajuster un copula archimédien à une série de données. GENEST et RIVEST [1993] ont développé une méthode pour identifier un copula archimédien. Soit (X, Y) un vecteur aléatoire et $\mathbf{C}$ le copula archimédien de générateur ϕ associé. Soit K la fonction définie par

$$K(u) = \Pr(\mathbf{C}(X, Y) \leq u) \quad (3.105)$$

On peut montrer (voir BARBE, GENEST, GHOUDI et RÉMILLARD [1996]) que

$$K(u) = u - \frac{\phi(u)}{\phi'(u)} \quad (3.106)$$

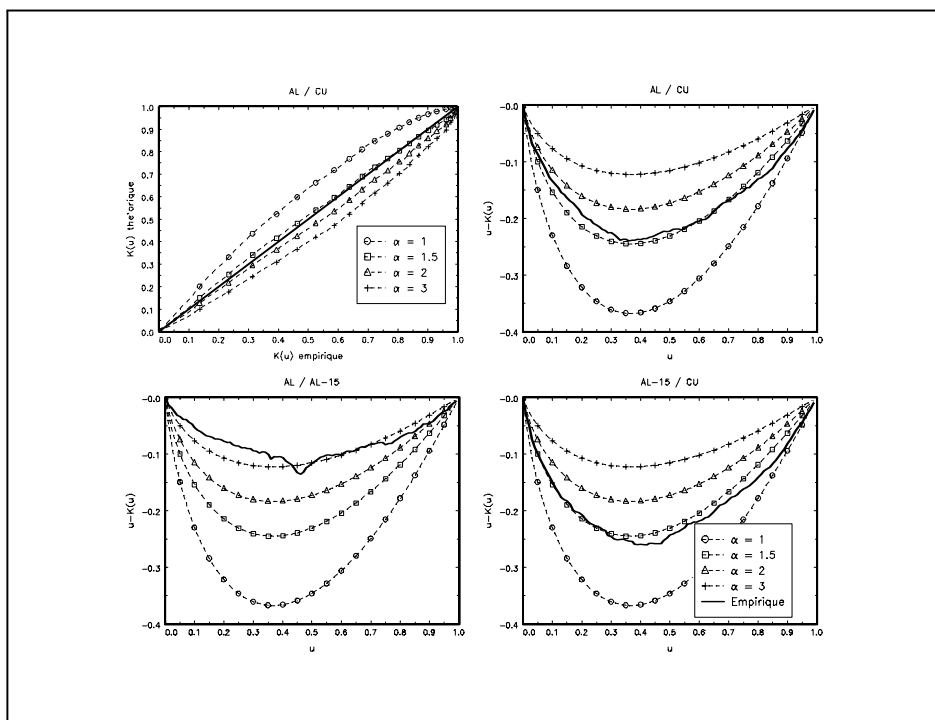
Une estimation non paramétrique de K est donnée par

$$\hat{K}(u) = \frac{1}{T} \sum_{t=1}^T \mathbf{1}_{[\mathcal{V}_i \leq u]} \quad (3.107)$$

où

$$\mathcal{V}_i = \frac{1}{T-1} \sum_{t=1}^T \mathbf{1}_{[X^t \leq X^i, Y^t \leq Y^i]} \quad (3.108)$$

L'idée est alors de choisir une famille de copula archimédiens qui s'ajuste bien à $\hat{K}$. Pour cela, on peut considérer un QQ-plot de K et $\hat{K}$ (FREES et VALDEZ [1997]). Une procédure graphique plus judicieuse est celle suggérée par GENEST et RIVEST [1993], qui consiste à comparer $u - K(u)$ et $u - \hat{K}(u)$. Sur le graphique (3.2), on voit ainsi que le paramètre $\alpha = 1.5$ de la copule de Franck convient relativement bien pour les couples AL/CU et AL-15/CU, alors que ce paramètre est beaucoup plus élevé pour AL/AL-15. Notons qu'il est possible de construire des tests beaucoup plus précis basés sur les distances entre les deux variables aléatoires $U - K(U)$ et $U - \hat{K}(U)$.



Graphique 3.2. Procédure d'identification des copules archimédiennes

3.3.3 Les copules extrêmes

3.3.3.1 Résultats généraux

Dans beaucoup d'applications de la statistique, l'étude des valeurs extrêmes est d'une très grande importance. Nous avons déjà vu dans le chapitre 2 l'étude des lois d'extrêmes univariées. L'étude du cas multivarié est cependant plus délicate, mais l'utilisation des copules la rend plus aisée.

Remarque 8 *La théorie des extrêmes multivariées est très technique. Pour la clarté de l'exposé, on ne traitera que le cas bivarié. Le lecteur pourra se reporter à Joe [1997] pour une généralisation au cas multivarié ainsi que pour les preuves.*

Considérons un vecteur aléatoire (X, Y) de fonction de distribution $\mathbf{H}$ et de marginales $\mathbf{H}_1$ et $\mathbf{H}_2$ continues. On sait d'après le théorème de Sklar qu'il existe une unique fonction copule telle qu'on vérifie

$$\mathbf{H}(x, y) = \mathbf{C}(\mathbf{H}_1(x), \mathbf{H}_2(y)) \quad (3.109)$$

Considérons une suite de vecteurs aléatoires $(X_1, Y_1), \dots, (X_n, Y_n)$ de vecteurs indépendants de lois $\mathbf{H}(x, y)$. Soit

$$(X_n^*, Y_n^*) = (\max(X_1, \dots, X_n), \max(Y_1, \dots, Y_n))$$

Définition 14 *On dit que $\mathbf{H}$ appartient au max-domaine d'attraction d'une distribution des valeurs extrêmes $\mathbf{H}^*$ avec des lois marginales non dégénérées s'il existe des suites $a_{1n} > 0$, $a_{2n} > 0$, b_{1n} et b_{2n} telles que*

$$\begin{aligned} \lim_{n \rightarrow \infty} \Pr \left[\frac{X_n^* - b_{1n}}{a_{1n}} \leq x, \frac{Y_n^* - b_{2n}}{a_{2n}} \leq y \right] &= \lim_{n \rightarrow \infty} \mathbf{H}^n(a_{1n}x + b_{1n}, a_{2n}y + b_{2n}) \\ &= \mathbf{H}^*(x, y) \end{aligned} \quad (3.110)$$

Théorème 17 *Soit $\mathbf{C}^*$ la copule correspondant à $\mathbf{H}^*$ et soient $\mathbf{H}_1^*$ et $\mathbf{H}_2^*$ les marginales de $\mathbf{H}^*$. $\mathbf{H}$ appartient au max-domaine d'attraction de $\mathbf{H}^*$ si et seulement si $\mathbf{H}_1$ et $\mathbf{H}_2$ appartiennent aux domaines d'attraction de $\mathbf{H}_1^*$ et $\mathbf{H}_2^*$ respectivement et si $\mathbf{C}$ satisfait*

$$\mathbf{C}^*(x, y) = \lim_{n \rightarrow \infty} \mathbf{C}(x^{1/n}, y^{1/n})^n \quad \forall 0 \leq x, y \leq 1 \quad (3.111)$$

D'après la théorie des extrêmes univariée exposée au chapitre 2, les marginales $\mathbf{H}_1^*$ et $\mathbf{H}_2^*$ appartiennent à l'un des trois types cités au chapitre 2. JOE [1997] énonce une caractérisation des copules extrêmes :

Théorème 18 *$\mathbf{C}(x, y)$ est une copule extrême si et seulement si*

$$\mathbf{C}(x^t, y^t) = \mathbf{C}^t(x, y) \quad (3.112)$$

Un autre résultat intéressant montré par BOUYÉ, NIKEGBALI, RIBOULET et RONCALLI [2000] est le suivant :

Proposition 1 *Pour toute copule extrême $\mathbf{C}$, on a*

$$\mathbf{C}^\perp \prec \mathbf{C} \prec \mathbf{C}^+ \quad (3.113)$$

Cela implique que les extrêmes ne présentent jamais de dépendance négative !

Maintenant nous allons énoncer un résultat qui peut se généraliser pour des dimensions supérieures, mais avec des contraintes supplémentaires.

Théorème 19 *Soit $\mathbf{C}^*$ une copule extrême bivariee. Alors il existe une fonction convexe A définie sur $[0, 1]$ et qui satisfait la condition $\max(t, 1-t) \leq A(t) \leq 1$ telle que*

$$\mathbf{C}^*(x, y) = \exp \left[\ln(xy) A \left(\frac{\ln(y)}{\ln(xy)} \right) \right] \quad \forall 0 \leq x, y \leq 1 \quad (3.114)$$

A est appelée la fonction de dépendance associée à $\mathbf{C}^*$.

Ce résultat nous enseigne que l'estimation d'un copula extrême peut se faire en deux étapes : tout d'abord on estime les marginales (avec les méthodes exposées au second chapitre) puis on estime A avec les méthodes paramétriques ou non paramétriques de la section précédente. Avant de donner des exemples de copules extrêmes, nous allons énoncer quelques résultats qui servent lors des estimations non paramétriques. Nous nous inspirons ici de ABDOUS, GHOUDI et KHOUDRAJI [1999]. Le lecteur peut consulter l'article pour les démonstrations.

Théorème 20 *Soit $\mathbf{C}^*$ un copula extrême bivariee dont la fonction de dépendance est A . Le copula $\mathbf{C}$ appartient au max-domaine d'attraction de $\mathbf{C}^*$ si et seulement si*

$$\lim_{u \rightarrow 0} \frac{1 - \mathbf{C} \left((1-u)^{1-t}, (1-u)^t \right)}{u} = A(t) \quad \forall t \in [0, 1] \quad (3.115)$$

L'idée à présent est de ramener l'estimation de A à celle d'une quantité dépendant du copula de départ $\mathbf{C}$. Pour cela, considérons un vecteur aléatoire (U, V) de marginales des lois uniformes sur $(0, 1)$ et de copule $\mathbf{C}$. Pour tout $t \in [0, 1]$, on pose

$$W_t = \begin{cases} 1 - \max(U^{1/(1-t)}, V^{1/t}) & \text{pour } t \in (0, 1) \\ 1 - U & \text{pour } t = 0 \\ 1 - V & \text{pour } t = 1 \end{cases} \quad (3.116)$$

On vérifie alors facilement que la fonction de distribution $\mathbf{G}_t$ de cette variable aléatoire satisfait

$$\mathbf{G}_t(u) = 1 - \mathbf{C} \left((1-u)^{1-t}, (1-u)^t \right) \quad (3.117)$$

On en déduit donc que $\mathbf{C}$ appartient au max-domaine d'attraction de $\mathbf{C}^*$ si et seulement si

$$\lim_{u \rightarrow 0} \frac{\mathbf{G}_t(u)}{u} = A(t) \quad \forall t \in [0, 1] \quad (3.118)$$

On voit sous cette forme que la fonction de dépendance $A(t)$ peut être vue comme la dérivée à droite en zéro de la fonction de distribution $\mathbf{G}_t$. Pour estimer $A(t)$, il suffit alors d'étudier le comportement au voisinage de zéro de $\mathbf{G}_t$.

