

Mémoire présenté le :
/ /
pour l'obtention du diplôme
de Statisticien Mention Actuariat
et l'admission à l'Institut des Actuaires

Par : Mathias MERCAT

Sujet : Relaxation de l'hypothèse d'indépendance dans la tarification des
traités "*Multiline Aggregate Excess of Loss*"

Confidentialité : NON OUI (Durée : 1 an 2 ans)

Les signataires s'engagent à respecter la confidentialité indiquée ci-dessus.

*Membre présents du jury de
l'Institut des Actuaires*

*Membres présents du jury de la
filière*

Entreprise :

Nom : QBE Re Brussels

Signature :

*Directeur de mémoire en
entreprise :*

Nom : Nicolas Daxhelet

Signature :

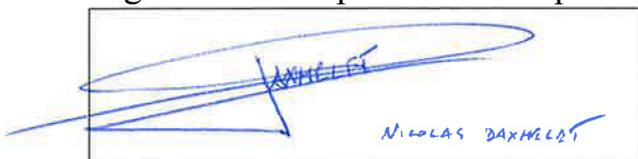
Invité :

Nom :

Signature :

**Autorisation de publication et de
mise en ligne sur un site de
diffusion de documents actuariels
(après expiration de l'éventuel
délai de confidentialité)**

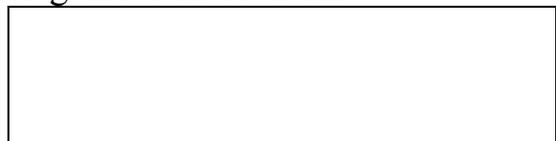
Signature du responsable entreprise



Secrétariat

Bibliothèque :

Signature du candidat





INSTITUT DE STATISTIQUE DE L'UNIVERSITÉ DE PARIS
ÉCOLE D'ACTUARIAT

RELAXATION DE L'HYPOTHÈSE D'INDÉPENDANCE DANS LA
TARIFICATION DES TRAITÉS "MULTILINE AGGREGATE EXCESS
OF LOSS"

Mathias Mercat

21 novembre 2014

Tuteur professionnel : M. Daxhelet

Tuteur académique : M. Nessi

RELAXATION DE L'HYPOTHÈSE D'INDÉPENDANCE DANS LA TARIFICATION DES TRAITÉS "MULTILINE AGGREGATE EXCESS OF LOSS"

Synthèse

Mathias Mercat

Institut de Statistique de l'Université de Paris - ISUP

Objectifs

Ce mémoire propose une méthodologie de relaxation de l'hypothèse d'indépendance entre les "*Lines of Business*" réassurées dans des traités de réassurances couvrant simultanément plusieurs branches. On s'intéresse tout particulièrement à la tarification des traités dit "*Multiline Aggregate Excess of Loss*". En tant que traités "*multiline*", ils sont sujets aux différents phénomènes de diversification. Il est donc tout naturel de se pencher sur les effets de la dépendance en cas de forte présomption contre l'hypothèse d'indépendance.

L'objectif principal est de décrire une méthodologie facilement implémentable, rapide à exécuter et s'intégrant directement aux processus déjà en place au sein du service de tarification d'une compagnie de réassurance. En d'autres termes, la méthode ne doit pas impacter les étapes de la tarification n'ayant pas attrait à la dépendance entre les "*Lines of Business*".

Traités "Multiline Aggregate Excess-of-Loss"

Dans un traité "*Multiline Aggregate Excess of Loss*", le réassureur indemnise la cédante lorsque la somme des pertes agrégées alimentant le traité dépasse un certain seuil appelé "*Global Annual Aggregate Deductible*" (GAAD). Ce montant d'alimentation, noté \mathbb{S} , est quant à lui l'agrégat de montants de sinistres éligibles sur différentes "*Lines of Business*" (comme représenté sur la figure 1). On a donc :

$$\mathbb{S} = \sum_{i=LoB_1}^{LoB_n} \mathbb{S}_i \text{ et } \mathbb{S}_i = \sum_{j=1}^{N_i} f_i(x_{i_j})$$

Avec N_i le nombre de sinistres considérés sur la "*Line of Business*" i et f_i la fonction d'éligibilité de la "*Line of Business*" i . Par exemple, $f_i(x_{i_j}) = \min(l_i, \max(x_{i_j} - d_i, 0))$ pour une tranche d'éligibilité $[d_i, l_i]$ classique.

Le réassureur limite sa prise en charge annuelle en ajoutant une limite au traité noté "*Global Annual Aggregate Limit*" (GAAL).

La prime pure d'un tel traité s'exprime alors sous la forme :

$$\mathbb{E} [\min (GAAL, \max (\mathbb{S} - GAAD, 0))]$$

Dans la suite, nous utiliserons la $GAAL$ $MAXL$ $GAAD$.

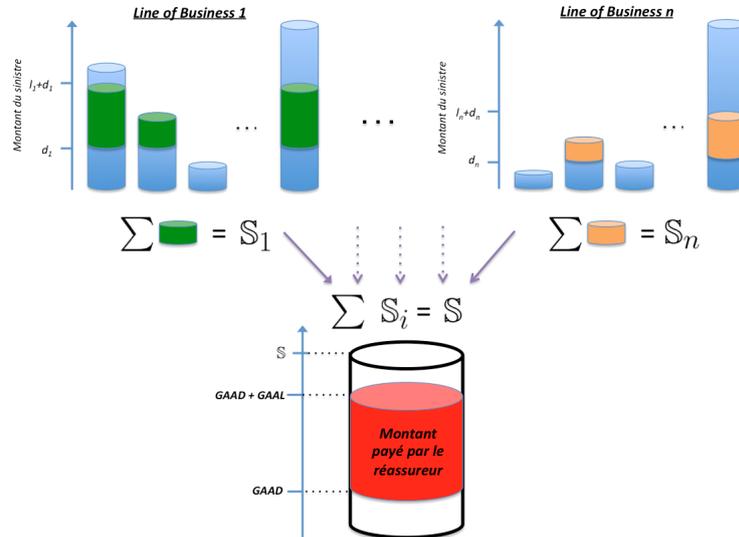


FIGURE 1 – Fonctionnement schématique d'un traité MAXL

Méthodes proposée

En réassurance non-proportionnelle, le réassureur ne s'intéresse qu'aux sinistres qui potentiellement affectent le traité. Concrètement cela veut dire que N_i représente le nombre de sinistres dépassant le seuil de prise en charge de la "Line of Business" i par le traité. Le réassureur n'est d'ailleurs informé par la cédante que de ces sinistres là. La mise en place d'une structure de dépendance doit tenir compte du caractère partiel du niveau d'information du réassureur.

Les montants S_i (sinistralités d'alimentation) sont le fruit d'effets non proportionnels, ils n'ont pas de réalité physique, donc, imposer une structure de dépendance sur ces valeurs n'aurait aucun sens. On s'intéresse alors à un autre forme d'agrégat de sinistralité, l'agrégat des montants *from-ground-up* des sinistres affectant les différentes "Line of Business".

$$\mathbb{T}_i = \sum_{j=1}^{N_i} x_{i_j}$$

On propose dans ce mémoire, d'étudier, modéliser et simuler selon la structure de dépendance régissant la relation multivariée entre les différents \mathbb{T}_i . Évidemment, la structure de dépendance est fonction des seuils à partir desquels on considère les sinistres. Les structures de dépendance proposées l'ont été car adaptées à cette problématique.

La méthode de tarification devient alors la suivante :

1 : La première étape demeure la simulation des sinistres individuels sur toutes les "*Lines of Business*" (fréquence/sévérité, période de retour,...). En commençant de la sorte nous nous assurons de l'intégrabilité du processus de mise sous dépendance à l'ensemble des méthodes déjà implémentées.

2 : La seconde étape consiste simplement à calculer les valeurs d'agrégats \mathbb{T}_i et \mathbb{S}_i .

3 : La troisième étape, celle de simulation de la structure de dépendance (copule), est sûrement celle nécessitant le plus de travail. On notera que plutôt que d'utiliser les réalisations de variables $\mathcal{U}_{[0,1]}$ on travaillera directement avec les statistiques de rang. On obtient alors simplement une structure de rang de la forme :

R_1	...	R_n
r_{11}	...	r_{n1}
...
r_{1N}	...	r_{nN}

4 : Lors de la dernière étape nous appliquons un algorithme de réorganisation des simulations de telle sorte que les rangs des \mathbb{T}_i aient la même structure que les rangs de la copule simulée. Ainsi le montant alimentant le traité au titre de la ligne i dans le p^e scénario sera le montant \mathbb{S}_i associé à la simulation la valeur \mathbb{T}_i qui a pour rang r_{ip} . Les scénarios des \mathbb{S}_i sont donc réorganisés selon une structure de dépendance sur les \mathbb{T}_i .

Dans le cadre d'un traité couvrant n "*Lines of Business*", la méthode de tarification à N scénarios peut se résumer sous la forme du pseudo-code algorithmique suivant :

```

Input:  $n$  le nombre de LoB
Input:  $N$  le nombre de scénario à simuler
for  $i \in \{1, \dots, n\}$  do
  | for  $s \in \{1, \dots, N\}$  do
  |   Simuler et mémoriser les montants des sinistres individuels from-ground-up (étape 1)
  | end
  |   Calculer et mémoriser les agrégats  $\mathbb{S}_i(s)$  et  $\mathbb{T}_i(s)$  (étape 2)
end
for  $i \in \{1, \dots, n\}$  do
  |   Simuler un échantillon de taille  $N$  de variables aléatoires  $U_i \sim \mathcal{U}_{[0,1]}$  selon une copule  $C$ 
  |   puis calculer  $r_{ip}$  rang de la  $p^e$  réalisation de  $U_i$  (étape 3)
end
for  $s' \in \{1, \dots, N\}$  do
  |   for  $i \in \{1, \dots, n\}$  do
  |     Affecter au scénario  $s'$  le montant  $\mathbb{S}_i(s)$  tel que  $s$  vérifie  $\text{Rang}(\mathbb{T}_i(s)) = r_{is'}$  (étape 4)
  |   end
end
for  $s' \in \{1, \dots, N\}$  do
  |   Calculer  $\mathbb{S}(s') = \sum_{i=1}^n \mathbb{S}_i(s')$ 
  |   Calculer  $\pi(s') = \min(GAAL, \max(\mathbb{S}_i(s') - GAAD, 0))$ 
end
La prime pure du traité MAXL est alors estimé par :  $\Pi = \mathbb{E}[\pi(s')]$ 

```

Structures de dépendances retenues

Une étape cruciale du processus de tarification consiste à simuler une structure de dépendance des \mathbb{T}_i afin de procéder à ensuite à l'étape de réarrangement. Trois approches ont été envisagées afin de faire face à cette structure de dépendance multivariée. Ces trois approches sont basées sur des structures copules.

La première d'entre elles, la famille de copule de Fréchet, est un modèle extrêmement simple de pondération convexe entre les structures d'indépendance et de comonotonicité $C_\theta = \theta C^\uparrow + (1 - \theta)C^\perp$ avec $\theta \in [0, 1]$. L'estimation du paramètre θ peut se faire à partir de la mesure de dépendance décrite dans Dhaene *et al.*, 2014 [13]

Étant donné que le niveau de dépendance va fortement dépendre des seuils à partir desquels on considère les sinistres sur chacune des branches, il est nécessaire, pour ce type de problématique, de pouvoir spécifier la dépendance au niveau de chaque paire de variables. En effet, compte tenu de la multiplication des effets de seuil, la modélisation directe par une copule multivariée est quasiment impossible à réaliser. De plus, le moindre changement de seuil sur une des variables nécessiterait une nouvelle modélisation complète. En utilisant des hypothèses "pairwise" nous avons à ré-étudier seulement les relations impliquant la branche modifiée. En procédant ainsi nous augmentons certes le nombre d'étude à réaliser mais nous simplifions considérablement le problème tout en lui donnant plus de liberté dans sa forme et son intensité.

Les deux autres approches partagent, donc cette caractéristique des relations "pairwise".

La seconde approche repose ainsi sur des copules multivariées classiques calibrables à partir d'hypothèses de dépendance "pairwise" : copule gaussienne multivariée, copule t et Individuated-t. Ce sont des modèles extrêmement classiques, largement décrits dans la littérature. La copule Individuated-t, sûrement la moins connue d'entre elles est une cousine de la copule-t, dans laquelle nous attribuons à chaque variable aléatoire son propre degré de liberté, donnant ainsi au modèle une marge de manœuvre dans la calibration des dépendances extrêmes.

Enfin la dernière approche, basée sur l'heuristique de simulation de Kaut décrite dans Kaut, 2013 [21], est la plus novatrice. Elle permet d'appréhender le caractère multidimensionnel de la problématique directement via l'étude complète des structures de dépendance "pairwise". En d'autres termes, dans cette méthode la structure multivariée de dimension d est spécifiée directement par l'ensemble des $\frac{d(d-1)}{2}$ copules bivariées. Cette méthode de simulation s'avère être particulièrement efficace et offre une liberté extrêmement appréciable dans la construction de la structure de dépendance.

La calibration de ces modèles est cruciale comme souvent dans ce type de problématiques. Nous proposons une méthode d'estimation basée sur l'historique de tous les traités non proportionnels en portefeuille. La problématique des seuils complexifie la problématique c'est pourquoi le volume de données nécessaire à une bonne estimation

est rapidement important.

L'intérêt de ces modèles et notamment celui basé sur l'heuristique de Kaut, est de pouvoir mélanger des hypothèses issues d'études sur des données et des hypothèses forfaitaires à dire d'experts, dont on peut certes douter de la robustesse mais qui sont souvent incontournables et aisément présentable en clientèle.

Résultats et conclusions

Un module reprenant la méthodologie décrite a été complètement implémenté au sein du processus de tarification et les premières observations sont que cette implémentation est simple et n'augmente pas, une fois la structure de dépendance connue, le temps de travail nécessaire à la tarification.

D'un point de vue tarifaire la dépendance positive impacte quasiment à chaque fois la prime pure. Lorsque les GAAD et GAAL sont calibrés pour couvrir la sinistralité d'alimentation à un niveau de quantile suffisamment élevé alors la dépendance positive a tendance à augmenter la prime pure.

Cet impact va de pair avec la déformation de la distribution de la sinistralité d'alimentation. En effet, la dépendance va déformer la loi de probabilité de \mathbb{S} . Si les GAAD et GAAL sont fixés pour couvrir \mathbb{S} entre deux niveaux de quantile il faut impérativement tenir compte de la dépendance dans la calibration du traité. Si le traité reste calibré comme il l'était sous l'hypothèse d'indépendance, la cédante se pensera couverte sur un intervalle de probabilité inexacte.

La première méthode offre un bon aperçu de l'impact sur la prime pure. Cependant, ce modèle n'offre qu'un ordre de grandeur et doit être complété par des modèles plus précis. De plus, la calibration du paramètre de ce modèle nécessite un historique de même profondeur sur toutes les "*Lines of Business*" ce qui dans les faits réduit considérablement l'historique exploitables.

L'utilisation de la copule gaussienne, bien que ne permettant pas l'introduction de dépendance de queue, semble en revanche largement suffisante pour appréhender avec assez de précision la nouvelle tarification. Les difficultés de calibration des copules t et Individuated- t font perdre à ces modèles beaucoup de crédibilité.

L'utilisation de l'heuristique de Kaut est véritablement pertinente si le niveau de connaissance de la dépendance est suffisant élevé. C'est à coup sûr le modèle offrant le plus de marge de manœuvre. Dans une situation où beaucoup d'hypothèses doivent être faites à dire d'experts, les résultats obtenus doivent être considérés avec prudence. La méthode peut cependant servir de contrôle ou de test par exemple pour mesurer l'impact de la modification de la structure de dépendance juste entre 2 "*Lines of Business*". La méthode permet aussi de tester quelle relation "*pairwise*" a le plus d'impact sur la prime pure.

De façon plus globale le fait d'imposer une structure de dépendance entre les sinistralités réelles T_i est critiquable du fait des effets liés aux seuils déclaratifs et aux besoins de données de calibration. Une étude de la dépendance devant être réalisée quasiment à chaque nouvelle tarification, on introduit, de fait, un risque opérationnel dans le tarif proposé. Cependant, dans un contexte de réassurance et compte tenu des contraintes d'intégrabilité du module dans le processus de tarification existant, cette méthode apparaît comme pertinente et relativement efficace.

Résumé

Mots-clefs : MAXL, Dépendance, Tarification, Réassurance, Copule, Heuristique de Kaut

La méthodologie proposée dans ce mémoire répond au besoin d'intégrer un "module" de prise en compte de la dépendance au sein du processus de tarification des traités dits "*Multiligne Aggregate Excess of Loss*". Ces traités proposant une couverture simultanée de plusieurs "*Lines of Business*", ils sont sujets aux différents phénomènes de diversification. Il est donc tout naturel de se pencher sur les effets de la dépendance en cas de forte présomption contre l'hypothèse d'indépendance.

Les conséquences tarifaires de la dépendance entre les différentes "*Lines of Business*" demeurent cependant difficilement prédictibles du fait des différents effets non proportionnels combinés au sein des traités "*Excess of Loss*". Seule l'implémentation concrète d'une méthode de simulation de scénarios de sinistralités tenant compte de la dépendance permet de conclure quant aux modifications en terme de diversification.

La méthode retenue lors de cette étude propose d'étudier et d'imposer lors de la simulation, une structure de dépendance entre les sommes, par "*Line of Business*", des sinistres *from-ground-up* dépassant un certain seuil. Les seuils en question correspondent dans la majorité des cas aux priorités internes des "*Lines of Business*" du traité MAXL.

Trois principales approches ont été envisagées pour la structure de dépendance. La première d'entre elle, la famille de copule de Fréchet, est un modèle extrêmement simple de pondération convexe entre indépendance et comonotonie. Elle permet de rapidement saisir la sensibilité du traité à la dépendance.

Étant donné que le niveau de dépendance va fortement dépendre des seuils à partir desquels on considère les sinistres sur chacune des branches, il est nécessaire, pour ce type de problématique, de pouvoir spécifier la dépendance au niveau de chaque paire de variables. Les deux autres approches partagent cette caractéristique.

La seconde approche repose ainsi sur des copules multivariées classiques calibrables à partir d'hypothèses de dépendance "*pairwise*" : copule gaussienne multivariée, copule t et Individuated- t . Enfin la dernière approche, basée sur l'heuristique de simulation de Kaut, est la plus novatrice. Elle permet d'appréhender le caractère multidimensionnel de la problématique directement via l'étude des structures de dépendance "*pairwise*". En d'autres termes, dans cette méthode la structure multivariée est spécifiée directement par l'ensemble des copules bivariées.

En fonction des caractéristiques des risques réassurés et des termes du traité, les conséquences de la dépendance sont diverses. Si l'augmentation de la prime pure est l'impact auquel l'actuaire s'attend le plus naturellement, l'existence de dépendance peut aussi remettre en cause la calibration du traité en terme de priorité et de portée.

Abstract

Key-words : MAXL, pricing, risk dependent risks, reinsurance, copula, Kaut heuristic

This master thesis describes the implementation of a pricing tool allowing to take into account dependencies among "Lines of Business" in the pricing of "Multiline Aggregate Excess of Loss" treaties. These treaties covering several "Lines of Business" are subjected to diversification effects. It is therefore obvious to study these effects when actuaries have strong presumptions against the independence hypothesis.

The effects of the dependence among risks on the pure premium are difficult to forecast because of the non-proportional effects involved by Excess of Loss treaties. Only the implementation of a pricing tool allowing to sample dependent loss scenarios can allow actuaries to assess diversification effects on the pure premium.

The method used in this study considers that sums of "from ground up" losses exceeding a specific threshold are dependent among "Lines of Business". In most of the cases, these thresholds correspond to internal deductible of the "Line of Business" in the terms of the MAXL treaty.

Three main approaches have been considered for the dependency structure. The first one, the Frechet copula family, is an extremely simple model of convex weighting between independence and comonotonicity. It rapidly allows determining the sensitivity of the treaty to dependency.

Given that the level of dependency will depend a lot on the threshold above which the losses are considered on each "Line of Business", it is necessary for this kind of issue to specify the dependency at the level of each pair of variable. The two others approaches share this feature.

The second approach is based on classical multivariate copulas which can be calibrated from pairwise dependency hypothesis. We will use the multivariate Gaussian, t and Individuated-t copulas. Finally, the third approach, based on the simulation heuristic of Kaut, is the most innovative. It allows to cope with the multidimensional feature of this issue directly through the study of the pairwise dependency structure. In other words, in this method the multivariate structure is directly specified by all the bivariate copula.

According to the characteristics of the reinsured risks and the terms of the treaty, the consequences of the dependency are diverse. If the increase of the pure premium is the impact that the actuary naturally expect, the existence of dependency can also challenge the calibration of the treaty in terms of limit and deductible.

Sommaire

Synthèse	i
Résumé	vii
Abstract	ix
1 La réassurance " <i>multiline</i> "	1
2 Cadre théorique de la dépendance multivariée	19
3 Proposition d'une méthodologie de tarification	47
4 Exemple de tarification	57
5 Conclusion	69

Chapitre 1

La réassurance "*multiline*"

Dans ce premier chapitre nous revenons rapidement sur les raisons d'être de la réassurance puis nous décrivons l'attrait, le fonctionnement et les grands principes de tarification des traités "*Multiline Aggregate Excess of Loss*".

Enfin nous introduisons les raisons qui, dans le cadre de la conception et la tarification des traités "*Multiline Aggregate Excess of Loss*" et plus généralement dans la mise en place d'un modèle interne, poussent les actuaires à s'intéresser à la dépendance entre les différentes "*Lines of Business*".

Sommaire

1.1	La réassurance	2
1.2	Traités "<i>multiline</i>" et diversification	3
1.3	"<i>Multiline Aggregate Excess of Loss</i>"	5
1.3.1	Définition	5
1.3.2	Fonctionnement théorique	6
1.3.3	Exemple de fonctionnement d'un MAXL à 3 " <i>Lines of Business</i> "	7
1.3.4	Un exemple pratique d'utilisation dans un programme	9
1.3.5	Le portefeuille QBE Re	10
1.3.6	Principe de tarification	10
1.3.6.1	Simulation des scénarios	10
1.3.6.2	Patterns de paiement	11
1.3.7	Les effets non-proportionnels Excess-of-Loss	12
1.4	Pourquoi s'attarder sur la dépendance entre les risques?	14
1.4.1	Quelques exemples intuitifs	14
1.4.2	A quel(s) effet(s) faut-il s'attendre?	15
1.4.2.1	Un effet difficile à prévoir	15
1.4.2.2	Illustration :	17

1.1 La réassurance

"Assurer les assureurs", voici sûrement la définition la plus simple et la plus claire de l'activité de réassurance. Une compagnie de réassurance propose ainsi d'indemniser, moyennant une prime, des assureurs - ou cédantes - contre tout ou partie d'un sinistre qu'ils pourraient observer parmi leurs polices d'assurance.

Les lois relatives à l'assurance du Code d'Hammourabi à Babylone au XVIII^e siècle avant J.C. sont considérées par beaucoup comme les premières formes d'assurance. Ce n'est que bien après et loin de la Mésopotamie et de l'Égypte que la réassurance est apparue. Les historiens¹ de l'économie font en fait naître l'activité de réassurance au XIV^e siècle lorsque l'assureur Gustave Cruciger décide de céder une partie du risque maritime qu'il assure. Les révolutions industrielles puis la mondialisation et la mise en place de différents régimes réglementaires encadrant la solvabilité des sociétés d'assurance ont peu à peu amené le marché de la réassurance à se structurer et devenir une activité économique importante et durablement installée dans le paysage financier international.

