

ROYAUME DU MAROC  
\*.\*.\*.\*.\*  
HAUT COMMISSARIAT AU PLAN  
\*.\*.\*.\*.\*

INSTITUT NATIONAL  
DE STATISTIQUE ET D'ECONOMIE APPLIQUEE

**INSEA**



## Projet de Fin d'étude

\*\*\*\*\*

*Apport de la théorie des copules, de la théorie bayésienne et des mesures de risque dans la résolution des problèmes en Actuariat.*

Préparé par : *Mr Gbongue Kanga Florent*

Sous la direction de : *Mr Abderrahim Oulidi (INSEA)*  
*Mr Ahmed Serhouchni (RMA WATANYA)*

*Soutenu publiquement comme exigence partielle en vue de l'obtention du*  
*Diplôme d'Ingénieur d'Etat*

**Option : Actuariat et Finance**

*Devant le jury composé de :*

- *Mr Abderrahim Oulidi (INSEA)*
- *Mr L. MOUTIK (INSEA)*
- *Mr Ahmed Serhouchni (RMA WATANYA)*

*JUIN 2012*

Résumé et mots clés

*Un contrat d'assurance est un contrat dans lequel l'assureur s'engage à fournir des prestations à l'assuré en contrepartie de ses primes versées. En acceptant ce contrat, l'assureur prend des risques qu'il doit estimer pour espérer honorer ses engagements futurs.*

*Le risque en assurance est une notion très délicate car il peut occasionner des pertes financières importantes, voire en extrême la faillite de l'assureur. Dans cette optique et du fait de la pression de l'autorité de contrôle, les compagnies d'assurance entreprennent des travaux de recherche ou demande l'avis d'expert externe dans le but de déterminer le best estimate de leur engagements futurs.*

*Ces engagements concernent :*

- Le calcul des provisions techniques qui représentent 70 % à 80 % du passif*
- Le calcul du Solvency Capital Requirement (SCR)*
- Le calcul d'un capital « add on » en cas d'insuffisance du SCR, etc...*

*Dans ce mémoire, nous avons abordé les thèmes clés de la théorie des sciences actuarielles à savoir le calcul du ratio de sinistralité, des provisions techniques et du capital agrégé. Cependant l'accent a été mis sur les **provisions techniques**, du fait de son importance en assurance.*

*Dans le souci d'obtenir la meilleure estimation des provisions techniques, une panoplie de méthodes stochastiques a été utilisée. En premier lieu, nous les avons calculées en utilisant le modèle GLM (Generalized Linear Model) qui est une extension du modèle linéaire, dans lequel les erreurs sont non gaussiennes et la distribution de la variable dépendante appartient à la famille exponentielle. Les distributions testées sont la loi log-normale, la loi de gamma et la loi de poisson pour la fonction lien « log ».*

*Le problème réside du fait que dans le modèle GLM, la forme paramétrique de la variable dépendante à une grande influence sur la prédiction des charges futures. Or les données à notre disposition ne nous donnent aucune information sur le comportement probabiliste de notre variable d'intérêt d'où la nécessité d'utiliser un modèle qui induit une indépendance entre la forme paramétrique de notre variable dépendante et les variables explicatives du modèle. Ceci nous a poussé à utiliser une régression non linéaire, non paramétrique pour prédire les charges futures. Parmi ces modèles non paramétriques figurent le modèle GAM*

(Generalized Additive Model). Dans ce dernier, il y'a une indépendance entre le comportement paramétrique de la variable dépendante et le comportement non linéaire des variables explicatives du modèle. Ce modèle est très utile pour la prédiction des charges futures mais la difficulté réside dans le choix des meilleurs paramètres (degré de liberté par exemple, etc...) des fonctions de lissage des variables explicatives. Il se fait de sorte à minimiser la quantité GCV (Generalized Cross Validation) et de ce fait d'obtenir un GCV score très faible. Plusieurs combinaisons de modèle GAM peuvent fournir un GCV score faible et par extension une baisse des provisions d'où la nécessité d'utiliser une méthode déterministe ou stochastique de provisionnement (Chain Ladder ou Mack Chain Ladder) comme référence pour ajuster les paramètres de la fonction de lissage. Aussi dans le choix du meilleur modèle GAM, les représentations graphiques et le critère AIC jouent un rôle déterminant dans le choix du meilleur modèle. Il est donc important de tenir compte de cela pour chaque modèle testé.

Parfois, il peut arriver que les résidus du modèle GAM soient autocorrolés, ceci nous permet de passer du modèle GAM vers un modèle GAM Mixed (GAMM) qui offre la possibilité à l'utilisateur d'insérer une structure de corrélation des erreurs.

Toute ces étapes sont indispensables pour déterminer le best estimate des provisions techniques qui répond à la vision à l'ultime de Solvabilité ii.

Après avoir utilisé les méthodes GAM et GAM Mixed dans le calcul des provisions techniques, nous nous sommes intéressés au calcul de ces provisions par la méthode bayésienne en nous inspirant des travaux du suisse Mario V. Wuthrich. Ici, nous avons calculé les provisions par la méthode **Bornhuetter-Ferguson et Benktander-Hovinen**.

Après les provisions techniques, des chapitres traitant le ratio de sinistralité et du capital agrégé et le problème d'allocation du capital agrégé viendront clôturer le mémoire.

En ce qui concerne le ratio de sinistralité, nous essayerons de calculer un ratio de sinistralité agrégé (pour trois branches) pour l'année de survenance  $i$ . Pour ce faire, nous nous sommes inspiré de l'article d'Andrew Tang et Emiliano A. Valdez intitulé « Economic Capital and the Aggregation of Risks using Copulas ».

Le capital agrégé obtenu dans ce mémoire est un capital pour les trois branches (B, C et D). Il a été obtenu en calculant la Value at Risk à 99 % sur un horizon d'une année.

L'allocation du capital agrégé consiste à allouer le capital obtenu sur l'ensemble des trois branches en fonction de la contribution au risque. Plusieurs méthodes d'allocation existent, cependant, nous avons présenté les plus importantes et facile à implémenter.

**Mots clés :** Provisions techniques, capital agrégé, ratio de sinistralité, allocation de capital

Dédicace

Je dédie ce mémoire en premier lieu à mon père qui a soutenu financièrement, moralement tout le long de mes cinq années au Maroc. Je ne saurais comment te rendre tout ce que tu as fait pour moi.

En deuxième lieu, je dédie aussi ce mémoire à ma mère dont mon amour pour elle est inestimable. Je ne peux que te dire Merci Maman pour ton amour et pour ta patience.

En troisième lieu, ce mémoire est dédié à mon seul frère Parfait que j'admire tant ainsi qu'à toutes mes trois sœurs.

Je ne saurai terminer cette dédicace sans remercier Micheline pour tout son soutien.

Remerciement

*Mes remerciements et ma gratitude vont à l'endroit de Monsieur Oulidi Abderrahim parce qu'il a toujours été présenté pour ses élèves et aussi pour la belle formation qui nous a offert en assurance et en réassurance tout le long des deux années passées avec lui. Il nous a inculqué en nous l'amour des sciences actuarielles. Je ne saurai comment lui être reconnaissant.*

*Mes remerciements vont aussi à l'endroit de mon encadrant externe Monsieur Ahned Serhrouchni qui m'a accueilli au sein de la RMA WATANYA et qui m'a procuré souvent des astuces du monde professionnel. J'aimerais aussi remercier tout le personnel du département Risk Management pour leur accueil chaleureux en occurrence Reda qui m'a beaucoup soutenu. Il a été pour moi un vrai ami.*

*Je ne saurai terminer ces remerciements sans remercier mon ami Fatou Fofana qui m'a permis de m'installer facilement à Casablanca et de travailler dans de bonnes conditions, ainsi que tous les professeurs de l'INSEA pour la formation de pointe qu'ils nous ont procurée tout le long de notre cycle.*

**Table des matières**

<i>Page de Titre .....</i>	<i>1</i>
<i>Résumé et mots clés.....</i>	<i>2</i>
<i>Dédicace.....</i>	<i>4</i>
<i>Remerciement.....</i>	<i>5</i>
<i>Table des matières.....</i>	<i>6</i>
<i>Liste des abréviations.....</i>	<i>9</i>
<i>Liste des figures.....</i>	<i>10</i>
<i>Liste des tableaux.....</i>	<i>12</i>
<i>Introduction.....</i>	<i>14</i>
<i>Chapitre 1 : Aspects introductifs.....</i>	<i>16</i>
<i>I. Solvabilité 2 .....</i>	<i>16</i>
<i>I.1. Route sur les trois piliers de Solvabilité 2 .....</i>	<i>16</i>
<i>I.2. Approche bilancielle.....</i>	<i>18</i>
<i>Chapitre 2 : Outils Mathématiques.....</i>	<i>19</i>
<i>I. Théorie des copules.....</i>	<i>19</i>
<i>I.1. Propriétés et notion de dépendance.....</i>	<i>19</i>
<i>I.2. Panorama des familles des copules.....</i>	<i>27</i>
<i>I.3. Estimation des paramètres d'une copule.....</i>	<i>34</i>
<i>I.4. Choix de la meilleure copula.....</i>	<i>38</i>
<i>I.5. Simulation des copules.....</i>	<i>40</i>

<b>II.</b>	<b><i>Théorie bayésienne.....</i></b>	<b><i>42</i></b>
<b>II.1.</b>	<b><i>Introduction .....</i></b>	<b><i>42</i></b>
<b>II.2.</b>	<b><i>Lois à postériori.....</i></b>	<b><i>42</i></b>
<b>II.3.</b>	<b><i>Application de la théorie bayésienne : Mise à jour des primes .....</i></b>	<b><i>44</i></b>
<b>III.</b>	<b><i>Modèles linéaires.....</i></b>	<b><i>45</i></b>
<b>IV.</b>	<b><i>Modèles linéaires généralisés.....</i></b>	<b><i>47</i></b>
<b>V.</b>	<b><i>Modèles additives généralisés.....</i></b>	<b><i>51</i></b>
<b>VI.</b>	<b><i>Bootstrap.....</i></b>	<b><i>52</i></b>
<b>VII.</b>	<b><i>Approche probabiliste du risque .....</i></b>	<b><i>54</i></b>
<b>VIII.</b>	<b><i>Mesure du risque.....</i></b>	<b><i>55</i></b>
<b><i>Chapitre 3 : Les provisions techniques en assurance non vie.....</i></b>		<b><i>61</i></b>
<b>I.</b>	<b><i>Méthodes basiques de calcul des provisions techniques.....</i></b>	<b><i>62</i></b>
<b>I.1.</b>	<b><i>Méthode de Chain Ladder.....</i></b>	<b><i>62</i></b>
<b>I.2.</b>	<b><i>Méthode de Bornhuetter-Ferguson.....</i></b>	<b><i>63</i></b>
<b>I.3.</b>	<b><i>Combinaison de la méthode Chain Ladder (CL) et de la méthode Bornhuetter-Ferguson (BF).....</i></b>	<b><i>64</i></b>
<b>I.4.</b>	<b><i>Application sur la branche Automobile Corporel .....</i></b>	<b><i>64</i></b>
<b>II.</b>	<b><i>Modèles non linéaires stochastiques de calcul des provisions.....</i></b>	<b><i>68</i></b>
<b>II.1.</b>	<b><i>Modèles de calcul des provisions techniques par la méthode GAM.....</i></b>	<b><i>69</i></b>
<b>II.2.</b>	<b><i>Modèles de calcul des provisions techniques par la méthode GLM.....</i></b>	<b><i>91</i></b>
<b><i>Chapitre 4 : Ratio de sinistralité d'une compagnie d'assurance non vie.....</i></b>		<b><i>97</i></b>
<b>I.</b>	<b><i>Introduction.....</i></b>	<b><i>97</i></b>
<b>II.</b>	<b><i>Présentation du modèle.....</i></b>	<b><i>97</i></b>
<b>III.</b>	<b><i>Application sur trois branches d'activités d'un assureur non vie.....</i></b>	<b><i>98</i></b>

<i>Chapitre 5 : Capital agrégé pour des lignes d'affaires.....</i>	<i>100</i>
<i>I. Introduction.....</i>	<i>100</i>
<i>II. Présentation du modèle.....</i>	<i>103</i>
<i>III. Application sur trois branches d'activités d'un assureur non vie.....</i>	<i>107</i>
<i>Chapitre 6 : Problème d'allocation du capital agrégé.....</i>	<i>109</i>
<i>I. Introduction.....</i>	<i>109</i>
<i>II. Modèle d'allocation de capital.....</i>	<i>111</i>
<i>Conclusion.....</i>	<i>114</i>
<i>Bibliographie.....</i>	<i>115</i>
<i>Annexe.....</i>	<i>116</i>

Liste des abréviations

**BF: Bornhuetter-Ferguson**

**BE: Best Estimate**

**BH: Benktander-Hovinen**

**CL: Chain Ladder**

**CTE: Conditional Tail Expectation**

**GAM: Generalized Additive Model**

**GLM: Generalized Linear Model**

**GCV: Generalized Cross Validation**

**MCR: Minimum Capital Requirement**

**RM: Risk Margin**

**SCR: Solvency Capital Requirement**

**VaR: Value at Risk**

Liste des figures

Figure 1 : Bilan sous Solvabilité 2 .....	16
Figure 2 : Simulation de 100000 observations d'une copule normale de dimension 2 avec $\Sigma = 0.8$ .....	27
Figure 3 : Simulation de 100000 observations d'une copule de Student de dimension 2 avec $\Sigma = 0.8$ et $v = 10$ .....	28
Figure 4 : Simulation de 100000 observations d'une copule de Clayton de paramètre $\alpha = 3$ et de dimension 2 avec une corrélation de Pearson $\Sigma = 0.8$ .....	30
Figure 5 : Simulation de 100000 observations d'une copule de Frank de paramètre $\alpha = 3$ et de dimension 2 avec une corrélation de Pearson $\Sigma = 0.45$ .....	31
Figure 6 : Simulation de 100000 observations d'une copule de Gumbel de paramètre $\alpha = 3$ et de dimension 2 avec une corrélation de Pearson $\Sigma = 0.85$ .....	32
Figure 7 : Triangle de liquidation de la branche Automobile Corporel .....	63
Figure 8 : Evolution des paiements cumulés en fonction de l'année de développement ....	64
Figure 9 : Evolution des paiements cumulés en fonction de l'année de survenance .....	64
Figure 10 : Structure de lissage de la variable « année de survenance » de la branche Accident de travail pour la loi log normale.....	
Figure 11 : Structure de lissage de la variable « année de développement » de la branche Accident de travail pour la loi log normale .....	74
Figure 12 : Structure de lissage simultané des 2 variables « année de survenance et année de développement » de la branche Accident de travail pour la loi log normale .....	75
Figure 13 : Diagnostics du modèle GAM de la branche AT pour la loi log normale .....	75
Figure 14 : Structure de lissage de la variable « année de survenance » de la branche Automobile Corporel pour la loi log normale	
Figure 15 : Structure de lissage de la variable « année de développement » de la branche Automobile Corporel pour la loi log normale	
Figure 16 : Structure de lissage simultané des 2 variables « année de survenance et année de développement » de la branche Automobile Corporel pour la loi log normale	
Figure 17 : Diagnostics du modèle GAM de la branche Automobile Corporel pour la loi log normale	