Nous allons à présent donner quelques exemples de copules qui serviront pour l'estimation paramétrique.

Exemple 16 *Nous indiquerons pour chaque exemple de copule la fonction de dépendance de celle-ci. On posera $\tilde{u} = -\ln u$.*

1. Le copula indépendant $\mathbf{C}^\perp$

$$\begin{aligned} \mathbf{C}^\perp(u_1, u_2) &= u_1 u_2 \\ A(w) &= 1 \end{aligned} \quad (3.119)$$

2. Le copula de Gumbel

$$\begin{aligned} \mathbf{C}^G(u_1, u_2) &= \exp \left[-(\tilde{u}_1^\alpha + \tilde{u}_2^\alpha)^{1/\alpha} \right] \quad \forall \alpha \in [1, \infty) \\ A(w) &= [w^\alpha + (1-w)^\alpha]^{1/\alpha} \end{aligned} \quad (3.120)$$

3. Un autre copula de Gumbel

$$\begin{aligned} \mathbf{C}(u_1, u_2) &= u_1 u_2 \exp \left[\alpha \frac{\tilde{u}_1 \tilde{u}_2}{\tilde{u}_1 + \tilde{u}_2} \right] & \alpha \in [0, 1] \\ A(w) &= \alpha w^2 - \alpha w + 1 \end{aligned} \quad (3.121)$$

4. Le copula de Galambos

$$\begin{aligned} \mathbf{C}(u_1, u_2) &= u_1 u_2 \exp \left[(\tilde{u}_1^{-\alpha} + \tilde{u}_2^{-\alpha})^{-1/\alpha} \right] & \forall \alpha \in [0, \infty) \\ A(w) &= 1 - \left[w^{-\alpha} + (1-w)^{-\alpha} \right]^{-1/\alpha} \end{aligned} \quad (3.122)$$

5. Le copula de Marshall-Olkin

$$\begin{aligned} \mathbf{C}(u_1, u_2) &= u_1^{1-\alpha_1} u_2^{1-\alpha_2} \min(u_1^{\alpha_1}, u_2^{\alpha_2}) & (\alpha_1, \alpha_2) \in [0, 1]^2 \\ A(w) &= \max(1 - \alpha_1 w, 1 - \alpha_2 (1-w)) \end{aligned} \quad (3.123)$$

6. La borne supérieure de Fréchet

$$\begin{aligned} \mathbf{C}^+ &= \min(u_1, u_2) \\ A(w) &= \max(w, 1-w) \end{aligned}$$

On remarque que toutes ces copules, sauf celle de Marshall-Olkin, sont symétriques⁶. Cela signifie que les variables aléatoires considérées jouent un rôle symétrique. Cette hypothèse est bien sûr discutable. GENEST, GHOUDI et RIVEST [1998] proposent une technique d'*asymétrisation* pour construire de nouveaux copules plus flexibles pour la modélisation. Si A_1 et A_2 sont deux fonctions de dépendance et $(p_1, p_2) \in [0, 1]^2$, la formule suivante définit une nouvelle fonction de dépendance

$$\begin{aligned} A(w) &= (p_1 w + p_2 (1-w)) A_1 \left(\frac{p_1 w}{p_1 w + p_2 (1-w)} \right) + \\ &((1-p_1) w + (1-p_2) (1-w)) A_2 \left(\frac{(1-p_1) w}{(1-p_1) w + (1-p_2) (1-w)} \right) \end{aligned} \quad (3.124)$$

En particulier, lorsque $p_1 = p_2 = p$, on obtient la fonction de dépendance

$$A(w) = p A_1(w) + (1-p) A_2(w) \quad (3.125)$$

Lorsque A_1 et A_2 sont les fonctions de dépendance de Gumbel et du copula indépendant, on a

$$A(w) = [p_1^\alpha w^\alpha + p_2^\alpha (1-w)^\alpha]^{\frac{1}{\alpha}} + (p_2 - p_1) w + (1-p_2) \quad (3.126)$$

et donc

$$\mathbf{C}(u_1, u_2) = \mathbf{C}^G(u_1^{p_1}, u_2^{p_2}) \mathbf{C}^\perp(u_1^{1-p_1}, u_2^{1-p_2}) \quad (3.127)$$

Cette relation peut d'ailleurs se généraliser. Par exemple, la copule de Marshall-Olkin est le produit de la copule supérieure de Fréchet et du copula indépendant, c'est-à-dire

$$\mathbf{C}(u_1, u_2) = \mathbf{C}^+(u_1^{\alpha_1}, u_2^{\alpha_2}) \mathbf{C}^\perp(u_1^{1-\alpha_1}, u_2^{1-\alpha_2}) \quad (3.128)$$

⁶ C'est-à-dire $\mathbf{C}(u, v) = \mathbf{C}(v, u)$

3.3.3.2 Méthode d'estimation paramétrique

On applique la méthode de maximum de vraisemblance ou la méthode IFM au vecteur des maxima (X_n^*, Y_n^*) en utilisant la relation

$$\mathbf{H}^*(x, y) = \mathbf{C}^*(\mathbf{G}_1(x), \mathbf{G}_2(y))$$

où $\mathbf{C}^*$ est un copule extrême et où $\mathbf{G}_i$ est une distribution **GEV**. On rappelle qu'une fonction de distribution est une distribution **GEV** si

$$\mathbf{G}(x) = \exp \left\{ - \left[1 + \xi \left(\frac{x - \mu}{\sigma} \right) \right]^{\frac{-1}{\xi}} \right\} \quad (3.129)$$

Notons qu'une étape très importante pour l'estimation paramétrique du copula est la valeur de départ qu'on donne aux paramètres. GHOUDI, KHOUDRAJI et RIVEST [1998] suggèrent d'utiliser les relations qui peuvent exister entre le tau de Kendall et les paramètres à estimer. En effet il existe pour les copules extrêmes des relations entre le tau de Kendall ou le rho de Spearman et le paramètre du copula.

Proposition 2 Soit $\mathbf{C}^*$ une copule extrême de fonction de dépendance A . Alors on a les relations suivantes :

$$\begin{aligned} \tau &= 4 \int w(1-w) d \ln A(w) \\ \varrho &= 12 \int \frac{1}{[A(w) + 1]^2} dw - 3 \end{aligned} \quad (3.130)$$

A titre d'exemple, on a

Famille de Copule	τ
$\mathbf{C}^\perp$	0
Gumbel	$1 - 1/\alpha$
Gumbel II	$8\alpha^{-\frac{1}{2}}(4 - \alpha)^{\frac{-1}{2}} \arctan \sqrt{\alpha(4 - \alpha)^{-1}} - 2$
Galambos	très compliqué
Marshall-Olkin	$\alpha_1 \alpha_2 (\alpha_1 - \alpha_1 \alpha_2 + \alpha_2)^{-1}$
$\mathbf{C}^+$	1

3.3.3.3 Estimation non paramétrique

On va exposer ici une méthode d'estimation non paramétrique de A . Considérons un échantillon $(X_1, Y_1), \dots, (X_n, Y_n)$ de n vecteurs aléatoires de même fonction de distribution $\mathbf{H}$, de copula associé $\mathbf{C}$ et de marginales $\mathbf{H}_1$ et $\mathbf{H}_2$. Soit $\mathbf{H}_n$ la distribution jointe empirique et $\mathbf{H}_{1n}$, $\mathbf{H}_{2n}$ les distributions marginales empiriques. Notons $\mathbf{C}_n$ le copula empirique donné par

$$\mathbf{C}_n(u, v) = \mathbf{H}_n(\mathbf{H}_{1n}^{-1}(u), \mathbf{H}_{2n}^{-1}(v)) \quad (3.131)$$

Notons également

$$\overline{\mathbf{C}}_n(u, v) = \mathbf{H}_n(\mathbf{H}_1^{-1}(u), \mathbf{H}_2^{-1}(v)) \quad (3.132)$$

On a vu plus haut que le copula $\mathbf{C}$ est dans le max-domaine d'attraction du copula extrême $\mathbf{C}^*$ si et seulement si la fonction $\mathbf{G}_t(u) = 1 - \mathbf{C}\left((1-u)^{1-t}, (1-u)^t\right)$ a un comportement linéaire au voisinage de zéro. Ainsi un estimateur naturel de $A(t)$ peut être obtenu en ajustant localement un polynôme à l'estimation empirique de $\mathbf{G}_t$. Dans la littérature, on sépare l'étude en deux cas : le cas où l'on connaît les marginales $\mathbf{H}_1$ et $\mathbf{H}_2$ et le cas où on ne les connaît pas. Comme on sait bien estimer les marginales lorsqu'elles sont extrêmes, on ne s'intéresse qu'à ce cas-là.