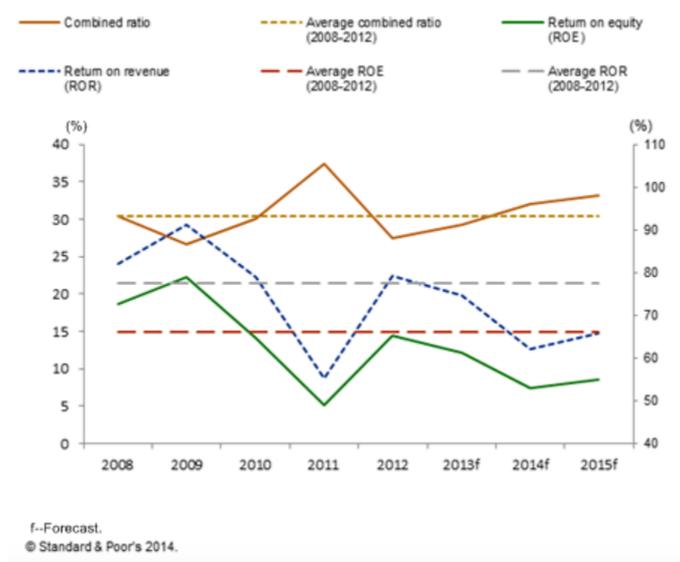


FIGURE 1.1 – S&P Global Reinsurance Earnings forecasts

La directive réassurance de l'Union Européenne (Directive 2005/68/CE du 16 novembre 2005) définit ainsi la réassurance moderne comme suit : "La réassurance constitue une activité financière essentielle, puisqu'elle permet aux assureurs directs, en facilitant une répartition plus large des risques au niveau mondial, d'augmenter leur capacité de souscription et de couverture, et de réduire leur coût en capital. En outre, elle joue un rôle fondamental en matière de stabilité financière, puisque, en tant qu'intermédiaires financiers et investisseurs institutionnels majeurs, les réassureurs contribuent, de façon

1. Le lecteur pourra se reporter au cours dispensé par M. Sibilleau à l'ISFA et dont le polycopié est disponible en ligne Sibilleau, 2007 [29]

décisive, à la solidité financière et à la stabilité des marchés de l'assurance directe et du système financier en général".

La réassurance n'a donc pas pour objet de changer la nature des risques à assurer ou de rendre une exposition au risque plus souhaitable, ou plus prévisible. Les objectifs sont en fait nombreux et le fait de fournir une alternative au capital, de limiter la responsabilité sur des risques spécifiques, de stabiliser l'expérience de sinistres, d'être protégés en cas de catastrophe et d'augmenter leur capacité de souscription, sont les principales motivations des assureurs recherchant de la réassurance.

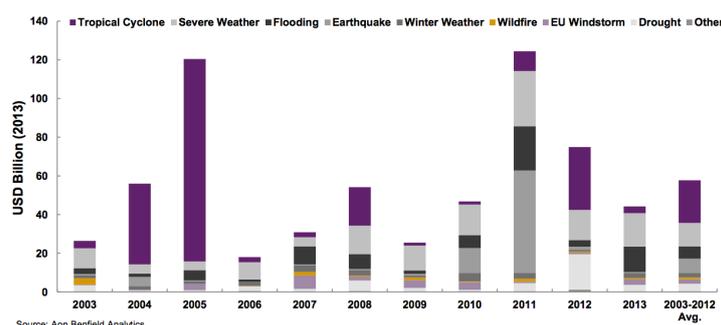


FIGURE 1.2 – Insured Losses by Year by Type (2003-2013)

1.2 Traités "multiline" et diversification

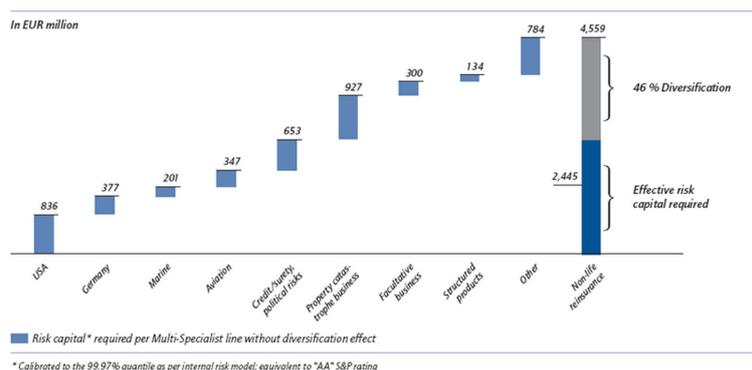
La recherche de produits innovants offrant des garanties optimales et adaptées aux besoins des cédantes, est depuis toujours au cœur de l'activité des sociétés de réassurance pratiquant une activité d'apporteur de solutions pérennes. Les différents cadres de régulations, aujourd'hui principalement incarnés par la directive européenne "Solvency II", renforcent toujours d'avantage le besoin des assureurs en terme de solution de gestion des risques et de management du capital.

La plupart des traités de réassurance sont établis par classes de risques. Dans ce cas, quelque soit le type de réassurance choisi, les différentes "*Lines of Business*" sont réassurées séparément en fonction des besoins de cession.

D'autres types de traités proposent des couvertures simultanées de plusieurs "*Lines of Business*". Ce sont principalement des traités "*Stop Loss*" (SL) ou "*Quote-part*" (QP) qui sont observés sur le marché mais, comme nous le verrons dans le cadre de ce mémoire, d'autres formes sont envisageables.

En mettant en place une couverture de type "*multiline*" l'assureur et le réassureur font *de facto* entrer en ligne de compte un effet diversification que l'on peut qualifier d'inter-"*Lines of Business*". La diversification des risques, concept fondateur de toute activité assurantielle, permet de réduire l'incertitude autour de la sinistralité agrégée.

L'intérêt potentiel de ce type de traité est donc évident. Cette conséquence directe de la loi des grands nombres est même un des principaux phénomènes recherchés dans la réassurance "multiline" car l'impact en terme de tarif et de gestion des risques est conséquent. Un traité "Stop Loss" sur l'ensemble d'un portefeuille permet par exemple à une cédante de protéger son *loss-ratio* global. Chacun comprendra que ce traité n'est en aucun cas équivalent à un ensemble de traités "Stop Loss" couvrant individuellement chacune des "Lines of Business" de la cédante.



source : Hannover Re

FIGURE 1.3 – Structure du SCR d'un réassureur : visualisation de l'effet de diversification

Imaginons agréger la sinistralité de 2 "Lines of Business". La cédante est prête à supporter un risque pour ces deux "Lines of Business" jusqu'à hauteur d'un certain montant. Ce qui excède ce montant sera pris en charge par un réassureur (ou un pool de réassureur) jusqu'à une certaine limite. Sans perte de généralité, l'agrégation pourrait se faire sur une tranche en excédent de sinistre. Nous construisons ainsi un MAXL. Un des intérêts des traités ces traités est donc de reproduire - en partie - les effets de diversification observés pour la structure des capitaux de solvabilités requis. L'objet du mémoire est de proposer une méthode permettant de réaliser la tarification sous certaines hypothèses de diversification autres que la simple indépendance. En tenant compte de cette dépendance et en mettant en relation ce programme avec le modèle interne de la cédante, il serait d'ailleurs possible en théorie de définir une couverture optimale pour la cédante, ce mémoire ne couvrira cependant que la tarification et non la recherche de programme optimaux..

S'ajoute à ces intérêts de diversification une série d'avantages liés à l'administration de ce type de traité. On peut évidemment penser à la réduction des coûts de gestion due à l'unicité du traité ainsi qu'à une plus grande collaboration d'entreprise que nécessite la mise en place d'un tel programme de réassurance (obligation de relation business de long terme, favorable à l'échange techniques)

1.3 “Multiline Aggregate Excess of Loss”

1.3.1 Définition

La société QBE Re commercialise depuis plusieurs années des traités de type “Multiline Aggregate Excess of Loss”. Le département de gestion des risques de la société définit ce type de traité de la façon suivante :

"In a "Multiline Aggregate Excess of Loss" treaty, the reinsurer indemnifies the cedant for an aggregate amount of losses in excess of a specified aggregate amount called the global annual aggregate deductible (GAAD). This aggregate amount is made up of the parts of individual losses coming from interior excess of loss layers covering different lines of business (eligible claim activity). The reinsurer can limit its intervention by adding an annual aggregate limit (GAAL)."

Dans la suite nous utiliserons la notation $GAAL \text{ MAXL } GAAD$.

Ces traités sont de puissants outils d’optimisation de la politique de réassurance d’une cédante. Ils ne sont pas particulièrement nouveaux mais restent considérés comme "exotiques". Bien que leur commercialisation soit en pleine expansion ils demeurent marginaux dans le portefeuille du réassureur QBE Re. Le traitement de ces traités n’a donc, comme chacun peut l’imaginer, pas encore donné lieu à une quelconque standardisation. Ces traités sont analysés au cas par cas avec la plus grande prudence en terme de modélisation et de construction de programme.

Nous attirons l’attention du lecteur sur le fait que la classification même de ces traités peut poser des problèmes de communication. En effet, ces traités sont souvent considérés à tort comme des "Stop Loss". L’élément clé faisant que ces traités ne peuvent en aucun cas être considérés comme tel, est l’existence de fonctions d’éligibilité internes - *interior excess of loss layers* - à chacune des "*Lines of Business*".

Pour information, les principales "*Lines of Business*" réassurées par QBE Re dans des traités MAXL sont :

- Responsabilité Civile Automobile
- Responsabilité Civile Générale
- Garantie Accident de la vie (par risque ou par événement)
- Cat-Nat
- Tempête et autres forces de la nature
- Grêle
- Poids de la neige
- Incendie et autres périls
- Terrorisme

1.3.2 Fonctionnement théorique

Soit un traité $GAALMAXL_{GAAD}$ couvrant n "*Lines of Business*". On utilisera dans la suite de ce mémoire les notations suivantes :

- N_i désigne le nombre de sinistres affectant la LoB_i dont le réassureur a connaissance. En effet, dans le cadre d'un traité de réassurance le seuil déclaratif du réassureur par la cédante est fixée contractuellement souvent en pourcentage de la priorité (50%, 70%, 100%).
- x_{i_j} désigne le montant "*From Ground Up*" du j^e sinistre affectant la LoB_i
- f_i désigne la fonction d'éligibilité du programme à appliquer aux sinistres de la LoB_i

Chaque "*Line of Business*" est ainsi caractérisée par une fonction d'éligibilité notée f_i dont la forme dépend des besoins de la cédante. Par exemple, pour un montant de sinistre x donné, on peut avoir :

- $f_i(x) = \min(l_i, \max(x - d_i, 0))$ avec l_i et d_i respectivement limite et priorité de la "*Line of Business*" i , pour une tranche d'éligibilité classique.
- $f_i(x) = x \mathbf{1}_{x > d_i}$ dans le cas d'un "English Deductible" de valeur d_i .
- $f_i(x) = \alpha x$ avec $\alpha \in [0, 1]$ pour une prise en charge proportionnelle

Pour chaque sinistre enregistré, le montant $f_i(x)$ va venir alimenter le traité MAXL. A la fin de la période de couverture, dans la plupart des cas 1 an², le montant total d'alimentation du traité s'exprime sous la forme :

$$\mathbb{S} = \sum_{i=1}^n \mathbb{S}_i = \sum_{i=1}^n \sum_{k=1}^{N_i} f_i(x_{i_k})$$

Le réassureur indemnise alors la cédante à hauteur de :

$$\min(GAAL, \max(\mathbb{S} - GAAD, 0))$$

2. Certaines compagnies de réassurance proposent des traités Multiline-Multiyear, la question de la modélisation de la dépendance temporelle a d'ailleurs été traitée dans Wang, 2013 [33]

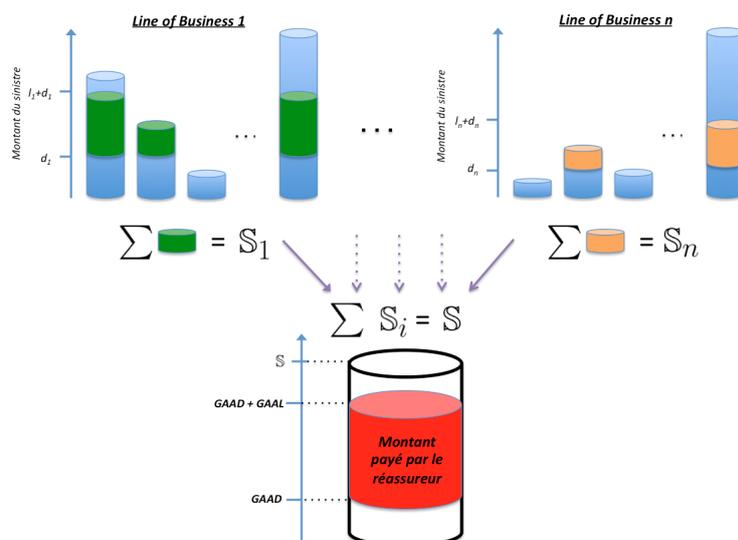


FIGURE 1.4 – Fonctionnement schématisé d'un traité MAXL

Si ces traités ne sont conceptuellement pas particulièrement complexes, de nombreuses difficultés se présentent lors de la tarification. C'est par exemple le cas lorsque le traité est alimenté non pas directement par des sinistres mais par des primes de reconstitutions d'autres traités³ de réassurance de la compagnie. La multiplication des effets non-proportionnels rend, elle, extrêmement complexe l'adéquation de ce type de traité avec la politique de gestion des risques et de réassurance souhaitée par la cédante. Une grande partie du travail de l'actuaire consiste à adapter les différents niveaux d'éligibilité du traité aux besoins de cession de la cédante.

D'autres problématiques de tarification proviennent elles de la notion même d'agrégat de différentes "Lines of Business". On peut penser par exemple aux difficultés liées à la réassurance simultanée de "long tail" et "short tail" business. Des clauses de commutation sont ajoutées à ces traités afin de limiter dans le temps cet effet.

1.3.3 Exemple de fonctionnement d'un MAXL à 3 "Lines of Business"

Considérons l'exemple de la société d'assurance XYZ qui propose à ses clients des protections sur les trois "Lines of Business" suivantes :

- Responsabilité Civile Automobile (MTPL)
- Responsabilité Civile Générale (GTPL)
- Incendie et autres périls (Fire)

3. Dans le cadre de ce mémoire nous ne nous intéresserons pas à ce type de couverture

Elle souhaite se couvrir via un traité ${}_{35M}\text{MAXL}_{15M}$ ayant les tranches d'éligibilité suivantes :

	Priorité	Portée
MTPL	5M	5M
GTPL	2.5M	2.5M
Fire	2M	5M

La sinistralité observée au cours de l'année est la suivante :

- MTPL : 0.5M, 12M, 2M, 8M
- GTPL : 3M, 0.2M, 8M
- Fire : 11M, 7M, 2M

Le traité MAXL va donc être alimenté au titre des différentes "Lines of Business" à hauteur de :

- $\mathbb{S}_{MTPL} = 0 + 5 + 0 + 3 = 8M$
- $\mathbb{S}_{GTPL} = 0.5 + 0 + 2.5 = 3M$
- $\mathbb{S}_{Fire} = 5 + 5 + 0 = 10M,$

Soit un total d'alimentation de $\mathbb{S} = 21M$. Le réassureur versera donc à la cédante la somme de $21M - 15M = 6M$.

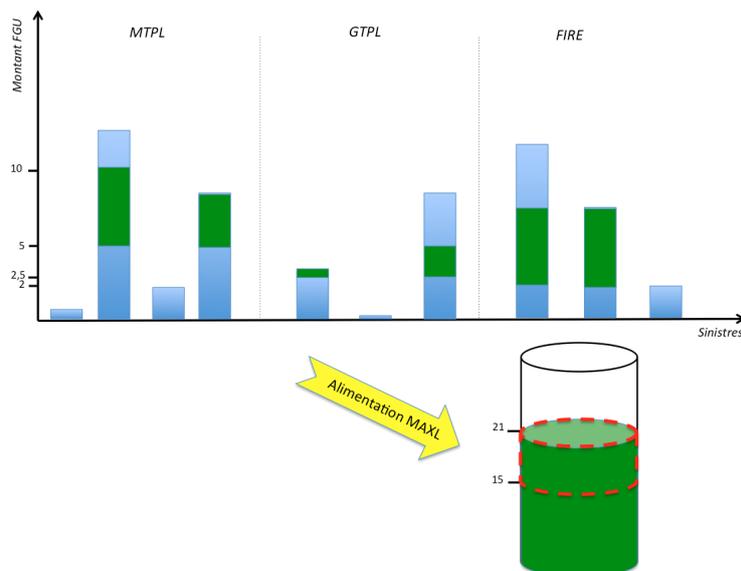


FIGURE 1.5 – Résumé du processus d'alimentation d'un traité MAXL

1.3.4 Un exemple pratique d'utilisation dans un programme

Les traités MAXL peuvent être utilisés dans de nombreuses situations. Ces traités offrent en effet de grandes marges de manœuvre en terme de clauses et de construction de programme. La couverture contre le risque de sur-fréquence est un besoin classique des assureurs et les traités MAXL, quand ils sont calibrés à cet effet, répondent parfaitement à ce besoin.

Imaginons un assureur souhaitant à la fois céder en réassurance XL une partie de ses risques et se protéger contre un risque de sur-fréquence. Une première solution consiste à placer des niveaux de rétention bas sur chacune des branches. Dans de nombreuses situations cela va être coûteux et ne constitue pas une réassurance optimale. Une autre solution consiste à protéger l'intégralité du portefeuille par un traité "Stop Loss", mais cela est difficile à placer et souvent très coûteux. L'utilisation d'un traité MAXL constitue alors une alternative particulièrement intéressante.

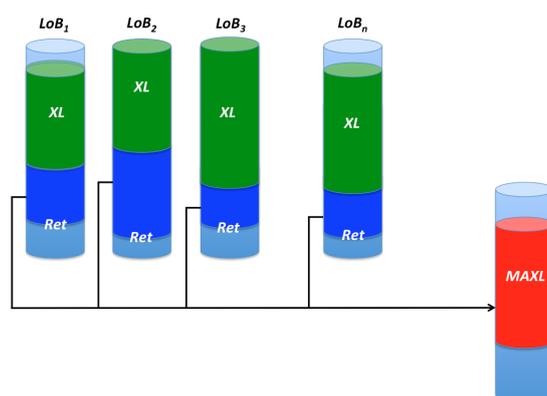


FIGURE 1.6 – Exemple d'un programme de réassurance avec MAXL

La figure 1.6 décrit un programme simple de réassurance dans lequel la cédante se couvre en XL sur chacune de ses "Lines of Business" (partie verte) et avec un traité MAXL sur une partie de la rétention (partie bleu foncé) de chacune de ces "Lines of Business". Ainsi ses objectifs sont atteints car les niveaux de rétention des traités XL peuvent être conservés assez hauts et moyennant une bonne calibration des tranches internes du traité MAXL, le risque de sur-fréquence est capté. Ce travail de calibrage relève généralement d'une analyse technique entre la cédante et le réassureur.

On peut remarquer qu'un traité MAXL couvrant l'intégralité de la rétention de la cédante peut alors, et c'est le cas uniquement dans cette situation, s'assimiler à un traité "Stop Loss" sur l'intégralité du portefeuille.

1.3.5 Le portefeuille QBE Re

Suite au renouvellement 2013, la société QBE Re disposait dans son portefeuille d'une petite quantité de traités MAXL. Les résultats de ces traités étant globalement assez bons, ils font l'objet d'une attention toute particulière de la part des membres des différents départements de la société et notamment celui de Recherche et Développement.

Priority	Cover	Excess ROE	Earned Premium	CR	Shared
220 000 000	110 000 000	16,98%	231,722	61,05%	3,00%
7 250 000	2 500 000	1,67%	142,830	70,03%	25,00%
9 000 000	10 000 000	33,88%	127 500	41,42%	15,00%
95 000 000	20 000 000	-17,82%	86 269	91,62%	7,50%
35 000 000	130 000 000	-2,38%	820 000	82,33%	10,00%
4 800 00	5 300 000	21,64%	199 130	45,85%	40,00%
40 000 000	20 000 000	7,89%	212 500	67,75%	12,50%
10 000 000	14 000 000	6,74%	104 229	52,27%	10,00%
500 000	1 500 000	-8,18%	271 488	85,94%	12,00%
16 000 000	11 500 000	12,22%	44 520	34,50%	7,00%
152 520 235	29 855 025	13,82%	636 042	68,93%	35,00%
75 000 000	80 000 000	-49,04%	85 750	182,33%	1,75%

TABLE 1.1 – Résumé des principaux traités MAXL en portefeuille chez QBE Re

1.3.6 Principe de tarification

1.3.6.1 Simulation des scénarios

La tarification de ces traités est réalisée par **simulation de scénarios sinistralité** via la méthode classique de Monte-Carlo.

Pour chacune des "*Lines of Business*" un modèle de sinistralité est déterminé suite à l'étude des différentes données disponibles et des hypothèses retenues dans les normes de souscription. Il est fréquent d'utiliser des modèles fréquence/sévérité de type Pareto/Poisson, Beta/Poisson, etc. Dans certains cas, le modèle utilisé est mixte c'est à dire que la partie inférieure de la distribution est une distribution empirique déterminée par l'expérience de la cédante et une partie supérieure est déterminée par les modélisations et hypothèses de la compagnie afin de prendre en compte le risque de sinistres graves. Pour les modèles de catastrophes naturelles ce sont, dans la plupart des cas, des logiciels dédiés - RMS notamment - qui sont utilisés.

Exemple :

Suite à une l'étude des sinistralités historiques il a été déterminé que, pour la tarification d'un des traités MAXL du portefeuille, la sinistralité au delà de 500 000 euros de la branche "Property by risk + non natural event" pouvait être modélisée par une loi fréquence/sévérité de type Pareto/Poisson. Avec $N \sim \text{Poisson}(\lambda = 4,277)$ et $\text{Severity} \sim \text{Pareto}(\alpha = 1,0233 \text{ et } A = 500000)$

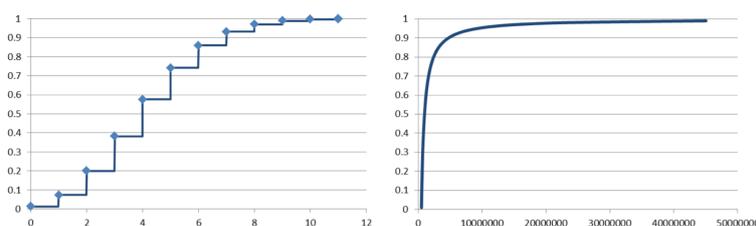


FIGURE 1.7 – Fonction de répartition des lois de fréquence et de sévérité ajustées au risque "Property by risk + non natural event"

Le scénario d'alimentation du traité MAXL est ensuite obtenu en agrégeant les sommes exigibles sur chacune des "*Lines of Business*". Une fois ces différents éléments du programme appliqués, la prime pure est estimée par le montant moyen que le réassureur s'attend à devoir payer à la cédante.

La pratique actuelle de tarification agrège la montant en appliquant une indépendance abusive. Une question se pose quand à l'éventuelle sous-estimation (sur-estimation) des risques en conservant cette hypothèse d'indépendance. **Nous verrons dans la suite des travaux que la méthode proposée dans ce mémoire consiste à ajouter une structure de dépendance lors de l'appariement des scénarios marginaux.** Un des principaux intérêts de la méthode de tarification proposée dans le troisième chapitre de ce mémoire est de prendre en compte cette dépendance entre "*Lines of Business*" sans modifier les modèles de simulation de sinistres par "*Lines of Business*".