Figure 18 : Structure de lissage de la variable « année de survenance » de la branche Automobile Matériel pour la loi log normale	
Figure 19 : Structure de lissage de la variable « année de développement » de la branche Automobile Matériel pour la loi log normale	
Figure 20 : Structure de lissage simultané des 2 variables « année de survenance et année de développement » de la branche Automobile Matériel pour la loi log normale	
Figure 21 : Diagnostics du modèle GAM de la branche Auto Matériel pour la loi log normale	
Figure 22 : Histogramme des trois branches dans un modèle GAM-Bootstrap pour la loi log normale.....	
Figure 23 : Distribution agrégée des trois branches dans un modèle GAM-Bootstrap pour la loi log normale.....	
Figure 24 : Représentation graphique de la dépendance entre la branche Automobile Corporel et Accident de travail .....	84
Figure 25 : Résultats de l'autocorrélation entre les résidus du modèle GAM .....	86
Figure 26 : Résultats du calcul des provisions techniques de la branche Automobile Corporel dans le modèle de <i>Mario V. Wuthrich</i> .....	94

**Liste des tableaux**

Tableau 1 : récapitulatif des paramètres de la densité de la famille exponentielle pour les distributions usuelles.....46

Tableau 2 : Capture d'écran des facteurs de développement de la méthode Chain Ladder pour la branche Automobile Corporelle ..... 63

Tableau 3 : Capture d'écran des facteurs de développement de la méthode Bornhuetter-Ferguson pour la branche Automobile Corporelle .....65

Tableau 4 : récapitulatif des provisions techniques pour les trois méthodes déterministes....65

Tableau 5 : Provisions techniques sans introduction du Bootstrap dans un modèle GAM pour la loi log normale : Données « as if ».....73

Tableau 6 : Provisions techniques sans introduction du Bootstrap dans un modèle GAM pour la loi log normale: Données non « as if ».....73

Tableau 7 : Provisions techniques avec introduction du Bootstrap dans un modèle GAM pour la loi log normale: Données « as if ».....79

Tableau 8 : Provisions techniques avec introduction du Bootstrap dans un modèle GAM pour la loi log normale : Données non « as if ».....79

Tableau 9 : Provisions techniques sans introduction du Bootstrap dans un modèle GAM pour la loi Gamma : Données non « as if ».....82

Tableau 10 : Provisions techniques sans introduction du Bootstrap dans un modèle GAM pour la loi Gamma : Données « as if ».....82

Tableau 11 : Provisions techniques avec introduction du Bootstrap dans un modèle GAM pour la loi Gamma : Données non « as if ».....82

Tableau 12 : Provisions techniques avec introduction du Bootstrap dans un modèle GAM pour la loi Gamma : Données « as if ».....83

Tableau 13 : Provisions techniques sans introduction du Bootstrap dans un modèle GAM pour la loi de poisson : Données non « as if ».....83

Tableau 14 : Provisions techniques sans introduction du Bootstrap dans un modèle GAM pour la loi de poisson : Données « as if ».....83

Tableau 15 : Provisions techniques sans introduction du Bootstrap dans un modèle GAM-Copule pour la loi log Gamma.....	86
Tableau 16 : Provisions techniques dans un modèle GAM-Copule-Bootstrap pour la loi de Gamma .....	86
Tableau 17 : Provisions techniques dans un modèle GAM Mixed pour la loi de Gamma...	87
Tableau 18 : Provisions techniques dans un modèle GAM Mixed-Bootstrap pour la loi de Gamma.....	87
Tableau 19 : Provisions techniques dans un modèle GLM pour la loi log normal.....	89
Tableau 20 : Provisions techniques dans un modèle GLM-Bootstrap pour la loi log normal.....	89
Tableau 21 : Provisions techniques dans un modèle GLM pour la loi de Gamma.....	90
Tableau 22 : Provisions techniques dans un modèle GLM-Bootstrap pour la loi de Gamma.....	90
Tableau 23 : Provisions techniques dans un modèle GLM pour la loi de poisson .....	90
Tableau 24 : Résultats des ratios de sinistralité agrégés par année de survenance.....	97
Tableau 25 : Résultats des capitaux agrégés par année de survenance .....	105

Introduction

Vu les évènements extrêmes qui ont marqué le monde de la finance ces dernières années, une nouvelle norme de solvabilité a vu le jour sous l'appellation de Solvabilité 2. Cette dernière s'inspire de Bâle 2 qui est la réglementation du système bancaire. Cette nouvelle norme oblige les assureurs à percevoir le risque d'une autre façon. C'est donc l'approche stochastique qui est encouragée en lieu et place de l'approche déterministe.

Pour implémenter un système de solvabilité au sein d'une entreprise d'assurance, deux approches sont possibles :

- ✚ L'approche standard
- ✚ L'approche par modèle interne

Ce mémoire reprend cette dernière approche qui est la mise en place d'un modèle interne au sein d'une compagnie d'assurance non vie. Une nouveauté introduite est la prise en compte de la dépendance entre les facteurs de risque.

Ceci nous pousse à introduire la fonction « copula » qui permet selon le théorème de SKLAR de créer un lien entre les marginales des facteurs de risque et de déduire la fonction de répartition jointe d'un vecteur de risque.

La prise en compte de la dépendance entre les risques est très importante pour un assureur. Pour preuve, nous remarquons que dans le modèle standard, la détermination du capital agréé (SCR) utilise une matrice de corrélation linéaire entre les différents risques, qui est en effet, une manière de tenir compte de la dépendance entre les risques (dépendance linéaire).

Une autre preuve est parue dans le monde (en septembre 2002) comme suit : *“la destruction des tours de New York a provoqué une secousse sans précédent chez les assureurs, mettant à mal tout leur système de fonctionnement. Aucun d'entre eux n'avait prévu qu'une telle corrélation entre les branches « vie » et « non vie » de l'assurance soit possible, mêlant les dommages aux biens, les pertes d'exploitation, les accidents du travail et les décès. Ce cumul et le coût qui en découle estimé entre 36 et 54 milliards de dollars (entre 36,7 milliards et 55,1 milliards d'euros) oblige la profession à reconsidérer sa manière d'appréhender les catastrophes”*. Alors que l'hypothèse d'indépendance était sous-jacente, et fondamentale, dans la plupart des modèles actuariels, ce risque de cumul a montré qu'il n'était plus possible de modéliser les risques indépendamment les uns des autres.

Au vu des arguments formulés ci-dessus, ce mémoire utilisera ce concept de dépendance entre les variables stochastiques pour résoudre de grandes problématiques rencontrées dans les sciences actuarielles.

Tout d'abord, dans le chapitre 1, nous rappellerons la structure de Solvabilité 2 et la notion de l'Enterprise Risk Management en assurance.

Ensuite dans le chapitre 2, nous présenterons les outils mathématiques nécessaires à la résolution des problèmes.

Les chapitres 3, 4, 5 et 6 traiteront respectivement les notions de provisions techniques, de Ratio de sinistralité, de capital agrégé et d'allocation de capital.

Enfin, une conclusion générale permettra d'achever le sujet.

*Chapitre 1 : Aspects introductifs*

*I. Solvabilité 2*

**I.1. Route sur les trois piliers de Solvabilité 2**

La crise des subprimes survenue en juillet 2007 a conforté la commission européenne dans son idée de changer la réglementation du marché de l'assurance dont le but principal est de se prémunir des risques dits systémiques (défaillances en chaîne des acteurs du système bancaire et assurantiel du monde financier). C'est dans cette optique que s'inscrit le système Solvabilité 2. Il oblige les assureurs à redéfinir cette notion de risque en mettant un système performant de gestion intégré du risque.

Ce système a une architecture basée sur trois piliers dénommés : Pilier 1, Pilier 2 et Pilier 3 dont les plus importants concernent les deux premiers piliers.

**Le pilier 1** a pour objectif de définir les normes quantitatives de calcul des provisions techniques et des fonds propres. Ces niveaux réglementaires sont définis pour les fonds propres sous l'appellation MCR et SCR

Le MCR (*Minimum Capital Requirement*) représente le niveau minimum de fonds propres en dessous duquel l'intervention de l'autorité de contrôle sera automatique.

Le SCR (*Solvency Capital Requirement*) représente le capital cible nécessaire pour absorber le choc provoqué par un risque majeur (par exemple : un sinistre exceptionnel, un choc sur les actifs...).

Les provisions techniques quant à elles se définissent comme étant la somme du **best estimate** et d'un complément appelé '**risk margin**'.

**Cette notion de « best estimate » est différente selon le contexte et la catégorie d'assurance. En assurance vie, Il représente l'actualisation des engagements futurs de l'assureur auxquels l'on retranche souvent les primes (cas du QIS 5). Ici il s'agira de déterminer les cash flows futurs en valorisant le passif (modélisation de la mortalité, des rachats etc...) et ensuite de les actualiser pour obtenir le best estimate.**

**En assurance non vie, le best estimate est obtenu en utilisant des méthodes stochastiques de provisionnement (GAM, GLM, Bayésien, Mack ChainLadder etc...).**

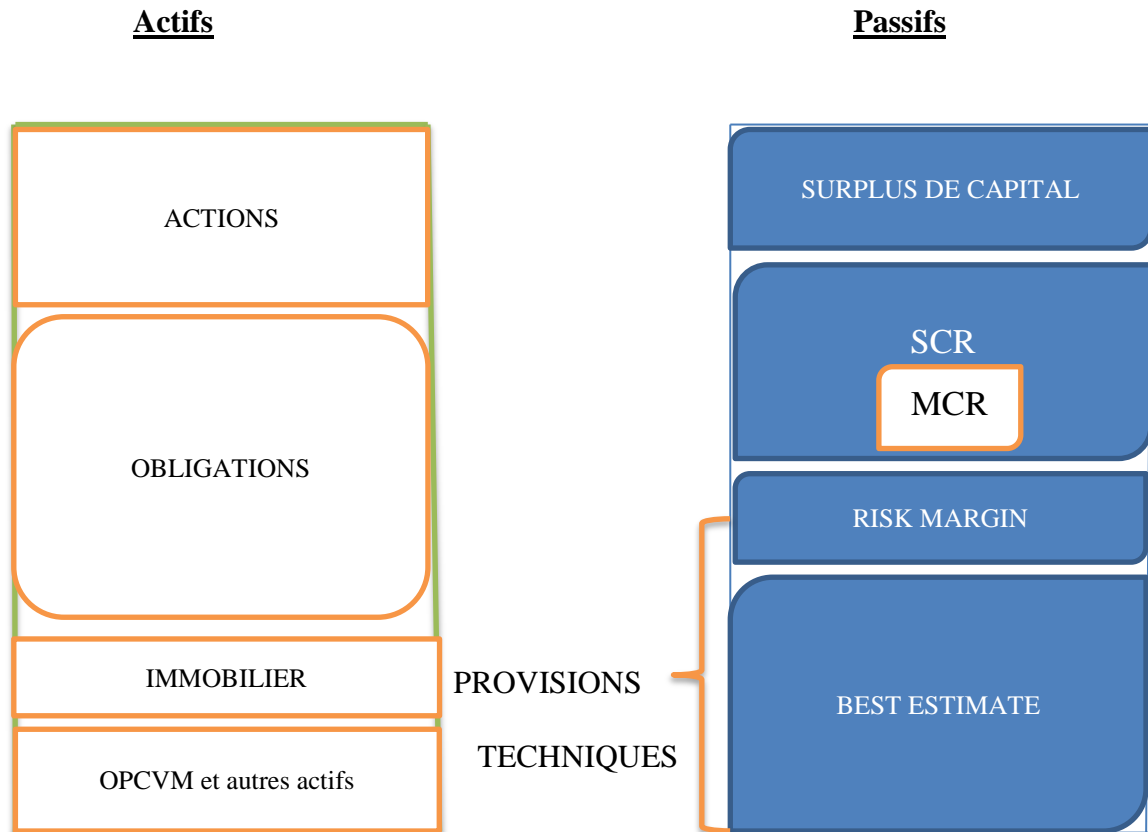
**Le pilier 2** a pour objectif de fixer des normes qualitatives de suivi des risques en interne aux sociétés et comment l'autorité de contrôle doit exercer ses pouvoirs de surveillance dans ce contexte. L'identification des sociétés "les plus risquées" est un objectif et les autorités de contrôle auront en leur pouvoir la possibilité de réclamer à ces sociétés de détenir un capital plus élevé que le montant suggéré par le calcul du SCR (capital add-on) et/ou de réduire leur exposition aux risques.

Vu les objectifs de ce pilier, une nouvelle notion a été introduite appelé « ORSA » ou l'Enterprise Risk Management (ERM) de l'assurance.

**Le pilier 3** a pour objectif de définir l'ensemble des informations détaillées auxquelles le public aura accès, d'une part, et auxquelles les autorités de contrôle pourront avoir accès pour exercer leur pouvoir de surveillance, d'autre part.

## I.2. Approche bilancielle

Le bilan sous Solvabilité 2 peut être vu par le schéma ci-dessous :



**FIGURE 1 : Bilan sous Solvabilité 2**

Dans le système de solvabilité 2, les actifs sont valorisés selon leur valeur du marché contrairement à solvabilité 1 dans lequel l'approche comptable est fortement privilégiée.

Plusieurs modèles financiers de valorisation des actifs existent. Le plus célèbre est le modèle de Black and Schole utilisé par les praticiens pour la valorisation des actifs risqués (actions, immobilier).

Chapitre 2 : Outils Mathématiques

I. Théorie des copules

Les copules sont des outils pour modéliser la dépendance entre les variables stochastiques. Cette notion « copula » fut utilisée pour la première fois dans le travail de SKLAR en 1959 comme solution d'un problème de probabilité énoncé par Maurice Fréchet dans le cadre des espaces métriques aléatoires et dérive du mot latin « copulare » qui signifie lien.

Cette notion révèle un caractère important dans le monde de l'assurance où l'on constate souvent une forte corrélation entre les branches de l'assureur comme c'est souvent le cas entre les branches 'accidents de travail et automobile corporel'. Ainsi ce constat peut avoir une grave conséquence sur la solvabilité de l'entreprise. Dans la suite de cette partie, nous allons approfondir cette notion de dépendance.