Supposons donc que l'on connaît les marginales $\mathbf{H}_1$ et $\mathbf{H}_2$. GHOUDI, KHOUDRAJI et RIVEST [1998] proposent pour chaque $(t, v) \in [0, 1]^2$ d'utiliser une méthode des moindres carrés pondérés pour ajuster localement un polynôme linéaire à la fonction

$$\overline{\mathbf{G}}_{t,n}(v) = 1 - \overline{\mathbf{C}}_n \left((1-v)^{1-t}, (1-v)^t \right) \quad (3.133)$$

A cette fin, on minimise par rapport à la variable a l'expression suivante :

$$\int_0^1 \mathcal{K}_h(u-v) [\overline{\mathbf{G}}_{t,n}(u) - au]^2 du \quad (3.134)$$

où h est une variable qui nous permet de quantifier le voisinage de zéro, $\mathcal{K}$ un noyau et $\mathcal{K}_h(u) = \mathcal{K}(u/h)/h$. La solution de ce problème donnée par

$$\hat{A}(t, v) = \frac{\int_0^1 u \mathcal{K}_h(u-v) \overline{\mathbf{G}}_{t,n}(u) du}{\int_0^1 u^2 \mathcal{K}_h(u-v) du} \quad (3.135)$$

fournit une estimation du comportement de $\mathbf{G}_t$ au voisinage de zéro. Une estimation triviale de $A(t)$ est $\hat{A}(t) = \hat{A}(t, 0)$. GHOUDI, KHOUDRAJI et RIVEST [1998] fournissent des résultats sur le comportement asymptotique de cet estimateur⁷.

3.3.4 Les copules archimax

Le but de cette section est de présenter une nouvelle famille de copules qui a été introduite par CAPÉRAA, FOUGÈRES et GENEST [2000] et qui englobe la plupart des familles connues des copules, notamment toutes les copules archimédiennes et toutes les copules extrêmes. Cette nouvelle famille offre plus de flexibilité pour la modélisation. En effet, on peut pour cette famille connaître a priori les différents max-domaines d'attraction. On pourra ainsi construire une famille de copules dont on se donne à l'avance le max-domaine d'attraction.

Définition 15 Une fonction bivariée est un copula archimax si et seulement si elle est de la forme

$$\mathbf{C}_{\phi,A}(x, y) = \phi^{-1} \left[(\phi(x) + \phi(y)) A \left(\frac{\phi(x)}{\phi(x) + \phi(y)} \right) \right] \quad (3.136)$$

pour tout $0 \leq x, y \leq 1$, avec

1. $A : [0, 1] \rightarrow [1/2, 1]$ tel que $\max(t, 1-t) \leq A(t) \leq 1$ pour tout $0 \leq t \leq 1$;
2. $\phi : (0, 1] \rightarrow [0, \infty)$, est une fonction convexe, décroissante qui vérifie $\phi(1) = 0$. On adoptera la convention $\phi(0) = \lim_{t \rightarrow 0^+} \phi(t)$ et $\phi^{-1}(s) = 0$ pour $s \geq \phi(0)$.

Ainsi, si on remplace x et y par des fonctions de distribution univariées $\mathbf{F}$ et $\mathbf{G}$, on obtient une fonction de distribution bivariée de marginales $\mathbf{F}$ et $\mathbf{G}$. On voit aisément que cette nouvelle famille de copula contient tous les copula archimédiens et tous les copula extrêmes. En effet, en choisissant $\phi(t) = \ln(1/t)$, on obtient

$$\mathbf{C}_{\phi,A}(x, y) = \mathbf{C}_A(x, y) = \exp \left[\ln(xy) A \left(\frac{\ln(y)}{\ln(xy)} \right) \right] \quad (3.137)$$

pour tout $0 \leq x, y \leq 1$. On reconnaît la forme générale des copules extrêmes bivariés. Si on choisit maintenant de prendre $A(t) \equiv 1$, on retrouve la forme générale des copules archimédiennes :

$$\mathbf{C}_{\phi,A}(x, y) = \mathbf{C}_\phi(x, y) = \phi^{-1}(\phi(x) + \phi(y)) \quad (3.138)$$

L'intérêt de la construction (3.136) réside dans le fait qu'avec un choix judicieux de ϕ et A , on obtient une copule qui sera dans le max-domaine d'attraction de n'importe quel "attracteur" $\mathbf{C}_{A^*}$ donné. Avant de donner un sens mathématique à cela, rappelons la définition de la variation régulière au voisinage de l'infini.

⁷Néanmoins, on remarque qu'en général la fonction estimée n'est pas convexe et ne vérifie pas la condition $\max(t, 1-t) \leq A(t) \leq 1$.

Définition 16 Une fonction $\Psi : (0, \infty) \rightarrow (0, \infty)$ est dite à variation régulière de degré ρ au voisinage de l'infini, et on note $\Psi \in \text{RV}_\rho$ si

$$\forall s > 0 \quad \lim_{t \rightarrow \infty} \frac{\Psi(st)}{\Psi(t)} = s^\rho \quad (3.139)$$

On peut maintenant énoncer le principal résultat de ce paragraphe :

Proposition 3 Soit $\mathbf{C}_{\phi,A}$ un copula archimax tel que $\phi(1 - 1/t) \in \text{RV}_{-m}$ pour un certain $m \geq 1$. Alors $\mathbf{C}_{\phi,A}$ appartient au max-domaine d'attraction du copula extrême $\mathbf{C}_{A^*}$ où A^* est donnée par

$$A^*(t) = [t^m + (1-t)^m]^{\frac{1}{m}} A^{\frac{1}{m}} \left(\frac{t^m}{t^m + (1-t)^m} \right) \quad 0 \leq t \leq 1 \quad (3.140)$$

Cet attracteur peut être considéré de deux points de vue. Le premier point de vue consiste à le considérer comme une copule archimax de générateur A^* et $\phi(t) = \ln(1/t)$, alors que l'autre point de vue le considère comme une copule archimax de générateurs A et $\phi^*(t) = \ln^m(1/t)$. De plus, A^* et A coïncident si et seulement si $m = 1$.

Démonstration 5 Voir CAPÉRAÀ, FOGÈRES et GENEST [2000].

Remarque 9 Il est alors très facile de montrer que la seule copule archimédienne et extrême est celle de Gumbel.

En faisant un raisonnement en sens inverse, on remarque que $\mathbf{C}_{\phi,A}$ appartient au max-domaine d'attraction de $\mathbf{C}_{A^*}$ si $\phi(1 - 1/t) \in \text{RV}_{-m}$ et

$$A(t) = \left[t^{1/m} + (1-t)^{1/m} \right]^m \left\{ A^* \left(\frac{t^{1/m}}{t^{1/m} + (1-t)^{1/m}} \right) \right\}^m \quad 0 \leq t \leq 1 \quad (3.141)$$

Maintenant il faut savoir si la relation (3.141) définit bien une fonction de dépendance. En posant $D(t) = \frac{d \ln(A^*(t))}{dt}$, $\mathbf{H}^*(t) = t + (1-t)D(t)$ et $h^*(t) = \frac{d\mathbf{H}^*(t)}{dt}$, CAPÉRAÀ, FOGÈRES et GENEST [2000] montrent que la relation (3.141) est vraie si et seulement si

$$\frac{h^*(t)}{\mathbf{H}^*(t) \{1 - \mathbf{H}^*(t)\}} \geq \frac{m}{t(1-t)} \quad \forall 0 \leq t \leq 1 \quad (3.142)$$

On peut au passage remarquer que cette relation reste valable pour tout $1 \leq m' \leq m$. Ainsi, si $A(t)$ est définie par (3.141), $\mathbf{C}_{\phi,A}$ appartient au max-domaine d'attraction de $\mathbf{C}_{A^*}$ pourvue que $\phi(1 - 1/t) \in \text{RV}_{-m'}$ pour $1 \leq m' \leq m$. Enfin, on peut aussi remarquer que pour des générateurs A et ϕ quelconques tels que $\phi(1 - 1/t) \in \text{RV}_{-m}$, on a

$$\lambda_U = \lim_{u \rightarrow 1^-} \frac{1 - 2u + \mathbf{C}_{\phi,A}(u, u)}{1 - u} = 2 - \{2A(1/2)\}^{1/m} \leq 1 \quad (3.143)$$

Ainsi, les copules archimax présentent une dépendance au niveau supérieur de la queue de distribution sauf si $m = 1$ et $A \equiv 1$.

On peut donner des exemples de copules archimax. On peut par exemple considérer le générateur de TAWN [1988] qui est défini pour $0 \leq \theta \leq 1$ par

$$A_\theta(t) = \theta t^2 - \theta t + 1 \quad \forall 0 \leq t \leq 1 \quad (3.144)$$

et le combiner au générateur de la famille de Clayton

$$\phi_{1,\alpha}(t) = \frac{t^{-\alpha} - 1}{\alpha} \quad \alpha > 0 \quad (3.145)$$

ou avec le générateur des copules de GENEST et GHOUDI [1994]

$$\phi_{2,\alpha}(t) = (1 - t^\alpha)^{1/\alpha} \quad 0 < \alpha \leq 1 \quad (3.146)$$

On peut remarquer que $\phi_{2,\alpha}(t) \in RV_{-1/\alpha}$. Ainsi, pour $\alpha \neq 1$, $\mathbf{C}_{\phi_{2,\alpha},A}(x,y)$ n'est pas dans le domaine d'attraction de $\mathbf{C}_A$. Enfin supposons que l'on veuille construire une copule dans le domaine de max-attraction de

$$A_{\alpha,\beta}^*(t) = \left[\alpha^l t^l + \beta^l (1-t)^l \right]^{\frac{1}{l}} + \left[(1-\alpha)^l t^l + (1-\beta)^l (1-t)^l \right]^{\frac{1}{l}} \quad \forall 0 \leq t \leq 1 \quad (3.147)$$

et où $0 \leq \alpha, \beta \leq 1$ et $l \geq 1$. Pour les valeurs $\alpha = 1/10$, $\beta = 3/10$, $m = 3/2$ et $l = 1.6$, la condition (3.142) est vérifiée. La copule originale A est donc celle définie par l'expression (3.141).

3.4 Applications financières

Dans un premier temps, nous généralisons le concept de scénario de crise au cas bi-dimensionnel. Dans un second temps, nous considérons l'évaluation d'un stress scénario multidimensionnel. Enfin, on présente une application originale des copules à la valeur en risque.