1.3.6.2 Patterns de paiement

Ces traités pouvant combiner des "long" et "short tails", l'estimation des "patterns" de paiement n'est pas évidente. Le plus simple reste d'appliquer des "patterns" par "*Lines of Business*".

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
RCAuto	0	8	13	19	30	40	50	60	70	80	90	100	100	100	100	100
RCPPro	0	1	4	5	8	9	21	29	38	50	60	70	80	90	95	100
Inc.	33	71	88	95	97	99	99	100	100	100	100	100	100	100	100	100
Dom. Auto	38	85	94	97	98	99	100	100	100	100	100	100	100	100	100	100

TABLE 1.2 – Exemple de "Pattern" de paiement pour 4 "*Lines of Business*" (en %)

La plupart des traités MAXL contiennent de plus une clause de commutation afin de rendre l'administration du mélange de "*long*" et "*short-tail*" plus simple. Pour rappel, une clause de commutation permet au réassureur de faire un paiement ponctuel qui remplit toutes les obligations auprès de la cédante après une période de temps donnée. L'exercice de la clause de commutation permet à l'entreprise de réassurance de se décharger de toutes les responsabilités en vertu d'une police et ainsi de limiter l'exposition

en cours.

Il est à noter que les phénomènes de dépendance entre "Lines of Business" impactent aussi les triangles de liquation. Ce type de problématique n'est pas directement traité dans ce mémoire mais le lecteur peut se reporter, pour plus d'information, aux travaux de Gillet et Serra, 2001 [17] à propos des effets de la dépendance entre différentes branches sur le calcul des provisions.

1.3.7 Les effets non-proportionnels Excess-of-Loss

Comme nous l'avons décrit précédemment, la prime pure d'un traité MAXL, notée Π , peut s'écrire sous la forme :

$$\Pi = \mathbb{E} [\min (GAAL, \max (\$ - GAAD, 0))]$$

Il apparait immédiatement que ce traité, comme tous les traités "Excess of Loss" combine, après l'agrégation de toutes les "Lines of Business" en un scénario d'alimentation, deux effets non-proportionnels. Le premier est un effet de priorité matérialisé par la présence d'un "Global Aggregate Deductible" (GAAD). Le second, effet de portée est quant à lui induit par la "Global Aggregate Annual Limit" (GAAL).

Afin de mieux visualiser les deux effets, la prime pure peut être réécrite de la sorte :

$$\Pi = \mathbb{E} [\max(0, (\$ - GAAD))] - \mathbb{E} [\max(0, (\$ - (GAAD + GAAL)))]$$

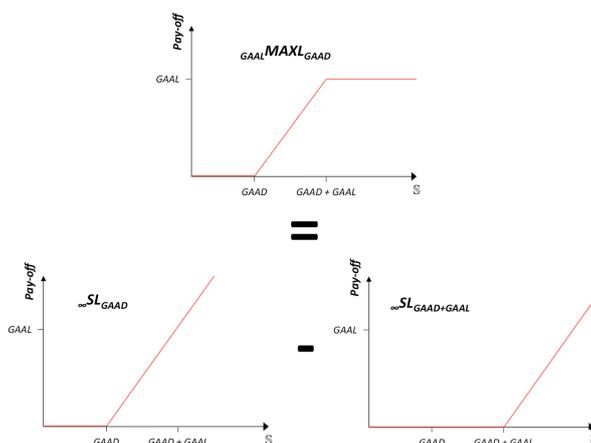


FIGURE 1.8 – Décomposition de la prime pure d'un traité MAXL

Cette nouvelle formulation fait apparaître séparément les deux effets non proportionnels. La prime pure du traité MAXL, s'exprime ainsi comme la différence de deux primes pures "Stop Loss". Attention par prime pure "Stop Loss" nous ne faisons pas référence à un traité "Stop Loss" à proprement parler mais simplement à la fonction

de pay-off $(S - C)^+ = \max(S - C, 0)$ avec C une constante.

Cette décomposition sera utile dans la suite du travail lorsque l'on étudiera l'impact de la relaxation de l'hypothèse d'indépendance sur la prime pure du traité MAXL. En effet nous pourrons visualiser cet impact sur chacun des effets non-proportionnels. De plus, comme nous le verrons dans le chapitre 1.4.2.1, la relation d'ordre dite "*Stop Loss*" a certaines propriétés liées à la mise sous dépendance particulièrement intéressantes.

Ce type de formulation est très fréquemment utilisé dans la création de produits dérivés. Par analogie avec ces produits financiers, un MAXL peut s'assimiler à une position "*Bull Call Spread*". Cette position s'obtenant en adoptant une position "*long*" sur un "*call-option*" de valeur d'exercice (GAAD) et une position "*short*" sur un "*call-option*" de valeur d'exercice (GAAD+GALL).

1.4 Pourquoi s'attarder sur la dépendance entre les risques ?

"L'assureur que je suis a été particulièrement sensible à l'introduction de la théorie de copules qui permet de modéliser la dépendance entre les différents risques auxquels sont confrontés les gestionnaires au sein de nos entreprises. Car, c'est une évidence, lorsque les catastrophes surviennent les sinistres ne sont pas indépendants"

Claude Bébéar, 31 mars 2003

Extrait de la préface de "**Mathématiques de l'assurance non-vie**"

par Michel Denuit et Arthur Charpentier

1.4.1 Quelques exemples intuitifs

Les éléments suivants présentent certaines idées intuitives laissant penser que certaines "*Lines of Business*" ne peuvent pas être considérées comme indépendantes.

- Les événements de type force de la nature impactent directement plusieurs "*Lines of Business*". Les branches dommages matériels (automobile, habitation,...), sinistres corporels graves, perte d'exploitations agricoles, "worker compensation" seront de près ou de loin, directement ou indirectement affectées par l'événement.
- Les événements naturels entre eux sont aussi dépendants. Les tempêtes de vent et la grêle ou les tempêtes et les inondations sont également en effet corrélées simplement du fait d'éléments physiques météorologiques. Dans le portefeuille QBE Re, la terrible tempête "Lothar et Martin" a causé des dommages en raison de la vitesse du vent élevée, mais aussi dans une moindre mesure en raison de la grêle et des inondations. Globalement, les mauvaises années en terme de tempête sont généralement d'assez mauvaises années en terme de grêle et d'inondation.
- Un dernier exemple lié aux événements naturels provient plus de la classification des sinistres qu'à la dépendance factuelle à proprement parler. Cela fut particulièrement le cas lors de la tempête Klaus entre les branches "Fire" et "Motor Own Damage (MOD)".
- En Responsabilité Civile Automobile et Responsabilité Civile Générale les sinistres graves sont souvent des sinistres corporels. Toute modification (législative, sociétale,...) dans la façon de calculer les indemnisations, entrainera, si elle n'est pas anticipée dans les modèles de tarification, une dépendance positive de ces deux "*Lines of Business*".
- Une dégradation du contexte économique global engendre aussi un certain nombre de changements de comportement des assurés et des personnes les conseillant (avocats,...). En temps de crise, les demandes de prise en charge par l'assureur augmentent assez significativement, à cela s'ajoute l'intégralité des actes frauduleux tels que les incendies volontaires ou les fausses déclarations. Cependant, si l'on se concentre sur les grands sinistres, ce point ne semble pas absolument essentiel.

On pourrait multiplier les explications de la sorte. Pour un réassureur, traitant de fait des sinistres importants, l'impact de ces dépendances peut être considérable et doit faire l'objet de recherches approfondies. Plusieurs éléments rendent le travail de modélisation difficile et le manque de données exploitables est comme souvent une des principales problématiques.

S'ajoute à ces difficultés le fait que le niveau de dépendance va considérablement dépendre du périmètre réassuré en terme de montant de sinistre et de localisation notamment. La dépendance entre les risques sur un territoire extrêmement concentré (grande ville,...) ne sera évidemment pas la même que dans des contrées plus reculées⁴. Les résultats doivent donc être transposés d'un traité à l'autre avec la plus grande prudence. La dépendance est aussi fonction du montant des sinistres que l'on considère, aucune dépendance n'est à prévoir (hormis l'impact des prix pratiqués sur le marché des garagistes) pour des petits accrochages automobiles, alors que la dépendance est forte lors d'une catastrophe touchant la branche automobile⁵.

Dans la suite de ce travail, nous ne considérerons que des dépendances positives. L'existence de dépendances négatives peut parfois se justifier comme entre les branches décès et invalidité mais dans le cadre de l'étude nous ne les intégrerons pas. Dans le pire de cas lorsqu'une dépendance négative est suspectée nous considérerons les risques comme indépendants.

1.4.2 A quel(s) effet(s) faut-il s'attendre ?

1.4.2.1 Un effet difficile à prévoir

L'effet de la dépendance sur ce type de traité est en fait extrêmement complexe à prévoir du fait de la multiplication des effets non proportionnels.

Intuitivement nous avons tendance à considérer qu'une dépendance positive, réduisant les effets de diversification, va entraîner une augmentation du risque et donc de la prime pure. Dans le cas d'un traité MAXL ce raisonnement n'est pas valable car l'effet combiné du GAAD et du GAAL n'est pas prévisible.

Afin de bien saisir l'origine de cette imprédictibilité, revenons sur la notion de relation d'ordre "Stop Loss" et de "positive cumulative dependence".

Définition 1.1. (*Stop-loss order*)

Étant donné 2 risques X et Y , on dit que X domine Y dans la relation d'ordre "Stop Loss", notée $Y \preceq_{SL} X$, si quelque soit la fonction ϕ , convexe, croissante d'espérance définie :

$$\mathbb{E}[\phi(X)] \leq \mathbb{E}[\phi(Y)]$$

4. Les disparités géographiques ne sont pas traitées dans ce mémoire mais devraient faire l'objet de recherches approfondies

5. Cet exemple montre l'importance d'utiliser des modèles de dépendance variable le long des distributions des sinistres : les copules répondent parfaitement à cet impératif

Définition 1.2. (Positive Cumulative Dependence PCD)

Les risques X_1, \dots, X_n sont dit PCD si $\forall j \in \{1, \dots, n\}, \forall (x_1, x_2) \in \mathbb{R}^2$

$$\mathbb{P} \left[\sum_{\substack{i=1 \\ i \neq j}}^n X_i > x_1, X_j > x_2 \right] \geq \mathbb{P} \left[\sum_{\substack{i=1 \\ i \neq j}}^n X_i > x_1 \right] \mathbb{P} [X_j > x_2]$$

Théorème 1.1. Si les risques X_1, \dots, X_n sont PCD et en notant $(X_i)^\perp$ et $(X_i)^\uparrow$ les mêmes risques marginaux respectivement en situation d'indépendance et de comonotonie, alors

$$\sum X_i^\perp \preceq_{SL} \sum X_i \preceq_{SL} \sum X_i^\uparrow$$

Pour la démonstration de ce résultat voir Denuit et al., 2001 [9]

Comme nous l'avons vu, la prime pure du traité MAXL ne s'exprime pas comme une fonction croissante et convexe de la somme des risques considérés, mais comme la différence de deux fonctions respectant ces propriétés.

$$\Pi = \mathbb{E} [\max(0, (\mathbb{S} - GAAD))] - \mathbb{E} [\max(0, (\mathbb{S} - (GAAD + GAAL)))]$$

$$\Pi = \mathbb{E} [\phi_{GAAD}(\mathbb{S})] - \mathbb{E} [\phi_{GAAD+GAAL}(\mathbb{S})]$$

Avec $\phi_a(X) = (X - a)^+$, ϕ est bien une fonction convexe et croissante et son espérance est définie lorsque celle de X l'est.

Ainsi les deux effets que l'on visualise dans cette équation vont être impactés par une dépendance positive entre les risques réassurés. On aura alors :

$$\mathbb{E} [\phi_{GAAD}(\mathbb{S}^\perp)] \leq \mathbb{E} [\phi_{GAAD}(\mathbb{S})] \leq \mathbb{E} [\phi_{GAAD}(\mathbb{S}^\uparrow)]$$

$$\mathbb{E} [\phi_{GAAD+GAAL}(\mathbb{S}^\perp)] \leq \mathbb{E} [\phi_{GAAD+GAAL}(\mathbb{S})] \leq \mathbb{E} [\phi_{GAAD+GAAL}(\mathbb{S}^\uparrow)]$$

La prime étant une différence de ces deux effets, il est impossible de prédire ce que sera l'effet global de la dépendance positive. Cet effet est en fait fonction de la forme et de la dépendance et de la tranche (GAAD, GAAD+GAAL). Une augmentation, une diminution voire aucun changement de la prime pure peuvent être observés en cas de dépendance positive. Cela rend le travail de l'actuaire particulièrement complexe et doit être correctement intégré à la politique de souscription de type de traité.

Il est en revanche évident qu'une dépendance positive entre les risques va impacter la distribution de \mathbb{S} or c'est souvent à partir de cette distribution que sont déterminés

les GAAD et GAAL. On place par exemple souvent le montant GAAD+GAAL à un quantile de 99,5%. La dépendance va ainsi avoir un impact quasi-certain sur la calibration des programmes de réassurance.

1.4.2.2 Illustration :

Afin d'illustrer ces phénomènes, prenons un exemple extrêmement simplifié de la somme de 3 lois normales $X_1, X_2, X_3 \sim \mathcal{N}(10, 3)$. Nous étudions \mathbb{S}^\perp , \mathbb{S}^\uparrow puis \mathbb{S} avec comme structure de dépendance une copule⁶ gaussienne et une matrice de corrélations

$$\begin{pmatrix} 1 & 0.2 & 0.5 \\ 0.2 & 1 & 0.3 \\ 0.5 & 0.3 & 1 \end{pmatrix}$$

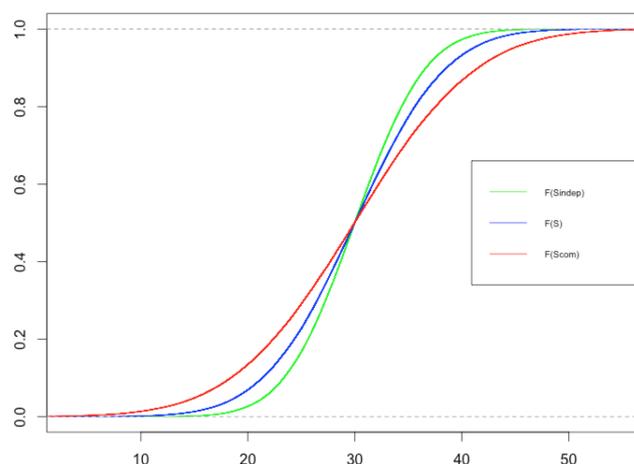


FIGURE 1.9 – Fonctions de répartition des variables \mathbb{S}^\perp , \mathbb{S}^\uparrow et \mathbb{S}

On observe que pour $\alpha < 0,5$ (*VaR* : *value-at-risk*) :

$$VaR_\alpha(\mathbb{S}^\uparrow) < VaR_\alpha(\mathbb{S}) < VaR_\alpha(\mathbb{S}^\perp)$$

et que l'inégalité s'inverse pour $\alpha > 0,5$

Si le programme était construit de telle sorte que $GAAD = 20$ et $GAAL = 15$ la dépendance aurait pour effet de diminuer la prime pure :

$$\Pi^\perp = 9.58 \text{ et } \Pi = 9,29$$

Si maintenant le traité réassurait avec un $GAAD = 35$ et un $GAAL = 15$, alors dans ce cas, la dépendance augmentait la prime pure :

$$\Pi^\perp = 0.46 \text{ et } \Pi = 0.87$$

6. Les structures de dépendance de type copule font l'objet du second chapitre de ce mémoire

Chapitre 2

Cadre théorique de la dépendance multivariée

Dans ce chapitre nous présentons de façon théorique les outils mathématiques qui seront utilisés dans la méthode de tarification. Un focus est notamment fait sur les trois grandes approches copules utilisées pour mettre en place la structure de dépendance entre les différentes "*Lines of Business*".

La première d'entre elle, la famille de copule de Fréchet, est un modèle extrêmement simple de pondération convexe entre indépendance et comonotonie. La seconde approche repose sur des copules multivariées classiques calibrables à partir d'hypothèses de dépendance "*pairwise*" : copule gaussienne multivariée, copule t et Individuated-t.

Enfin la dernière approche est la plus novatrice, elle permet d'appréhender le caractère multidimensionnel de la problématique qu'à travers l'étude des dépendance "*pairwise*". En d'autres termes dans cette méthode la structure multivariée est spécifiée par l'ensemble des copules de dimension 2.

Sommaire

2.1	Les copules	21
2.2	Bornes de Fréchet et Indépendance	23
2.3	Mesure de dépendance bivariable	24
	2.3.0.3 Le coefficient de corrélation de Pearson	24
	2.3.0.4 Le τ de Kendall	25
2.3.1	La dépendance de queue	25
2.4	Une mesure de dépendance pour l'agrégation de risques	27
2.4.1	Principales propriétés	27
2.4.2	Estimation	28
2.5	Les différentes approches de modélisation copule	30
2.5.1	Première approche : Famille de copule de Fréchet	30
2.5.2	Seconde approche : copules multivariées	31
2.5.2.1	Copule Gaussienne multivariée	31

2.5.2.1.1	Caractéristiques	31
2.5.2.1.2	Exemple	31
2.5.2.1.3	Intérêt	32
2.5.2.2	Copule t multivariée	32
2.5.2.2.1	Définition	33
2.5.2.2.2	Dépendance de queue	33
2.5.2.2.3	Simulation	33
2.5.2.2.4	Exemple	33
2.5.2.2.5	Intérêt	34
2.5.2.3	Copule Individuated-t	34
2.5.2.3.1	Définition	34
2.5.2.3.2	Simulation	35
2.5.2.3.3	Dépendance de queue	36
2.5.2.3.4	Exemple	36
2.5.2.3.5	Intérêts	37
2.5.3	Troisième approche : Dépendance "pairwise" et heuristique de Kaut	37
2.5.3.1	L'heuristique de Kaut	38
2.5.3.1.1	Cas n=2	40
2.5.3.1.2	Extension au cas multivarié	41
2.5.3.2	Copules bivariées usuelles	42
2.5.3.3	Exemple illustratif	45

2.1 Les copules

En latin, le terme *copula* désigne un lien, un nœud ou une alliance, en d'autres termes un outil ou une idée bâtissant une relation entre différents éléments. Dans sa version mathématique, une copule est une fonction dont l'objet est précisément de lier des lois marginales de variables aléatoires entre elles et ainsi construire une loi jointe.

La théorie des copules est assez récente et a connu dans les 30 dernières années un essor considérable, elle trouve son origine dans des travaux de théorie des espaces métriques aléatoires de M. Fréchet au milieu du 20^e siècle. En 1959¹, M. Sklar est le premier à employer le terme "copule" pour décrire la structure de dépendance d'une distribution multivariée. Dans un désormais célèbre théorème portant son nom, il présente la séparation des lois marginales et de la structure de dépendance sous un angle nouveau et ouvre ainsi le champ à de nombreuses recherches et applications.

Ces outils sont toutefois restés pas ou peu utilisés jusqu'à ce qu'au début des années 80, M. Genest et M. MacKay, publient *The joy of copulas : Bivariate distributions with uniform marginals*, Genest et Mackay, 1986 [15].

Les copules sont aujourd'hui particulièrement utiles en modélisation financière et/ou assurantielle et font l'objet de nombreux travaux de recherche.

Définition 2.1. Une copule C , de dimension n , est une fonction croissante et continue à droite de l'hypercube $[0, 1]^n$ dans l'intervalle $[0, 1]$, satisfaisant les propriétés suivantes :

$$(i) C(u_1, \dots, i_{i-1}, 0, u_{i+1}, \dots, u_n) = 0 \text{ pour } i = 1, \dots, n$$

$$(ii) C(1, \dots, 1, u_i, 1, \dots, 1) = u_i \text{ pour } i = 1, \dots, n$$

$$(iii) \text{ Pour tout } \alpha, \beta \in [0, 1]^n \text{ avec } \forall i \alpha_i < \beta_i :$$

$$\forall u \in [0, 1]^n \Delta_{\alpha_1, \beta_1} \Delta_{\alpha_2, \beta_2} \dots \Delta_{\alpha_n, \beta_n} C(u) \geq 0$$

Notation : soit F_X la fonction de répartition d'une variable aléatoire X

$$\Delta_{\alpha_i, \beta_i} F_X(x) = F_X(x_1, \dots, x_{i-1}, \beta_i, x_{i+1}, \dots, x_n) - F_X(x_1, \dots, x_{i-1}, \alpha_i, x_{i+1}, \dots, x_n)$$

Théorème 2.1. (Sklar - 1959) Toute fonction de répartition F_X d'une distribution multivariée (vecteur aléatoire) $X = (X_1, \dots, X_n)$ de marginales F_1, \dots, F_n peut être représentée en terme de copule C sous la forme :

$$F_X(x) = C(F_1(x_1), \dots, F_n(x_n)), \quad x \in \mathbb{R}^n$$

1. Sklar, 1959 [30]

Si de plus les marginales sont continues alors la fonction copule C est unique et peut être explicitement décrite sous la forme :

$$C(u) = F_X (F_1^{-1}(u_1), \dots, F_n^{-1}(u_n)), \quad u \in [0, 1]^n$$

Si les marginales ne sont pas continues, l'utilisation des inverses généralisées permet de conserver la forme précédente mais l'on perd l'unicité de la structure de la fonction copule.

Ce théorème s'avère jouer un rôle central dans les travaux de modélisations de risques. En effet, il permet en quelque sorte d'aborder la modélisation d'une distribution de risques multivariés en deux étapes. La première consistant à identifier les distributions marginales de chacun des risques considérés et la seconde à analyser la structure de dépendance sous-jacente (fonction copule).

Proposition 2.1. *La fonction de densité associée à une fonction copule C est définie presque partout sur $[0, 1]^n$ (par presque partout on entendra partout sauf sur des éléments de mesure de Lebesgue nulle) et s'exprime sous la forme :*

$$c(u) = \frac{\partial^n}{\partial u_1 \dots \partial u_n} C(u)$$

Proposition 2.2. *La fonction de densité d'une distribution multivariée X de fonction copule C et de marginales F_1, \dots, F_n s'exprime alors comme :*

$$f_X(x) = \left(\prod_{i=1}^n f_i(x_i) \right) c(F_1(x_1), \dots, F_n(x_n))$$

Théorème 2.2. *(Invariance) Soit X un vecteur aléatoire de dimension n de fonction copule C et t_1, \dots, t_n des fonctions croissantes et continues. Alors le vecteur aléatoire :*

$$(t_1(X_1), \dots, t_n(X_n))$$

a la fonction C pour copule.

Le théorème précédent, dit théorème d'invariance, présente une caractéristique particulièrement attrayante des structures de dépendance de type copule. En effet, la structure de dépendance, entièrement décrite par la fonction copule, est invariante par transformation continue et croissante des lois marginales. L'inflation et la super-inflation de certaines branches n'ont donc pas d'impact sur la structure de dépendance à proprement parler.