I.1. Propriétés et notion de dépendance

I.1.1. Propriétés des copules

1. Introduction

Une copule  $C$  de dimension  $d$  est une fonction de répartition multivariée définie sur l'hypercube  $[0,1]^d$  et dont les marginales sont uniformes sur  $[0,1]$ .

En clair, c'est une fonction de  $[0,1]^d$  vers  $[0,1]$  vérifiant les trois propriétés suivantes :

1.  $c(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_d)$  est croissante au sens large en chaque composante  $\mathbf{u}_i$ .
2. Pour tout vecteur  $u$  dans  $[0,1]^d$ ,  $C(u) = 0$  si au moins l'une des coordonnées de  $u$  est égale à 0, et  $C(u) = \mathbf{u}_i$  si toutes les coordonnées de  $u$  sont égales à 1 sauf  $\mathbf{u}_i$ .
3. **Propriété de supermodularité :**

Pour  $(\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_d), (\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_d) \in [0,1]^d$  tels que  $\mathbf{a}_i \leq \mathbf{b}_i, i = 1 \dots d$ , alors :

$$\sum_{i_1=1}^2 \dots \sum_{i_d=1}^2 (-1)^{i_1+\dots+i_n} \times C(\mathbf{u}_{1,i_1}, \mathbf{u}_{2,i_2}, \dots, \mathbf{u}_{d,i_d})$$

Où  $u_{i,1} = a_i$  et  $u_{i,2} = b_i$  pour tout  $i$  dans  $\{1 \dots d\}$ .

En dimension 2, nous pouvons réécrire cette définition comme suit :

Une copule bivariée  $C$  est une fonction non décroissante et continue à droite définie de  $[0,1]^2$  vers  $[0,1]$  et satisfaisant les conditions suivantes :

- a.  $\lim_{u_i \rightarrow 0} C(u_1, u_2) = 0 \quad \forall i = 1, 2 ;$

b.  $\lim_{u_1 \rightarrow 1} C(u_1, u_2) = u_2$  et  $\lim_{u_2 \rightarrow 1} C(u_1, u_2) = u_1$  ;

c.  $C$  est super-modulaire tel que l'inégalité

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

est valide pour  $u_1 \leq v_1, u_2 \leq v_2$

### Théorème 1 : Continuité d'une copule

Soit une copule  $C$  de dimension  $d$  tel que pour tout  $u$  et  $v$  dans  $[0,1]^d$ ,

$$|C(u) - C(v)| \leq \sum_{i=1}^d |v_i - u_i| , \text{ alors } C \text{ est uniformément continue sur } [0,1]^d .$$

SKLAR (1959) établit un lien entre la marginale jointe et les marginales individuelles. Ce lien est appelé « la fonction copule » ce qui nous donne la possibilité d'utiliser la copule si l'on veut déterminer la distribution d'un ensemble de variables aléatoires.

### Théorème 2 : Théorème de SKLAR

L'idée de Sklar est que la structure de dépendance entre les variables stochastiques  $X_1, X_2, \dots, X_d$  est décrit par une copule si et seulement si la distribution jointe est donnée par :

$$F(x_1, x_2, \dots, x_d) = C(F_1[x_1], F_2[x_2], \dots, F_d[x_d]). \quad (\text{Eq 1})$$

Où  $F_i$  désigne la fonction de distribution marginale de  $x_i$  et  $F$  la fonction de répartition jointe.

En faisant l'hypothèse que les  $F_i, i = 1, 2, \dots, d$  sont continues, alors cela implique que la copule  $C$  est unique.

### Corollaire 1 : Inversion du théorème de SKLAR

Une copule peut être aussi vue comme une fonction qui associe les quantiles de l'une des variables aléatoires avec les quantiles des autres variables aléatoires.

En effet, en réécrivant l'équation énoncée dans le théorème de SKLAR, nous obtenons le résultat suivant :

$$C(u_1, u_2, \dots, u_d) = F(F_1^{-1}[u_1], F_2^{-1}[u_2], \dots, F_d^{-1}[u_d]),$$

Avec  $F_i^{-1}$  la fonction de répartition inverse définie comme suit :

$$F_i^{-1}(u_i) = \{x_i / F_i(x_i) = u_i\} \quad \forall i \in [1, d]$$

A partir de cette expression, nous déduisons que la copule n'est autre qu'une fonction de répartition dont les marginales sont uniformes. Elle peut être écrite comme suit :

Soit  $U_1, U_2, \dots, U_d$  des variables aléatoires uniformes, alors :

$$C(u_1, u_2, \dots, u_d) = P(U_1 \leq u_1, \dots, U_d \leq u_d)$$

**Théorème 3 : Invariance de la fonction copule**

Soit  $T_1, \dots, T_d$  des fonctions continues et croissantes, alors  $(X_1, X_2, \dots, X_d)^t$  et  $(T(X_1), T(X_2), \dots, T(X_d))^t$  ont la même copule C.

**2. Densité d'une copule et distribution conditionnelle**

La fonction de densité d'une copule est donnée par :

$$c(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_d) = \frac{\partial^d C(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_d)}{\partial \mathbf{u}_1 \dots \partial \mathbf{u}_d} = \frac{f(F_1^{-1}[\mathbf{u}_1], \dots, F_d^{-1}[\mathbf{u}_d])}{f_1(F_1^{-1}[\mathbf{u}_1]) \dots f_d(F_d^{-1}[\mathbf{u}_d])}$$

Avec  $f$  la fonction de densité jointe de  $F$  ;  $f_1, \dots, f_d$  les densités marginales de  $F_1, \dots, F_d$  et  $F_1^{-1}, \dots, F_d^{-1}$  les fonctions de distributions marginales inverses ;

Si et seulement si les hypothèses suivantes sont satisfaites :

- La copule C est une copule absolument continue et différentiable.
- F est une fonction de répartition jointe absolument continue.
- Les fonctions de répartition marginales  $F_1[x_1], F_2[x_2], \dots, F_d[x_d]$  sont continues et strictement croissantes.

Denuit et al (2005) ont introduit une propriété importante qui permet de déterminer la densité jointe d'une suite de variables aléatoires à partir de la densité de la copule.

Intuitivement, cela suppose une expression close de la densité de la copule ce qui n'est pas le cas pour certaines familles de copules.

Cette propriété s'énonce comme suit :

**Propriété 1**

Soit  $F_1, \dots, F_d$  des fonctions de répartition continues avec des densités marginales respectives  $f_1, \dots, f_d$ . La distribution jointe de  $(X_1, X_2, \dots, X_d)^t$  est donnée par :

$$f(x_1, x_2, \dots, x_d) = \prod_{i=1}^d f_i(x_i) \times c(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_d) \quad (\text{Eq 2})$$

Avec :  $c(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_d) = \frac{\partial^d C(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_d)}{\partial \mathbf{u}_1 \dots \partial \mathbf{u}_d}$

**Interprétation de cette propriété**

Cette équation signifie qu'en cas de dépendance, la densité jointe de  $(X_1, X_2, \dots, X_d)^t$  est égale au produit de la densité jointe de  $(X_1, X_2, \dots, X_d)^t$  en cas d'indépendance par la

**densité de la copule C de dimension d.** Autrement dit, si  $(X_1, X_2, \dots, X_d)^t$  sont indépendants, alors  $c=1$ , ce qui implique que :

$$f(x_1, x_2, \dots, x_d) = \prod_{i=1}^d f_i(x_i)$$

La distribution conditionnelle de la copule quant à elle se définit comme suit :

$$\begin{aligned} c(u_d/u_1, u_2, \dots, u_{d-1}) &= \frac{\partial^{d-1} C(u_1, u_2, \dots, u_d)}{\partial u_1 \dots \partial u_{d-1}} / \frac{\partial^{d-1} C(u_1, u_2, \dots, u_{d-1})}{\partial u_1 \dots \partial u_{d-1}} \\ &= P(U_d \leq u_d / U_1 = u_1, \dots, U_{d-1} = u_{d-1}) \end{aligned}$$

**NB :**

Les copules peuvent être utilisées pour combiner des risques quand leurs distributions marginales sont estimées individuellement. Aussi la distribution jointe obtenue préserve les mêmes caractéristiques que les marginales. Dans cette optique Rosenberg and Schuermann (2004) ont montré qu'une dépendance normale peut être préservé en utilisant une copule normale pour joindre des distributions marginales individuelles différentes (weibull, log-normal par exemple).

### **3. Bornes de Fréchet-Hoeffding**

Dès 1940, Hoeffding travailla sur les distributions multivariées. Certes, il n'a pas introduire cette notion de copule directement mais ces travaux ont contribué à déduire beaucoup de résultats intéressants.

Notons :

$$C^-(u_1, u_2, \dots, u_d) = \max\{\sum_{i=1}^d u_i + 1 - d; 0\}$$

$$C^+(u_1, u_2, \dots, u_d) = \min\{u_1, u_2, \dots, u_d\}$$

Considérons une copule  $C(u_1, u_2, \dots, u_d)$  de dimension d, alors nous avons l'inégalité suivante :

$$C^-(u_1, u_2, \dots, u_d) \leq C(u_1, u_2, \dots, u_d) \leq C^+(u_1, u_2, \dots, u_d)$$

Ou les fonctions copules  $C^-$  et  $C^+$  sont appelées respectivement la borne inférieure et la borne supérieure de **Fréchet-Hoeffding**.

#### 4. Copule de survie

La copule de survie  $\tilde{\mathcal{C}}(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_d)$  est une copule dont l'expression s'écrit comme suit :

$$\tilde{\mathcal{C}}(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_d) = \bar{\mathcal{C}}(\mathbf{1} - \mathbf{u}_1, \mathbf{1} - \mathbf{u}_2, \dots, \mathbf{1} - \mathbf{u}_d)$$

Avec :  $\bar{\mathcal{C}}(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_d) = P(U_1 > \mathbf{u}_1, U_2 > \mathbf{u}_2, \dots, U_d > \mathbf{u}_d)$

Notons que la fonction  $\tilde{\mathcal{C}}(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_d)$  est une copule car elle remplit les critères liés à la définition. Par contre  $\bar{\mathcal{C}}(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_d)$  n'est pas une copule.

#### 1.1.2. Notion de dépendance

##### 1. Introduction

On dit que deux variables  $X$  et  $Y$  sont indépendants si la valeur connue de l'une des variables n'a aucun effet sur la distribution de probabilité de l'autre variable.

Par contre, ils sont dépendants si l'une des variables peut influencer l'autre variable.

Soit  $F_X(x)$  et  $F_Y(y)$  les fonctions de répartition de  $X$  et  $Y$ .

$X$  et  $Y$  sont dépendants  $\overset{\text{implique}}{\iff} F_{X,Y}(x, y) \neq F_X(x) \times F_Y(y)$ ,

$\forall (x, y) \in \mathbb{R}^2$  Et  $F_{X,Y}(x, y) = P(X \leq x, Y \leq y)$ .

##### 2. Mesures de la dépendance

Scarsini (1984) a défini des propriétés intéressantes sur la mesure d'association entre deux risques.

Planchet (2011) quant à lui définit une mesure de dépendance comme étant une application qui associe à deux variables aléatoires un réel permettant de quantifier la force de la dépendance qui les lie.

##### Propriété 2

Une mesure de dépendance  $\delta$  est une mesure de concordance si elle possède les propriétés suivantes :

1.  $\delta(X_1, X_2) = \delta(X_2, X_1)$  -----> Symétrie
2.  $-1 \leq \delta(X_1, X_2) \leq 1$  -----> Normalisation
3.  $\delta(X_1, X_2) = 1 \overset{\text{implique}}{\iff} (X_1, X_2) =_{loi} (F_1^{-1}(U), F_2^{-1}(U))$  Avec  $U \sim U[0,1]$
4.  $\delta(X_1, X_2) = -1 \overset{\text{implique}}{\iff} (X_1, X_2) =_{loi} (F_1^{-1}(U), F_2^{-1}(U))$  Avec  $U \sim U[0,1]$

5. Si  $f$  est strictement monotone,

$$\delta(f(X_1), X_2) = \begin{cases} \delta(X_1, X_2) & \text{si } f \uparrow \\ -\delta(X_1, X_2) & \text{si } f \downarrow \end{cases}$$

### ✚ Coefficient de corrélation de Pearson

Le coefficient de corrélation de Pearson est une mesure d'association entre deux variables aléatoires permettant de capturer le degré de la relation linéaire qui existe entre ceux-ci. Pour deux variables  $X_1$  et  $X_2$  possédant des variances finies, Il se définit comme suit :

$$\rho_{\text{pears}}(X_1, X_2) = \frac{\text{Cov}(X_1, X_2)}{\sqrt{\text{Var}(X_1)} \times \sqrt{\text{Var}(X_2)}}$$

Les variances de  $X_1$  et  $X_2$  doivent être finies pour que le coefficient de corrélation de Pearson soit défini. Ceci est donc un inconvénient pour les mesures de dépendance et cause des problèmes lorsqu'on travaille avec des distributions à queue lourde. Par conséquent les actuaires non vie qui modélisent les pertes entre les branches d'activités avec des distributions de variances infinies doivent être conscient de cela.

### Propriété

Soit  $t_1$  et  $t_2$  deux transformations strictement croissantes, alors :

$$\rho_{\text{pears}}(t_1(X_1), t_2(X_2)) \neq \rho_{\text{pears}}(X_1, X_2)$$

### ✓ Copule et $\rho_{\text{pears}}$

Soit  $F_1$  et  $F_2$  les fonctions de répartition de  $X_1$  et  $X_2$ , et  $C$  une copule alors

$$\rho_{\text{pears}} = \frac{1}{\sqrt{\text{Var}(X_1)} \times \sqrt{\text{Var}(X_2)}} \int_0^1 \int_0^1 (C(u_1, u_2) - u_1 u_2) dF_1^{-1}(u_1) dF_2^{-1}(u_2)$$

Nous remarquons que dans cette expression,  $\rho_{\text{pears}}$  dépend de la copule et des marginales.

### ✚ Coefficient de corrélation de rang de Kendall

C'est une mesure d'association non paramétrique basé sur le nombre de concordances et de discordances dans un échantillon d'observations pairées.