3.4.1 Modélisation bivariée des scénarios de crise

On considère la base de données LME que nous avons introduite dans le second chapitre. On suppose que la dépendance est représentée par une copule de Gumbel⁸. En utilisant la méthode d'estimation IFM, on obtient une estimation du paramètre α ainsi qu'une estimation du tau de Kendall. On donne également le ratio de vraisemblance LR qui correspond à l'hypothèse nulle de l'indépendance. On note χ^- et χ^+ respectivement les minimums et maximums par blocs des rendements⁹. Pour l'opposé des minima $(-\chi_1^-, -\chi_2^-)$, on trouve les résultats suivants

$\hat{\alpha}_{\text{IFM}}$	AL-15	CU	LR	AL-15	CU	τ	AL-15	CU
AL	2.397	1.159	AL	62.09	3.003	AL	0.583	0.137
AL-15		1.092	AL-15		1.011	AL-15		0.084

Pour les maxima (χ_1^+, χ_2^+) , on trouve

$\hat{\alpha}_{\text{IFM}}$	AL-15	CU	LR	AL-15	CU	τ	AL-15	CU
AL	3.015	1.065	AL	86.39	0.485	AL	0.668	0.061
AL-15		1.040	AL-15		0.160	AL-15		0.039

Pour un niveau de confiance de 99%, on ne rejette pas l'indépendance pour les couples de maxima (AL, CU) et (AL-15, CU). Pour les couples de minima, la dépendance n'est pas rejetée pour le couple (AL, AL-15).

Remarque 10 Dans une étude plus détaillée, BOUYÉ, NIKEGBALI, RIBOULET et RONCALLI [2000] constatent qu'il y a une différence entre le cas des minima et celui des maxima : "Les minima apparaissent plus corrélés que les maxima. D'un point de vue économique, cela signifie que la chute des prix présente une plus grande corrélation entre les produits que la hausse de ceux-ci."

Les scénarios de crise en bivarié peuvent être vus comme la représentation d'une zone de défaillance (voir DE HAAN, PENG, SINHA et DRAISMA [1997]). On a

$$\begin{aligned} \Pr(\chi_1^+ > \chi_1, \chi_2^+ > \chi_2) &= 1 - \Pr(\chi_1^+ \leq \chi_1) - \Pr(\chi_2^+ \leq \chi_2) + \Pr(\chi_1^+ \leq \chi_1, \chi_2^+ \leq \chi_2) \\ &= 1 - \mathbf{F}_1(\chi_1) - \mathbf{F}_2(\chi_2) + \mathbf{C}(\mathbf{F}_1(\chi_1), \mathbf{F}_2(\chi_2)) \\ &= \overline{\mathbf{C}}(\mathbf{F}_1(\chi_1), \mathbf{F}_2(\chi_2)) \end{aligned} \quad (3.149)$$

⁸ La log-vraisemblance individuelle a pour expression

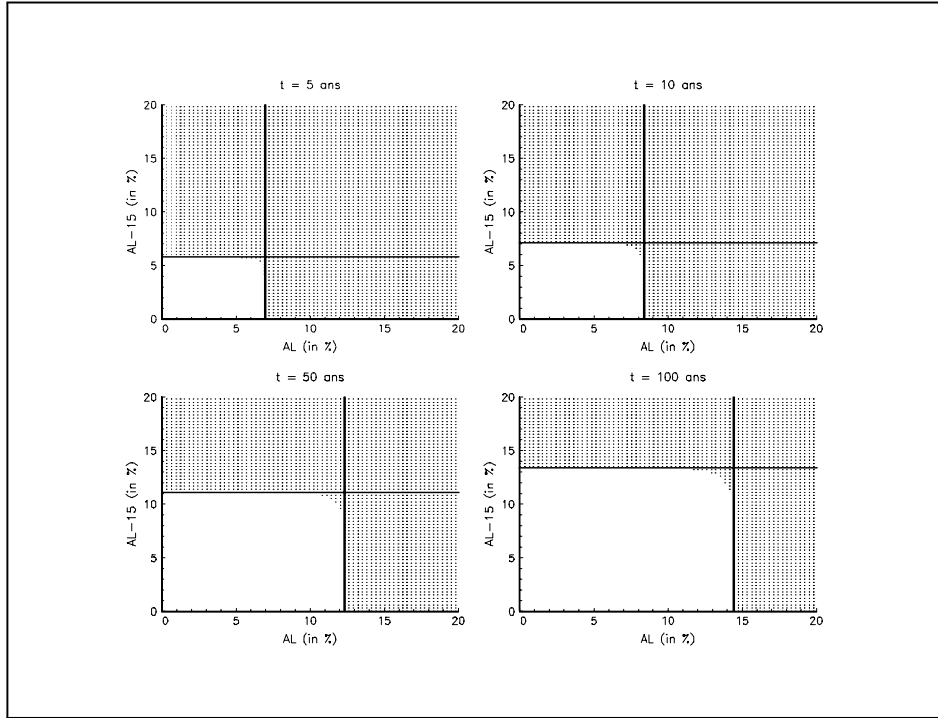
$$\begin{aligned} \ell(u_1, u_2; \alpha) &= -(\tilde{u}_1^\alpha + \tilde{u}_2^\alpha)^{\frac{1}{\alpha}} + \tilde{u}_1 + \tilde{u}_2 + (\alpha - 1) \ln(\tilde{u}_1 \tilde{u}_2) + (\alpha^{-1} - 2) \ln(\tilde{u}_1^\alpha + \tilde{u}_2^\alpha) \\ &\quad + \ln \left[(\tilde{u}_1^\alpha + \tilde{u}_2^\alpha)^{\frac{1}{\alpha}} + \alpha - 1 \right] \end{aligned} \quad (3.148)$$

⁹ Nous avons pris comme période de référence pour la longueur d'un bloc 44 jours de trading.

Soit t le temps de retour. La zone de défaillance est représentée par

$$\left\{ (\chi_1, \chi_2) \in \mathbb{R}^2 \mid \overline{\mathbf{C}}(\mathbf{F}_1(\chi_1), \mathbf{F}_2(\chi_2)) < \frac{1}{t} \right\} \quad (3.150)$$

On a représenté les zones de défaillance sur les graphiques (3.3) à (3.8). Les lignes verticales et horizontales en gras représentent les stress de scénarios unidimensionnels. La zone de défaillance correspond à l'ensemble grisé. Contrairement au cas unidimensionnel, on n'obtient pas un seul scénario de crise mais un ensemble. On rappelle les valeurs des stress unidimensionnels obtenues à la page 23. Par exemple, on obtient 6.96% et 7.53% pour une position courte en AL et CU et pour un temps de retour de 5 ans. Construire un scénario de crise bidimensionnel à partir des deux scénarios unidimensionnels est biaisé, puisque dans ce cas, le temps de retour est beaucoup plus élevé. Même des valeurs de 5% et 6% surestiment le temps de retour. Néanmoins, ce biais disparaît si la dépendance est parfaite, ou importante (voir par exemple le cas des portefeuilles AL/AL-15). **Ceci pose un problème, puisque généralement, on considère des facteurs de risque indépendants, et non parfaitement corrélés.**



Graphique 3.3. Scénarios de crise pour un portefeuille court en AL et AL-15

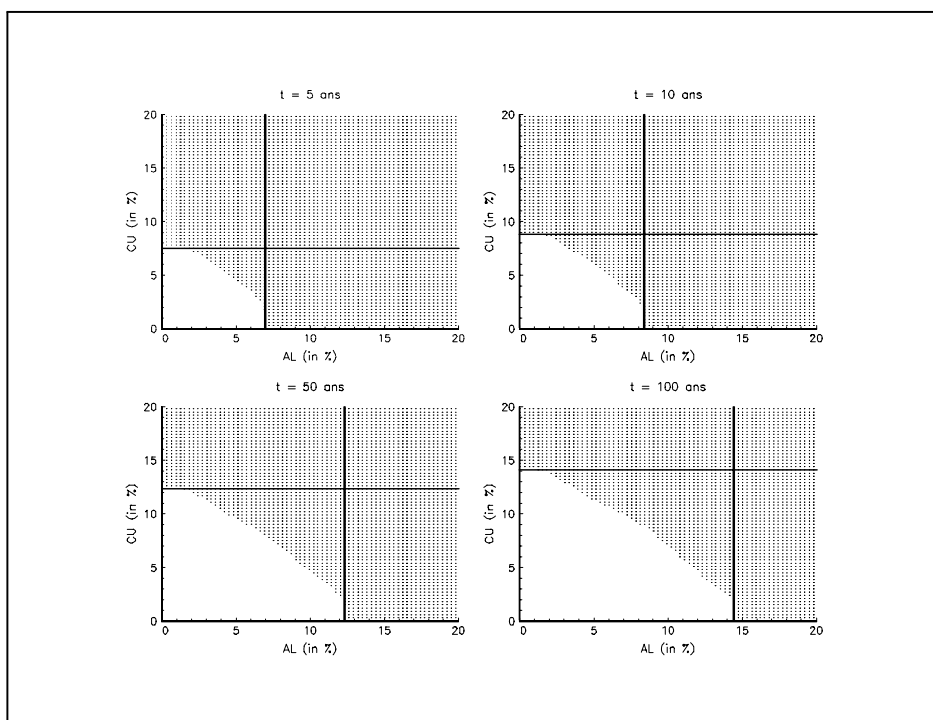
3.4.2 Quantification d'un scénario de crise multidimensionnel

On a introduit dans la sous-section précédente la notion de zone de défaillance. Pour le cas multivarié, celle-ci est définie de manière analogue (BOUYÉ, NIKEGHBALI, RIBOULET et RONCALLI [2000]) :