2.2 Bornes de Fréchet et Indépendance

Nous décrivons dans la suite de ce chapitre une série de copules utiles à nos modélisations. Cependant, quelques modèles de dépendance méritent d'être mis en lumière tant ils jouent un rôle important dans l'approche de notre travail. En premier lieu les bornes de Fréchet. Elles représentent des situations extrêmes et particulièrement maniables de dépendance. On a ainsi la copule borne supérieure de Fréchet, notée C_U ou C^\uparrow qui est définie comme :

$$C^\uparrow(u) = \min\{u_1, \dots, u_n\}, \quad u \in [0, 1]^n$$

Proposition 2.3. *Un vecteur aléatoire est comonotonique si et seulement si sa fonction copule C vérifie $C = C^\uparrow$*

Théorème 2.3. *Toute copule C vérifie l'inégalité :*

$$\max \left\{ \sum_{i=1}^n u_i - (n-1), 0 \right\} \leq C(u) \leq C^\uparrow, \quad \forall u \in [0, 1]^n$$

Proposition 2.4. *La borne inférieure de l'inégalité précédente est une généralisation de la copule bivariée $C_L = \max\{u_1 + u_2 - 1, 0\}$. Lorsque la dimension est strictement supérieure à 2, cette borne inférieure ne satisfait pas toujours les conditions pour être une fonction copule.*

Dans le cadre de notre étude, nous ne considérons que des dépendances positives (nous n'envisageons pas le cas des risques se compensant les uns les autres). La comonotonie doit être vue comme le cas le plus extrême de dépendance. Elle correspond à une situation dans laquelle les événements sont simulés au même niveau de risque (ou période de retour) sur toutes les branches considérées. En d'autres termes pour un scénario donné s , $\forall i \in \{1, \dots, n\}$, $F_i(X_i(\text{scénario } s)) = p(\text{scénario } s)$, avec $p(\text{scénario } s)$ niveau de quantile simulé pour toutes les marginales du scénario s .

Une autre copule va alors jouer un rôle particulièrement déterminant. Il s'agit de la copule d'indépendance. Lorsqu'une méthode de tarification ne tient pas compte de la dépendance entre les différentes "Lines of Business" c'est en fait implicitement cette copule qui est utilisée. Si l'on suppose qu'aucune dépendance négative n'est envisageable alors la somme de différents risques est optimalement diversifiée sous la copule d'indépendance.

Définition 2.2. *La copule indépendante, notée C^\perp , est définie par :*

$$C^\perp = \prod_{i=1}^n u_i, \quad u \in [0, 1]^n$$

Proposition 2.5. *Si un vecteur aléatoire possède pour copule la copule C^\perp alors les variables aléatoires X_i sont mutuellement indépendantes.*

2.3 Mesure de dépendance bivariée

C'est en 1984 que M. Scarsini a clairement défini les caractéristiques qu'une mesure d'association entre deux variables aléatoires devait respecter. L'idée sous-jacente au concept de concordance peut intuitivement s'appréhender en considérant que les variables aléatoires X et Y sont positivement concordantes si les grandes valeurs de X vont de pair avec les grandes valeurs de Y .

Définition 2.3. *La fonction $r(.,.)$ associant un nombre réel à tout couple de variables aléatoires (X, Y) est une mesure de dépendance si elle vérifie les conditions suivantes :*

- (i) $r(X, Y) = r(Y, X)$ (symétrie)
- (ii) $-1 \leq r(X, Y) \leq 1$ (normalisation)
- (iii) $r(X, Y) = 1 \Leftrightarrow X$ et Y sont comonotones
- (iv) $r(X, Y) = -1 \Leftrightarrow X$ et Y sont anticomonotones
- (v) Si g est une fonction réelle strictement croissante $r(g(X), g(Y)) = r(X, Y)$

Ajouter à ces propriétés celle selon laquelle X et Y sont indépendantes si et seulement si $r(X, Y) = 0$ serait intéressant mais est malheureusement en contradiction avec la condition (v).

Dans les paragraphes suivants nous discutons de plusieurs mesures couramment utilisées et particulièrement utiles pour la calibration des différents modèles de dépendance.

2.3.0.3 Le coefficient de corrélation de Pearson

Le coefficient de corrélation de Pearson capture la dépendance linéaire au sein d'un couple de variables aléatoires. Il ne respecte cependant que les conditions (i) et (ii) et n'est donc pas une mesure de concordance. Cependant ce coefficient est particulièrement utile pour la calibration des modèles gaussiens.

Définition 2.4. *Pour un couple de variables aléatoires ayant des marginales à variance finie le coefficient de corrélation de Pearson's ρ_p s'exprime sous la forme :*

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

Il est connu que cette mesure doit être utilisée avec la plus grande prudence notamment car une corrélation de Pearson nulle n'implique pas nécessairement l'indépendance des variables.

Proposition 2.6. *En terme de copule le coefficient de corrélation de Pearson s'exprime sous la forme :*

$$\rho_p(X, Y) = \frac{1}{\sqrt{\text{Var}(X)\text{Var}(Y)}} \int_0^1 \int_0^1 (C(u_1, u_2) - u_1 u_2) dF_1^{-1}(u_1) dF_2^{-1}(u_2)$$

Proposition 2.7. ρ_p dépend de la structure copule ET des distributions marginales

2.3.0.4 Le τ de Kendall

Dans le cadre général, et notamment lorsque la dépendance ne relève pas d'un caractère linéaire, la corrélation de Pearson n'est pas un indicateur de la dépendance efficace. Les actuaires préfèrent, dans de nombreux cas, utiliser des corrélations dites de rang. Le coefficient de corrélation de Kendall - ou τ de Kendall - est une mesure d'association non paramétrique basée sur le nombre de paires concordantes et discordantes d'un échantillon.

On dit que deux paires d'observations de (X, Y) sont concordantes si une domination en X va de paire avec une domination en Y . Dans le cas inverse on dit que les paires sont discordantes.

On peut donc écrire :

$$\mathbb{P}(\text{concordance}) = \mathbb{P}[(X - X')(Y - Y') > 0]$$

$$\mathbb{P}(\text{discordance}) = \mathbb{P}[(X - X')(Y - Y') < 0]$$

Définir une mesure de dépendance à partir de ces probabilités assure que la mesure ne dépendra que de la copule sous-jacente et non des marginales. On se protège ainsi des effets indésirables des observations extrêmes et autres "outliers".

Définition 2.5. *Le coefficient τ de Kendall est défini pour un couple de variables aléatoires (X, Y) comme :*

$$\tau = \mathbb{P}(\text{concordance}) - \mathbb{P}(\text{discordance})$$

Proposition 2.8. *Le coefficient τ de Kendall vérifie les propriétés (i), (ii), (iii), (iv), (v)*

Proposition 2.9. *X et Y sont indépendants $\rightarrow \tau(X, Y) = 0$*

Proposition 2.10. *Si le couple de variables aléatoires (X, Y) a pour copule C alors :*

$$\tau(X, Y) = 4 \int_0^1 \int_0^1 C(x, y) dC(x, y) - 1$$

Le ρ de Spearman, l'indice de Gini ou le coefficient de Blomqvist sont des mesures couramment employées dans la littérature. Cependant, dans le cadre de cette étude nous utiliserons uniquement la corrélation linéaire et le τ de Kendall.

2.3.1 La dépendance de queue

La mesure de dépendance de queue est spécifique à l'étude des dépendances au niveau des valeurs extrêmes. Chacun comprendra donc son importance dans le cadre d'une activité de réassurance.

Contrairement à la corrélation linéaire, au τ de Kendall et aux autres mesures de dépendance évoquées qui mesurent la dépendance sur l'ensemble de la distribution, le coefficient de dépendance de queue est une mesure locale de la dépendance sur les extrémités de distribution.

Définition 2.6. *Le coefficient de dépendance de queue inférieure ou lower tail dependence coefficient de deux variables aléatoires X et Y , de fonctions de répartition respectives F_X et F_Y , est défini lorsqu'il existe par*

$$\lambda_L(X, Y) = \lim_{\alpha \rightarrow 0^+} \mathbb{P} [X \leq F_X^{-1}(\alpha) \mid Y \leq F_Y^{-1}(\alpha)]$$

Définition 2.7. *Le coefficient de dépendance de queue supérieure ou upper tail dependence coefficient de deux variables aléatoires X et Y , de fonctions de répartition respectives F_X et F_Y , est défini lorsqu'il existe par*

$$\lambda_U(X, Y) = \lim_{\alpha \rightarrow 1^-} \mathbb{P} [X > F_X^{-1}(\alpha) \mid Y > F_Y^{-1}(\alpha)]$$

On dit que les variables aléatoires X et Y sont asymptotiquement dépendantes au niveau supérieur de la queue de distribution si $\lambda_U \in]0; 1]$ et asymptotiquement indépendantes au niveau supérieur de la queue de distribution si $\lambda_U = 0$. Lorsque $\lambda_U \in]0; 1]$, une interprétation simple peut être donnée lorsque l'on étudie la sinistralité extrême concomitante sur deux branches de risque : sachant qu'un sinistre extrême est survenu dans une branche, il existe une probabilité non nulle qu'un sinistre d'une intensité relativement comparable survienne concomitamment sur l'autre branche.

Proposition 2.11. *Soient X et Y deux variables aléatoires de copule C , alors on a :*

$$\lambda_L(X, Y) = \lim_{u \rightarrow 0^+} \frac{C(u, u)}{u} \text{ et } \lambda_U(X, Y) = \lim_{u \rightarrow 1^-} \frac{1 - 2u + C(u, u)}{1 - u}$$

2.4 Une mesure de dépendance pour l'agrégation de risques

Mesurer la dépendance entre deux variables aléatoires est quelque chose de relativement intuitif et les outils présentés dans la partie précédente ne posent pas énormément de problèmes conceptuels. Dans le cas multivarié, la problématique n'est absolument plus la même. Le choix de la mesure de dépendance doit impérativement être fait en conscience des questions que l'on souhaite traiter. La problématique au cœur de l'étude étant avant tout une question d'agrégation de risques, la mesure de dépendance que l'on utilise a été choisie pour ses caractéristiques dans l'étude de la somme de variables aléatoires.

Cette mesure a été proposée dans Dhaene *et al.*, 2014 [13] et répond parfaitement à nos besoins. En effet, si la plupart des mesures de dépendance proposée dans la littérature repose sur la structure copule ou la distribution jointe, les auteurs proposent ici de directement s'intéresser à la somme des risques considérés $\mathbb{T} = \mathbb{T}_1 + \dots + \mathbb{T}_d$.

Définition 2.8. *La mesure de dépendance, ρ_c , d'un vecteur aléatoire $(\mathbb{T}_i)_{i \in \{1, \dots, d\}}$ aux marginales non dégénérées est alors définie comme :*

$$\begin{aligned} \rho_c((\mathbb{T}_i)_{i \in \{1, \dots, d\}}) &= \frac{\text{Var}(\mathbb{T}) - \text{Var}(\mathbb{T}^\perp)}{\text{Var}(\mathbb{T}^\uparrow) - \text{Var}(\mathbb{T}^\perp)} \\ &= \frac{\sum_{i=1}^d \sum_{j < i} \text{Cov}(\mathbb{T}_i, \mathbb{T}_j)}{\sum_{i=1}^d \sum_{j < i} \text{Cov}(\mathbb{T}_i^c, \mathbb{T}_j^c)} \end{aligned}$$

2.4.1 Principales propriétés

Définition 2.9. *Un vecteur aléatoire $X = (X_1, \dots, X_d)$ est dit inférieur au vecteur aléatoire $Y = (Y_1, \dots, Y_d)$ pour la relation d'ordre de concordance \leq_c , notée $X \leq_c Y$, si à la fois :*

$$\mathbb{P}(X_1 \leq t_1, \dots, X_d \leq t_d) \leq \mathbb{P}(Y_1 \leq t_1, \dots, Y_d \leq t_d)$$

et

$$\mathbb{P}(X_1 > t_1, \dots, X_d > t_d) \leq \mathbb{P}(Y_1 > t_1, \dots, Y_d > t_d)$$

$$\forall (t_1, \dots, t_d) \in \mathbb{R}^d$$

Ainsi, pour tout vecteur aléatoire $X = (X_1, \dots, X_d)$ et $Y = (Y_1, \dots, Y_d)$, ρ_c a les propriétés suivantes :

Proposition 2.12. *Si X a une structure de dépendance comonotonique alors*

$$\rho_c(X) = 1$$

Proposition 2.13. *Si les éléments de X sont indépendants alors*

$$\rho_c(X) = 0$$

Proposition 2.14. *Si $X \leq_c Y$ alors $\rho_c(X) \leq \rho_c(Y)$*

Proposition 2.15. *Pour toute permutation (i_1, \dots, i_d) de $(1, \dots, d)$ on a*

$$\rho_c(X_{i_1}, \dots, X_{i_d}) = \rho_c(X_1, \dots, X_d)$$

$$\rho_c(-X_1, \dots, -X_d) = \rho_c(X_1, \dots, X_d)$$

Théorème 2.4. *Si la copule sous-jacente d'un vecteur aléatoire \mathbf{X} s'écrit sous la forme :*

$$C(x) = \alpha C^\uparrow(x) + (1 - \alpha)C^\perp(x)$$

Alors $\rho_c(X) = \alpha$

Pour la démonstration de ce résultat voir Dhaene et al., 2014 [13].

2.4.2 Estimation

Dans cette section nous présentons une méthode d'estimation de ρ_c à partir d'un échantillon. Cette estimation repose sur l'utilisation des estimateurs classiques des deux premiers moments d'une distribution.

Ainsi, si l'on considère $\{(t_{i1}, \dots, t_{id})\}_{i=1, \dots, n}$, un échantillon de taille n d'une distribution de dimension d , nous avons :

$$\forall j \in 1, \dots, d, \bar{t}_j = \frac{1}{n} \sum_{i=1}^n t_{ij}$$

$$\forall i \in 1, \dots, n, \mathbb{t}_i = \sum_{j=1}^d t_{ij}$$

$$\bar{\mathbb{t}} = \frac{1}{n} \sum_{i=1}^n \mathbb{t}_i = \sum_{i=1}^n \sum_{j=1}^d t_{ij} = \sum_{j=1}^d \bar{t}_j$$

Étant donné que $Var(\mathbb{T}^\perp) = \sum_{j=1}^d Var(\mathbb{T}_j)$ (variance en situation d'indépendance), $Var(\mathbb{T})$ et $Var(\mathbb{T}^\perp)$ peuvent respectivement s'estimer par :

$$\frac{1}{n-1} \sum_{i=1}^n (\mathbb{t}_i - \bar{\mathbb{t}})^2 \quad \text{et} \quad \sum_{j=1}^d \frac{1}{n-1} \sum_{i=1}^n (t_{ij} - \bar{t}_j)^2$$

L'estimation de $Var(\mathbb{T}^\uparrow)$ est un peu plus problématique. L'idée pour estimer cette valeur est de préalablement transformer les scénarios d'observations de telle sorte à obtenir une structure sous-jacente comonotone. Il suffit pour cela de "trier" les scénarios par ordre croissant. De façon formelle cela revient à considérer l'échantillon trié : $\{t_{(i)1}, \dots, t_{(i)d}\}_{i=1, \dots, n}$ où $t_{(i)j}$ est la i^e statistique d'ordre de la variable \mathbb{T}_j . Ainsi $\{t_{(i)1} + \dots + t_{(i)d}\}_{i=1, \dots, n}$ constitue un échantillon de \mathbb{T}^\uparrow .

D'après les propriétés des structures comonotones on sait que :

$$\forall p \in [0, 1], F_{\mathbb{T}^\uparrow}^{\leftarrow}(p) = \sum_{j=1}^d F_{\mathbb{T}_j}^{\leftarrow}(p)$$

En remplaçant la fonction quantile par son estimateur classique et en considérant $p = \frac{i-0.5}{n}$ on obtient :

$$\begin{aligned} \forall i \in \{1, \dots, n\} \quad t_i &= \hat{F}_{\mathbb{T}^\uparrow}^{\leftarrow} \left(\frac{i-0.5}{n} \right) \\ &= \sum_{j=1}^d F_{\mathbb{T}_j}^{\leftarrow} \left(\frac{i-0.5}{n} \right) \\ &= \sum_{j=1}^d t_{(i)j} \end{aligned}$$

$$\text{Comme : } \bar{t}^\uparrow = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^d t_{(i)j} = \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^d t_{ij} = \bar{t}$$

On a alors $Var(\mathbb{T}^\uparrow)$ qui s'estime par la valeur :

$$\frac{1}{n-1} \sum_{i=1}^n \sum_{j=1}^d (t_{(i)j} - \bar{t})^2$$

Au final, un estimateur de la mesure de dépendance multivariée proposée dans Dhaene *et al.*, 2014 [13] est :

$$\hat{\rho}_c = \frac{\sum_{i=1}^n \left[\left(\sum_{j=1}^d t_{ij} - \bar{t} \right)^2 - \sum_{j=1}^d (t_{ij} - \bar{t}_j)^2 \right]}{\sum_{i=1}^n \left[\left(\sum_{j=1}^d t_{(i)j} - \bar{t} \right)^2 - \sum_{j=1}^d (t_{ij} - \bar{t}_j)^2 \right]}$$

L'estimation de cette mesure est théoriquement assez simple mais nécessite dans les fait un historique profond et crédible pour toutes les "*Lines of Business*".

2.5 Les différentes approches de modélisation copule

2.5.1 Première approche : Famille de copule de Fréchet

Comme nous l'avons vu dans les sections précédentes, l'indépendance et la dépendance parfaite (comonotonie) représentent des situations de dépendance aux propriétés intéressantes. L'idée de cette première approche est de proposer une structure de dépendance intermédiaire entre l'indépendance et la comonotonie. Ce premier modèle permettra très rapidement de saisir la sensibilité globale du traité à la relaxation de l'hypothèse d'indépendance.

Proposition 2.16. *La famille de copules de Fréchet est définie par :*

$$C_\theta = \theta C^\uparrow + (1 - \theta) C^\perp$$

Avec :

- $\theta \in [0, 1]$
- $C^\perp(u_1, \dots, u_n) = u_1 \times \dots \times u_n$ copule d'indépendance
- $C^\uparrow(u_1, \dots, u_n) = \min(u_1, \dots, u_n)$ borne supérieure de Fréchet

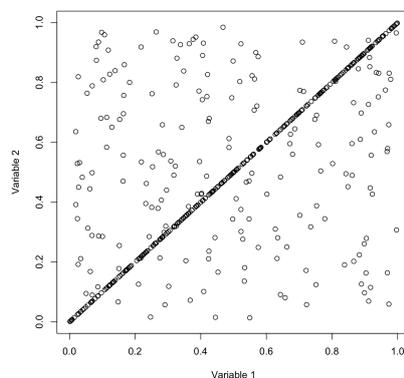


FIGURE 2.1 – Simulation d'une copule bivariee de la famille de Fréchet pour $\theta = 0,6$

L'idée d'utiliser ce modèle est de pouvoir visualiser rapidement l'évolution des principaux indicateurs (prime pure, quantiles,...) lorsque la valeur de θ varie entre 0 et 1. Une première approche envisageable est alors de réaliser automatiquement des simulations pour quelques valeurs de θ et d'analyser les variations d'indicateurs.

Proposition 2.17. *La valeur de θ peut être estimée par la mesure de dépendance décrite dans Dhaene et al., 2014 [13] et Théorème 2.4*

$$\hat{\theta} = \rho_c$$

Cette propriété est particulièrement pratique notamment lorsque l'on souhaitera comparer d'autres modèles de dépendance à ce modèle de mixture entre indépendance et comonotonie. En effet, sous d'autres modèles nous serons toujours en mesure de calculer la valeur de ρ_c , nous pourrions alors directement dire que la structure de dépendance mise en place (notamment les hypothèses prises pour construire ce modèle) est comparable en terme de mesure de Dhaene à la copule de la classe de Fréchet de paramètre $\theta = \rho_c$.

2.5.2 Seconde approche : copules multivariées

Il existe dans la littérature une grande quantité de copules multivariées aux propriétés intéressantes. Cependant peu d'entre elles offrent une grande marge de manœuvre en terme de calibration. En effet, beaucoup de copules multivariées ne permettent pas de différencier le niveau de dépendance entre paires de variables (c'est le cas des copules monoparamétriques qui imposent le même niveau de dépendance entre toutes les variables aléatoires). Par exemple, la copule de survie de Clayton multivariée fait partie de ces exemples, certes ses propriétés sont intéressantes mais ne semblent pas adaptées pour un modèle où l'on souhaite rendre dépendants des risques dont le niveau de dépendance "pairwise" n'a aucune raison d'être identique pour toutes les paires de variables. Nous nous sommes alors intéressés à trois copules classiques dont la calibration repose sur des hypothèses "pairwise", la copule gaussienne multivariée, la copule t de Student multivariée ainsi qu'une extension de cette dernière permettant d'ajuster le niveau de dépendance de queue en fonction des couples de variables, la copule individuée-t.

2.5.2.1 Copule Gaussienne multivariée

2.5.2.1.1 Caractéristiques Soit Σ une matrice $n \times n$ dite de variance-covariance de diagonale unitaire, c'est à dire définie telle que $\sigma_{kk} = 1$. Si l'on note H_Σ la distribution d'une loi $\mathcal{N}(0, \Sigma)$ alors la copule gaussienne multivariée est donnée par :

$$C_\Sigma(u) = H_\Sigma(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_n))$$

La densité copule peut être obtenue à partir de l'identité :

$$\frac{1}{\sqrt{(2\pi)^n |\Sigma|}} \exp\left(-\frac{1}{2}x^t \Sigma^{-1} x\right) = c_\Sigma(\Phi(x_1), \dots, \Phi(x_n)) \prod_{i=1}^n \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x_i^2\right)$$

Qui donne :

$$c_\Sigma(\Phi(x_1), \dots, \Phi(x_n)) = |\Sigma|^{-\frac{1}{2}} \exp\left(-\frac{1}{2}x^t \Sigma^{-1} x + \frac{1}{2} \sum_{i=1}^n x_i^2\right)$$

2.5.2.1.2 Exemple Simulation de 10000 scénarios selon une copule gaussienne de matrice de corrélations :

$$\begin{pmatrix} 1,0 & 0,1 & 0,8 \\ 0,1 & 1,0 & 0,5 \\ 0,8 & 0,5 & 1,0 \end{pmatrix}$$

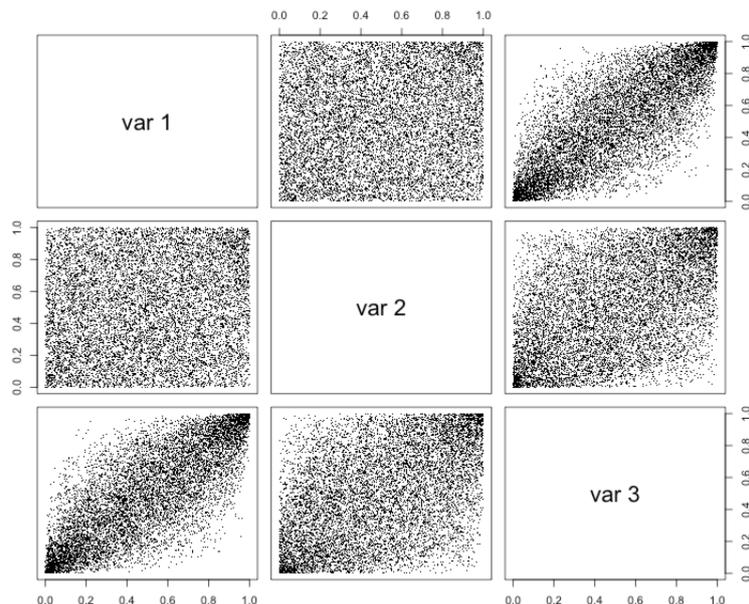


FIGURE 2.2 – Simulation d'une copule Gaussienne de dimension 3

2.5.2.1.3 Intérêt

Le principal intérêt de cette copule est sa facilité de calibration et d'utilisation. L'impossibilité de tenir compte des dépendances de queue est cependant un véritable problème surtout dans un contexte de réassurance non-proportionnelle.