Autrement dit, soit  $(X_1, X_2)$  et  $(X'_1, X'_2)$  deux paires d'observations iid. On déduit :

$$P(\text{concordance}) = P([X_1 - X'_1][X_2 - X'_2] > 0)$$

Et

$$P(\text{discordance}) = P([X_1 - X'_1][X_2 - X'_2] < 0)$$

Le coefficient de corrélation de rang de Kendall noté  $\rho_{\text{Kendall}}$  se définit comme suit :

$$\rho_{Kendall}(X_1, X_2) = P(\text{concordance}) - P(\text{disconcordance})$$

Si les marginales de  $X_1$  et  $X_2$  sont continues, alors le  $\rho_{Kendall}$  peut être réécrit comme suit :

$$\rho_{Kendall}(X_1, X_2) = 2P([X_1 - X'_1][X_2 - X'_2] > 0) - 1$$

Où  $(X'_1, X'_2)$  est une copie indépendante de  $(X_1, X_2)$  et les deux couples sont mutuellement indépendants.

Une autre propriété intéressante est l'invariance du coefficient de corrélation de rang de Kendall c'est-à-dire que  $\forall t_1$  et  $t_2$  deux transformations strictement croissantes,

$$\rho_{Kendall}(t_1(X_1), t_2(X_2)) = \rho_{Kendall}(X_1, X_2)$$

✓ Copule et  $\rho_{Kendall}$

Comme :

$$\rho_{Kendall}(X_1, X_2) = 2P([X_1 - X'_1][X_2 - X'_2] > 0) - 1 = 4P(X_1 \leq X'_1, X_2 \leq X'_2) - 1$$

$$\text{Alors : } \rho_{Kendall}(X_1, X_2) = 4 \int_0^1 \int_0^1 C(u_1, u_2) dC(u_1, u_2) - 1 = 4E(C(U_1, U_2)) - 1$$

Avec :  $U_i \sim \text{Uniforme}(0,1) \quad \forall i = 1, 2$

✓  $\rho_{Kendall}$  pour une copule Archimédienne

Soit  $\varphi$  le générateur associé à la copule archimédienne.

$$\rho_{Kendall}^\varphi = 4 \int_0^1 \frac{\varphi(t)}{\varphi'(t)} dt + 1$$

La démonstration de ce résultat peut être vue dans Genest and Mackay (1986b).

En appliquant cette relation à la copule de Clayton  $\varphi(t) = \frac{t^{-\alpha}-1}{\alpha}$ , nous obtenons que :

$$\rho_{Kendall}^\varphi = \frac{\alpha}{\alpha + 2}$$

Avec  $\alpha$  le paramètre de la copule de Clayton.

✚ Coefficient de corrélation de rang de Spearman

Le coefficient de corrélation de Spearman  $\rho_{Spearman}$  pour une paire de variables continues  $X_1, X_2$  est identique au coefficient de corrélation de Pearson pour les rangs de  $F_1$  et  $F_2$ .

$$\rho_{Spearman}(X_1, X_2) = \rho_{pears}(F_1(X_1), F_2(X_2))$$

Comme dans le cas du coefficient de corrélation de Kendall,

$$\rho_{Spearman}(X_1, X_2) = 3(P([X_1 - X_1^\circ][X_2 - X_2^\circ] > 0) - P([X_1 - X_1^\circ][X_2 - X_2^\circ] < 0))$$

Où  $(X_1^\circ, X_2^\circ)$  est une version indépendante de  $(X_1, X_2)$ . En outre,  $(X_1^\circ, X_2^\circ)$  est un vecteur aléatoire avec des composants indépendants tel que  $X_1 =_{loi} X_1^\circ, X_2 =_{loi} X_2^\circ$  et  $(X_1, X_2)$  et  $(X_1^\circ, X_2^\circ)$  sont mutuellement indépendants.

Une autre propriété intéressante est l'invariance du coefficient de corrélation de rang de Spearman c'est-à-dire que  $\forall t_1$  et  $t_2$  deux transformations strictement monotones,

$$\rho_{Spearman}(t_1(X_1), t_2(X_2)) = \rho_{Spearman}(X_1, X_2)$$

✓ Copule et  $\rho_{Spearman}$

De la relation qui lie le coefficient de corrélation de Spearman et le coefficient de corrélation de Pearson, nous avons :

$$\begin{aligned} \rho_{Spearman}(X_1, X_2) &= \rho_{Pearson}(F_1(X_1), F_2(X_2)) \\ &= 12 \int_0^1 \int_0^1 (C(u_1, u_2) - u_1 u_2) du_1 du_2 \end{aligned}$$

✚ Coefficient de corrélation de Gini

Le coefficient de corrélation de Gini est utilisé en économie pour mesurer les différences de revenus entre deux populations.

Techniquement, il y'a une sorte de « distance » entre la structure de dépendance du vecteur  $X = (X_1, X_2)$  et la dépendance monotone représentée par la borne inférieure et supérieure de la copule de Fréchet.

Ainsi le coefficient de corrélation de Gini est défini par :

$$\begin{aligned} \rho_{Gini}(X_1, X_2) &= \rho_{Gini}(F_1(X_1), F_2(X_2)) \\ &= 2 \int_0^1 \int_0^1 (|u_1 + u_2 - 1| - |u_1 - u_2|) dC(u_1, u_2) \end{aligned}$$

Avec C une copule du couple  $(X_1, X_2)$ .

### 3. Dépendance de queue ou « Tail dependence »

C'est une notion très importante dans l'étude de la dépendance asymptotique entre deux variables aléatoires. Elle nous permettra d'apprécier le niveau de la dépendance dans les valeurs extrêmes (upper tail dependence) et dans les petites valeurs (lower tail dependence). Ces coefficients seront exprimés à partir des copules.

✚ upper tail dependence

L'objet est donc l'étude de la dépendance dans la queue commune de la distribution bi-variée.

Considérons deux variables stochastiques  $X$  et  $Y$  dont les fonctions de répartition respectives sont  $F_X$  et  $F_Y$ .

Le coefficient « upper tail dependence » est :

$$\lambda_u = \lim_{u \rightarrow 1^-} P(Y > F_Y^{-1}(u) \mid X > F_X^{-1}(u))$$
$$\lambda_u = \lim_{u \rightarrow 1^-} \frac{C(u, u) - 2u + 1}{1 - u}, \quad \lambda_u \in [0, 1]$$

Nous remarquons que ce coefficient est une fonction de la copule d'où il est invariant par transformation croissante.

Si :

- $\lambda_u \in [0, 1]$  Alors il existe une dépendance asymptotique
- $\lambda_u = 0$ , il y'a indépendance asymptotique

### lower tail dependence

Le coefficient « lower tail dependence » s'écrit comme suit :

$$\lambda_l = \lim_{u \rightarrow 0^+} P(Y \leq F_Y^{-1}(u) \mid X \leq F_X^{-1}(u))$$
$$\lambda_l = \lim_{u \rightarrow 0^+} \frac{C(u, u)}{u}, \quad \lambda_l \in [0, 1]$$

C'est aussi une fonction dépendante de la copule  $C$  d'où la propriété d'invariance par transformation croissante s'applique.

## I.2. Panorama des familles des copules

Il existe dans la littérature, deux grandes familles de copules à savoir les copules elliptiques et les copules archimédiennes. Cependant d'autres familles de copules existent. Parmi ces familles, nous pouvons citer les copules :

- à valeurs extrêmes
- de Marshall-Olkin
- de Fréchet
- de Farlie-Gumbel-Morgenstern

Nous allons dans cette sous-section présenter les différentes familles de copules en commençant par la famille des copules elliptiques.

### I.2.1. Les copules elliptiques

Soit  $\mathcal{M}_d(\mathfrak{R})$  l'ensemble des matrices carrées réelles de taille  $d^2$ . Une loi continue est dite elliptique de paramètre de position  $\mu = (\mu_1, \dots, \mu_d) \in \mathfrak{R}^d$  et de matrice de forme symétrique définie  $\Sigma \in \mathcal{M}_d(\mathfrak{R})$  si sa densité  $f$  peut s'écrire pour tout vecteur

$$x = (x_1, \dots, x_d) \in \mathfrak{R}^d :$$

$$f(x) = (\det \Sigma)^{-\frac{1}{2}} g((x - \mu) \Sigma^{-1} (x - \mu)')$$

Avec  $z' = (x - \mu)'$  la transposé de  $z = (x - \mu)$  et  $g$  une fonction à valeurs positives vérifiant  $\int g(xx') dx = 1 \quad \forall x \in \mathfrak{R}^d$ .

#### 1. La copule gaussienne

La copule gaussienne ne présente pas de dépendance de queue et n'est donc pas adaptée à des valeurs extrêmes. L'importance de cette copule réside dans le fait qu'elle est sous-jacente à la distribution normale multi-variée. En effet, modéliser la structure de dépendance d'un échantillon par une copule gaussienne est cohérent avec la mesure de cette dépendance par le coefficient de corrélation linéaire. La fonction de distribution de la copule gaussienne  $d$ -dimensionnelle, s'écrit pour tout  $(u_1, u_2, \dots, u_d)^t \in [0,1]^d$  :

$$C(u_1, u_2, \dots, u_d) = \Phi_{\Sigma}(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_d))$$

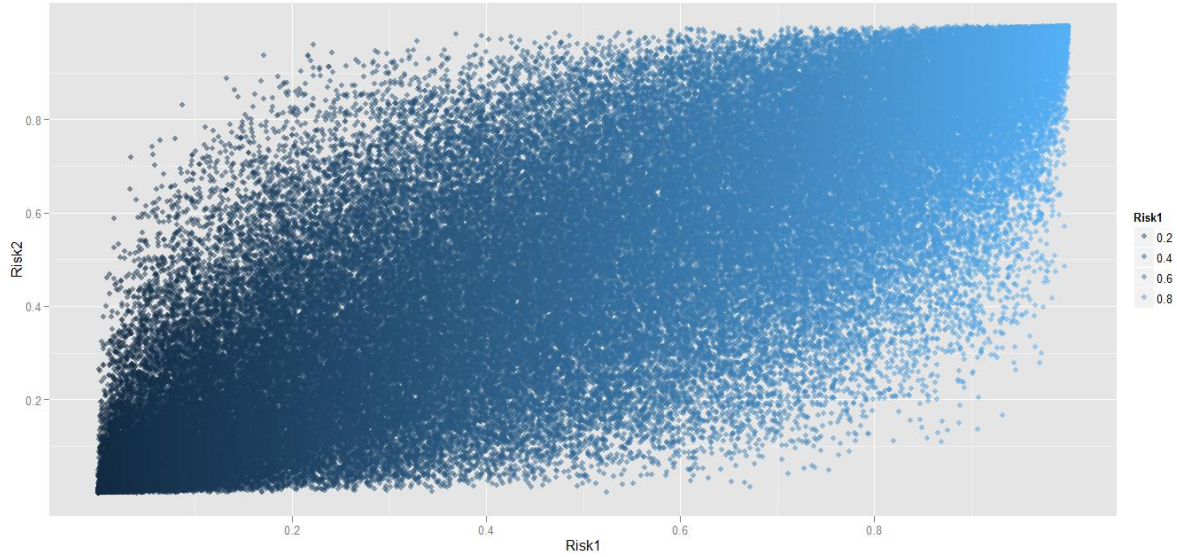
Avec  $\Phi^{-1}$  la fonction de répartition inverse de la loi normale centrée réduite.

Comme  $\Phi_{\Sigma}(x) = \frac{\exp(-\frac{1}{2} x \Sigma^{-1} x')}{(2\pi)^{d/2} \det(\Sigma)^{\frac{1}{2}}}$  est la fonction de répartition d'une loi normale multivariée centrée et réduite, alors la densité de la copule gaussienne de dimension  $d$  se définit comme suit :

$$c(u_1, u_2, \dots, u_d) = \frac{1}{\det(\Sigma)^{\frac{1}{2}}} \exp\left(-\frac{1}{2} \beta (\Sigma^{-1} - I_d) \beta'\right)$$

Où  $I_d$  est la matrice unité de  $\mathcal{M}_d(\mathfrak{R})$  et  $\beta = (\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_d))$

### Graphique



**Figure 2 : Simulation de 100000 observations d'une copule normale de dimension 2 avec  $\Sigma = 0.8$**

## 2. La copule de Student

La copule de Student (t copula) est une copule qui dérive de la distribution multivariée de Student. Cette structure de dépendance capte les dépendances extrêmes positives et négatives. Elle est construite de la même manière que la copule gaussienne mais à partir de la distribution de Student centrée réduite. La fonction de densité de la copule de Student en dimension  $d$  se définit comme suit :

$$\forall (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_d) \in [0, 1], \quad c(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_d) = \frac{f_{v,\Sigma}(\mathbf{t}_v^{-1}(\mathbf{u}_1), \dots, \mathbf{t}_v^{-1}(\mathbf{u}_d))}{\prod_{i=1}^d f_v(\mathbf{t}_v^{-1}(\mathbf{u}_i))}$$

**Avec :**

$f_{v,\Sigma}$  : La densité de la loi de student multivariée centrée et réduite, de matrice de corrélation  $\Sigma$  et de nombre de degré de liberté  $v$ .

$\mathbf{t}_v^{-1}$  : La fonction de répartition inverse de la loi de Student centrée et réduite univariée à  $v$  degré de liberté.

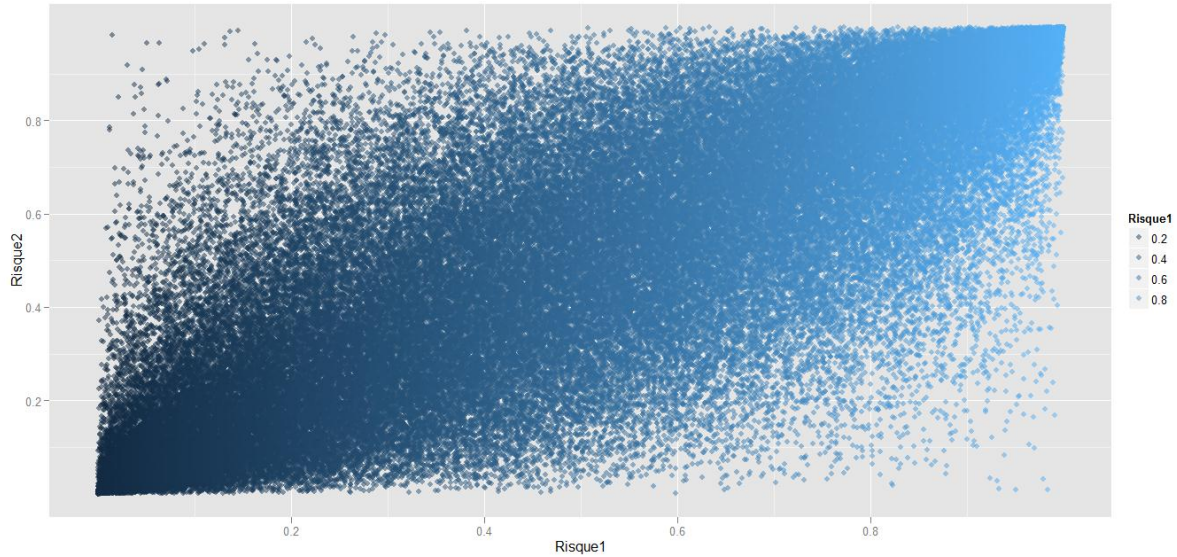
$f_v$  : La densité de la loi de student univariée centrée et réduite à  $v$  degré de liberté et  $\Sigma = \mathbf{1}$ .