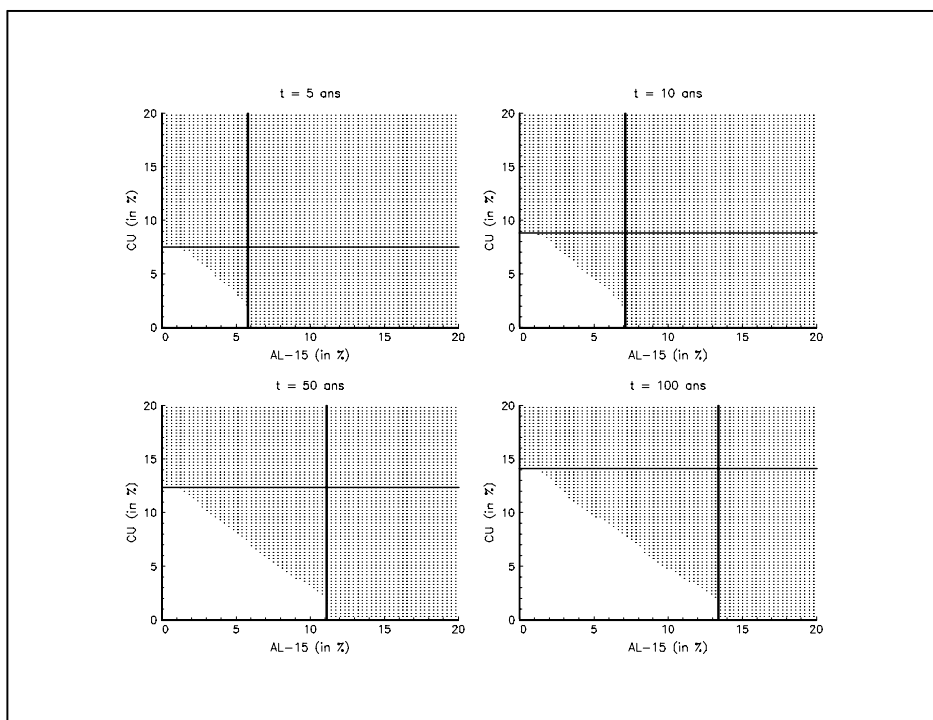
$$\left\{ (\chi_1, \dots, \chi_N) \in \mathbb{R}^N \mid u_1 = \mathbf{F}_1(\chi_1), \dots, u_N = \mathbf{F}_N(\chi_N), \overline{\mathbf{C}}(u_1, \dots, u_N) < \frac{1}{t} \right\} \quad (3.151)$$

avec

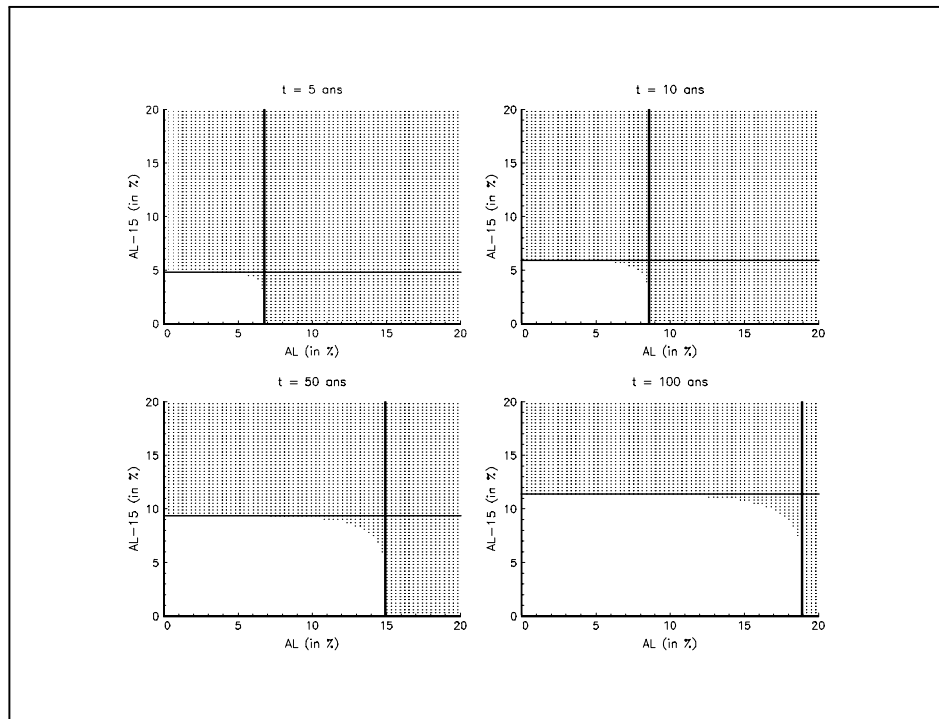
$$\overline{\mathbf{C}}(u_1, \dots, u_N) = \sum_{n=0}^N \left[(-1)^n \sum_{\mathbf{u} \in \mathcal{Z}(N-n, N)} \mathbf{C}(\mathbf{u}) \right] \quad (3.152)$$



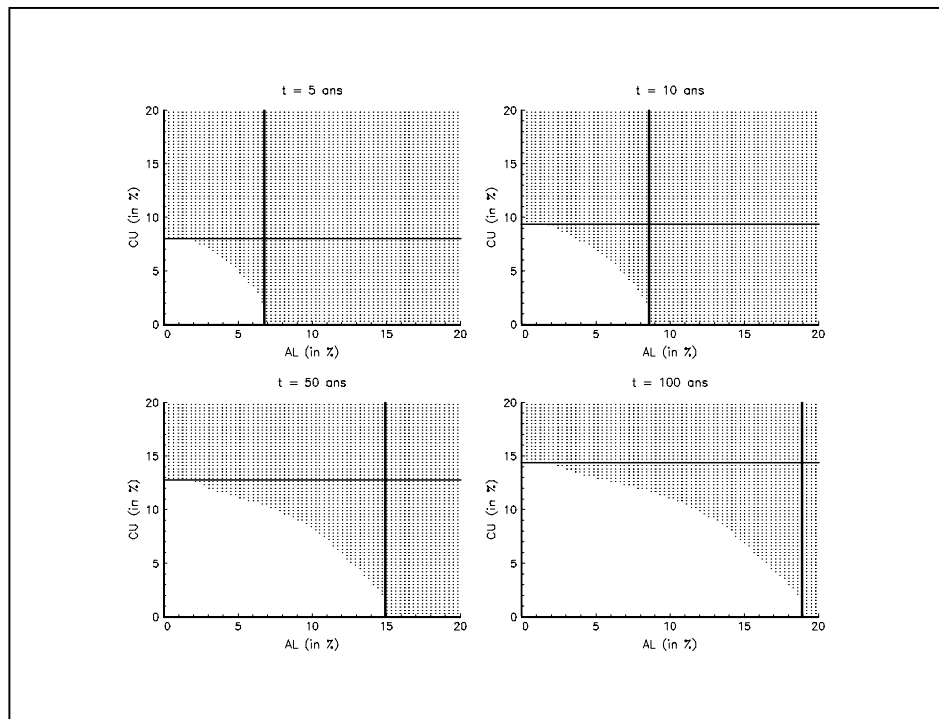
Graphique 3.4. Scénarios de crise pour un portefeuille court en AL et CU



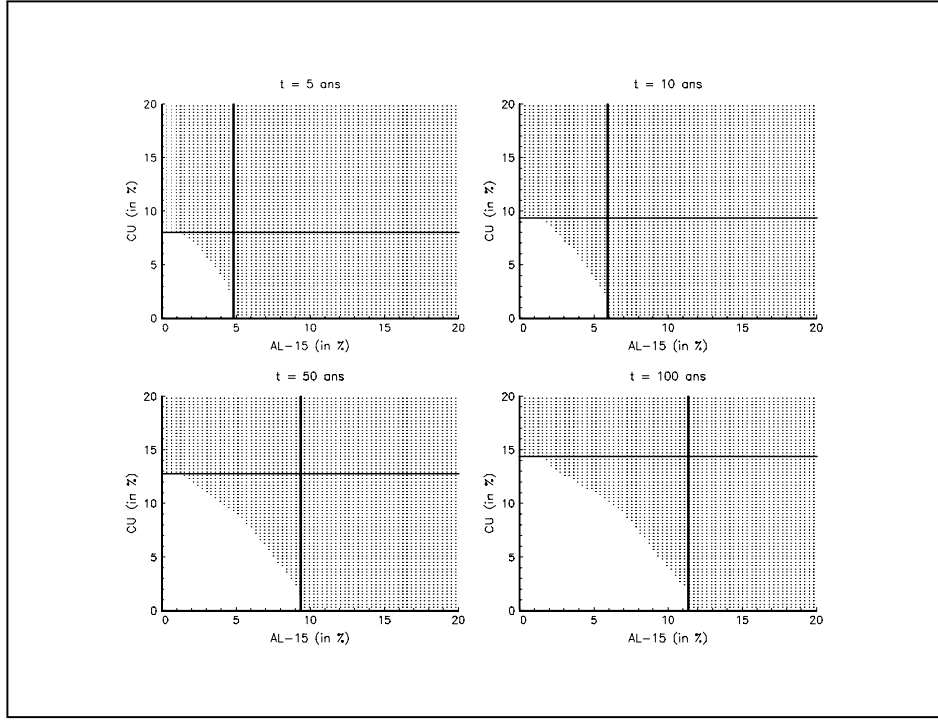
Graphique 3.5. Scénarios de crise pour un portefeuille court en AL-15 et CU



Graphique 3.6. Scénarios de crise pour un portefeuille long en AL et AL-15



Graphique 3.7. Scénarios de crise pour un portefeuille long en AL et CU



Graphique 3.8. Scénarios de crise pour un portefeuille long en AL-15 et CU

où $\mathcal{Z}(M, N)$ représente l'ensemble $\left\{ \mathbf{u} \in [0, 1]^N \mid \sum_{n=1}^N \chi_{\{1\}}(u_n) = M \right\}$. En dimension 2, on a

$$\begin{aligned} \overline{\mathbf{C}}(u_1, u_2) &= \mathbf{C}(1, 1) - \mathbf{C}(u_1, 1) - \mathbf{C}(1, u_2) + \mathbf{C}(u_1, u_2) \\ &= 1 - u_1 - u_2 + \mathbf{C}(u_1, u_2) \end{aligned} \quad (3.153)$$

En dimension 3, on obtient

$$\begin{aligned} \overline{\mathbf{C}}(u_1, u_2, u_3) &= \mathbf{C}(1, 1, 1) - \mathbf{C}(u_1, 1, 1) - \mathbf{C}(1, u_2, 1) - \mathbf{C}(1, 1, u_3) \\ &\quad + \mathbf{C}(u_1, u_2, 1) + \mathbf{C}(u_1, 1, u_3) + \mathbf{C}(1, u_2, u_3) - \mathbf{C}(u_1, u_2, u_3) \\ &= 1 - u_1 - u_2 - u_3 + \mathbf{C}(u_1, u_2) + \mathbf{C}(u_1, u_3) + \mathbf{C}(u_2, u_3) - \mathbf{C}(u_1, u_2, u_3) \end{aligned} \quad (3.154)$$