2.5.2.2 Copule t multivariée

Quelques notations :

- $\mathbf{Z} = (Z_1, \dots, Z_n)'$ vecteur distribué selon une distribution Gaussienne $\Phi_{\Sigma}(\mathbf{z})$ (Σ moyennes nulles, et de variance unitaire)
- $\mathbf{U} = (U_1, \dots, U_n)'$ est un vecteur aléatoire défini sur $[0, 1]^n$
- S est une variable aléatoire de loi $\mathcal{U}_{[0,1]}$ indépendante de \mathbf{Z}
- $W = \sqrt{\left(\frac{\nu}{\chi_{\nu}^{-1}(S)}\right)}$ avec $\chi_{\nu}^{-1}(\cdot)$ la fonction de répartition inverse d'une loi du Chi-2 à ν degré de liberté. Pour des questions de simplicité de lecture la fonction de répartition de W sera notée G_{ν}
- $t_{\nu}(\cdot)$ fonction de distribution de la loi *standard* t à ν degré de liberté

2.5.2.2.1 Définition

Définition 2.10. *Étant donné les notations décrites précédemment, le vecteur aléatoire $\mathbf{X} = W \times \mathbf{Z}$ est distribué selon la loi t multivariée. Ainsi le vecteur aléatoire*

$$U = (t_\nu(X_1), \dots, t_\nu(X_n))$$

est distribué selon une copule standard t .

En d'autres termes si l'on note Σ la matrice de variance covariance, et $t_{\nu, \Sigma}$ la loi de Student multivariée, nous avons :

$$t_{\nu, \Sigma}(x_1, \dots, x_n) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_n} \frac{\Gamma\left(\frac{\nu+n}{2}\right) |\Sigma|^{-1/2}}{\Gamma\left(\frac{\nu}{2}\right) (\nu\pi)^{n/2}} \left(1 + \frac{1}{\nu} x^t \Sigma x\right)^{\frac{-\nu+n}{2}} dx_1 \dots dx_n$$

La copule t multivariée standard est alors donnée sous la représentation :

$$C_{\nu, \Sigma} = t_{\nu, \Sigma}(t_\nu^{-1}(u_1), \dots, t_\nu^{-1}(u_n))$$

Proposition 2.18. *La densité copule s'obtient alors facilement :*

$$c_{\nu, \Sigma} = \frac{\Gamma\left(\frac{\nu+n}{2}\right) |\Sigma|^{-1/2}}{\Gamma\left(\frac{\nu}{2}\right)} \left(\frac{\Gamma\left(\frac{\nu}{2}\right)}{\Gamma\left(\frac{\nu+1}{2}\right)}\right)^n \left(1 + \frac{1}{\nu} \zeta^t \Sigma^{-1} \zeta\right)^{\frac{-\nu+n}{2}} \prod_{i=1}^n \left(1 + \frac{\zeta_i}{\nu}\right)^{-\frac{\nu+1}{2}}$$

Avec $\zeta = (\zeta_1, \dots, \zeta_n) = (t_\nu^{-1}(u_1), \dots, t_\nu^{-1}(u_n))$

2.5.2.2.2 Dépendance de queue

Proposition 2.19. *Pour une copule standard t multivariée, quelque soit le couple de variables considéré, la dépendance de queue sera égale à :*

$$\lambda_L(\rho, \nu) = \lambda_U(\rho, \nu) = 2t_\nu\left(-\sqrt{\frac{(\nu+1)(1-\rho)}{1+\rho}}\right)$$

2.5.2.2.3 Simulation La simulation de la copule multivariée t standard respecte exactement le même processus de simulation que celui de la copule *Individuated-t* dans le cas où tous les degrés de liberté sont considérés comme identiques.

Nous encourageons donc le lecteur à consulter le paragraphe suivant sur la copule *Individuated-t*.

2.5.2.2.4 Exemple

Le graphique suivant représente la simulation de 10000 scénarios selon la même matrice de corrélation que pour la copule gaussienne mais avec une copule t à $\nu = 2$.

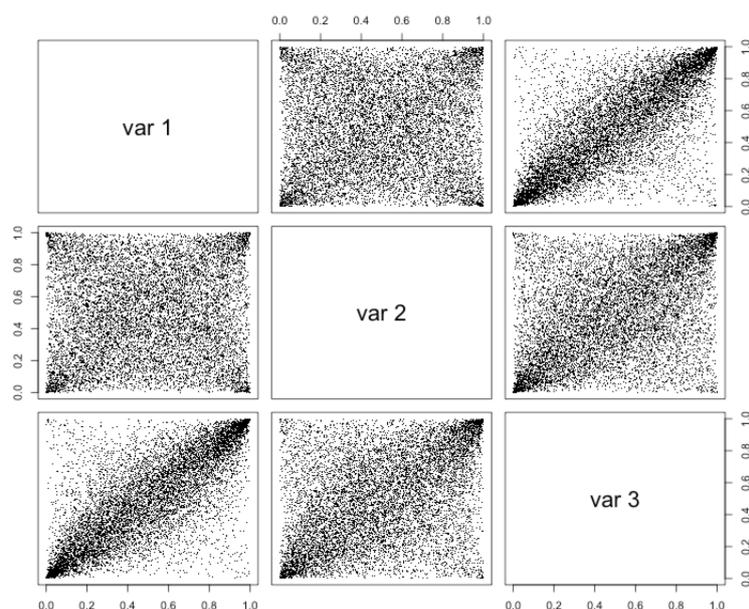


FIGURE 2.3 – Simulation d’une copule t de dimension 3 à $\nu = 2$

2.5.2.2.5 Intérêt

Cette copule a pour principal intérêt d’être relativement simple à calibrer à partir d’hypothèses sur les dépendances bivariées. Elle permet de tenir compte des dépendances de queue, ce que la copule gaussienne ne permet pas.

2.5.2.3 Copule Individuated-t

Comme nous l’avons vu la copule t est souvent utilisée en modélisation et risk management car elle permet d’intégrer aux modèles des dépendances de queue tout en restant relativement simple à simuler et calibrer. Cependant, l’utilisation de cette copule dans sa forme standard avec un unique degré de liberté limite considérablement la capacité du modélisateur à ajuster les dépendances de queue dans le cas de modèle multivarié. Dans Luo et Shevchenko, 2010 [26] les auteurs partent d’une copule grouped- t^2 dans laquelle chacun des groupes de risques n’est constitué que d’une seule variable aléatoire. En d’autres termes l’Individuated- t copule est une copule cousine de la copule t standard dans laquelle chacune des variables aléatoires a son propre degré de liberté.

2.5.2.3.1 Définition

Définition 2.11. (*Rappels sur la copule Grouped-t*) Si l’on partitionne les n variables aléatoires en m groupes distincts de taille n_1, \dots, n_m , on notera alors ν_k le degré de

2. La grouped- t copula est une copule cousine de la copule t standard dans laquelle les risques sont regroupés en groupes et un degré de liberté est attribué à chacun des groupes

liberté associé au sous groupe k . Alors la copule du vecteur

$$\mathbf{X} = (W_1 Z_1, \dots, W_1 Z_{n_1}, W_2 Z_{n_1+1}, \dots, W_2 Z_{n_1+n_2}, \dots, W_m Z_n)'$$

avec $W_k = G_{\nu_k}^{-1}(S)$, $k = 1, \dots, m$ est la copule grouped-t, c'est à dire que

$$\mathbf{U} = (t_{\nu_1}(X_1), \dots, t_{\nu_1}(X_{n_1}), t_{\nu_2}(X_{n_1+1}), \dots, t_{\nu_2}(X_{n_1+n_2}), \dots, t_{\nu_m}(X_n))'$$

est un vecteur aléatoire distribué selon une copule grouped-t. En d'autres termes la sous-copule de chaque groupe est une copule standard t avec son propre degré de liberté.

L'idée de coeur de la construction de l'Individuated-t copula, décrite dans Luo et Shevchenko, 2010 [26], est de construire des groupes constitués d'une unique variable aléatoire.

Le vecteur \mathbf{X} devient alors :

$$\mathbf{X} = (W_1 Z_1, \dots, W_n Z_n)'$$

et donc le vecteur

$$\mathbf{U} = (t_{\nu_1}(X_1), \dots, t_{\nu_n}(X_n))'$$

est distribué selon une copule Individuated-t copula.

En notant $\nu = (\nu_1, \dots, \nu_n)$, la densité de cette copule s'exprime sous la forme :

$$c_{\nu, \Sigma}(u) = \int_0^1 \phi_{\Sigma}(z_1(u_1, s), \dots, z_n(u_n, s)) \prod_{k=1}^n (w_k(s))^{-1} ds \left(\prod_{k=1}^n f_{\nu_k}(x_k = t^{-1}(u_k)) \right)^{-1}$$

Avec :

- $z_k(u_k, s) = t_{\nu_k}^{-1}/w_k(s)$
- $w_k(s) = G_{\nu_k}^{-1}(s)$
- ϕ_{Σ} densité de la loi normale multivariée
- f_{ν} densité de la loi t univariée.

2.5.2.3.2 Simulation

Le processus de simulation selon une copule Individuated-t est extrêmement simple.

- 1- Simuler les variables Z_1, \dots, Z_n selon une distribution normale multivariée $\Phi_{\Sigma}(z)$
- 2- Simuler un nombre $S \sim \mathcal{U}_{[0,1]}$ indépendant des Z_i
- 3- Calculer $W_k = G_{\nu_k}^{-1}(S)$
- 4- Calculer $\mathbf{X} = (W_i Z_i)_{i=1..n}'$ et $\mathbf{U} = (t_{\nu_i}(X_i))_{i=1..n}'$

Le vecteur \mathbf{U} est alors distribué selon une copule individuated-t.

2.5.2.3.3 Dépendance de queue

Cette copule possède de nombreuses caractéristiques intéressantes décrites dans Luo et Shevchenko, 2010 [26]. Parmi ces dernières, celle relative aux dépendances de queue est particulièrement intéressante dans le cadre de cette étude.

Pour tout sous couple de variable (X_i, X_j) on a

$$\lambda_L = \lim_{q \rightarrow 0^+} \frac{C_{ij}(q, q)}{q}$$

et

$$\lambda_U = \lim_{q \rightarrow 1^-} \frac{\bar{C}_{ij}(q, q)}{1 - q}$$

Du fait de la symétrie radiale de la copule Individuated-t on a $\lambda_U = \lambda_L = \lambda$ Par définition et en utilisant la densité copule exprimée au paragraphe précédent on a :

$$\lambda(\rho, \nu_i, \nu_j) = \lim_{q \rightarrow 0^+} \frac{1}{q} \int_0^1 \Phi_\rho \left(\frac{t_{\nu_i}^{-1}(q)}{w_i(s)}, \frac{t_{\nu_j}^{-1}(q)}{w_j(s)} \right) ds$$

Dans Luo et Shevchenko, 2010 [26], les auteurs proposent une limite analytique de cette expression et donnent le résultat sous la forme :

$$\lambda(\rho, \nu_i, \nu_j) = \Omega(\rho, \nu_i, \nu_j) + \Omega(\rho, \nu_j, \nu_i)$$

Avec :

$$\Omega(\rho, \nu_i, \nu_j) = \int_0^\infty g_{\nu_i+1}(t) F_N \left(-\frac{B_{\nu_i, \nu_j} t^{\frac{\nu_i}{2\nu_j}} - \rho\sqrt{t}}{\sqrt{1-\rho}} \right) dt$$

$$B_{\nu_1, \nu_2} = \left(\frac{2^{\frac{\nu_2}{2}} \Gamma \left[\frac{1+\nu_2}{2} \right]}{2^{\frac{\nu_1}{2}} \Gamma \left[\frac{1+\nu_1}{2} \right]} \right)^{\frac{1}{\nu_2}} \quad \text{et} \quad g_\nu(t) = \frac{1}{2^{\frac{\nu}{2}} \Gamma \left(\frac{\nu}{2} \right)} e^{-\frac{t}{2}} t^{\frac{\nu}{2}-1}$$

Ce résultat est particulièrement intéressant car il permet au sein d'une copule relativement simple de différencier les niveaux de dépendance de queue au sein de chaque couple variables aléatoires.

2.5.2.3.4 Exemple

Le graphique suivant représente la simulation de 10000 scénarios selon la même matrice de corrélation que pour la copule gaussienne et t standard mais avec des degrés de liberté valant respectivement 2,2 et 20. En comparant avec la copule t standard, on observe la même forme de dépendance entre X_1 et X_2 . Les structures de dépendance

faisant entrer en ligne de compte X_3 sont quant à elles modifiées, on observe une diminution de la dépendance dans les queues de distribution comme cela était prévisible.

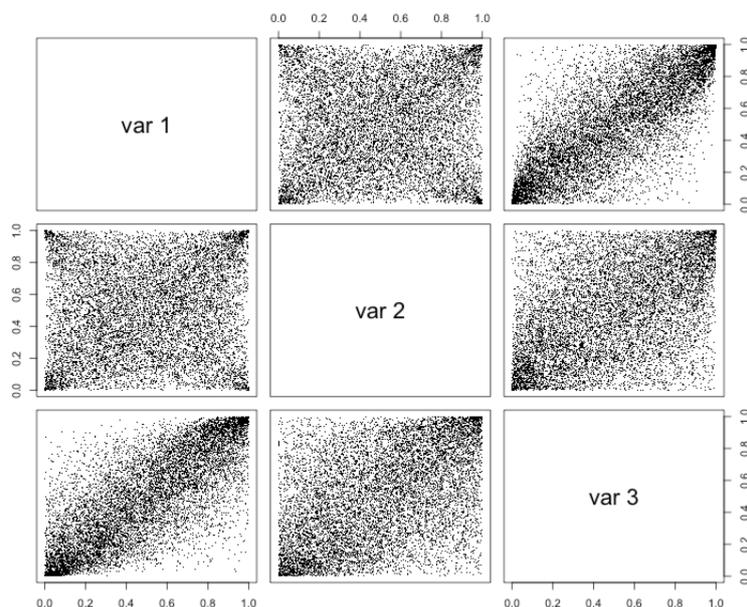


FIGURE 2.4 – Simulation d'une copule Individuated-t de dimension 3

2.5.2.3.5 Intérêts

Les intérêts de cette copule sont ceux de la copule t standard auxquels il faut ajouter la possibilité d'ajustement des dépendances de queue.

2.5.3 Troisième approche : Dépendance "pairwise" et heuristique de Kaut

Dans ce chapitre nous avons cherché à répondre à la question suivante :

Comment simuler une structure de dépendance multivariée à partir d'hypothèses relatives aux dépendances bivariées ?

La plupart des structures de dépendance multivariées classiquement utilisées, imposent une unique forme de dépendance entre toutes les variables. On peut penser aux copules archimédiennes en dimension d avec lesquelles tous les sous groupes de dimension $\leq d$ ont une structure de dépendance analogue à la copule d'origine. Bien souvent, même l'intensité de la relation est la même quelque soit le couple de variables envisagé c'est le cas par exemple pour les copules à un seul paramètre en dimension d (Gumbel, Clayton,...).

Dans le cadre de l'étude cela pose un véritable problème. Compte tenu des difficultés déjà rencontrées pour expliquer, modéliser et quantifier les dépendances et du fait qu'il

Il y a de bonnes raisons de penser que les structures de dépendance vont être fortement dépendantes des seuils considérés pour l'agrégation des risques, il est inenvisageable de justifier la mise en place d'une structure multidimensionnelle décrivant la dépendance entre les différentes "*Lines of Business*".

L'approche que l'on souhaite mettre en place dans cette partie est donc d'utiliser uniquement des hypothèses dites "*pairwise*", en d'autres termes un ensemble de structures copules bivariées. Procéder ainsi permet d'étudier la dépendance uniquement en dimension 2 et ainsi donner énormément de liberté au modèle tout en réduisant considérablement la difficulté de la quantification et de la modélisation de la dépendance.

	Fire	GTPL	MTPL
Fire	-	Copule Survie de Clayton	Indépendance
GTPL		-	Copule Normale
MTPL			-

TABLE 2.1 – Exemple d'hypothèses "pairwise" pour 3 "*Lines of Business*" classiques

En procédant de la sorte, on conserve une très grande souplesse dans les hypothèses envisageables. Une fois toutes les relations "*pairwise*" fixées de façon cohérente³, l'immense chantier restant consiste à mettre en place une méthode de simulation adaptée. Il n'existe pas à ce jour de méthode permettant de simuler directement d variables aléatoires étant données les $\frac{d(d-1)}{2}$ hypothèses de dépendance "*pairwise*".

Dans Kaut, 2013 [21], une heuristique est proposée par le chercheur du SINTEF⁴ Technology and Society (Norvège).

Cet algorithme permet de tendre vers la structure cible. Cette méthode apporte une solution assez simple et particulièrement pratique à un problème fréquent. L'auteur ouvre là une porte pour l'innovation en proposant en fait une véritable alternative à toutes les méthodes de modélisation et simulation de scénarios existantes jusqu'à ce jour (vine copulas, structures d'agrégation hiérarchique, etc.)

2.5.3.1 L'heuristique de Kaut

Les éléments techniques présentés dans cette section sont inspirés des travaux de Monsieur Michal Kaut, (Kaut, 2013 [21]).

Comme expliqué dans les précédents paragraphes dédiés à la notion de dépendance en théorie des probabilités, la fonction copule sous-jacente d'une distribution multiva-

3. Nous souhaitons insister sur le fait que toutes les structures ne sont pas envisageables par exemple si L1 est supposée comonotone avec L2, L2 comonotone avec L3 alors L1 et L3 ne peuvent pas être supposées anti-comonotones

4. *Stiftelsen for industriell og teknisk forskning*

riée peut-être vue comme le porteur de l'information relative à la dépendance entre les variables aléatoires considérées. Plus que la valeur prise par chacune des variables $\mathcal{U}_{[0,1]}$, ce sont les rangs des réalisations les unes par rapport aux autres qui portent l'information. Dans la suite de la présentation de l'heuristique de Kaut nous ne travaillerons qu'avec les rangs des réalisations. Le rang 1 sera attribué au minimum et S au maximum (échantillon de taille S).

A partir de là, la réalisation de taille M d'une copule n -dimensionnelle peut-être vue comme :

$$\mathcal{C} = \{\mathbf{r} = (r_1, \dots, r_n) : 1 \leq r_i \leq S, \forall i \leq n\}$$

Pour la fonction de répartition (*CDF*) évaluée en un point $\mathbf{r} \in \mathcal{C}$, nous retiendrons les notations suivantes

$$\begin{aligned} C_{rank}(\mathbf{r}) &= C\left(\frac{r_1}{S}, \dots, \frac{r_n}{S}\right) \\ &= \mathbb{P}\left(\left[0, \frac{r_1}{S}\right] \times \dots \times \left[0, \frac{r_n}{S}\right]\right) \end{aligned}$$

Avec C_{rank} and C les fonctions copules empiriques en terme de rang et de valeur sur $[0, 1]^n$.

La copule cible, notée C^* , n'est alors pas décrite en une unique formulation mais sous la forme de $\frac{n(n-1)}{2}$ relations "pairwise". En d'autres termes $\frac{n(n-1)}{2}$ copules bivariables.

Afin d'évaluer la qualité de simulation d'un échantillon, nous pouvons comparer les fonctions de répartition empirique et cible. En d'autres termes nous pouvons calculer en chacun des points de \mathcal{C} , la distance entre les deux valeurs, soit $\forall \mathbf{r} \in \mathcal{C}$:

$$\begin{aligned} dev(\mathbf{r}) &= dev(r_1, \dots, r_n) \\ &= C_{rank}(\mathbf{r}) - C_{rank}^*(\mathbf{r}) \end{aligned}$$

La qualité globale de la simulation peut quant à elle s'évaluer en tant que fonction de l'ensemble des valeurs prises par dev sur \mathcal{C} . Dans Kaut, 2013 [21], l'auteur considère deux fonctions possibles, nous nous intéresserons principalement à la première des deux :

$$d_{avg}(\mathcal{C}, C^*) = \frac{1}{S^n} \sum_{r_1=1}^S \dots \sum_{r_n=1}^S |dev(r_1, \dots, r_n)| \quad (2.1)$$

$$d_{max}(\mathcal{C}, C^*) = \max_{\substack{1 \leq r_i \leq S \\ i=1, \dots, S}} |dev(r_1, \dots, r_n)| \quad (2.2)$$

D'un point de vue informatique, un algorithme nécessitant le calcul de ces fonctions requiert l'évaluation de ces fonctions sur une grille de taille S^n , ce qui n'est évidemment pas réalisable dans un intervalle de temps raisonnable (le nombre d'opérations évoluant en $\mathcal{O}(S^n)$). Le fait de travailler uniquement avec les relations copules "pairwise" va nous permettre de ramener le nombre d'opérations en $\mathcal{O}(n^2 S^2)$. Le prix de cette

méthode est qu'il sera impossible d'introduire des dépendances d'ordres supérieurs⁵.

Concentrons nous dans un premier temps sur l'heuristique dans le cas $n = 2$. Le cas multivarié sera ensuite présenté comme une généralisation, le raisonnement étant globalement le même.

2.5.3.1.1 Cas $n=2$

L'idée au cœur de l'algorithme en dimension 2 consiste à associer les rangs des deux variables de telle sorte à obtenir une structure convergeant vers la structure cible.

L'heuristique repose principalement sur l'observation du fait que l'évaluation de la fonction de répartition $C_{rank}(i, j)$ ne dépend que des points (i', j') du sous ensemble de la grille de S^2 tels que $i' \leq i$ et $j' \leq j$. Ainsi nous pouvons construire l'échantillon en associant le rang i de l'une au rang j de l'autre de telle sorte que la distance $dev(i \rightarrow j)$ engendrée par cette création de scénario soit minimale. La distance $dev(i \rightarrow j)$ créée par l'association des rangs concernés est la somme des valeurs absolues des distances aux points de la grille $(., j)$:

$$\begin{aligned} dev(i \rightarrow j) &= \sum_{l=1}^S |dev(l, j)| \\ &= \sum_{l=1}^S |C_{rank}(l, j) - C_{rank}^*(l, j)| \end{aligned}$$

Ce calcul est rendu possible par l'utilisation du fait que :

$$C_{rank}(i, j) = C_{rank}(i, j - 1) + \begin{cases} \frac{1}{S} & \text{si on couple } j \text{ avec } i' \leq i \\ 0 & \text{sinon} \end{cases}$$

On obtient ainsi :

$$\begin{aligned} dev(i \rightarrow j) &= \sum_{l=1}^S |C_{rank}(l, j - 1) + \frac{1}{S} \mathbb{1}_{i \leq l} - C_{rank}^*(l, j)| \\ &= \sum_{l=1}^{i-1} |C_{rank}(l, j - 1) - C_{rank}^*(l, j)| + \\ &\quad \sum_{l=i}^S |C_{rank}(l, j - 1) + \frac{1}{S} - C_{rank}^*(l, j)| \\ dev((i - 1) \rightarrow j) &= \sum_{l=1}^{i-2} |C_{rank}(l, j - 1) - C_{rank}^*(l, j)| + \\ &\quad \sum_{l=i-1}^S |C_{rank}(l, j - 1) + \frac{1}{S} - C_{rank}^*(l, j)| \end{aligned}$$

5. Décrire les dépendances via l'ensemble des copules bivariées est cependant largement suffisant dans le cadre de cette étude

Par soustraction on obtient alors :

$$\begin{aligned} dev(i \rightarrow j) = dev((i-1) \rightarrow j) &+ | C_{rank}(i-1, j-1) - C_{rank}^*(i-1, j) | \\ &- | C_{rank}(i-1, j-1) + \frac{1}{S} - C_{rank}^*(i-1, j) | \end{aligned}$$

Initialisé par :

$$dev(0 \rightarrow j) = \sum_{l=1}^S | C_{rank}(l, j-1) + \frac{1}{S} - C_{rank}^*(l, j) |$$

Cette formulation récursive nous permet de calculer $dev(i \rightarrow j)$ pour toutes les valeurs de i en conservant l'heuristique en un nombre d'opérations assimilable à $\mathcal{O}(S^2)$.