En dimension  $d$ , la densité de la loi de student s'écrit comme suit :

$$f_{v,\Sigma} = \frac{\Gamma\left(\frac{v+d}{2}\right)}{\Gamma\left(\frac{v}{2}\right) \sqrt{(\pi v)^d \det(\Sigma)}} \times \left( \mathbf{1} + \frac{\mathbf{x}\Sigma^{-1}\mathbf{x}'}{v} \right)^{-(v+d)/2}$$

$\Gamma$  étant la fonction gamma. Sous R, la commande à utiliser est : `gamma ()`.

### Graphique



**Figure 3: Simulation de 100000 observations d'une copule de Student de dimension 2 avec  $\Sigma = 0.8$  et  $\nu = 10$**

### I.2.2. Les copules archimédiennes

GENEST et MACKAY [1986] ont introduit cette notion de copule archimédienne. Cette famille regroupe un certain nombre de copule dont la copule de Clayton, Frank et Gumbel.

L'idée sous-jacente d'une copule archimédienne de générateur  $\varphi$  est que la transformation  $\omega(\mathbf{u}) = \exp(-\varphi(\mathbf{u}))$  appliquée aux marginales rend « les composantes indépendantes ».

Ceci veut dire que :

$$\omega(C(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_d)) = \prod_{i=1}^d \omega(\mathbf{u}_i)$$

En clair, les copules archimédiennes sont construites à partir du générateur  $\varphi$ . Pour preuve, l'expression générale d'une copule archimédienne de générateur  $\varphi$  se définit comme suit :

$$\begin{cases} C(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_d) = \varphi^{-1}(\varphi(\mathbf{u}_1) + \dots + \varphi(\mathbf{u}_d)) & \text{si } \sum_{i=1}^d \varphi(\mathbf{u}_i) \leq \varphi(\mathbf{0}) \\ C(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_d) = \mathbf{0} & \text{sinon} \end{cases}$$

Sa densité quant à elle peut être écrite sous la forme suivante :

$$\begin{cases} c(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_d) = (\varphi^{-1})^d(\varphi(\mathbf{u}_1) + \dots + \varphi(\mathbf{u}_d)) \prod_{i=1}^d \varphi'(\mathbf{u}_i) & \text{si } \sum_{i=1}^d \varphi(\mathbf{u}_i) \leq \varphi(\mathbf{0}) \\ c(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_d) = \mathbf{0} & \text{sinon} \end{cases}$$

Le choix du générateur  $\varphi$  a donc un impact sur la structure de dépendance et correspond à une copule particulière. De plus, il doit être choisi de sorte qu'il soit de classe  $\mathbf{C}^2$  et que nous ayons les relations suivantes :

$$\varphi(\mathbf{1}) = \mathbf{0}, \quad \varphi'(\mathbf{u}) \leq \mathbf{0} \quad \text{et} \quad \varphi''(\mathbf{u}) \geq \mathbf{0}$$

### 1. Copule de Clayton

Connue aussi sur le nom de « copule de Cook-Johnson », la copule de Clayton est une copule archimédienne définie par son générateur  $\varphi(\mathbf{u}) = \frac{\mathbf{u}^{-\alpha} - 1}{\alpha} \quad \forall \alpha > \mathbf{0}$  et  $\mathbf{u} \in [0, 1]$ .

L'expression de cette copule s'écrit :

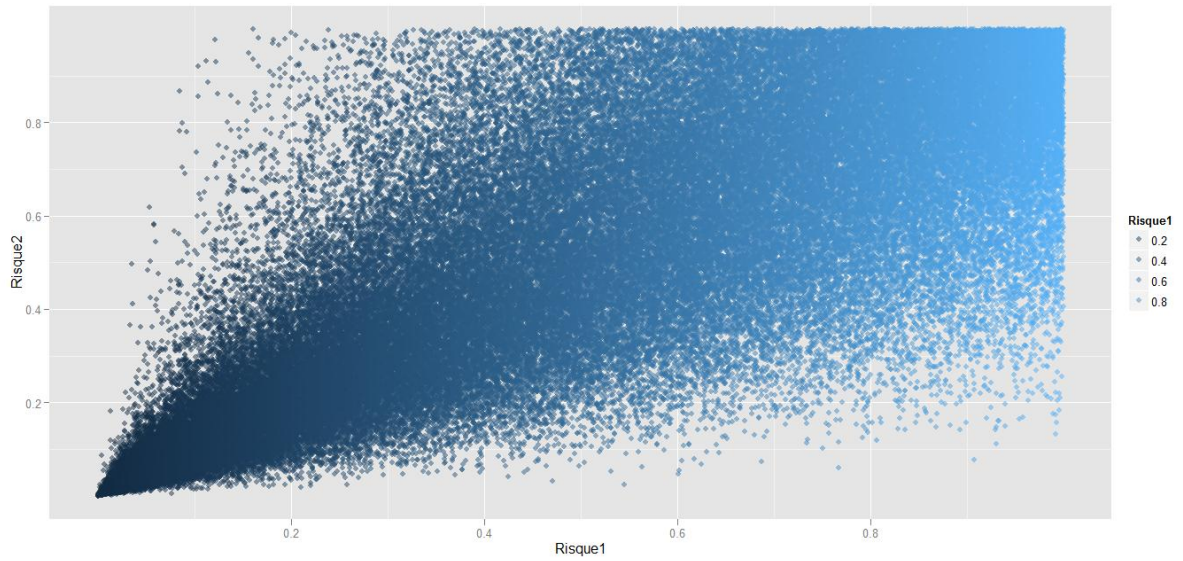
$$C(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_d) = \left( \mathbf{1} - d + \sum_{i=1}^d \mathbf{u}_i^{-\alpha} \right)^{-d - \frac{1}{\alpha}}$$

Et sa densité est :

$$c(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_d) = \left( \mathbf{1} - d + \sum_{i=1}^d \mathbf{u}_i^{-\alpha} \right)^{-d - \frac{1}{\alpha}} \prod_{j=1}^d \left( \mathbf{u}_j^{-\alpha-1} (j\alpha - \alpha + \mathbf{1}) \right)$$

La copule de Clayton présente une dépendance asymptotique à gauche (sur les valeurs négatives) ce qui n'est pas le cas de la copule gaussienne.

**Graphique**



**Figure 4 : Simulation de 100000 observations d’une copule de Clayton de paramètre  $\alpha = 3$  et de dimension 2 avec une corrélation de Pearson  $\Sigma = 0.8$**

**2. Copule de Frank**

La copule de Frank ne présente pas de dépendance de queue. Le générateur de cette copule archimédienne est :  $\varphi(u) = -\ln\left(\frac{e^{-\alpha u}-1}{e^{-\alpha}-1}\right) \forall \alpha \neq 0 \text{ et } u \in ]0, 1]$

En partant de l’expression générale d’une copule archimédienne, nous définissons la copule de Frank d-variée comme suit :

$$C(u_1, u_2, \dots, u_d) = -\frac{1}{\alpha} \times \ln\left(1 + \frac{1}{(e^{-\alpha}-1)^{d-1}} \prod_{i=1}^d (e^{-\alpha u_i}-1)\right)$$

Cette copule étant d-différentiable d’où la densité de la copule de Frank admet une formule close. Pour le cas bi-varié, nous avons l’expression suivante :

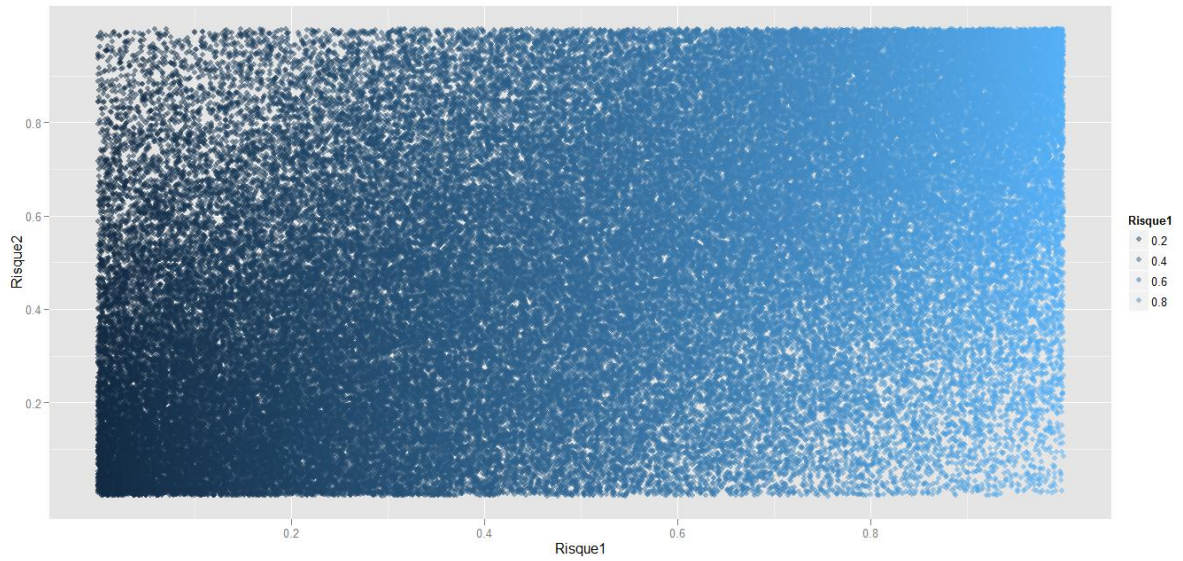
$$c(u_1, u_2) = \frac{\alpha(1-e^{-\alpha})e^{-\alpha(u_1+u_2)}}{((1-e^{-\alpha})-(1-e^{-\alpha u_1})(1-e^{-\alpha u_2}))^2}$$

La densité de Frank en dimension d s’obtient en partant de l’expression générale

$$c(u_1, u_2, \dots, u_d) = (\varphi^{-1})^d(\varphi(u_1) + \dots + \varphi(u_d)) \prod_{i=1}^d \varphi'(u_i)$$

Pour  $\varphi^{-1}(u) = -\frac{\ln((e^{-\alpha}-1)e^{-u}+1)}{\alpha}$  et  $\varphi'(u) = \frac{\alpha e^{-\alpha u}}{(e^{-\alpha u}-1)}$

**Graphique**



**Figure 5: Simulation de 100000 observations d’une copule de Frank de paramètre  $\alpha = 3$  et de dimension 2 avec une corrélation de Pearson  $\Sigma = 0.45$**

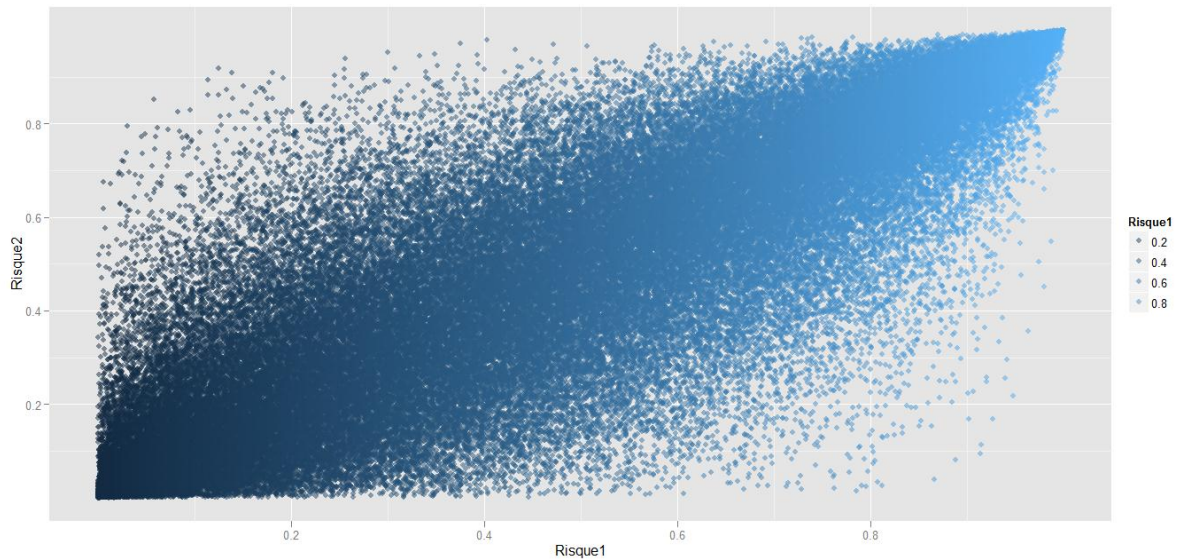
**3. Copule de Gumbel**

Contrairement à la copule gaussienne la copule de Gumbel, parfois appelée copule de Gumbel-Hougaard, permet de modéliser les dépendances extrêmes. En effet, la copule de Gumbel appréhende les dépendances positives et possède la caractéristique de pouvoir représenter des risques dont la structure de dépendance est plus accentuée sur la queue supérieure. Elle appartient à la famille des copules archimédiennes et son générateur s’écrit :  $\varphi(\mathbf{u}) = (-\ln(\mathbf{u}))^\alpha \forall \alpha > 1$  et  $\mathbf{u} \in ]0, 1]$ . Cette copule s’énonce comme suit :

$$C(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_d) = \exp\left(-\left[\sum_{i=1}^d (-\ln(\mathbf{u}_i))^\alpha\right]^{1/\alpha}\right)$$

Cette copule étant comme dans le cas de la copule de Frank et de Clayton d-différentiable d’où la densité de la copule Gumbel admet une formule close.

**Graphique**



**Figure 6: Simulation de 100000 observations d'une copule de Gumbel de paramètre  $\alpha = 3$  et de dimension 2 avec une corrélation de Pearson  $\Sigma = 0.85$**

**I.3. Estimation des paramètres d'une copule**

Nous allons présenter ici les différentes méthodes de calibrage d'une copule. En réalité, la copule est une fonction comme toute autre fonction à la différence que celle-ci possède des propriétés permettant de lier d'autres fonctions possédant des propriétés différentes. Etant une fonction, le calibrage de la copule est une étape très importante pour son utilisation. Comme dans le cas du calibrage des fonctions de probabilités classiques, les méthodes de calibrage des copules seront aussi les mêmes méthodes en occurrence la méthode paramétrique, semi-paramétrique et non paramétrique.