Remarque 11 *La construction de stress scénarios multidimensionnels est très difficile car la zone de défaillance fournit un ensemble dans l'espace $\mathbb{R}^N$. Se pose alors le problème du choix d'un scénario particulier parmi une infinité de scénarios possibles. C'est pourquoi, en général, les scénarios de crises multidimensionnels dans une banque sont fournis par les économistes. L'intérêt des copules est alors de quantifier les scénarios économiques et de comprendre leur cohérence (c'est-à-dire de déceler les incohérences du scénario). Ceci se fait alors par le biais du temps de retour implicite associé.*

Définition 17 *Le temps de retour implicite pour un vecteur $(\chi_1, \dots, \chi_N)$ est donné par la formule*

$$\mathbf{t}(\chi_1, \dots, \chi_N) = \overline{\mathbf{C}}^{-1}(\mathbf{F}_1(\chi_1), \dots, \mathbf{F}_N(\chi_N)) \quad (3.155)$$

Dans ce qui suit, on utilise pour modéliser la structure de dépendance de nos données LME la copule¹⁰

$$\mathbf{C}^{\text{GT}} = \exp \left[- \left((\tilde{u}_1^{\alpha_2} + \tilde{u}_2^{\alpha_2})^{\frac{\alpha_1}{\alpha_2}} + \tilde{u}_3^{\alpha_1} \right)^{\frac{1}{\alpha_1}} \right] \quad (3.158)$$

On considère le scénario suivant pour un portefeuille court dans les trois actifs :

$$\chi^{(1)} = \begin{bmatrix} +5\% \\ +5\% \\ +5\% \end{bmatrix}$$

Le temps de retour associé à notre copule $\mathbf{C}^{\text{GT}}$ est alors 20 ans. Si on suppose que les copules¹¹ sont $\mathbf{C}^\perp$ ou $\mathbf{C}^+$, on trouve respectivement des temps de retour de 170 ans et 3 ans. Ces résultats sont bien cohérents puisqu'il est normal qu'on trouve un temps de retour plus grand dans le cas indépendant et qu'au contraire on trouve un temps de retour beaucoup plus petit si on désire modéliser le cas de la parfaite dépendance.

On propose maintenant un second scénario de crise :

$$\chi^{(2)} = \begin{bmatrix} +4\% \\ +5.5\% \\ +5\% \end{bmatrix}$$

Comment savoir si ce second scénario est “plus fort” que le premier ? Si on calcule les temps de retour pour les trois copules $\mathbf{C}^{\text{GT}}$, $\mathbf{C}^\perp$ et $\mathbf{C}^+$, on trouve respectivement comme temps de retour implicite $\mathbf{t}_{(2)}^{\text{GT}} = 25$ ans, $\mathbf{t}_{(2)}^\perp = 118$ ans et $\mathbf{t}_{(2)}^+ = 4$ ans. On remarque qu'on a les relations suivantes $\mathbf{t}_{(2)}^{\text{GT}} > \mathbf{t}_{(1)}^{\text{GT}}$ et $\mathbf{t}_{(2)}^+ > \mathbf{t}_{(1)}^+$ mais $\mathbf{t}_{(2)}^\perp < \mathbf{t}_{(1)}^\perp$. Dans le cas où $\chi_n^{(1)} < \chi_n^{(2)}$, pour tout $n = 1, \dots, N$, on peut vérifier que $\mathbf{t}_{(2)} > \mathbf{t}_{(1)}$ pour toute copule. Mais lorsque toutes ces inégalités ne sont pas vérifiées, la comparaison est moins immédiate. Le concept de temps de retour apparaît donc très utile pour comparer l'intensité des scénarios de crise.

Remarque 12 *On a toujours nécessairement*

$$\mathbf{t}^+ \leq \mathbf{t}^C \leq \mathbf{t}^\perp \quad (3.159)$$

On peut essayer maintenant d'établir des liens entre les scénarios de crise univariés et les scénarios de crise multivariés. Si on choisit le cas univarié correspondant à un temps de retour de 5 ans, on a le scénario suivant :

$$\chi^{(3)} = \begin{bmatrix} +6,96\% \\ +5,79\% \\ +7,53\% \end{bmatrix}$$

¹⁰Pour suggérer un tel copule, on tient compte des résultats des tests de dépendance obtenus dans le cas bivarié. L'expression de la log-vraisemblance individuelle est alors

$$\begin{aligned} \ell(u_1, u_2, u_3; \alpha_1, \alpha_2) &= - \left((\tilde{u}_1^{\alpha_2} + \tilde{u}_2^{\alpha_2})^{\frac{\alpha_1}{\alpha_2}} + \tilde{u}_3^{\alpha_1} \right)^{\frac{1}{\alpha_1}} + \tilde{u}_1 + \tilde{u}_2 + \tilde{u}_3 + (\alpha_2 - 1) \ln(\tilde{u}_1 \tilde{u}_2) \\ &\quad + (\alpha_1 - 1) \ln(\tilde{u}_3) + (\alpha_1 \alpha_2^{-1} - 2) \ln A + (\alpha_1^{-1} - 2) \ln B + \ln \varkappa(u_1, u_2, u_3; \alpha_1, \alpha_2) \end{aligned} \quad (3.156)$$

with

$$\begin{aligned} \varkappa(u_1, u_2, u_3; \alpha_1, \alpha_2) &= A^{\frac{\alpha_1}{\alpha_2} B^{\frac{2}{\alpha_1} - 1}} + (\alpha_2 - \alpha_1) B^{\frac{1}{\alpha_1}} + (\alpha_1 - 1) A^{\frac{\alpha_1}{\alpha_2} B^{\frac{1}{\alpha_1} - 1}} \\ &\quad + (\alpha_1 - 1) (2\alpha_1 - 1) A^{\frac{\alpha_1}{\alpha_2} B^{-1}} + (\alpha_1 - 1) (\alpha_2 - \alpha_1) \\ A &= \tilde{u}_1^{\alpha_2} + \tilde{u}_2^{\alpha_2} \\ B &= A^{\frac{\alpha_1}{\alpha_2}} + \tilde{u}_3^{\alpha_1} \end{aligned} \quad (3.157)$$

¹¹Dans le cas des extrêmes, $\mathbf{C}^\perp$ est la borne basse de la structure de dépendance. Il n'est donc pas possible d'employer le copule $\mathbf{C}^-$.

Dans le cas multivarié, on obtient des temps de retour implicite : $\mathbf{t}_{(3)}^{\text{GT}} = 182$ ans, $t_{(3)}^{\perp} = 4038$ ans et $\mathbf{t}_{(3)}^+ = 5$ ans. Comme le font remarquer BOUYÉ, NIKEGBALI, RIBOULET et RONCALLI [2000], “il faut faire très attention si on veut construire des scénarios de crise multivariés à partir de ceux obtenus en univariés. Dans le cas où on suppose une parfaite dépendance entre les différents actifs du portefeuille, cela ne pose pas de problème ; dans le cas contraire, on risque d’obtenir des scénarios qui n’ont pas beaucoup de sens et qui ne sont pas intéressants pour la gestion des risques.”

L’exemple précédent suggère l’idée d’un **indicateur de la contribution de la structure de dépendance sur les scénarios de crise**. BOUYÉ, NIKEGBALI, RIBOULET et RONCALLI [2000] suggèrent la mesure suivante :

$$\pi = \frac{\mathbf{t}^{\perp}(\chi_1, \dots, \chi_N) - \mathbf{t}(\chi_1, \dots, \chi_N)}{\mathbf{t}^{\perp}(\chi_1, \dots, \chi_N) - \mathbf{t}^+(\chi_1, \dots, \chi_N)} \quad (3.160)$$

On a $\pi \in [0, 1]$ Cet indicateur est égal à un si il y a parfaite dépendance et il est égal à zéro dans le cas de l’indépendance. Pour nos trois exemples ci-dessus, on a : $\pi_{(1)} = 89\%$, $\pi_{(2)} = 81\%$, $\pi_{(3)} = 95\%$. On voit que la structure de dépendance a moins d’impact sur le second scénario.

On peut également s’interroger sur l’impact d’un actif en particulier dans un scénario de crise multivarié. On va tenter de réaliser cela à partir du scénario $\chi^{(1)}$. Cela s’apparente à de l’analyse de sensibilité. Rappelons que $\mathbf{t}^{\text{GT}}(+5\%, +5\%, 5\%) = 25$ ans. Sur la figure (3.9), on a représenté les $\mathbf{t}^{\text{GT}}(\chi_1, \dots, \chi_n + \Delta\chi_n, \dots, \chi_N)$, c’est-à-dire les temps de retour lorsqu’un paramètre varie alors que les autres restent fixés. On a également représenté les temps de retour implicites $\mathbf{t}^{\perp}(\chi_1, \dots, \chi_n + \Delta\chi_n, \dots, \chi_N)$ et $\mathbf{t}^+(\chi_1, \dots, \chi_n + \Delta\chi_n, \dots, \chi_N)$. On remarque que c’est CU qui a la plus grosse contribution. Enfin, remarquons que l’on peut mesurer la contribution du n -ième élément dans le scénario de crise multivarié en prenant $\chi_n = -\infty$. Dans notre exemple, on obtient

$\mathbf{t}^{\text{GT}}(\chi_1, \dots, \chi_n, \dots, \chi_N) - \mathbf{t}^{\text{GT}}(\chi_1, \dots, \chi_n = -\infty, \dots, \chi_N)$	AL	AL-15	CU
	1	11	17

On voit bien que CU (et dans une moindre mesure AL-15) contribue fortement à la valeur de 25 ans du temps de retour implicite. Ce n’est pas le cas de AL, puisque nous avons

$$\mathbf{t}^{\text{GT}}(\text{AL}, \text{AL} - 15, \text{CU}) \simeq \mathbf{t}^{\text{GT}}(\text{AL} - 15, \text{CU})$$

3.4.3 Un retour sur la valeur en risque

Pour finir ce chapitre, nous allons étudier comment nous pouvons appliquer les copules dans un problème d’agrégation de risque de type *Value-at-Risk*. Soient deux variables aléatoires X et Y de fonctions de distribution $\mathbf{F}_1$ et $\mathbf{F}_2$ respectivement. On suppose qu’on ne connaît pas la structure de dépendance du modèle (le copule). On va énoncer un théorème dû à Makarov (voir FRANK, NELSEN et SCHWEIZER [1987]) et qui va nous permettre d’avoir une borne supérieure pour $\text{VaR}(X + Y)$.