Ainsi le principe algorithmique consiste à construire l'échantillon en associant les rangs de chacune des variables de telle sorte que la distance au point considéré soit minimale. Le pseudo-code algorithmique est alors le suivant :

```

 $\mathcal{I} \leftarrow \{1, \dots, S\};$ 
 $dev^* \leftarrow \infty;$ 
for  $j \in \{1, \dots, S\}$  do
  for  $i \in \mathcal{I}$  do
    if  $dev(i, j) < dev^*$  then
       $i^* \leftarrow i;$ 
       $dev^* \leftarrow dev(i, j)$ 
    end
  end
   $scenario(i^*, j);$ 
   $\mathcal{I} = \mathcal{I} - \{i^*\}$ 
end

```

Algorithm 1: Heuristique de Kaut dans le cas bivarié

2.5.3.1.2 Extension au cas multivarié

L'extension au cas multivarié est réalisée en commençant par 2 variables et en ajoutant les variables suivantes les unes après les autres. Cela se comprend mieux en raisonnant à la manière d'une récurrence. Imaginons que nous avons déjà généré un m variables aléatoires $\mathcal{U}_{[0,1]}$ dont nous connaissons les rangs, nous allons alors entamer l'étape "d'ajout" de la $m+1$ ^e en utilisant les m copules cibles $C_{1,m+1}, C_{2,m+1}, \dots, C_{m,m+1}$.

Contrairement au cas bivarié on ne peut pas procéder par un simple rapprochement de lignes et colonnes, car on doit prendre en compte le fait que les rangs des m premières variables ont déjà été assignés à un scénario donné. Nous allons en fait assigner les rangs de la $m+1$ ^e variables aux différents scénarios.

Pour cela Michal Kaut introduit des notations complémentaires, r_s^k désigne le rang de la variable k assigné au scénario s , et C_r^k C_r^{*k} respectivement les fonctions de répartition de l'échantillon et de la cible de la copule bivariable des variables k et $m + 1$. Enfin dev^k la fonction de distance avec :

$$dev^k(i, j) = C_r^k(i, j) - C_r^{*k}(i, j)$$

L'assignement d'un rang à un scénario est fait de façon analogue au cas bivarié modulo quelques adaptations. Chaque rang de la $m + 1$ ^e marginale est lié au scénario $s \in \mathcal{S}$ qui engendre la plus petite distance. La principale modification est que la distance considérée n'est plus la simple valeur dev mais la somme des toutes les dev^k . C'est à dire que l'on considère pour associer un rang à un scénario, la valeur $dev_s = \sum_{k=1}^m dev^k(r_s^k, j)$.

```

S ← {1, ..., S};
dev* ← ∞;
for j ∈ {1, ..., S} do
  for s ∈ S do
    for k ∈ {1, ..., m} do
      | Calculer devk(rsk, j)
    end
    devs ← ∑i=1m devk(rsk, j) if devs < dev* then
      | s* ← s;
      | dev* ← devs
    end
  end
  rjm+1 ← s*;
  S = S - {s*}
end

```

Algorithm 2: Ajout de la $m + 1$ ^e marginale dans l'Heuristique de Kaut multivarié

On peut remarquer que le cas multivarié n'est que la généralisation du cas bivarié et que ce dernier s'obtient facilement en considérant $m = 1$ et $\forall s r_s^1 = s$.

Une question reste cependant en suspend : quelle est la sensibilité de l'heuristique à l'ordre dans lequel les variables aléatoire (les rangs en fait) sont simulées. Plusieurs tests ont été effectués sans pouvoir conclure de façon pertinente. Étant donnée la structure algorithmique on peut s'attendre à ce que l'ordre de simulation ait un impact sur la qualité. Aucune règle n'est, à ce jour, recommandée pour réaliser les simulations.

2.5.3.2 Copules bivariées usuelles

Cette méthodologie requiert de définir au préalable l'intégralité des dépendances "pairwise". Pour cela on utilise une série de copules usuelles dont les caractéristiques sont résumées dans le tableau ci-dessous.

De nombreuses méthodes de "fitting" existent et sont décrites dans la littérature. Pour la sélection de la meilleure copule on peut notamment citer des tests basés sur

l'écart entre la copule empirique et la copule paramétrique comme le test de Cramer Van Mises. L'estimation du paramètre de la copule se fait de façon paramétrique ou non, les méthodes classiques étant celles du maximum de vraisemblance et la méthode CML (Canonical Maximum Likelihood). Ces méthodes sont largement décrites dans la littérature et résumées dans le mémoire d'actuariat de Lagenebre, 2009 [23].

Dans le cadre de cette étude, cependant, l'estimation à proprement parler de structure copule a été impossible du fait du manque de profondeur d'historique. Les hypothèses testées seront des hypothèses à dire d'experts.

Nom de la Copule	$C(u, v)$	τ_{Kendall}	λ_U	λ_L	Exemple
Clayton	$(u_1^{-\theta} + u_2^{-\theta} - 1)^{-\frac{1}{\theta}}$	$\frac{\theta}{2 + \theta}$	0	$\frac{1}{2 - \frac{1}{\theta}}$	
Survie de Clayton	$w_1 + w_2 - 1 + ((1 - u_1)^{-\theta} + (1 - u_2)^{-\theta} - 1)^{-\frac{1}{\theta}}$	$\frac{\theta}{2 + \theta}$	$\frac{1}{2 - \frac{1}{\theta}}$	0	
Gumbel	$\exp\left(-\left[(-\ln(u_1))^\theta + (-\ln(u_2))^\theta\right]^{\frac{1}{\theta}}\right)$	$1 - \frac{1}{\theta}$	$2 - 2\frac{1}{\theta}$	0	
Normal	$\Phi_\rho(\Phi^{-1}(u_1), \Phi^{-1}(u_2))$	$\frac{2}{\pi} \arcsin(\rho)$	0	0	
t	$t_{\rho, \nu}(t_\nu^{-1}(u_1), t_\nu^{-1}(u_2))$	$\frac{2}{\pi} \arcsin(\rho)$	$2t_\nu\left(-\sqrt{\frac{(\nu+1)(1-\rho)}{1+\rho}}\right)$	$\frac{2}{\pi} \arcsin(\rho)$	
Frank	$-\frac{1}{\theta} \ln\left(1 + \frac{(e^{-\theta u_1} - 1)(e^{-\theta u_2} - 1)}{(e^{-\theta} - 1)}\right)$	$1 - \frac{4}{\theta} \left(1 - \frac{1}{\theta} \int_0^\theta \frac{t}{e^t - 1} dt\right)$	0	0	

TABLE 2.2 – Résumé des principales copules utilisées dans le cadre de la modélisation des dépendances "pairwise"

2.5.3.3 Exemple illustratif

Voici la simulation selon l'heuristique de Kaut d'une structure de dépendance à 4 variables aléatoires respectant les dépendances "pairwise" suivantes :

	V_1	V_2	V_3	V_4
V_1	-	Clayton(1)	Indépendant	Normal(0,2)
V_2		-	Normal(0,4)	Gumbel(1.5)
V_3			-	Clayton(2)
V_4				-

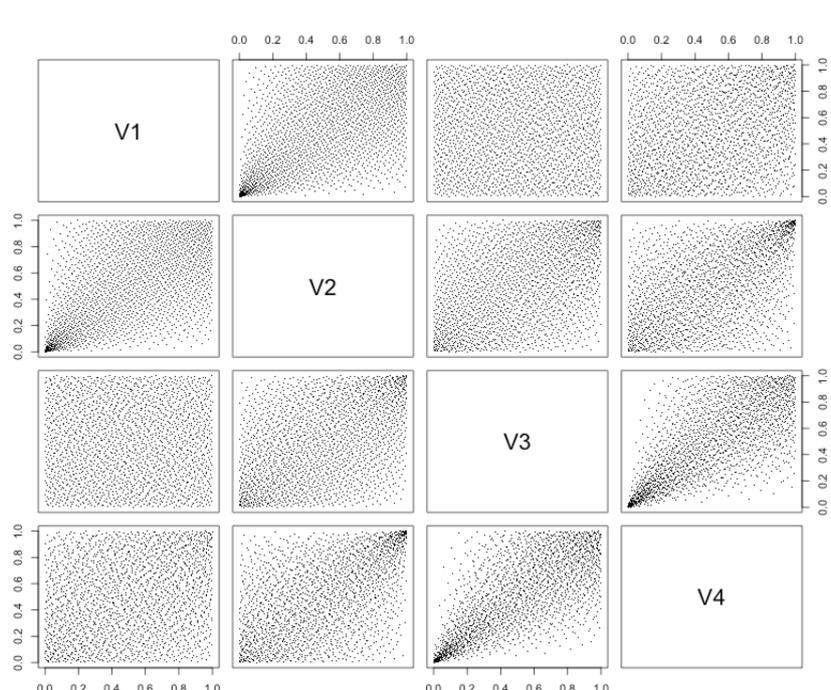


FIGURE 2.5 – Simulation d'une structure de dépendance à 4 variables selon l'heuristique de Kaut (affichage de 2000 réalisations)

Les résultats obtenus sur un échantillon de 10 000 simulations sont étudiés en utilisant le test développé dans Genest *et al.*, 2009 [16] (procédure `gofCopula`⁶ dans le logiciel R), le tableau suivant montre que toutes les p -value sont largement supérieures à 10% et qu'il n'y a donc jamais de présomption contre l'hypothèse nulle (hypothèse selon laquelle la copule empirique = copule théorique). C'est là un moyen efficace de vérifier combien la simulation avec l'heuristique de Kaut semble respecter les

6. *Parametric bootstrap goodness-of-fit test*

hypothèses souhaitées⁷. Ces tests, absolument nécessaires du fait de l'utilisation d'une heuristique, permettent notamment de prendre conscience parfois d'incohérences dans les hypothèses retenues pour la simulation.

	V_1	V_2	V_3	V_4
V_1	-	$p = 0,548$	$p = 0,733$	$p = 0,545$
V_2		-	$p = 0.784$	$p = 0.456$
V_3			-	$p = 0.653$
V_4				-

7. Un test statistique ne permet certes jamais de valider une hypothèse mais étant donné le nombre de simulation et la valeur des *p-value* il y a de bonnes raisons de penser que la simulation selon l'heuristique a dans ce cas correctement fonctionné

Chapitre 3

Proposition d'une méthodologie de tarification

"The independence hypothesis is so common to be made that many authors forget to mention it."

Hans Bühlmann, 1963

3.1 Motivation

L'objet principal de l'étude est la mise en place d'une méthodologie de tarification relaxant l'hypothèse d'indépendance entre les différentes "*Lines of Business*" réassurées dans un même traité MAXL¹. En effet, si l'hypothèse d'indépendance est particulièrement pratique pour la tarification et l'étude de ce type de traité, il est clair qu'un risque de modèle est encouru dès que la présomption de dépendance entre "*Lines of Business*" est forte.

Nous avons donc souhaité analyser la faisabilité de la relaxation de l'hypothèse d'indépendance dans ce contexte bien particulier. Dans une démarche de recherche et développement, l'idée est de mettre en place un outil de prise en compte de la dépendance entre "*Lines of Business*" sans totalement bouleverser le processus de tarification par simulation de scénarios. En d'autres termes nous ne souhaitons pas que les modèles marginaux des différentes "*Lines of Business*" soient modifiés.

Relaxer l'hypothèse d'indépendance implique de prendre des décisions quant à la façon dont la dépendance entre les "*Lines of Business*" va être intégrée au processus de simulation des sinistres. Plusieurs modèles sont à disposition mais une série de contraintes techniques, pratiques et/ou manageriales restreint particulièrement les choix réellement envisageables.

Pour bien comprendre comment nous en sommes arrivés à la méthode, reprenons à la base le fonctionnement du traité. Pour chaque "*Line of Business*" i des sinistres de

1. Les traités étudiés ici sont les traités MAXL mais la plupart des démarches sont applicables à l'intégralité des traités "multiline".

montant x_{i_j} - individuels ou événements selon les termes du traité - sont observés (simulés dans le cadre d'une tarification). A partir de ces sinistres, nous pouvons calculer deux agrégats de sinistralité de la "*Line of Business*". Le premier est la sinistralité réelle totale des sinistres considérés par le réassureur. Par l'expression "considérés par le réassureur", on entend que, dans le cadre d'un traité de réassurance non-proportionnelle, seuls les sinistres excédant un certain seuil sont considérés et donc simulés pour la tarification. Ainsi, dans les faits, le réassureur n'a pas connaissance de l'intégralité des sinistres² mais seulement de ceux qui potentiellement affectent le traité. On considérera dans la suite les notations suivantes pour les deux agrégats de sinistralités.

L'agrégat noté \mathbb{T}_i de la "*LoB_i*" correspondra à la somme *from-ground-up* des sinistres excédant le seuil de la "*Line of Business*" i .

$$\mathbb{T}_i = \sum_{k=1}^{N_i} x_{i_k}$$

Le second agrégat est, comme nous l'avons déjà vu, la sinistralité d'alimentation du MAXL par la "*Line of Business*" i .

$$\mathbb{S}_i = \sum_{k=1}^{N_i} f_i(x_{i_k})$$

Le programme $GAALMAXLGAAD+GAAL$ s'applique ensuite à la sinistralité d'alimentation totale.

$$\mathbb{S} = \sum_i \mathbb{S}_i$$

Sous hypothèse d'indépendance, nous avons vu que le principe de tarification est conceptuellement assez simple. Les sinistres sont simulés indépendamment sur chacune des lignes et agrégés par la suite. La suite de ce chapitre présente le "module" de prise en compte de la dépendance.

3.2 Contraintes et méthodes non-retenues

L'idée intuitive, pour prendre en compte les phénomènes de dépendance, serait donc de rendre individuellement les sinistres dépendants. Cela pourrait se faire, par exemple, via des lois de fréquences ou de montants de sinistres dépendants. Sur le papier, ces concepts sont intéressants mais imposent des changements drastiques dans l'approche de simulation.

En effet, le modèle fréquence/sévérité par "*Line of Business*" ne pourrait plus être employé directement. Il faudrait dans un cas créer des lois de fréquence multivariées or les copules discrètes ne disposent pas des propriétés d'unicité nécessaires au traitement de ce type de problématique. Dans l'autre cas il faudrait considérer une unique variable

2. La limite d'information du réassureur par la cédante est fixée contractuellement souvent en pourcentage de la priorité (50%, 70%)

fréquence à laquelle serait associé un vecteur de sinistralité. En d'autres termes, considérer qu'un événement créerait automatiquement un sinistre sur toutes les "*Lines of Business*" (la probabilité d'observer un montant nul sur une des "*Lines of Business*" étant évidemment conséquente). Malheureusement, ces modèles déforment *de facto* les lois de probabilité régissant les comportements marginaux de sinistralité ce qui, pour des raisons de cohérence de tarification et de faisabilité³, n'a pas été considéré comme envisageable. Les modèles à chocs communs, traités pour les modèles collectifs à m classes d'affaires dans LeBlanc, 2000 [24] ont eux aussi été envisagés mais étant donné la nécessité d'intégration du module de traitement de la dépendance aux processus de tarification existant ces méthodes n'ont pas été retenues.

Dernier point et pas des moindres, l'intégralité de ces méthodes proposant de travailler directement sur la simulation des sinistres, les utiliser rendrait impossible l'intégration dans le "module" des "Lines of Business" simulées à partir de logiciels dédiés comme RMS.

	Indépendance	Idée 1	Idée 2
Modèle	$(f_i, s_i)^\perp$	$\mathcal{L}(f_i), (s_i)^\perp$	$f, \mathcal{L}(s_i)$
<i>Exemple</i> $n=2$	$f_1 \sim \text{Poisson}$ $f_1 \sim \text{Poisson}$ $s_1 \sim \text{Pareto}$ $s_2 \sim \text{Pareto}$	$(f_1, f_2) \sim \text{Poisson bivariée}$ $s_1 \sim \text{Pareto}$ $s_2 \sim \text{Pareto}$	$f \sim \text{Poisson}$ $(s_1, s_2) \sim \text{Pareto bivariée}$

TABLE 3.1 – Résumé de méthodes envisagées mais non retenues

3.3 Méthode proposée

L'approche que nous avons finalement décidée de suivre diffère quelque peu des approches précédentes et nécessite un certain nombre d'hypothèses dont l'utilisateur doit avoir conscience afin de toujours garder un esprit critique sur les résultats obtenus.

Nous avons en effet décidé d'intégrer la dépendance à un niveau post-simulation des sinistres. L'idée de base est de considérer que si des "*Lines of Business*" ont des raisons d'engendrer des sinistres dépendants alors cette dépendance s'observera sur la somme des sinistres considérés par le réassureur \mathbb{T}_i .

Du fait des programmes internes à appliquer par "*Lines of Business*", ces sinistralités réelles (\mathbb{T}_i) ne nous intéressent pas réellement pour la tarification, cependant ce sont sur ces variables que la dépendance existe dans la réalité. Intégrer la dépendance sur les \mathbb{S}_i n'aurait en effet aucun sens car ces sommes sont le fruit d'effets non-proportionnels et n'ont donc aucune réalité physique. Nous intégrons donc la dépendance entre les \mathbb{T}_i pour ensuite nous intéresser aux répercussions sur le montant total alimentant le traité MAXL \mathbb{S}_i (notamment l'impact sur la prime pure du traité).

3. Il faudrait redéfinir l'intégralité des modèles de tarification ce qui n'est pas envisageable dans le cadre de cette étude

Évidemment le niveau de dépendance dépendra du seuil au delà duquel on observe les sinistres. L'estimation des paramètres de dépendance en fonction de ce seuil constitue une difficulté majeure et nécessitera parfois de réaliser des hypothèses forfaitaires quant au niveau de dépendance à considérer.

Le principe de tarification devient donc le suivant :

1 : La première étape demeure la simulation des sinistres individuels sur toutes les "Lines of Business" (fréquence/sévérité, période de retour,...). En commençant de la sorte, nous nous assurons de l'intégrabilité du processus de mise sous dépendance à l'ensemble des méthodes déjà implémentées.

Simulation	Sinistres								
	Line of Business 1				...	Line of Business n			
1	x_{1_1}	x_{1_2}	-	-	...	x_{n_1}	x_{n_2}	x_{n_3}	x_{n_4}
2	x_{1_1}	-	-	-	...	-	-	-	-
3	-	-	-	-	...	x_{n_1}	-	-	-
...
N	x_{1_1}	x_{1_2}	x_{1_3}	-	...	x_{n_1}	x_{n_2}	-	-

Lecture : Dans le 3^e scénario le 1^{er} sinistre sur la "Line of Business" n a pour montant x_{n_1}

2 : La seconde étape consiste simplement à calculer les valeurs d'agrégats \mathbb{T}_i et \mathbb{S}_i .

Simulation	Line of Business 1		...	Line of Business n	
	\mathbb{T}_1	\mathbb{S}_1	...	\mathbb{T}_n	\mathbb{S}_n
1	$\sum_j x_{1_j}$	$\sum_j f_1(x_{1_j})$...	$\sum_j x_{n_j}$	$\sum_j f_n(x_{n_j})$
...
N	$\sum_j x_{1_j}$	$\sum_j f_1(x_{1_j})$...	$\sum_j x_{n_j}$	$\sum_j f_n(x_{n_j})$

3 : La troisième étape, celle de simulation de la structure de dépendance, est sûrement celle nécessitant le plus de travail. On notera que plutôt que d'utiliser les réalisations des variables $\mathcal{U}_{[0,1]}$, on travaillera directement avec les statistiques de rangs.

R_1	...	R_n
r_{11}	...	r_{n1}
...
r_{1N}	...	r_{nN}

Lecture : r_{ij} est rang de la j^e réalisation de la i^e variable

4 : Lors de la dernière étape avant analyse des résultats, nous appliquons un algorithme de réorganisation des simulations de telle sorte que les rangs des \mathbb{T}_i aient la même structure que les rangs de la copule simulée.

Ainsi, le montant alimentant le traité au titre de la ligne i dans le p^e scénario sera le montant \mathbb{S}_i associé à la simulation dont la valeur \mathbb{T}_i a pour rang r_{ip}

Les scénarios des \mathbb{S}_i sont donc réorganisés selon une structure de dépendance sur les \mathbb{T}_i . A proprement parler, on ne contrôle pas la structure de dépendance des \mathbb{S}_i car ces derniers n'ont pas de réalité physique et combinent de multiples effets non proportionnels rendant toute anticipation de la dépendance hasardeuse.

Pour résumer, la méthode retenue peut s'écrire sous la forme du pseudo code algorithmique suivant.

```

Input:  $n$  le nombre de LoB
Input:  $N$  le nombre de scénario à simuler
for  $i \in \{1, \dots, n\}$  do
  for  $s \in \{1, \dots, N\}$  do
    Simuler et mémoriser les montants des sinistres individuels
    from-ground-up (étape 1)
  end
  Calculer et mémoriser les agrégats  $\mathbb{S}_i(s)$  et  $\mathbb{T}_i(s)$  (étape 2)
end
for  $i \in \{1, \dots, n\}$  do
  Simuler un échantillon de taille  $N$  de variables aléatoires  $U_i \sim \mathcal{U}_{[0,1]}$  selon une
  copule  $C$  puis calculer  $r_{ip}$  rang de la  $p^e$  réalisation de  $U_i$  (étape 3)
end
for  $s' \in \{1, \dots, N\}$  do
  for  $i \in \{1, \dots, n\}$  do
    Affecter au scénario  $s'$  le montant  $\mathbb{S}_i(s)$  tel que  $s$  vérifie
     $Rang(\mathbb{T}_i(s)) = r_{is'}$  (étape 4)
  end
end
for  $s' \in \{1, \dots, N\}$  do
  Calculer  $\mathbb{S}(s') = \sum_{i=1}^n \mathbb{S}_i(s')$ 
  Calculer  $\pi(s') = \min(GAAL, \max(\mathbb{S}_i(s') - GAAD, 0))$ 
end

```

La prime pure du traité MAXL est alors estimé par : $\Pi = \mathbb{E}[\pi(s')]$

Algorithm 3: Méthode de tarification avec prise en compte de la dépendance entre les \mathbb{T}_i

3.4 Les modèles de dépendance

3.4.1 Trois approches possibles

Comme nous venons de le voir, la méthodologie retenue consiste à imposer la dépendance entre les sommes *from-ground-up* des sinistres excédant un seuil⁴.

La première structure de dépendance que nous avons considérée est un modèle convexe entre indépendance et comonotonie. Il permet d'avoir une première idée de la sensibilité du traité à la relaxation de l'hypothèse d'indépendance.

Pour ajuster la dépendance en fonction des caractéristiques des marginales simulées, il est alors une quasi-nécessité de pouvoir calibrer les dépendances de façon "*pairwise*". En effet, travailler directement sur des structures multivariées rendrait la calibration totalement impossible à réaliser que ce soit d'ailleurs à partir des données ou d'hypothèses forfaitaires à dire d'experts.

C'est pourquoi nous avons retenu les deux approches suivantes. La première est basée sur l'utilisation de copules multivariées dont la calibration repose uniquement sur l'étude des corrélations linéaires et dans certains cas des dépendances de queue souhaitées. Nous utilisons pour cela la copule gaussienne, la copule t ainsi que la copule Individuated-t.

La copule gaussienne dispose de propriétés la rendant extrêmement simple à utiliser. Cependant, elle ne permet pas de prendre en compte les dépendances extrêmes. C'est pourquoi on s'intéresse, lorsque l'on a de fortes présomptions de dépendance extrême, à la copule t standard. Cette copule dispose de la même facilité de calibration que la copule gaussienne et, modulo un travail supplémentaire sur le degré de liberté, permet d'obtenir des dépendances extrêmes. Or, en utilisant un unique degré de liberté, cette copule ne permet pas bien de différencier le niveau de dépendance extrême en fonction du couple de variables aléatoires considérées. C'est pourquoi nous avons recours à la copule Individuated-t qui, en affectant un degré de liberté à chacune des variables aléatoires, permet d'ajuster le niveau de dépendance extrême en fonction du couple considéré.