### II.3.1. Estimation non paramétrique

#### 1. Rappel : fonction de répartition empirique

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

$$F_K(x) = \frac{1}{K} \sum_{i=1}^K I_{\{x_i \leq x\}}$$

Pour un  $K$ -échantillon  $(x_1, \dots, x_K)$  de la loi  $F$ . En dimension  $d$ , si l'on se donne

$(\mathbf{x}_1^t, \mathbf{x}_2^t, \dots, \mathbf{x}_d^t)_{1 \leq t \leq T}$ , un  $T$ -échantillon du vecteur (de dimension  $d$ )  $X$ , on peut généraliser l'expression de la fonction de répartition empirique en posant :

$$F_K(x_1, \dots, x_d) = \frac{1}{K} \sum_{i=1}^K I_{\{x_1^t \leq x_1, \dots, x_d^t \leq x_d\}}$$

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

#### 2. Copule empirique

Soit  $(\mathbf{x}_1^t, \mathbf{x}_2^t, \dots, \mathbf{x}_d^t)_{1 \leq t \leq T}$ , un  $T$ -échantillon du vecteur (de dimension  $d$ )  $X$ . En observant que

$C(u_1, \dots, u_d) = F(F_1^{-1}(u_1), \dots, F_d^{-1}(u_d))$ , nous introduisons la notion de copule empirique à partir des versions empiriques des fonctions de répartition de cette expression :

$$C_T(u_1, \dots, u_d) = F_T(F_{1,T}^{-1}(u_1), \dots, F_{d,T}^{-1}(u_d)),$$

Avec  $F_{j,T}^{-1}(u) = \inf\{t | F_{j,T}(t) \geq u\}$ . Cette notion a été introduite par DEHEUVELS [1979].

#### Remarque

$C$  est la fonction de répartition d'un vecteur dont les marginales sont uniformes sur  $[0,1]$  ;

cette propriété n'est plus vraie pour  $C_T$  qui est, conditionnellement aux observations, la

distribution d'un vecteur dont les marginales sont réparties uniformément sur l'ensemble

discret  $\{1, \frac{1}{T}, \dots, \frac{T-1}{T}\}$ . Il suffit de connaître les valeurs de la copule empirique en ces points

discrets ; on a en particulier le lien avec les statistiques d'ordre suivant :

$$C_T\left(\frac{t_1}{T}, \dots, \frac{t_d}{T}\right) = \frac{1}{T} \sum_{t=1}^T I_{\{x_1^t \leq x_1^{(t_1)}, \dots, x_d^t \leq x_d^{(t_d)}\}}$$

Cette expression peut être également écrite à partir des rangs des observations (ce qui est

intuitif, puisque la copule est invariante pour toute transformation croissante des marginales) :

$$C_T \left( \frac{t_1}{T}, \dots, \frac{t_d}{T} \right) = \frac{1}{T} \sum_{t=1}^T I_{\{r_1^t \leq t_1, \dots, r_d^t \leq t_d\}}$$

DEHEUVELS [1979] a obtenu des résultats asymptotiques sur  $C_T$  et  $T^{-1/2}(C_T - C)$ , généralisant les résultats classiques sur la fonction de répartition empirique.

Les copules empiriques sont notamment utiles pour fournir des estimateurs non paramétriques de mesures de dépendance. Par exemple on peut ainsi proposer l'estimateur  $\rho$  de Spearman :

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

### II.3.2. Estimation paramétrique et semi-paramétrique

#### 1. Méthode des moments

C'est une méthode simple qui permet d'exprimer les paramètres de la copule à partir de la relation existante entre cette dernière et les mesures de concordances (tau de Kendall, Spearman).

Par exemple, pour le tau de Kendall, nous avons :  $\rho_{Kendall} = \frac{c-d}{c+d}$ , avec c (resp d) le nombre de paires concordantes (resp discordantes) dans l'échantillon.

Pour les copules elliptiques, nous avons une relation entre la matrice de corrélation  $\Sigma$  et le tau de Kendall comme suit :

$$\Sigma = \sin \left( \frac{\pi}{2} \rho_{Kendall} \right)$$

En ce qui concerne les copules archimédiennes, nous avons :

$$\color{blue}{\oplus} \text{ Copule de Clayton } \alpha = \frac{2\rho_{Kendall}}{1-\rho_{Kendall}}$$

$$\color{blue}{\oplus} \text{ Copule de Gumbel } \alpha = \frac{1}{1-\rho_{Kendall}}$$

#### 2. Maximum de vraisemblance

La densité jointe d'un vecteur aléatoire  $X = (x_1, x_2, \dots, x_d)$  se définit comme suit :

$$L(\theta) = f(x_1, x_2, \dots, x_d) = \prod_{i=1}^d f_i(x_i) \times c(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_d)$$

En appliquant la fonction logarithme, nous obtenons la fonction log de vraisemblance ci-dessous :

$$Ln(L(\theta)) = \sum_{i=1}^d \ln(f_i(x_i)) + \ln(c(\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_d))$$

Nous déduisons que pour un échantillon  $(\mathbf{x}_1^t, \mathbf{x}_2^t, \dots, \mathbf{x}_d^t)_{1 \leq t \leq T}$  de taille T, le log de vraisemblance est :

$$Ln(L(\theta)) = \sum_{t=1}^T \sum_{i=1}^d \ln(f_i(\mathbf{x}_i^t; \theta)) + \sum_{t=1}^T \ln(c(F_1(\mathbf{x}_1^t), F_2(\mathbf{x}_2^t), \dots, F_d(\mathbf{x}_d^t); \theta))$$

Le paramètre  $\theta$  est égale à :

$$\theta = ArgMax(Ln(L(\theta)))$$

$\theta$  Étant un vecteur de paramètres composé des paramètres des lois marginales et du paramètre de la copule.

### **3. Méthode IFM: Inference Function for Marging**

C'est une méthode proposé par Joe et XU [1996] pour palier à la complexité de la méthode du maximum de vraisemblance. Elle consiste en deux étapes :

#### **Étape 1 : Calibrer les paramètres des distributions marginales**

$$\hat{\theta}_k = ArgMax \left( \sum_{t=1}^T \ln(f_k(\mathbf{x}_k^t; \theta)) \right), \forall k \in [1, d]$$

#### **Étape 2 : A partir des paramètres obtenus à l'étape 1, déduire les paramètres de la copule.**

$$\hat{\alpha} = ArgMax \left( \sum_{t=1}^T \ln(c(F_1(\mathbf{x}_1^t; \hat{\theta}_1), F_2(\mathbf{x}_2^t; \hat{\theta}_2), \dots, F_d(\mathbf{x}_d^t; \hat{\theta}_d); \alpha)) \right)$$

La méthode proposée par Joe et Xu est plus pratique que l'estimation directe par le maximum de la vraisemblance classique. En effet, elle permet de réduire considérablement la complexité des calculs et rend ainsi le calibrage de plusieurs modèles multi-variés faisable et convergent. Par ailleurs, il est important de préciser que la comparaison directe de l'IFM et de la méthode d'estimation par le maximum de vraisemblance classique est un point assez difficile à cause du temps de calcul nécessaire pour obtenir les estimations. Cependant, de nombreux tests comparatifs effectués dans XU [1996] sur plusieurs modèles multi-variés montrent que l'IFM est une méthode très efficace et donne des résultats proches de la méthode classique.

#### 4. Méthode CML: Canonical Maximum Likelihood

C'est une méthode recommandée par BOUYE et al [2000]. Sa mise en œuvre ne nécessite pas l'estimation des paramètres des distributions marginales.

Il s'agira ici de transformer les observations  $(x_1^t, x_2^t, \dots, x_d^t)_{1 \leq t \leq T}$  en uniforme  $(\hat{u}_1^t, \hat{u}_2^t, \dots, \hat{u}_d^t)_{1 \leq t \leq T}$  par le biais de la fonction de répartition empirique univariée.

A partir de cette transformation, le paramètre de la copule se déduit comme suit :

$$\hat{\alpha} = \text{ArgMax} \left( \sum_{t=1}^T \ln(c(\hat{u}_1^t, \hat{u}_2^t, \dots, \hat{u}_d^t; \alpha)) \right)$$

Cette méthode est la plus utilisée en pratique du fait de sa simplicité. Son grand avantage réside dans l'estimation paramétrique de la copule indépendamment de la forme paramétrique des marginales. De plus sa mise en œuvre demande un temps de calcul limité.

#### I.4. Choix de la meilleure copula

Le choix de la meilleure copule est une étape très importante. Il existe une panoplie de méthode de choix de la meilleure copule. Dans cette section, je présenterai celle implémenté dans le package « copula » sous R. Cependant les intéressés peuvent consulter le mémoire de Cadoux et Loiseau et bien d'autres documents traitant de la théorie des copules.

Le test de validation du choix de la meilleure copule sous R est basé sur un processus empirique défini comme suit :

$$K_n(u) = \sqrt{n} \{C_n(u) - C_{\theta_n}(u)\}, u \in [0,1]^d \quad (1)$$

Ou  $C_n$  est la copule empirique et  $C_{\theta_n}$  est un estimateur de la copule  $C$  sous l'hypothèse  $H_0 : C \in \mathcal{C}_\theta$ . L'estimateur de  $\theta_n$  de  $\theta$  dans (1) est aussi un estimateur basé sur les rangs.

La statistique du test se définit comme suit :

$$S_n = \int K_n(u)^2 d C_n(u) = \sum_{i=1}^n \left( C_n(\hat{U}_i) - C_{\theta_n}(\hat{U}_i) \right)^2 \quad (2)$$

Une p-value approximative de  $S_n$  peut être obtenue à partir d'une procédure basée sur la méthode bootstrap paramétrique.

**Cette procédure se définit comme suit :**

- ◆ Calculer  $C_n$  à partir des pseudo-observations  $\hat{U}_1, \dots, \hat{U}_n$  et estimer  $\theta$  à partir de  $\hat{U}_1, \dots, \hat{U}_n$  au moyen d'un estimateur  $\theta_n$  basé sur les rangs.
- ◆ Calculer la statistique de test  $S_n$  définie en (2)
- ◆ Pour un nombre assez élevé  $N$ , répéter les étapes suivantes pour chaque  $k \in \{1, \dots, N\}$  :
  - a. Générer un échantillon aléatoire  $X_1^{(k)}, \dots, X_n^{(k)}$  à partir d'une copule  $C_{\theta_n}$  et calculer les pseudo-observations  $\hat{U}_1^{(k)}, \dots, \hat{U}_n^{(k)}$  associées.
  - b. Soit :

$$C_n^{(k)}(u) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}_{\{\hat{U}_i^{(k)} \leq u\}}, u \in [0,1]^d$$

Et calculer un estimateur  $\theta_n^{(k)}$  de  $\theta$  à partir de  $\hat{U}_1^{(k)}, \dots, \hat{U}_n^{(k)}$  en utilisant le même estimateur basé sur les rangs comme dans l'étape 1.

- c. Calculer une réalisation indépendante approximative de  $S_n$  sous l'hypothèse  $H_0$  par :

$$S_n^{(k)} = \sum_{i=1}^n \left( C_n^{(k)}(\hat{U}_i^{(k)}) - C_{\theta_n^{(k)}}(\hat{U}_i^{(k)}) \right)^2$$

- d. Une p-value approximative pour le test est donnée par :

$$\frac{1}{N} \sum_{i=1}^N \mathbf{1}_{\{S_n^{(k)} \geq S_n\}}$$

Cette procédure est implantée sous R via le package copula à partir de la commande « gofCopula ».

### I.5. Simulation des copules

Les copules fournissent un cadre d'analyse bien adapté à la « fabrication » des distributions conjointes possédant une structure de dépendance donnée.

En effet, soit  $X = (X_1, \dots, X_d)$  est un vecteur gaussien centré de matrice de variance-covariance fixée ; on peut remplacer les lois marginales gaussiennes par d'autres lois (par exemple pour tenir compte d'un caractère leptocurtique affirmé des observations) tout en conservant la même structure de dépendance.

Pour simuler des réalisations d'un vecteur  $X = (X_1, \dots, X_d)$  à partir de la décomposition « copule » de  $F$ , il faut être capable de simuler des réalisations d'un vecteur  $U$  ayant des marginales uniformes et dont la distribution est la copule  $C$ , puisqu'on a l'égalité en loi :

$$X = \left( F_1^{-1}(U_1), \dots, F_d^{-1}(U_d) \right).$$

La suite de ce document illustre quelques méthodes qui permettent de mener à bien cette tâche.

#### I.5.1. La méthode des distributions

On suppose que l'on se trouve dans une situation où la loi conjointe de  $X$  est plus facile à simuler directement que la copule  $C$  ; c'est par exemple le cas de la copule gaussienne : un vecteur de dimension  $d$  est aisé à simuler (via la décomposition de CHOLESKY de la matrice de variance-covariance), alors que la copule gaussienne n'est pas simple à simuler directement. La méthode consiste à simuler des réalisations de  $X$  (de loi  $F$ ) et à appliquer ensuite la transformation  $U = (F_1(X_1), \dots, F_d(X_d))$ .

### I.5.2. La méthode des distributions conditionnelles

#### 1. Présentation générale

On se donne  $(U_1, U_2)$  un vecteur dont la distribution est  $C$  ; on simule  $(v_1, v_2)$  uniformes et indépendantes.

On pose alors  $u_1 = v_1$ , puis  $u_2 = C_{u_1}^{-1}(v_2)$  avec  $C_{u_1}(u_2) = P(U_2 \leq u_2 | U_1 = u_1)$ .

Cette expression se calcule à partir de la copule  $C$  :

$$C_{2|1}(u_1, u_2) = P(U_2 \leq u_2 | U_1 = u_1) = \frac{\partial C}{\partial u_1}(u_1, u_2)$$

#### 2. Cas particulier des copules archimédiennes

Dans le cas particulier d'une copule archimédienne, la distribution conditionnelle s'exprime à l'aide du générateur de la copule et conduit à :

$$u_2 = \varphi^{-1} \left( \varphi \left( \varphi'^{-1} \left( \varphi' \left( \frac{v_1}{v_2} \right) \right) \right) - \varphi(v_1) \right)$$

### I.5.1. Les méthodes adaptées spécifiquement à une copule

Comme pour la simulation de loi univariées, des méthodes « ad hoc » adaptées à une copule particulière peuvent être élaborées, par exemple pour la copule de Clayton :

$$C(u, v) = (u^{-\theta} + v^{-\theta} - 1)^{-\frac{1}{\theta}}$$

On peut utiliser l'algorithme suivant :

- Simuler  $x_1$  et  $x_2$ , deux réalisations indépendantes d'une loi  $exp(1)$ .
- Simuler une réalisation  $x$  issue d'une variable aléatoire de loi  $\Gamma(1, \theta)$

- Prendre  $u_2 = \left(1 + \frac{x_2}{x}\right)^{-\theta}$  et  $u_1 = \left(1 + \frac{x_1}{x}\right)^{-\theta}$

## II. Théorie bayésienne

### II.1. Introduction

Considérons le contrat  $j$  auquel correspondent les montants annuels de sinistres  $Y_{j1}, \dots, Y_{jT}$  et le paramètre de risque  $\Theta_j$ . Ce paramètre de risque n'est pas observable mais pourra être partiellement récompensé à partir à partir des sinistres observés.