Théorème 21 *Pour tout $z \in \mathbb{R}$, on a*

$$\Pr[X + Y \leq z] \geq \sup_{x+y=z} \mathbf{C}^-(\mathbf{F}_1(x), \mathbf{F}_2(y)) := \Psi(z) \quad (3.161)$$

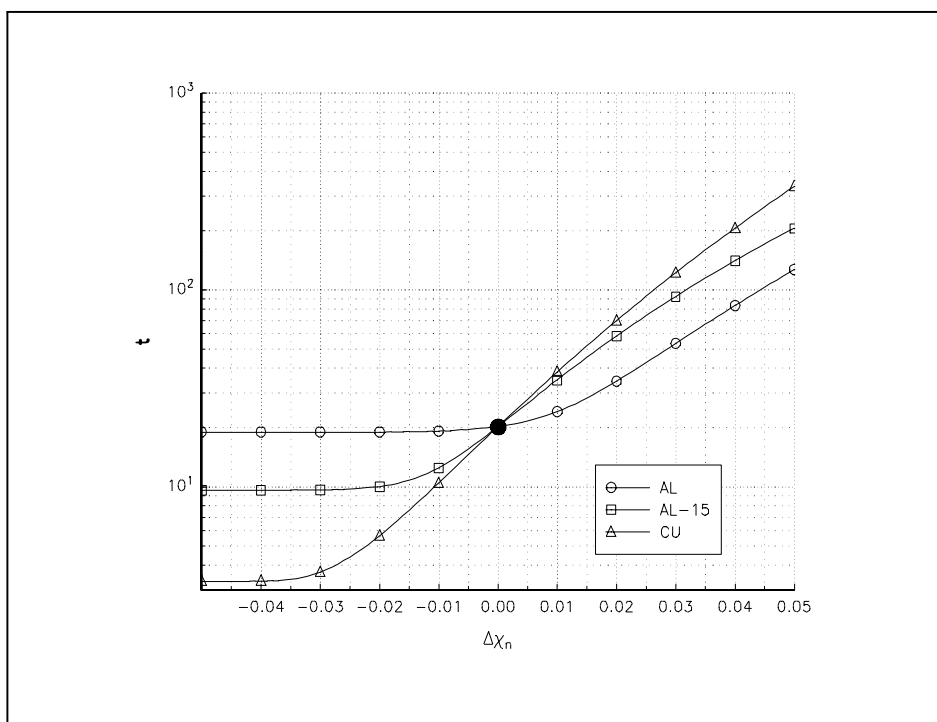
Soit $\Psi^{-1}(\alpha) = \inf\{z \mid \Psi(z) \geq \alpha\}$, $\alpha \in (0, 1)$, l’inverse généralisée de Ψ . Alors

$$\Psi^{-1}(\alpha) = \inf_{M(u,v)=\alpha} \{\mathbf{F}_1^{-1}(u) + \mathbf{F}_2^{-1}(v)\} \quad (3.162)$$

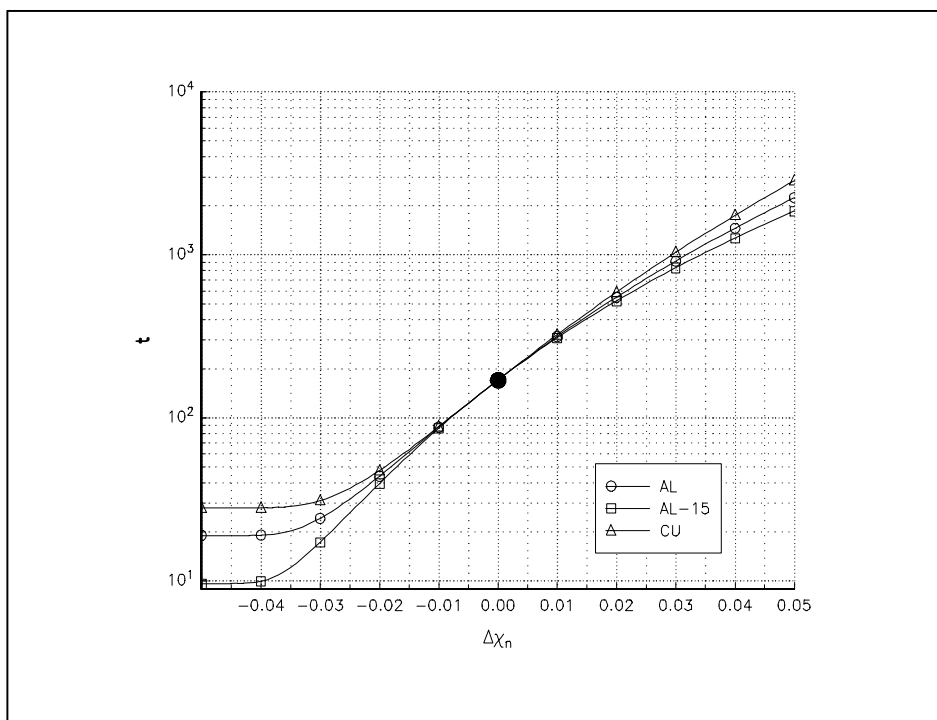
On en déduit immédiatement que (EMBRECHTS, MCNEIL et STRAUMANN [1999])

$$\text{VaR}_{\alpha}(X + Y) \leq \Psi^{-1}(\alpha) \quad (3.163)$$

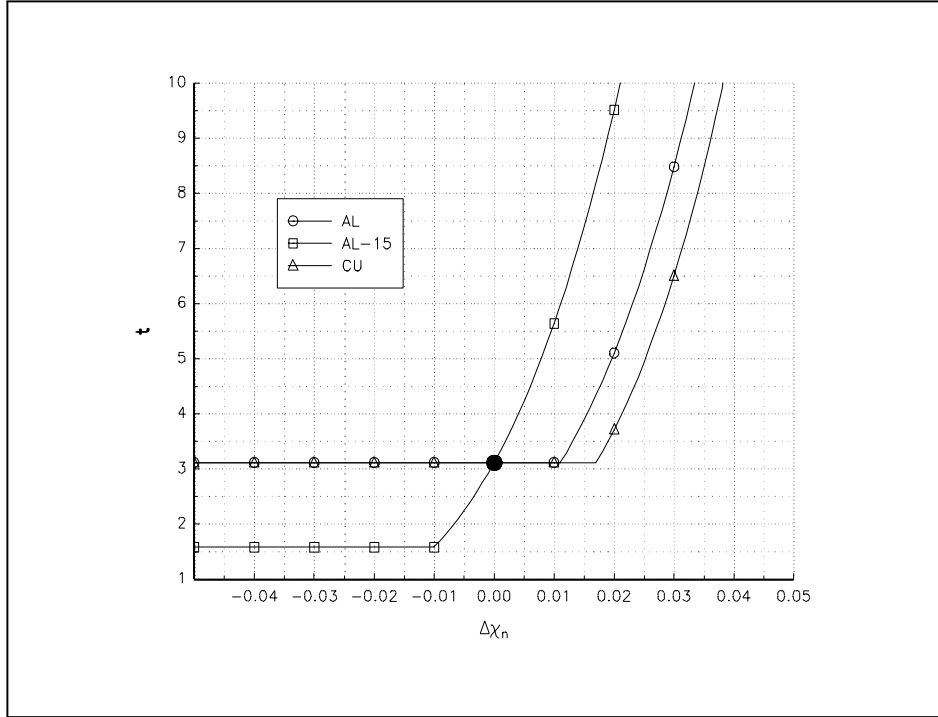
Cette borne est la meilleure possible.



Graphique 3.9. Contribution des stress de scénarios unidimensionnels avec le copula $\mathbf{C}^{\text{GT}}$



Graphique 3.10. Contribution des stress de scénarios unidimensionnels avec le copula $\mathbf{C}^{\perp}$



Graphique 3.11. Contribution des stress de scénarios unidimensionnels avec le copula $\mathbf{C}^+$

Remarque 13 Dans le cas où X et Y sont obtenues l'une de l'autre par une transformation monotone, on a $VaR_\alpha(X + Y) = VaR_\alpha(X) + VaR_\alpha(Y)$. Mais il existe des cas où $VaR_\alpha(X + Y)$ est beaucoup plus grande que la somme des quantiles univariés $VaR_\alpha(X) + VaR_\alpha(Y)$. Comme nous l'avons vu lors du premier chapitre, cela est dû au fait que la VaR n'est pas une mesure de risque sous-additive. On peut ainsi en un certain sens considérer que la différence $\Psi^{-1}(\alpha) - VaR_\alpha(X) - VaR_\alpha(Y)$ représente le défaut de sous-additivité de la VaR pour certaines valeurs de α et certaines distributions marginales.

Conclusion

Nous avons au cours de cette étude établi quelques résultats relatifs aux mesures de risque. Nous avons souligné les qualités que devait avoir une bonne mesure de risque et avons ensuite présenté la *Value-at-Risk* telle qu'elle est aujourd'hui utilisée par les banques. Nous avons ensuite montré l'importance de construire d'autres mesures de risques qui sont fondées sur la loi des extrêmes.

L'utilisation de la *VaR GEV* et plus généralement l'utilisation de la théorie des extrêmes peut s'avérer très utile pour obtenir des indications sur les risques extrêmes ou encore pour élaborer des scénarios de crise. De plus, la théorie des extrêmes peut s'imposer si la réglementation bancaire évolue sur les seuils imposés pour le calcul de la *VaR*. **Dans tous les cas, elle présente un intérêt évident pour les projets d'allocation de fonds propres sur base économique.**

Nous avons enfin utilisé la théorie des copules pour placer l'étude du risque dans un contexte multidimensionnel. La théorie des copules apparaît très riche et très prometteuse quant à son utilisation dans le métier bancaire — voir par exemple l'application des copules à la gestion de portefeuille de MÜLLER et SCARSINI [1999]. Nous avons, en ce qui nous concerne, obtenu quelques résultats nouveaux grâce à des méthodes récentes utilisant les copules. L'application des copules à la gestion des risques est naissante : de nombreuses choses restent encore à faire et de nombreuses améliorations peuvent encore être apportées quant à l'utilisation pratique des copules. Une piste intéressante, parmi tant d'autres, serait d'utiliser les résultats publiés récemment sur les copules et les distances afin d'optimiser le choix des copules lors de leur utilisation pour la modélisation.