La dernière approche, repose sur la simulation via l'heuristique de Kaut. Cette heuristique permet de simuler une structure de dépendance n -dimensionnelle respectant $n(n-1)/2$ hypothèses de dépendances (ou copules) bivariées. On ramène ainsi l'étude d'une structure multivariée à un ensemble de structures bivariées. Cette méthode est assez novatrice en cela qu'elle permet au tarificateur d'utiliser uniquement des copules de dimension 2 et ainsi réaliser proprement le travail de modélisation de la dépendance en fonction des seuils de sinistralité considérés. Certes, beaucoup plus de copules sont à étudier, mais dans les faits, le travail est nettement plus simple. Les marges de manœuvre, en termes de forme et d'intensité de dépendance, offertes par cette méthodologie sont comme chacun peut l'imaginer extrêmement intéressantes.

4. Ce seuil étant celui à partir duquel les sinistres sont simulés.

3.4.2 Paramétrisation

3.4.2.1 Cadre théorique

Cette section décrit le cadre théorique permettant d'estimer les paramètres des modèles de dépendance. Dans de nombreuses situations les données disponibles ne sont pas suffisantes et l'information à disposition doit être complétée par des hypothèses à dire d'experts.

Comme tout modèle, les copules selon lesquelles nous souhaitons simuler doivent être sélectionnées et calibrées. Nous avons retenu une méthode de calibration reposant sur l'étude des traités non-proportionnels déjà en portefeuille et pour lesquelles le seuil d'information contractuel rend l'information exploitable. En d'autres termes, si pour la tarification du traité MAXL les sinistres simulés sur la LoB_i sont ceux excédant le $Seuil_i$ alors on sectionnera dans la base de données des traités, ceux couvrant la LoB_i et dont le niveau le seuil déclaratif⁵ du réassureur est inférieur à $Seuil_i$.

En cas de données disponibles insuffisantes, des hypothèses simplificatrices ou forfaitaires peuvent être retenues à dire d'experts pour réaliser les tarifications. La méthodologie reposant sur l'heuristique de Kaut est en ce sens extrêmement intéressante car on peut imaginer par exemple que dans un modèles à 4 "Lines of Business" impliquant 6 relations bivariées, 4 soient bien connues⁶ et que 2 ne le soient absolument pas. Le modèle permet alors de conserver les 4 hypothèses connues et d'imposer des hypothèses forfaitaires⁷ pour les deux autres.

Dans le cas où une base de travail a pu être constituée nous allons calculer pour chaque année d'exercice et pour chaque branche le ratio "S/P" global agrégé au delà d'un seuil. Soit :

$$S/P(LoB_i, Seuil_i, \text{année } j) = \frac{\sum_{\text{Traités}} \text{sinistres supérieurs au } Seuil_i}{\sum_{\text{Traités}} EPI}$$

Le montant de sinistre considéré étant le montant extrapolé en ultime.

La dernière étape consiste à projeter ce ratio sur la prime perçue par l'assureur (EPI) au titre des contrats couvrant la "Line of Business" i réassurée dans le traité MAXL. Ainsi on estime pour chaque années j , le montant $T_i(j, Seuil_i)$ par :

$$S/P(LoB_i, Seuil_i, \text{année } j) \times EPI(MAXL_{LoB_i})$$

L'ensemble de ces manipulations consiste en quelque sorte à mettre en *as-if* les sinistralités⁸ historiques agrégées sur tous les traités. On dispose alors d'une base de données constituée de l'extrapolation sur historique de la somme des sinistres dépassant un seuil donné pour le portefeuille réassurée dans le traité MAXL.

5. Le réassureur reçoit dans la plupart des cas l'information de sinistre lorsque l'estimation de celui-ci dépasse 50 ou 70% de la priorité

6. Car étudiées lors d'une autre étude ou issues de données de marché

7. En supposant l'indépendance ou la comonotonicité par exemple

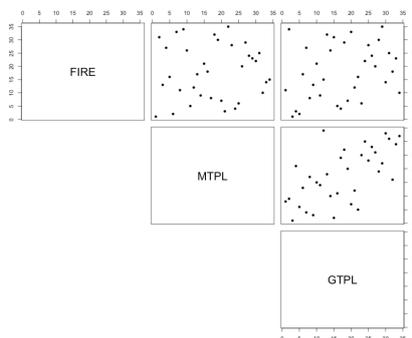
8. La sinistralité considérée ici étant celle ne prenant en compte que les sinistres excédant un seuil

Remarque : si la cédante est en mesure de fournir un historique étoffé de la sinistralité des *"Lines of Business"* réassurées alors le même travail peut être réalisée en utilisant uniquement les données de la cédante considérée plutôt que l'agrégat de tous les traités comme décrit précédemment.

A partir de cette table de données, les paramètres (ρ de Spearman, τ de Kendall) s'estiment sans grande difficulté. Cependant, aucun réassureur ne disposera d'un historique suffisamment long pour estimer des structures copules complètes. L'idée, notamment utile pour l'implémentation de l'heuristique de Kaut, est alors de réaliser une série d'hypothèses à dire d'experts quant à la copule sous-jacente à la structure de dépendance entre les *"Lines of Business"* et de visualiser l'effet de ces hypothèses notamment en comparaison des résultats obtenus avec la copule Gaussienne.

3.4.2.2 Exemple

Le scatterplot ci-dessous présente graphiquement les structures de dépendance entre entre les rangs des \mathbb{T}_i pour les branches Incendies et autres périls, responsabilité civile automobile et responsabilité civile générale sur un historique de 33 ans. Les limites considérées pour les trois *"Lines of Business"* étant 500 000 € pour les responsabilités et 1 000 000 € pour l'incendie.⁹



On observe pour la base de données ainsi constituée, la matrice des τ de Kendall suivante :

	Incendie	RCA	RCG
Incendie	1,000	0,019	0,214
RCA		1,000	0,440
RCG			1,000

TABLE 3.2 – Matrice des τ de Kendall des \mathbb{T}_i de l'exemple

9. Par soucis de confidentialité des résultats, ces seuils ne sont pas ceux réellement utilisés mais cela n'entrave pas la compréhension de la démarche

L'estimation de la mesure de dépendance multivariée construite pour l'agrégation de risque est périlleuse car nécessite d'avoir un historique de même longueur sur toutes les "*Lines of Business*". Dans le cas de cet exemple l'ordre de grandeur est :

$$\hat{\rho}_c \approx 17\%$$

A ces niveaux de seuil, on observe clairement une dépendance positive entre les deux responsabilités civiles. L'incendie semble indépendant de la RC Auto, ce qui semble assez intuitif. La RC Générale et l'Incendie montrent quand à eux une dépendance certes faible mais qui semble toutefois exister.

Tous ces résultats doivent bien évidemment être lus en conscience du peu de données sur lesquels ils reposent.

3.5 Limites et critiques

L'approche proposée a le mérite d'être facilement implémentable et les outils qui l'accompagnent offrent des marges de manœuvre de modélisations assez larges. Cependant, cette approche est critiquable à bien des égards.

Le fait que la dépendance soit introduite entre des sinistralités au delà d'un seuil fait qu'une étude de la dépendance doit être réalisée quasiment à chaque nouvelle tarification, introduisant, de fait, un risque opérationnel dans le tarif proposé.

L'autre grande difficulté est relative aux données disponibles. En effet, c'est un cadre théorique qui est présenté dans ce mémoire. Pour des lignes telles que la responsabilité civile générale et automobile on peut penser que l'approche est crédible. Pour d'autres risques, les expositions et longueurs d'historiques rendent impossibles toute estimation robuste et imposent des hypothèses à dire d'experts.

Les outils proposés pour les structures de dépendance offrent de grandes marges de manœuvre. Cependant, dans de nombreuses situations le tarificateur est dans l'impossibilité de prendre une décision relative au niveau de dépendance. Dans cette situation, une hypothèse envisageable serait de retenir le niveau le plus prudent. Or rien n'indique en théorie quel est le niveau de dépendance le plus sûr pour le réassureur. Dans ce cas, la meilleure approche semble de réaliser une série de tarifications selon des hypothèses de dépendance différentes puis de décider, après analyse, à quel niveau le réassureur souhaite placer la prime pure.

Chapitre 4

Exemple de tarification

Dans ce chapitre nous allons traiter la mise en place concrète de la structure de dépendance pour un traité MAXL couvrant 4 "*Lines of Business*" fréquemment réassurées au sein de ce type de traité.

Pour des soucis de clarté et de confidentialité, les modèles de simulations ont volontairement été exprimés dans leur forme la plus simplifiée et ce sans nuire aux conclusions et observations relatives à la mise en place de la structure de dépendance entre les risques considérés.

La tarification est effectuée via la simulation de 10 000 scénarios de sinistralité.

Sommaire

4.1	Le traité	58
4.2	Les garanties	58
4.3	Les modèles de simulation de sinistres	60
4.4	La structure de dépendance	62
4.5	Tarification	67

4.1 Le traité

Le traité que l'on souhaite tarifier est un traité MAXL couvrant les 4 "*Lines of Business*" suivantes :

- ◇ **Responsabilité civile générale.** On considérera que la responsabilité civile est engagée, soit en raison de l'inexécution d'un contrat, soit en raison d'un acte volontaire ou non, entraînant pour la personne qui est fautive ou qui est légalement présumée fautive, l'obligation de réparer le dommage qui a été subi par une ou plusieurs autres.¹
- ◇ **Responsabilité civile automobile.** Cette garantie responsabilité civile automobile couvre la responsabilité civile du conducteur du véhicule ou de toute personne, autorisée ou non, conduisant le véhicule en question, ainsi que celle de tous les passagers. Concrètement, elle couvre les conséquences des dommages causés "aux tiers", c'est-à-dire à toutes les personnes autres que le conducteur lui-même (passagers, piétons, autres conducteurs, etc.).
- ◇ **Incendies et autres périls** tels que l'explosion, le vol avec ou sans effraction, le bris de glaces et d'enseignes, les pertes d'exploitation et/ou perte de la valeur vénale à la suite d'un accident, d'un incendie ou d'une explosion ayant détruit des installations professionnelles, les dégâts des eaux, les risques annexes de responsabilité (recours des voisins et des tiers, risques locatifs...) et de dommages pécuniaires.
- ◇ Le conséquences des événements de type **Force de la Nature** et leurs manifestations notamment les tempête, la grêle et le gel.

4.2 Les garanties

Le traité que l'on étudie a été calibré avec les tranches d'éligibilité internes du MAXL suivantes :

<i>"Line of Business"</i>	Priorité	Portée
RC Générale	580 000	1 420 000
RC Auto	500 000	2 000 000
Incendie et autres périls	250 000	400 000
Force de la nature	200 000	3 000 000

TABLE 4.1 – Tranches d'éligibilité internes du MAXL (en €)

1. Définition tirée de "www.dictionnaire-juridique.fr"

Lors du renouvellement 2013, la tarification a été effectuée pour une garantie MAXL :

GAAD	GAAL
4 000 000	9 000 000

Nous verrons par la suite que les calibrations du GAAD et GAAL peuvent être directement impactées par le changement de structure de dépendance. En effet, ces montants sont souvent placés à un certain niveau de quantile de \mathbb{S} . Or la dépendance va, comme nous le verrons, impacter la Value-at-Risk de \mathbb{S} modifiant ainsi le niveau de couverture souhaité par la cédante.

Ici le traité a été calibré de telle sorte que sous hypothèse d'indépendance $VaR_{\mathbb{S}}(GAAD) \approx 40\%$ et $VaR_{\mathbb{S}}(GAAD + GAAL) = 99,5\%$

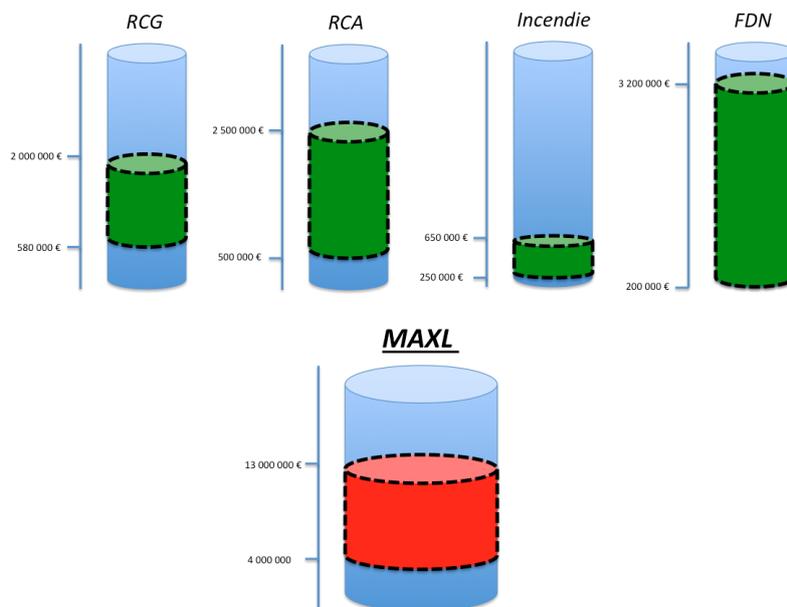


FIGURE 4.1 – Tranches d'éligibilité par "Line of Business" et couverture MAXL

4.3 Les modèles de simulation de sinistres

Les sinistres en Responsabilité Civile Générale seront considérés à partir du seuil de 580 000 €.

- La fréquence est modélisée par une loi de Poisson de paramètre $\lambda_{RCG} = 2,3453$
- La sévérité est quant à elle modélisée par une loi de Pareto de *shape* $\alpha_{RCG} = 2,1454$ et de *scale* $A_{RCG} = 580\,000$

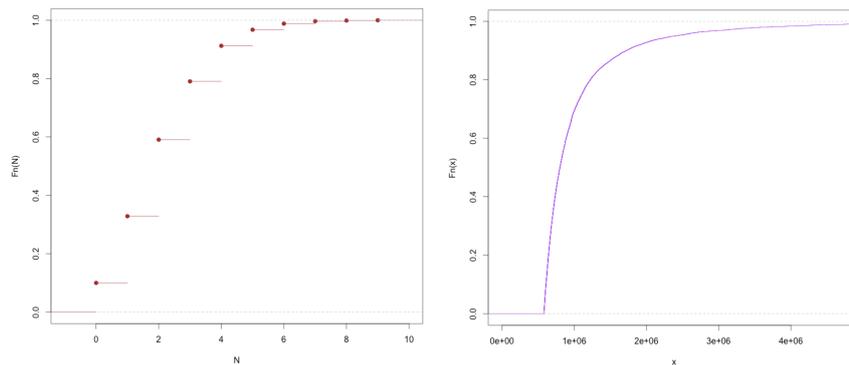


FIGURE 4.2 – Fonction de répartition des lois de fréquence (gauche) et sévérité (droite) pour la LoB : RCG

Les sinistres en Responsabilité Civile Automobile seront considérés à partir du seuil de 500 000 €.

- La fréquence est modélisée par une loi de Poisson de paramètre $\lambda_{RCA} = 2,9345$
- La sévérité est quant à elle modélisée par une loi de Pareto de *shape* $\alpha_{RCA} = 1,4353$ et de *scale* $A_{RCA} = 500\,000$

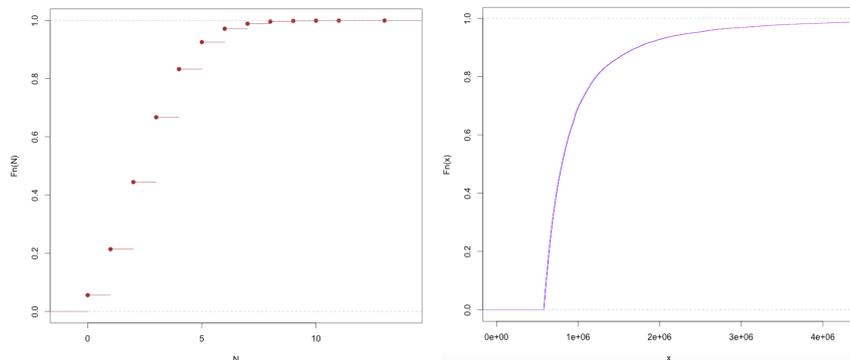


FIGURE 4.3 – Fonction de répartition des lois de fréquence (gauche) et sévérité (droite) pour la LoB : RCA

Les sinistres en Incendie et autres périls seront considérés à partir du seuil de 250 000 €.

- La fréquence est modélisée par une loi de Poisson de paramètre $\lambda_{Incendie} = 7,4584$
- La sévérité est quant à elle modélisée par une loi de Pareto de *shape* $\alpha_{Incendie} = 3,2225$ et de *scale* $A_{Incendie} = 250\ 000$

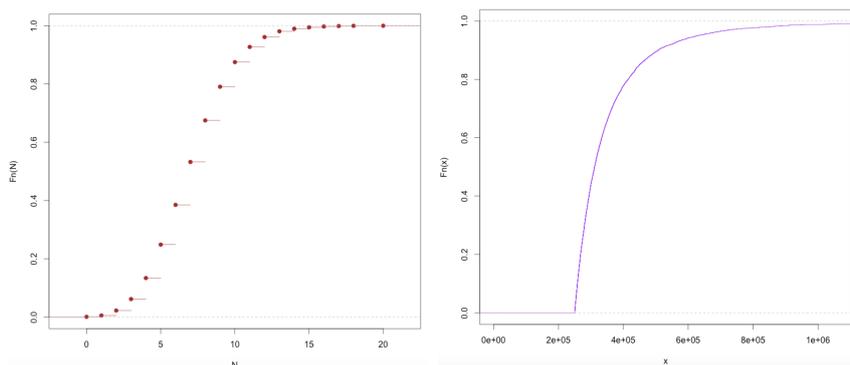


FIGURE 4.4 – Fonction de répartition des lois de fréquence (gauche) et sévérité (droite) pour la LoB : Incendie et autres périls

Les sinistres Forces de la Nature seront considérés à partir du seuil de 250 000 €.

- La fréquence est modélisée par une loi de Poisson de paramètre $\lambda_{FDN} = 1,2493$
- La sévérité est quant à elle modélisée via des scénarios issus du logiciel RMS. La loi de probabilité de la sévérité est connue de façon empirique.

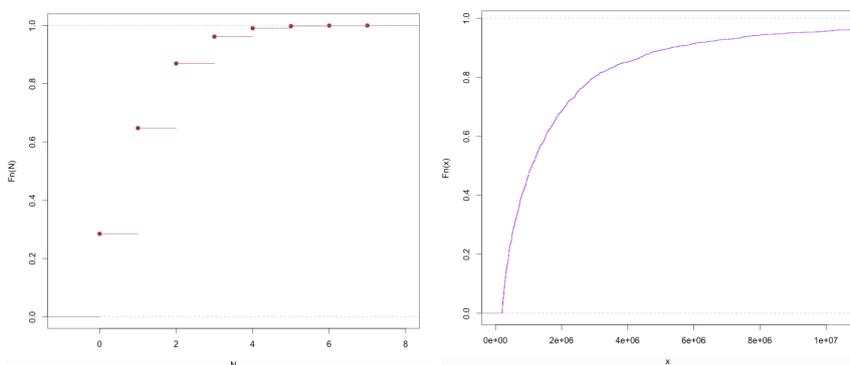


FIGURE 4.5 – Fonction de répartition des lois de fréquence (gauche) et sévérité (droite) pour la LoB : Force de la Nature

"Line of Business"	Moyenne	Écart-type	Médiane	VaR _{90%}	VaR _{95%}	VaR _{99,5%}
	Variable : \mathbb{T}_i					
RCG	2 577 232	2 769 804	2 080 783	5 094 725	6 505 519	13 071 197
RCA	4 592 978	14 151 954	2 925 916	8 580 359	12 009 594	44 682 152
Incendie	2 718 718	1 379 949	2 604 475	4 165 630	4 715 165	6 446 677
Force de la Nature	3 380 628	7 696 658	1 153 780	7 973 383	13 393 891	50 382 180
	Variable : \mathbb{S}_i					
RCG	907 124	875 061	644 147	2 127 440	2 646 941	4 046 103
RCA	1 692 212	1 461 407	1 368 350	3 765 400	4 496 773	6 659 783
Incendie	744 355	399 550	692 707	1 282 611	1 477 453	2 052 521
Force de la Nature	1 608 570	1 886 116	854 760	4 210 568	5 552 414	10 061 611

TABLE 4.2 – Statistiques descriptives des agrégats \mathbb{T}_i et \mathbb{S}_i (en €)

4.4 La structure de dépendance

Afin de réaliser les tarifications selon les différentes approches nous retenons comme paramètres de dépendance les résultats suivants² :

La mesure de dépendance multivariée, estimateur du paramètre de mélange de la copule mixture indépendance/comonotonicité, est mesurée à :

$$\hat{\rho}_c = 11,4\%$$

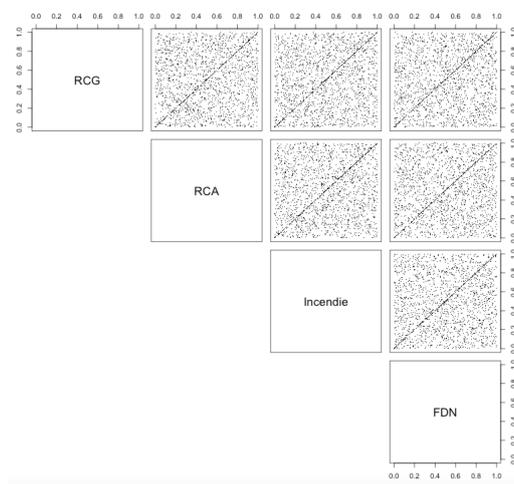


FIGURE 4.6 – Structure de dépendance simulée avec une copule mixture de paramètre de mélange $\rho = 11,4\%$

2. Résultats issus de la méthodologie décrite au chapitre 3 et d'hypothèses à dire d'experts

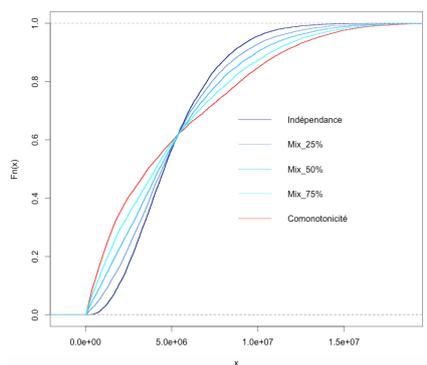


FIGURE 4.7 – Fonction de répartition de \mathbb{S} pour quelques valeurs de ρ

Afin de calibrer les copules gaussienne, t et Individuated-t on utilisera les matrices de corrélations et de τ de Kendall.

	RCG	RCA	Incendie	FDN
RCG	-	$\rho = 29\%$	$\rho = 14\%$	$\rho = 7\%$
RCA	$\tau = 44\%$	-	$\rho = 0\%$	$\rho = 0\%$
Incendie	$\tau = 22\%$	$\tau = 0\%$	-	$\rho = 17\%$
FDN	$\tau = 10\%$	$\tau = 0\%$	$\tau = 27\%$	-

TABLE 4.3 – Mesures de dépendance bvariée entre les valeurs de \mathbb{T}_i

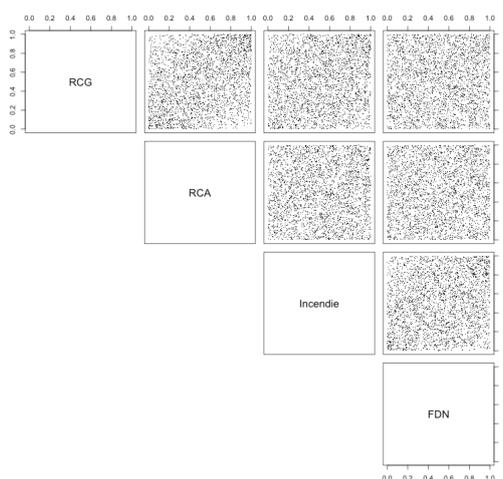


FIGURE 4.8 – Structure de dépendance simulée avec la copule gaussienne

La calibration des degrés de liberté des copule t et Individuated-t est particulière-

ment complexe. Étant donné la profondeur d'historique toute calibration reposant, par exemple, sur l'estimation par la méthode classique du maximum de vraisemblance est inenvisageable.