Donnons-nous des lois entièrement spécifiées pour le paramètre de risque et pour les variables de sinistres conditionnellement au paramètre de risque :

- La loi marginale de  $\Theta_j$  admet une densité  $\mu(\theta_j)$ ,
- Les variables  $Y_{j1}, \dots, Y_{jT}$  sont conditionnellement à  $\Theta_j$ , indépendantes entre elles et distribuées identiquement avec pour densité  $f(y_{jt}|\theta_j)$ .

En remarquant que, pour tous évènements  $A$  et  $B$ , on a  $P[A \cap B] = P[A|B] \times P[B]$ , la densité des variables  $Y_{j1}, \dots, Y_{jT}, \Theta_j$  est donnée par :

$$\mu(\theta_j) \prod_{t=1}^T f(y_{jt}|\theta_j)$$

### II.2. Lois à postériori

L'utilisation de la règle de Bayes va nous permettre, à partir de l'information disponible en date  $T$ , d'enrichir notre connaissance de  $\Theta_j$  et de  $Y_{j,T+1}$ .

**Rappel :** (Règle de Bayes)

Quels que soient les évènements  $A$  et  $B$ , on a :

$$P[B|A] = \frac{P[A|B] \times P[B]}{P[A]}$$

**Proposition 1**

Si le couple  $(X, Y)$  à valeurs dans  $\mathbb{R}^2$  admet une densité  $h$ , les densités conditionnelles existent et sont données par :

$$f(x|y) = \frac{h(x,y)}{g(y)} \text{ et } g(x|y) = \frac{h(x,y)}{f(x)}$$

D'après la règle de Bayes, elles peuvent se réécrire :

$$f(x|y) = \frac{g(y|x)f(x)}{\int g(y|x)f(x) dx} \text{ et } g(y|x) = \frac{f(x|y)g(y)}{\int f(x|y)g(y) dy}$$

**Proposition 2**

La densité de la loi a postérieure du paramètre de risque  $\theta_j$  est donnée par :

$$\mu(\theta_j | \sigma(Y_{j1}, \dots, Y_{jT})) = \frac{\mu(\theta_j) \prod_{t=1}^T f(y_{jt} | \theta_j)}{\int \mu(\theta) \prod_{t=1}^T f(y_{jt} | \theta) d\theta}$$

On parle de densité de paramètre de risque a postérieure puisqu'il s'agit de la densité conditionnée par l'information fournie par les sinistres passés  $\sigma(Y_{j1}, \dots, Y_{jT})$ . Parallèlement, cet historique de sinistres va nous permettre de préciser la distribution du risque futur  $Y_{jT+1}$ .

**Proposition 3**

La densité prédictive sur le risque futur  $Y_{jT+1}$  est donnée par :

$$f(y_{j,T+1} | \sigma(Y_{j1}, \dots, Y_{jT})) = \frac{\int \mu(\theta) \prod_{t=1}^{T+1} f(y_{jt} | \theta) d\theta}{\int \mu(\theta) \prod_{t=1}^T f(y_{jt} | \theta) d\theta}$$

La prime de crédibilité (ou de Bayes) pour la période  $[T, T + 1]$  est donc donnée par :

$$\begin{aligned} p(y_{j1}, \dots, y_{jT}) &= E[Y_{j,T+1} | Y_{j1} = y_{j1}, \dots, Y_{jT} = y_{jT}] \\ &= \int y_{j,T+1} f(y_{j,T+1} | \sigma(Y_{j1}, \dots, Y_{jT})) dy_{j,T+1} \end{aligned}$$

Notons que la prime de Bayes est solution de :

$$\min_p E \left[ (Y_{j,T+1} - p)^2 \mid Y_{j1} = y_{j1}, \dots, Y_{jT} = y_{jT} \right]$$

Cette prime est donc optimale au sens des moindres carrés. Le principal inconvénient de ces formules réside dans la grande dépendance de ces lois au choix de la loi a priori  $\mu(\theta_j)$ . Toutefois, cette forte dépendance aux lois a priori peut présenter un avantage lorsque l'on travaille avec peu de données et donc lorsque l'approche fondée sur l'historique des sinistres n'est pas stable et que l'on souhaite donc donner au poids plus important au tarif a priori.

### II.3. Application de la théorie bayésienne : Mise à jour des primes

L'approche bayésienne devient très intéressantes lorsqu'il s'agit de mettre à jour les primes puisqu'elle nous fournit des formules de mise à jour simples.

#### **Proposition 4**

La loi jointe de  $\Theta_j$ ,  $Y_{j,T+1}$  sachant  $\sigma(Y_{j1}, \dots, Y_{jT})$  est donnée par :

$$f(y_{j,T+1} | \theta_j) \mu(\theta_j | \sigma(Y_{j1}, \dots, Y_{jT}))$$

La proposition suivante exprime la densité de la loi du paramètre de risque conditionnellement à la connaissance d'une charge annuelle de sinistres supplémentaire.

#### **Proposition 5**

La nouvelle densité de la loi a posteriori du paramètre de risque  $\Theta_j$  est donnée par :

$$\mu(\theta_j | \sigma(Y_{j1}, \dots, Y_{j,T+1})) = \frac{f(y_{j,T+1} | \theta_j) \mu(\theta_j | \sigma(Y_{j1}, \dots, Y_{jT}))}{\int f(y_{j,T+1} | \theta) \mu(\theta | \sigma(Y_{j1}, \dots, Y_{jT})) d\theta}$$

Cette mise à jour permet d'exprimer simplement la mise à jour de la densité prédictive du risque futur.

**Proposition 6**

La nouvelle densité prédictive sur le risque futur  $Y_{j,T+2}$  est donnée par :

$$f(y_{j,T+2} | \sigma(Y_{j1}, \dots, Y_{jT})) = \int f(y_{j,T+2} | \theta) \mu(\theta | \sigma(Y_{j1}, \dots, Y_{j,T+1})) d\theta$$

Remarquons qu'il existe des choix coordonnés de la loi  $f$  et de la loi a priori  $\mu$  qui permettent des mises à jour explicites des lois a posteriori successives (cf DROESBEKE et al (2002)).

**III. Modèles linéaires**

**III.1. Modèle linéaire simple**

Dans cette section, nous allons introduire cette notion de modèle linéaire simple.

Si nous disposons d'une base de donnée composés de n observations  $x_i, y_i$  correspondantes aux variables respectives  $X_i, Y_i$ ,  $Y_i$  étant la variable dépendante.

Le modèle linéaire se formule comme suit :

$$Y_i = \mu_i + \varepsilon_i$$

Avec  $\mu_i = E(Y_i) = x_i\beta$  et  $\varepsilon_i \sim N(0, \sigma)$

L'utilisation d'un modèle requiert le calibrage du modèle et la vérification des hypothèses du modèle. Pour le modèle linéaire, le calibrage du modèle est obtenu en minimisant la quantité

$$S = \sum_{i=1}^n (Y_i - \mu_i)^2 = \sum_{i=1}^n (Y_i - x_i\beta)^2$$

Comme S est différentiable par rapport à  $\beta$ , nous obtenons l'expression suivante de  $\beta$  :

$$\hat{\beta} = \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2}$$

$\hat{\beta}$  est une variable aléatoire dont l'espérance est  $\beta$  et sa variance correspond à  $\frac{\sigma^2}{\sum_{i=1}^n x_i^2}$ .

Ceci implique que :

$$\hat{\beta} \sim N\left(\beta; \frac{\sigma^2}{\sum_{i=1}^n x_i^2}\right)$$

Un autre aspect important est de vérifier la significativité des paramètres du modèle. Ceci est basé sur le test de student. En effet, il s'agira de construire la statistique de test  $T = \frac{\hat{\beta} - \beta_0}{\sigma_{\hat{\beta}}}$  sous l'hypothèse que  $\beta = \beta_0$ . T suit une loi de student à n-1 degré de liberté ce qui nous permet à partir de la table de Student de conclure sur l'acceptation ou le rejet du test.

Après avoir déterminé les paramètres du modèle linéaire simple, il reste à vérifier les hypothèses du modèle qui se résume principalement sur la validation du test sur la normalité des erreurs et sur quelques représentations graphiques.

### III.2. Modèle linéaire multiple

Dans le cas du modèle linéaire multiple, nous sommes en présence de plusieurs variables et plusieurs observations. C'est souvent le cas du calcul des provisions techniques en tenant compte de deux variables explicatives à savoir « année de survenance » et « année de développement ».

Soit une base de données composée de n observations, k variables indépendants résumés dans le vecteur X et une variable explicative Y.

Le modèle se formule comme suit :

$$Y = X\beta + \varepsilon$$

Avec :

Y : La variable dépendante représentant une matrice de dimension  $n \times 1$

X : La variable indépendante représentant une matrice de dimension  $n \times (k + 1)$

$\beta$  : Les coefficients du modèle représentant une matrice de dimension  $(k + 1) \times 1$

$\varepsilon$  : Les erreurs du modèle représentant une matrice de dimension  $n \times 1$

Comme dans le cas du modèle linéaire simple, le calibrage du modèle est obtenu en

minimisant la quantité  $S = \sum_{i=1}^n \varepsilon_i^2 = \varepsilon' \varepsilon = (y - X\beta)'(y - X\beta)$

Nous obtenons :

$$\hat{\beta} = (X'X)^{-1}X'y$$

$\hat{\beta}$  est appelé l'estimateur des moindres carrés ordinaires de  $\beta$ .

Tout comme dans le modèle linéaire simple,  $\hat{\beta}$  est une variable aléatoire dont sa distribution peut être approchée par une loi normale.

$$\hat{\beta} \sim N(\beta; \Sigma),$$

$\Sigma = \sigma^2(X'X)^{-1}$  est la matrice de covariance de  $\hat{\beta}$ .

Un autre aspect important à noter est le test de Fisher qui permet de choisir les meilleures variables qui expliquent le modèle.

Enfin, la dernière étape dans la construction d'un modèle est la validation des hypothèses du modèle, qui est comme dans le cas du modèle linéaire simple, la validation de l'hypothèse de normalité, la représentation graphique des valeurs prédites par rapport aux résidus et d'autres représentations graphiques.

Nous verrons plus tard, une application avec le logiciel R.

#### **IV. Modèles linéaires généralisés**

Le modèle linéaire généralisé connu sous l'expression anglo-saxon « Generalized linear models (GLM) (*Nelder and Wedderburn, 1972*) » est un modèle qui fournit à la variable dépendante une distribution autre que la distribution normale et introduit dans la structure du modèle un degré de non linéarité.

Il se formule comme suit :

$$g(\mu_i) = X_i\beta$$

Ou :

$$\mu_i \equiv E(Y_i)$$

g est une fonction monotonique « plane » appelé fonction lien.

$X_i$  est la i-ème ligne de la matrice X et  $\beta$  représente les paramètres du modèle à déterminer.

En utilisant le modèle GLM, l'on suppose automatique que la variable à prédire est indépendante et suit une distribution appartenant à la famille exponentielle.

Cette dernière comprend des distributions discrètes (Poisson, Binomiale, quasipoisson etc..) et des distributions continues (normale, Gamma etc...), voir **McCullagh et Nelder (1989)** pour plus de précision.

### IV.1. Famille exponentielle

Dire que la variable explicative Y appartient à la famille exponentielle, c'est dire que cette variable possède la fonction de densité suivante :

$$f_{\theta}(y) = \exp\left(\frac{y\theta - b(\theta)}{a(\phi)} + c(y, \phi)\right)$$

Avec b, c et a des fonctions arbitraires,  $\phi$  est le paramètre de dispersion et  $\theta$  est connu sous le nom de 'paramètre canonique'.

Ce mémoire ayant un souci pratique, nous allons résumer pour quelques distributions connues, les valeurs des paramètres ci-dessus :

	Normal	Poisson	Binomial	Gamma	Inverse Gaussian
$f(y)$	$\frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right)$	$\frac{\mu^y \exp(-\mu)}{y!}$	$\binom{n}{y} \left(\frac{\mu}{n}\right)^y (1-\frac{\mu}{n})^{n-y}$	$\frac{1}{\Gamma(\nu)} \left(\frac{\nu}{\mu}\right)^{\nu} y^{\nu-1} \exp\left(-\frac{\nu y}{\mu}\right)$	$\sqrt{\frac{\gamma}{2\pi y^3}} \exp\left[-\frac{\gamma(y-\mu)^2}{2\mu^2 y}\right]$
Range	$-\infty < y < \infty$	$y = 0, 1, 2, \dots$	$y = 0, 1, \dots, n$	$y > 0$	$y > 0$
$\theta$	$\mu$	$\log(\mu)$	$\log\left(\frac{\mu}{n-\mu}\right)$	$-\frac{1}{\mu}$	$-\frac{1}{2\mu^2}$
$\phi$	$\sigma^2$	1	1	$\frac{1}{\nu}$	$\frac{1}{\gamma}$
$a(\phi)$	$\phi(-\sigma^2)$	$\phi(-1)$	$\phi(-1)$	$\phi\left(-\frac{1}{\nu}\right)$	$\phi\left(-\frac{1}{\gamma}\right)$
$b(\theta)$	$\frac{\theta^2}{2}$	$\exp(\theta)$	$n \log(1 + e^{\theta})$	$-\log(-\theta)$	$-\sqrt{-2\theta}$
$c(y, \phi)$	$-\frac{1}{2} \left[ \frac{y^2}{\phi} + \log(2\pi\phi) \right]$	$-\log(y!)$	$\log \binom{n}{y}$	$\nu \log(\nu y) - \log(y\Gamma(\nu))$	$-\frac{1}{2} \left[ \log(2\pi y^3 \phi) + \frac{1}{\phi y} \right]$
$V(\mu)$	1	$\mu$	$\mu(1 - \mu/n)$	$\mu^2$	$\mu^3$
$g_c(\mu)$	$\mu$	$\log(\mu)$	$\frac{\mu}{n-\mu}$	$\frac{1}{\mu}$	$\frac{1}{\mu^2}$
$D(y, \hat{\mu})$	$(y - \hat{\mu})^2$	$2y \log\left(\frac{y}{\hat{\mu}}\right) - 2(y - \hat{\mu})$	$2 \left[ y \log\left(\frac{y}{\hat{\mu}}\right) + (n-y) \log\left(\frac{n-y}{n-\hat{\mu}}\right) \right]$	$2 \left[ \frac{y-\hat{\mu}}{\hat{\mu}} - \log\left(\frac{y}{\hat{\mu}}\right) \right]$	$\frac{(y-\hat{\mu})^2}{\hat{\mu}^3 y}$