Références

- [1] Amendment to the capital accord to incorporate market risks, *Basle Committee on Banking Supervision*, Janvier 1996, N° **24**
- [2] A new capital adequacy framework, *Basle Committee on Banking Supervision*, Juin 1999, N° **50**
- [3] ABDOUS, B., K. GHOUDI et A. KHOUDRAJI [1999], Non-parametric estimation of the limit dependence function of multivariate extremes, *Extremes*, **3**, 243-265
- [4] ARTZNER, A., F. DELBAEN, J-M. EBER et D. HEATH [1997], Thinking coherently, *Risk magazine*, **10**, Novembre, 68-71
- [5] ARTZNER, A., F. DELBAEN, J-M. EBER et D. HEATH [1999], Coherent measures of risk, *Mathematical Finance*, **9**, 203-228
- [6] BARBE, P., C. GENEST, K. GHOUDI et B. RÉMILLARD [1996], On Kendall's process, *Journal of Multivariate Analysis*, **58**, 197-229
- [7] BASSI, B , P. EMBRECHTS et M. KAFETZAKI [1997], Risk Management and Quantile Estimation, Departement Mathematik, ETHZ, Zürich
- [8] BAUD, N., A. FRACHOT, P. IGIGABEL, P. MARTINEU et T. RONCALLI [1999], An Analysis Framework for Bank Capital Allocation, Groupe de Recherche Opérationnelle, Crédit Lyonnais, *Working Paper*
- [9] BOUCHAUD, J-P. et M. POTTERS [1997], Théorie des Risques Financiers, Aléa Saclay, Diffusion Eyrolles, Paris
- [10] BOULIER, J-F, A. BRABANT, R. DALAUD et A-L DIEU [1997], Risques de marché : Vue de profil, Direction de la Recherche et de l'Innovation, Crédit Commercial de France, *Quants*, **28**
- [11] BOUYÉ, E., A. NIKEGHBALI, G. RIBOULET et T. RONCALLI [2000], Copulas for Finance — A reading guide and some applications, Groupe de Recherche Opérationnelle, Crédit Lyonnais, *Working Paper*
- [12] CAPÉRAA, P., FOUGÈRES, A-L. et C. GENEST [1997], Estimation of bivariate extreme value copulas. *Biometrika*, **84**, 567-577
- [13] CAPÉRAA, P., A-L. FOUGÈRES et C. GENEST [2000], Bivariate distributions with given extreme value attractor, *Journal of Multivariate Analysis*, **72**, 30-49
- [14] COLES, S.G et J.TAWN [1999], Statistical methods for extreme values, *A course presented at the 1998 RSS conference*, Strathclyde, September 1998
- [15] COSTINOT A, G. RIBOULET et T. RONCALLI [2000], Note technique sur l'utilisation de la théorie des extrêmes unidimensionnelles dans la construction des stress scenarii, Groupe de Recherche Opérationnelle, Crédit Lyonnais, *document interne*
- [16] DE HAAN, L., L. PENG, A.K. SINHA et G. DRAISMA [1997], The Erasmus University Contribution to the Neptune project, Task T-400, Erasmus University Rotterdam, *Technical Report*

- [17] DEHEUVELS, P. [1979], La fonction de dépendance empirique et ses propriétés — Un test non paramétrique d'indépendance, *Acad. Roy. Belg. Bull. Cl. Sci.*, **65**, 274-292
- [18] DELBAEN, F. [1998], Coherent Risk Measures on General Probability Spaces, ETHZ, *Working Paper*
- [19] EMBRECHTS, P., C. KLÜPPELBERG et T. MIKOSCH [1997], Modelling Extremal Events for Insurance and Finance, Springer-Verlag, Berlin
- [20] EMBRECHTS, P., A.J. MCNEIL et D. STRAUMANN [1999], Correlation and dependency in risk management : properties and pitfalls, Departement of Mathematik, ETHZ, Zürich, *Working Paper*
- [21] FELLER, W. [1968], An Introduction to Probability Theory and its Application, New York, John Wiley & Sons
- [22] FRANK, M.J., R.B. NELSEN, et B. SCHWEIZER [1987], Best-possible bounds for the distribution of a sum — a problem of Kolmogorov, *Probability Theory and Related Fields*, **74**, 199-211
- [23] FREES, E.W. et E.A. VALDEZ, [1998], Understanding relationships using copulas, *North American Actuarial Journal*, **2**, 1-25
- [24] GENEST, C. et K. GHOUDI [1994], Une famille de lois bidimensionnelles insolite, *Comptes Rendus de l'Académie des Sciences de Paris*, **318**, série I, 351-354
- [25] GENEST, C., K. GHOUDI et L.P. RIVEST [1998], Discussion of "Understanding relationships using copulas, by Edward Frees and Emiliano Valdez", *North American Actuarial Journal*, **3**, 143-149.
- [26] GENEST, C. et J. MACKAY [1986], The joy of copulas : Bivariate distributions with uniform marginals, *American Statistician*, **40**, 280-283
- [27] GENEST, C. et L. RIVEST [1993], Statistical inference procedures for bivariate Archimedean copulas, *Journal of the American Statistical Association*, **88**, 1034-1043
- [28] GHOUDI, K., A. KHOUDRAJI et L.P. RIVEST, [1998], Propriétés statistiques des copules de valeurs extrêmes bidimensionnelles, *Canadian Journal of Statistics*, **26**, 187-197
- [29] HOEUNG S., G. RIBOULET et T. RONCALLI [1999], Les risques de marché — Réglementations, mesures et pratiques, Groupe de Recherche Opérationnelle, Crédit Lyonnais, *document interne*
- [30] JOE, H. [1997], Multivariate Models and Dependence Concepts, Chapman & Hall, London
- [31] JOE, H. et T. HU [1996], Multivariate distributions from mixtures of max-infinitely divisible distributions, *Journal of Multivariate Analysis*, **57**, 240-265
- [32] JOE, H. et J.J. XU [1996], The estimation method of inference functions for margins for multivariate models, Department of Statistics, University of British Columbia, *Technical Report*, **166**
- [33] KIMELDORF, G. et A. SAMPSON [1989], A framework for positive dependence, *Ann. Inst. Statist. Math.*, **41**, 31-45
- [34] MARSHALL, A.W. et I. OLKIN [1983], Domains of attraction of multivariate extreme value distributions, *Ann. Probab.*, **11**, 168-177
- [35] MCNEIL, A.J. [1999], Extreme value theory for risk managers, Departement Mathematik, ETHZ, Zürich, *Preprint*
- [36] MCNEIL, A.J. [1998], Calculating quantile risk measures for financial return series using extreme value theory, *ASTIN Bulletin*, **27**, 117-137
- [37] MÜLLER, A. et M. SCARSINI [1999], Stochastic comparison of random vectors with fixed copulae, *WIOR-Report*, **554**
- [38] NELSEN, R.B. [1999], An Introduction to Copulas, Springer Verlag, New York
- [39] OAKES, D. [1989], Bivariate Survival Models Induced by Frailties, *Journal of the American Statistical Association*, **84**, 487-493
- [40] RESNICK, S.I. [1987], Extreme Values, Point Processes and Regular Variation, Springer-Verlag, New York
- [41] SCHMOCK, U. et D. STRAUMANN [1999], Allocation of risk capital, RiskLab, Department of Mathematics, ETH Zürich, *Working Paper*
- [42] SCHWEIZER, B. et A. SKLAR [1983], Probabilistic Metric Spaces, Elsevier North-Holland, New York
- [43] SCHWEIZER, B. et E. WOLFF [1981], On nonparametric measures of dependence for random variables, *Annals of Statistics*, **9**, 879-885

- [44] STRAETMANS, S. [1999], Extreme financial returns and their comovements, Erasmus University Rotterdam's Thesis, *Tinbergen Institute Research Series*, **181**
- [45] STRAETMANS, S. [2000], Spillovers in financial markets, Free University Amsterdam, *Working Paper*
- [46] SKLAR, A. [1959], Fonctions de repartition à n dimensions et leurs marges, *Publ. Inst. Stat. Univ. Paris*, **8**, 229-231
- [47] TASCHE, D. [1999], Risk contributions and performance measurement, Zentrum Mathematik, TU München, *Working Paper*, July
- [48] TAWN, J.A. [1988], Bivariate extreme value theory : models and estimation, *Biometrika*, **75**, 397-415
- [49] WANG, S.S. [1999], Aggregation of correlated risk portfolios : Models & algorithms, CAS Committee on Theory of Risk, *preprint*
- [50] YUN, S. [1997], On domains of attraction of multivariate extreme value distribution under absolute continuity, *Journal of Multivariate Analysis*, **63**, 277-295

Table des figures

2.1	Distributions de Gumbel et Fréchet	11
2.2	Convergence de la distribution exponentielle vers la Gumbel	12
2.3	Quantiles de la distribution GEV	13
2.4	Cours de l'aluminium	18
2.5	Cours du cuivre	19
2.6	Cours à terme de l'aluminium	19
2.7	Estimations paramétrique et non paramétrique des rendements	20
2.8	Densité des lois d'extrêmes estimées	24
2.9	Relation entre l'échelle de risque et le temps de retour	24
3.1	Comparaison des fonctions $\lambda_U(u)$	34
3.2	Procédure d'identification des copules archimédiennes	42
3.3	Scénarios de crise pour un portefeuille court en AL et AL-15	50
3.4	Scénarios de crise pour un portefeuille court en AL et CU	51
3.5	Scénarios de crise pour un portefeuille court en AL-15 et CU	51
3.6	Scénarios de crise pour un portefeuille long en AL et AL-15	52
3.7	Scénarios de crise pour un portefeuille long en AL et CU	52
3.8	Scénarios de crise pour un portefeuille long en AL-15 et CU	53
3.9	Contribution des stress de scénarios unidimensionnels avec le copula $\mathbf{C}^{\text{GT}}$	56
3.10	Contribution des stress de scénarios unidimensionnels avec le copula $\mathbf{C}^{\perp}$	56
3.11	Contribution des stress de scénarios unidimensionnels avec le copula $\mathbf{C}^+$	57