L'idée développée ici a alors été de définir le degré de liberté en fonction des niveaux de dépendances de queue ciblés (à dire d'experts faute de mieux). On sélectionne donc les degrés de liberté qui donnent des dépendances de queue les plus proches de celles ciblées.

Pour cette tarification on utilisera pour la copule-t, $\nu = 7$ et pour la copule Individuated-t, $\nu_{RCG} = 11$, $\nu_{RCA} = 11$, $\nu_{Incendie} = 3$, $\nu_{FDN} = 3$.

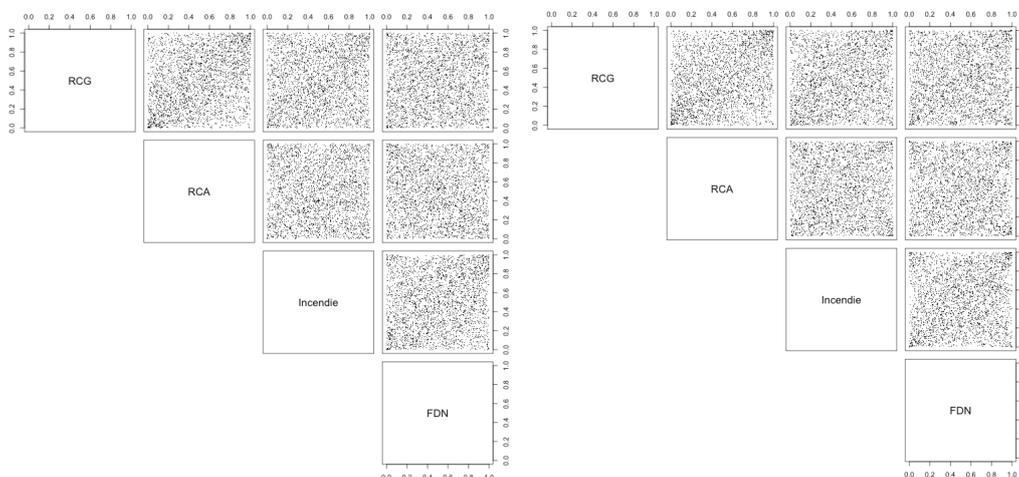


FIGURE 4.9 – Structure de dépendance simulée avec les copules t et Individuated-t

Dans le cadre de la simulation via l'heuristique de Kaut nous utilisons la structure des copules bivariées³ suivante :

	RCG	RCA	Incendie	FDN
RCG	-	Gumbel	Gaussienne	Gumbel
RCA		-	Indépendant	Indépendant
Incendie			-	Gumbel
FDN				-

TABLE 4.4 – Structure des Copules bivariées

3. Ces copules sont calibrées à partir des valeurs de τ de Kendall. $\theta_{copule} = f(\tau)$

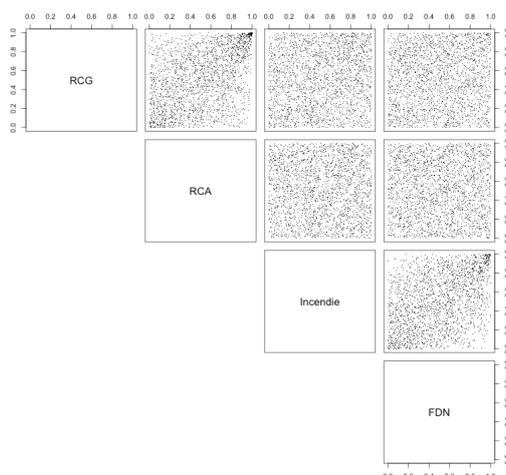


FIGURE 4.10 – Structure de dépendance simulée avec la structure des copules bivariées et l’heuristique de Kaut

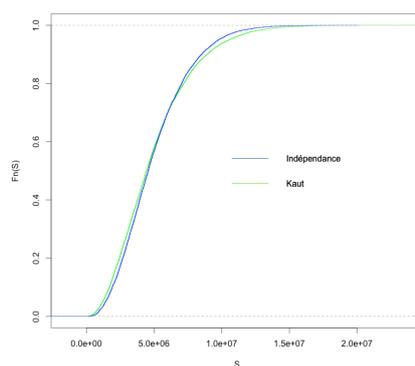


FIGURE 4.11 – Fonction de répartition de \mathbb{S} simulé en situation d’indépendance et avec l’algorithme de Kaut

Comment ces dépendances entre les \mathbb{T}_i se transmettent-elles aux \mathbb{S}_i ?
 Les matrices suivantes présentent les τ de Kendall des \mathbb{S}_i

	Structure : \perp			
	RCG	RCA	Incendie	FDN
RCG	-	0%	0%	0%
RCA		-	0%	0%
Incendie			-	0%
FDN				-

Structure : \uparrow				
	RCG	RCA	Incendie	FDN
RCG	-	73%	63%	74%
RCA		-	64%	78%
Incendie			-	66%
FDN				-

Structure : Mixture				
	RCG	RCA	Incendie	FDN
RCG	-	6,3%	7,9%	6,9%
RCA		-	6,7%	7,7%
Incendie			-	7,1%
FDN				-

Structure : Gauss				
	RCG	RCA	Incendie	FDN
RCG	-	16,3%	8,1%	4,2%
RCA		-	0,0%	0,0%
Incendie			-	10,3%
FDN				-

Structure : t				
	RCG	RCA	Incendie	FDN
RCG	-	18,3%	7,1%	3,2%
RCA		-	0,0%	0,0%
Incendie			-	9,9%
FDN				-

Structure : Indiv-t				
	RCG	RCA	Incendie	FDN
RCG	-	17,4%	7,6%	2,2%
RCA		-	0,0%	0,0%
Incendie			-	7,9%
FDN				-

Structure : Kaut				
	RCG	RCA	Incendie	FDN
RCG	-	19,9%	6,6%	5,2%
RCA		-	0,0%	0,0%
Incendie			-	12,9%
FDN				-

On observe que globalement la dépendance se transmet aux S_i dans une moindre mesure. Cela s'explique facilement par le fait que les très grands scénarios de sinistres vont être captés par les portées internes du MAXL. Les tranches d'éligibilité ont ici en quelque sorte un effet absorbeur de dépendance. En plus de limiter l'exposition à de grands sinistres, les prises en charges capées limitent les risques d'effets d'accumulation due à la dépendances entre les "*Lines of Business*".

4.5 Tarification

Le tableau suivant présente la synthèse des résultats des tarifications obtenus avec les différentes copules :

	$\Pi = \mathbb{E}(\pi)$		$\sigma(\pi)$	
	Valeur	Δ	Valeur	Δ
Indépendance	1 509 508	-	2 011 409	-
Comonotonicité	2 106 690	+39,6%	2 982 977	+48,3%
Mixture	1 579 223	+4,6%	2 136 925	+6,2%
Gaussienne	1 581 942	+4,8%	2 100 323	+4,4%
t	1 571 504	+4,1%	2 111 261	+5%
Individuated-t	1 570 890	+4,1%	2 113 029	+5,1%
Heuristique de Kaut	1 585 737	+5,1%	2 213 051	+10,0%

Δ indique la variation par rapport à l'hypothèse d'indépendance, exprimée en pourcentage de la valeur sous hypothèses d'indépendance.

	$F_{\mathbb{S}}(GAAD)$		$F_{\mathbb{S}}(GAAD + GAAL)$		$VaR_{99,5\%}(\mathbb{S})$	
	Valeur	Δ	Valeur	Δ	Valeur	Δ
Indépendance	41,02%	-	99,5%	-	13 011 134	-
Comonotonicité	52,77%	+12,72%	94,29%	-5,21%	17 532 941	+34,8%
Mixture	41,76%	+1,76%	98,75%	-0,75%	14 649 201	+12,6%
Gaussienne	41,63%	+1,6%	99,12%	-0,38%	13 792 554	+6%
t	41,75%	+1,55%	98,95	-0,55%	13 989 160	+7,6%
Individuated-t	41,37%	+1,37%	89,99	-0,51%	14 417 389	+10,8%
Heuristique de Kaut	43,13%	+3,13%	98,4%	-1,01%	15 062 794	+14,8%

Avec, $F_{\mathbb{S}}(X) = \mathbb{P}(\mathbb{S} < X)$

Δ indique la variation par rapport à l'hypothèse d'indépendance, exprimée en variation nette pour les niveaux de quantile et en pourcentage de la valeur sous hypothèse d'indépendance pour la $VaR_{99,5\%}(\mathbb{S})$.

Cette tarification met en évidence une série d'observations :

- La méthodologie de prise en compte de la dépendance est opérationnelle. Une fois les niveaux de dépendance déterminés, le temps ajouté à la simulation est infinitésimale (les algorithmes de simulations de copules et de réarrangement étant extrêmement rapides).
- En fonction de la structure choisie, la dépendance ne se transmet pas totalement des T_i aux S_i bien qu'en terme d'ordre de grandeur les résultats restent relativement stables.

- Déterminer le niveau de dépendance est une tâche particulièrement complexe. La possibilité d'étudier la dépendance au niveau des paires de variables est un véritable pré-requis à la méthode. A un moment ou un autres, des hypothèses à dire d'experts sont nécessaires.
- La dépendance a un impact certain sur la prime pure. Pour des niveaux de dépendance relativement modérés, l'impact n'est pas extrêmement violent.
- La concentration de la dépendance dans les queues de distributions (copule t et Indiv-t) semble engendrer une variation de prime pure inférieure à celle obtenue avec une copule sans dépendance extrême (Gaussienne). Cela est une conséquence, de l'effet du GAAL. En effet en concentrant ensemble les grandes sinistralités on obtient plus de scénarios dans lesquels la prise en charge par le réassureur est limitée GAAL et donc mécaniquement plus de scénario engendrant une prise en charge inférieure. La prime pure est donc inférieure à celle obtenue avec un niveau de dépendance équivalent mais moins concentré dans les queues de distribution.
- En revanche, l'incertitude sur l'estimation de la prime pure, semble elle, assez nettement impactée dans tous les modèles. Ce sujet devrait être approfondi dans le cas de prime calculée à partir de méthode intégrant cette volatilité (cf. transformée de Wang).
- Un effet central de la dépendance est de modifier le niveau de couverture effectif du traité MAXL. En effet, si le traité est calibré via la distribution du risque en situation d'indépendance cela peut entraîner d'importants écarts si l'objectif est de couvrir la cédante à hauteur de 99,5% de la distribution de S . En gardant la couverture calibrée en situation d'indépendance pour faire correspondre l'éligibilité du MAXL à la tranche [40% - 99,5%], on observe qu'en situation de dépendance, la couverture n'est plus celle-ci. La portée du traité doit alors être assez nettement augmentée si l'on souhaite couvrir jusqu'à un niveau de quantile 99,5%.
- D'un point de vue méthodologique, il semble que l'utilisation de la copule gaussienne donne une assez bonne approximation des effets de la dépendance. L'utilisation de méthodes plus sophistiquées est certes intéressante d'un point de vue méthodologique mais, compte tenu des difficultés de calibration, pas nécessairement incontournable.

Chapitre 5

Conclusion

La méthode proposée dans ce mémoire permet d'introduire une structure de dépendance entre les différentes "*Lines of Business*" réassurées dans un même traité MAXL. L'implémentation de la méthodologie est simple et n'augmente pas, une fois la structure de dépendance connue, le temps de travail nécessaire à la tarification.

Pour diverses raisons la dépendance est introduite entre les sommes, par "*Line of Business*", des sinistres *from-ground-up* dépassant certains seuils. Les seuils en question correspondent dans la majorité des cas aux priorités internes des "*Lines of Business*" du traité MAXL.

Différentes tarifications ont été effectuées avec ce module et nous en tirons les conclusions suivantes.

La dépendance positive impacte la prime pure. Lorsque les GAAD et GAAL sont calibrés pour couvrir la sinistralité d'alimentation à un niveau de quantile suffisamment élevé alors la dépendance positive a tendance à augmenter la prime pure.

Cet impact va de pair avec les conséquences de la dépendance sur la distribution de la sinistralité d'alimentation. En effet, la dépendance va déformer la loi de probabilité de \mathbb{S} . Si les GAAD et GAAL sont fixés pour couvrir \mathbb{S} entre deux niveaux de quantile il faut impérativement tenir compte de la dépendance dans la calibration du traité.

L'étude présente trois différentes approches, la première est extrêmement pratique car nécessite l'utilisation d'un unique paramètre. Cependant l'estimation de ce paramètre est complexe et le modèle quelque peu simpliste. Il est un bon début pour sentir l'ordre de grandeur de l'impact de la dépendance mais n'est pas suffisant.

L'utilisation de la copule gaussienne, bien que ne permettant pas l'introduction de dépendance de queue, semble en revanche largement suffisante pour appréhender avec assez de précision la nouvelle tarification. Les difficultés de calibration des copules t et Individuated-t font perdre à ces modèles beaucoup de crédibilité.

L'utilisation de l'heuristique de Kaut et des relations "*pairwise*" est extrêmement

séduisante. Si le niveau de connaissance de la dépendance est suffisant, c'est à coup sûr le modèle offrant la plus de marge de manœuvre. Dans une situation où beaucoup d'hypothèses doivent être faites à dire d'experts, les résultats obtenus doivent être considérés avec prudence.

Globalement cette étude montre combien il est important, pour les sociétés de réassurance, de disposer de bases de données proprement construites et rapidement exploitables. La structure, le fonctionnement et les possibilités d'exploitation des "*data warehouse*", restent malheureusement trop souvent inadaptées à ce type d'étude.

Table des matières

Synthèse	i
Résumé	vii
Abstract	ix
1 La réassurance "multiline"	1
1.1 La réassurance	2
1.2 Traités "multiline" et diversification	3
1.3 " <i>Multiline Aggregate Excess of Loss</i> "	5
1.3.1 Définition	5
1.3.2 Fonctionnement théorique	6
1.3.3 Exemple de fonctionnement d'un MAXL à 3 " <i>Lines of Business</i> "	7
1.3.4 Un exemple pratique d'utilisation dans un programme	9
1.3.5 Le portefeuille QBE Re	10
1.3.6 Principe de tarification	10
1.3.6.1 Simulation des scénarios	10
1.3.6.2 Patterns de paiement	11
1.3.7 Les effets non-proportionnels Excess-of-Loss	12
1.4 Pourquoi s'attarder sur la dépendance entre les risques?	14
1.4.1 Quelques exemples intuitifs	14
1.4.2 A quel(s) effet(s) faut-il s'attendre?	15
1.4.2.1 Un effet difficile à prévoir	15
1.4.2.2 Illustration :	17
2 Cadre théorique de la dépendance multivariée	19
2.1 Les copules	21
2.2 Bornes de Fréchet et Indépendance	23
2.3 Mesure de dépendance bivariée	24
2.3.0.3 Le coefficient de corrélation de Pearson	24
2.3.0.4 Le τ de Kendall	25
2.3.1 La dépendance de queue	25
2.4 Une mesure de dépendance pour l'agrégation de risques	27
2.4.1 Principales propriétés	27
2.4.2 Estimation	28
2.5 Les différentes approches de modélisation copule	30

2.5.1	Première approche : Famille de copule de Fréchet	30
2.5.2	Seconde approche : copules multivariées	31
2.5.2.1	Copule Gaussienne multivariée	31
2.5.2.1.1	Caractéristiques	31
2.5.2.1.2	Exemple	31
2.5.2.1.3	Intérêt	32
2.5.2.2	Copule t multivariée	32
2.5.2.2.1	Définition	33
2.5.2.2.2	Dépendance de queue	33
2.5.2.2.3	Simulation	33
2.5.2.2.4	Exemple	33
2.5.2.2.5	Intérêt	34
2.5.2.3	Copule Individuated-t	34
2.5.2.3.1	Définition	34
2.5.2.3.2	Simulation	35
2.5.2.3.3	Dépendance de queue	36
2.5.2.3.4	Exemple	36
2.5.2.3.5	Intérêts	37
2.5.3	Troisième approche : Dépendance "pairwise" et heuristique de Kaut	37
2.5.3.1	L'heuristique de Kaut	38
2.5.3.1.1	Cas n=2	40
2.5.3.1.2	Extension au cas multivarié	41
2.5.3.2	Copules bivariées usuelles	42
2.5.3.3	Exemple illustratif	45
3	Proposition d'une méthodologie de tarification	47
3.1	Motivation	47
3.2	Contraintes et méthodes non-retenues	48
3.3	Méthode proposée	49
3.4	Les modèles de dépendance	52
3.4.1	Trois approches possibles	52
3.4.2	Paramétrisation	53
3.4.2.1	Cadre théorique	53
3.4.2.2	Exemple	54
3.5	Limites et critiques	55
4	Exemple de tarification	57
4.1	Le traité	58
4.2	Les garanties	58
4.3	Les modèles de simulation de sinistres	60
4.4	La structure de dépendance	62
4.5	Tarification	67
5	Conclusion	69

Bibliographie

- [1] ANDRIANTAVY, C. (2007). Tempêtes : prise en compte de la dépendance spatiale avec la théorie des copules. *Mémoire d'actuariat*.
- [2] AUBRON, E. (2009). Étude de l'impact des dépendances sur le capital Économique. *Mémoire d'actuariat*.
- [3] BELGUISE, O. (2001). Tempêtes : Etude des dépendances entre les branches auto et incendie avec la théorie des copulas. *Mémoire d'actuariat*.
- [4] DAUL, S., GIORGI, E. D., LINDSKOG, F. et MCNEIL, A. (2003). The grouped t-copula with an application to credit risk.
- [5] DE PRIL, N., DE VYLDER, F. et GOOVAERTS, M. (1982). Bounds on modified stop-loss premiums in case of known mean and variance of the risk variable. *ASTIN Bulletin International Actuarial Association*.
- [6] DEMARTA, S. et MCNEIL, A. J. (2005). The t copula and related copulas. *International Statistical Review*, 73(1):111–129.
- [7] DENUIT, C., GENEST, C. et MESFIOUI, M. (2003). Stop-loss bounds of functions of possibly dependent risks in the presence of partial information on their marginals. Discussion paper.
- [8] DENUIT, M., DHAENE, J., GOOVAERTS, M. et KAAS, R. (2006). Actuarial theory for dependent risks. *John Wiley and Sons Ltd*.
- [9] DENUIT, M., DHAENE, J. et RIBAS, C. (2001). Does positive dependence between individual risks increase stop-loss premiums? *Insurance : Mathematics and Economics*, 28(3):305 – 308.
- [10] DHAENE, J. et DENUIT, M. (1999). The safest dependence structure among risks. *Insurance : Mathematics and Economics*, 25(1):11 – 21.
- [11] DHAENE, J., DENUIT, M., GOOVAERTS, M., KAAS, R. et VYNCKE, D. (2002a). The concept of comonotonicity in actuarial science and finance : applications. *Insurance : Mathematics and Economics*, 31(2):133 – 161.
- [12] DHAENE, J., DENUIT, M., GOOVAERTS, M., KAAS, R. et VYNCKE, D. (2002b). The concept of comonotonicity in actuarial science and finance : theory. *Insurance : Mathematics and Economics*, 31(1):3 – 33. Special Issue : Papers presented at the 5th IME Conference, Penn State University, University Park, PA, 23-25 July 2001.

- [13] DHAENE, J., LINDERS, D., SCHOUTENS, W. et VYNCKE, D. (2014). A multivariate dependence measure for aggregating risks. *Journal of Computational and Applied Mathematics*, 263(0):78 – 87.
- [14] DURANTE, F., SÁNCHEZ, J. F. et SEMPI, C. (2013). Multivariate patchwork copulas : A unified approach with applications to partial comonotonicity. *Insurance : Mathematics and Economics*, 53(3):897 – 905.
- [15] GENEST, C. et MACKAY, J. (1986). The joy of copulas : Bivariate distributions with uniform marginals. *The American Statistician*, 40(4):280–283.
- [16] GENEST, C., REMILLARD, B. et BEAUDOIN, D. (2009). Goodness-of-fit tests for copulas : A review and a power study. *Insurance : Mathematics and Economics*, 44(2):199–213.
- [17] GILLET, A. et SERRA, B. (2001). Effets de la dépendance entre différentes branches sur le calcul des provisions. *Thème ASTIN : Contrôle des risques*.
- [18] HØYLAND, K., KAUT, M. et WALLACE, S. (2003). A heuristic for moment-matching scenario generation. *Computational Optimization and Applications*, 24(2-3):169–185.
- [19] HÜRLIMANN, W. (2007). An aggregate claims model between independence and comonotone dependence. *New Dimensions in Fuzzy Logic and Related Technologies*.
- [20] KAAS, R., LAEVEN, R. J. et NELSEN, R. B. (2009). Worst var scenarios with given marginals and measures of association. *Insurance : Mathematics and Economics*, 44(2):146 – 158.
- [21] KAUT, M. (2013). A copula-based heuristic for scenario generation. *Computational Management Science*, 25.
- [22] KRAUTH, G. (2007). Provisionnement et corrélation entre branches. *Mémoire d'actuariat*.
- [23] LAGENEBRE, V. (2009). Prise en compte des dépendances entre risques par la théorie des copules : Impact sur le besoin en fonds propres de l'assureur. *Mémoire d'actuariat*.
- [24] LEBLANC, R. (2000). Etude de l'effet de la dépendance dans le modèle collectif de risque. *Faculté des études supérieures de l'Université Laval*.
- [25] LI, J. (2005). Modelling dependency between different lines of business with copulas. *Centre for Actuarial Studies, University of Melbourne, Australia*.
- [26] LUO, X. et SHEVCHENKO, P. V. (2010). The t copula with multiple parameters of degrees of freedom : bivariate characteristics and application to risk management. *Quantitative Finance*, 10(9):1039–1054.
- [27] PAUL EMBRECHTS, A. J. M. et STRAUMANN, D. (2002). Correlation and dependence in risk management : Properties and pitfalls. *Cambridge University Press*.

-
- [28] SCHMID, F. et SCHMIDT, R. (2007). Multivariate extensions of spearman's rho and related statistics. *Statistics and Probability Letters*, 77(4):407 – 416.
- [29] SIBILLEAU (2007). Réassurance. *Cours de l'ISFA*.
- [30] SKLAR, M. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publication de l'ISUP*.
- [31] WALHIN, J.-F. (2001). On the practical multiline excess of loss pricing. *CAS Forum*.
- [32] WALHIN, J.-F. (2003). On the optimality of multiline excess of loss covers. *Secura Belgian Re*.
- [33] WANG, P. (2013). Risk modeling of multi-year, multi-line reinsurance using copulas. *Journal of Insurance Issues*, 36(1):58–81.
- [34] WANG, S. et DHAENE, J. (1998). Comonotonicity, correlation order and premium principles. *Insurance : Mathematics and Economics*, 22:235–242.
- [35] WANG, S. S. et D, P. (1997). Aggregation of correlated risk portfolios : Models and algorithms. *Preprint Casualty Actuarial Society (CAS)*, pages 848–939.