**Tableau 1 : récapitulatif des paramètres de la densité de la famille exponentielle pour les distributions usuelles**

**IV.2. Des résultats importants**

Si  $Y$  appartient à la famille exponentielle, alors

$$f_{\theta}(y) = \exp\left(\frac{y\theta - b(\theta)}{a(\phi)} + c(y, \phi)\right) \xrightarrow{\text{implique}} \log[f_{\theta}(y)] = l(\theta) = \frac{y\theta - b(\theta)}{a(\phi)} + c(y, \phi)$$

En dérivant  $l(\theta)$  par rapport à  $\theta$ , on a :

$$\frac{\partial l}{\partial \theta} = \frac{y - b'(\theta)}{a(\phi)}$$

$$E\left(\frac{\partial l}{\partial \theta}\right) = \frac{1}{a(\phi)} [E(y) - b'(\theta)] = 0 \xrightarrow{\text{implique}} \mathbf{E(y) = b'(\theta)}$$

**Nous déduisons que l'espérance de la variable  $Y$  est :  $E(y) = b'(\theta)$**

La dérivée seconde de  $l(\theta)$  se définit comme suit :

$$\frac{\partial^2 l}{\partial \theta^2} = \frac{b''(\theta)}{a(\phi)} \xrightarrow{\text{implique}} E\left(\frac{\partial^2 l}{\partial \theta^2}\right) = E\left(\frac{b''(\theta)}{a(\phi)}\right) = -E\left(\left[\frac{\partial l}{\partial \theta}\right]^2\right)$$

$$\frac{b''(\theta)}{a(\phi)} = \frac{E\left[(y - b'(\theta))^2\right]}{a(\phi)^2}$$

D'où nous concluons que :

$$\mathbf{Var(y) = b''(\theta) \times a(\phi) = Var(E(y))\phi}$$

$a(\phi) = \frac{\phi}{\omega}$ ,  $\omega$  étant un paramètre constant.

**IV.3. Calibrage des paramètres du modèle**

Soit un échantillon composé de  $n$  observations  $y_1, \dots, y_n$  de la variable  $Y$ .

$g(\mu_i) = X_i\beta$  et  $Y_i \sim f_{\theta_i}(y_i)$ .

La fonction de vraisemblance se définit comme suit :

$$L(\beta) = \prod_{i=1}^n f_{\theta_i}(y_i)$$

Et la log de vraisemblance est :

$$l(\beta) = \log(L(\beta)) = \sum_{i=1}^n \log(f_{\theta_i}(y_i)) = \sum_{i=1}^n \left[ \frac{y_i \theta_i - b_i(\theta_i)}{a_i(\phi)} + c_i(y_i, \phi) \right]$$

Comme  $a_i(\phi) = \frac{\phi}{\omega_i}$  alors :

$$l(\beta) = \sum_{i=1}^n \left[ \omega_i \left( \frac{y_i \theta_i - b_i(\theta_i)}{\phi} \right) + c_i(y_i, \phi) \right]$$

De manière implicite, la fonction log de vraisemblance dépend du paramètre  $\beta$  par l'intermédiaire de  $\mu_i$ . Cette fonction est différentiable par rapport à  $\beta$  et ceci revient à dériver cette fonction pour tous les éléments  $j$  du vecteur  $\beta$ . Nous avons :

$$\frac{\partial l}{\partial \beta_j} = \frac{1}{\phi} \sum_{i=1}^n \left[ \omega_i \left( y_i \frac{\partial \theta_i}{\partial \beta_j} - b_i'(\theta_i) \frac{\partial \theta_i}{\partial \beta_j} \right) \right]$$

$$\frac{\partial \theta_i}{\partial \beta_j} = \frac{\partial \theta_i}{\partial \mu_i} \times \frac{\partial \mu_i}{\partial \beta_j} \quad \text{et} \quad \frac{\partial \mu_i}{\partial \theta_i} = b_i''(\theta_i) \quad \text{d'où} \quad \frac{\partial \theta_i}{\partial \mu_i} = \frac{1}{b_i''(\theta_i)}$$

Ceci implique que :

$$\frac{\partial l}{\partial \beta_j} = \frac{1}{\phi} \sum_{i=1}^n \left[ \frac{(y_i - b_i'(\theta_i))}{b_i''(\theta_i)/\omega_i} \frac{\partial \mu_i}{\partial \beta_j} \right]$$

La détermination du paramètre  $\beta$  se fait en résolvant l'équation  $\frac{\partial l}{\partial \beta_j} = 0$ , soit :

$$\frac{\partial l}{\partial \beta_j} = S = \sum_{i=1}^n \left[ \frac{(y_i - \mu_i)}{V(\mu_i)} \frac{\partial \mu_i}{\partial \beta_j} \right] = 0 \quad \forall j.$$

#### IV.4. Résidus de Pearson

Un autre aspect important de la méthode GLM concernant les résidus de Pearson définis par la quantité ci-dessous :

$$\varepsilon_i^P = \frac{y - \hat{\mu}_i}{\sqrt{V(\hat{\mu}_i)}}$$

Cette quantité est très utilisée en pratique car elle permet d'apprécier la qualité des prévisions. Souvent, elle est aussi utilisée pour faire du bootstrap.

### **V. Modèles additives généralisés**

Un modèle de régression additive est un modèle dans lequel la variable dépendante  $Y$  se déduit en fonction des  $p$  variables indépendantes en posant :

$$Y = \alpha + \sum_{j=1}^p f_j(x_j)$$

Contrairement au modèle linéaire dans lequel,  $Y = \alpha + \sum_{j=1}^p \beta_j x_j$ .

Nous constatons un changement par rapport au modèle linéaire marqué par l'introduction de la fonction  $f_j$ .

Cette dernière peut être représentée par des fonctions splines, des fonctions polynomiales, des fonctions non linéaires comme la fonction logarithme, la fonction exponentielle etc. La forme des  $f_j$  est estimée à partir des données.

Les modèles de régression additives peuvent aussi inclure des modèles semi-paramétriques comme par exemple :

$$y_i = \alpha + \beta_1 x_{1i} + f_2(x_{2i}) + \dots + f_k(x_{ki}) + \varepsilon_i$$

Ils peuvent aussi contenir des modèles dans lesquels il existe des interactions entre les variables comme par exemple :

$$y_i = \alpha + f_{12}(x_{1i}, x_{2i}) + f_3(x_{3i}) + \dots + f_k(x_{ki}) + \varepsilon_i$$

**Le modèle additive généralisé (Hastie et Tibshirani, 1986, 1990)** est un modèle dont le prédicteur linéaire est composé d'une somme de fonctions  $f$ .

Dans ce modèle, le prédicteur se formule de la manière générale suivante :

$$g(\mu_i) = X_i^* \beta + f_1(x_{1i}) + f_2(x_{2i}) + f_3(x_{3i}, x_{4i}) + \dots$$

Avec :

$\mu_i \equiv E(Y_i)$  et  $Y_i \sim$  famille exponentielle.

En clair, dans le modèle GLM, le prédicteur linéaire s'écrit comme suit :

$$g(\mu_i) = \alpha + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki}$$

Alors que dans le modèle additive généralisé, le prédicteur linéaire aura la forme suivante :

$$g(\mu_i) = \alpha + f_1(x_{1i}) + f_2(x_{2i}) + \dots + f_k(x_{ki})$$

Tous ces modèles appartiennent à la famille des modèles de régression non paramétriques généralisés et sont discutés dans Fox (2000a, 2000b).

## **VI. Bootstrap**

Le principe du bootstrap est une méthode de ré-échantillonnage visant à retenir la mécanique de Monte Carlo afin d'estimer la loi du phénomène observée.

Elle a été introduite par EFRON [1979] et est particulièrement utilisée dans le cas de petits échantillons de données afin d'améliorer la qualité des estimations.

L'usage du bootstrap dans la résolution de problèmes actuariels est large, en particulier en tarification, et dans l'évaluation des provisions de sinistres. Ceci s'explique par le large éventail d'application de la méthodologie : intervalles de confiance, tests d'hypothèse, modèles linéaires puis généralisés, etc.

La prédiction d'un niveau « suffisant » de réserve, qui permet à la société d'assurance de faire face à ses engagements, est un sujet primordial de la science actuarielle. Depuis quelques années, la technique du bootstrap est communément utilisée pour analyser la variabilité des montants de sinistre et obtenir des erreurs de prédiction pour différentes méthodes de provisionnement, principalement pour les méthodes basées sur Chain Ladder et sur les modèles linéaires généralisés.

Sa mise en œuvre est simple et consiste à remarquer que, pour un échantillon de taille suffisante, la fonction de répartition de la loi sous-jacente peut être approchée par la fonction de répartition empirique :

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

Evaluer des statistiques par simulation se ramène alors à générer des échantillons à l'aide de la distribution empirique. Or un tirage dans la distribution empirique s'obtient simplement par un tirage avec remise des  $n$  valeurs dans l'échantillon initial. On obtient ainsi au plus  $n^n$  « bootstrapés » à partir desquels on va calculer les estimateurs empiriques des statistiques d'intérêt.

Nous verrons dans la suite du mémoire que dans le cas des modèles linéaires, la méthode du bootstrap se base sur les résidus de Pearson ou sur les résidus ajustés de Pearson.

Pratiquement, à partir du triangle des incréments, on constitue un triangle des résidus. Ce triangle sert ensuite de base à la création de  $k$  (nombre arbitraire ou en relation avec les permutations possibles du triangle) nouveaux triangles de résidus : les valeurs du triangle initial sont « cassées » aléatoirement dans chaque nouveau triangle. Plus intuitivement, chaque nouveau triangle correspond à « une permutation » du triangle originel.

La formule  $\varepsilon_{ij}^P = \frac{y_{ij} - \hat{\mu}_{ij}}{\sqrt{v(\hat{\mu}_{ij})}}$  permet de constituer  $k$  triangles des incréments à partir des  $k$

triangles des résidus.

L'étape suivante consiste à estimer les paramètres du modèle linéaire et à obtenir  $k$

estimations de la statistique désirée  $\left(\hat{\mu}_{ij}, \hat{E}(R)\right)$ . Dès lors, il est possible d'établir par

exemple, un intervalle de confiance de la moyenne de réserve prévue.

Cette technique a été brièvement présentée car elle sera utilisée dans la suite du mémoire pour le calcul des provisions techniques dans le cas du modèle linéaire généralisé (GLM) et du modèle additif généralisé (GAM) pour les lois Gamma, Poisson et Quasipoisson.

Cependant elle peut aussi s'étendre pour les modèles GLMM (Generalized Linear Model Mixed) et GAMM (Generalized Additive Model Mixed) et à d'autres lois de probabilités.

## **VII. Approche probabiliste du risque**

**Un risque peut être décrit comme un événement qui peut ou ne peut avoir lieu, et dont sa survenance entraîne des conséquences financières négatives.**

Ce caractère aléatoire lié au risque nous permet d'utiliser la théorie de la probabilité pour le modéliser.

En effet la théorie des probabilités nous offre une panoplie de distributions permettant d'approximer l'évolution, le comportement du risque.

Il existe deux grandes familles de distributions de probabilité à savoir :

- ✚ Les distributions discrètes (binomial, poisson etc.)
- ✚ Les distributions continues (normal, gamma, pareto, etc.)

En pratique, la détermination de la distribution d'une variable aléatoire stochastique passe par trois étapes à savoir :

### **1<sup>ère</sup> étape : Hypothèse**

Dans cette étape, l'on suppose que notre variable est issue de :

- la famille des lois discrètes à savoir poisson, binomial, binomial négative
- la famille des lois continues à savoir Gamma, log-normale, Pareto

### **2<sup>ème</sup> étape : Estimation**

A partir de l'hypothèse sur la distribution de cette variable formulée à la première étape, la deuxième étape sert à estimer les paramètres de la distribution.

### **3<sup>ème</sup> étape : Validation**

Cette étape consiste à valider le choix de la distribution par un test statistique et des représentations graphiques.

En pratique, les tests statistiques utilisés sont :

- Le test de Kolmogorov Smirnov qui s'appuie sur l'écart minimal entre la fonction de répartition théorique et celle de la distribution empirique.
- Le test de Cramer Von-Mises dont la statistique de test mesure l'intégrale du carré de la différence entre les fonctions de répartition théorique et observée.

- Le test d'Anderson-Darling qui utilise le fait que si la loi empirique et la loi théorique coïncident, la transformée de la distribution observée par la fonction de répartition théorique suit une loi uniforme sur  $[0,1]$ .

Les représentations graphiques permettent de choisir la loi de probabilité qui s'ajustent le mieux à nos données. Elles concernent en général :

- + L'histogramme
- + La densité non paramétrique
- + Le QQ-Plots

Nous avons en gros présenté les différentes étapes utilisées en pratiques pour appréhender le comportement du risque selon une approche probabiliste. En réalité, il est difficile de déterminer une distribution qui s'adapte le mieux à nos données avec un test d'ajustement satisfaisant. Vu cette contrainte liée en partie à la qualité des données, les actuaires sont souvent obligés de faire des hypothèses sur la distribution du risque. En général, les hypothèses portent en ce qui concernent les charges de sinistres sur les lois log-normale, Pareto et souvent Gamma.

Cette étape est très importante et le lecteur s'en rendra compte dès le chapitre 3 concernant le calcul des provisions techniques.

### **VIII. Mesures du risque**

La conception de la notion de mesure de risque dépend du contexte dans laquelle elle a été utilisée. En effet, en Finance, les risques financiers peuvent être mesurés par leurs forces ou leurs conséquences financières sur une activité économique spécifique. Par contre, dans un cadre de solvabilité, la mesure du risque est matérialisée par un capital requis à immobiliser. Pour comprendre encore plus cette notion de mesure de risque, Tsanakas (2007) définit trois domaines d'application des mesures de risques :

- ✚ Comme représentation de l'aversion au risque dans les modèles d'évaluation des actifs avec paradigme de premier plan, l'utilisation de la variance comme mesure de risque dans la théorie de portefeuille de Markowitz.
- ✚ Comme des outils de calcul des primes en assurance. Dans ce contexte, les mesures de risques sont appelées selon l'expression anglo-saxon « premium calculation principles » dans la littérature actuarielle classique.
- ✚ Comme des quantificateurs du capital économique que le détenteur d'un portefeuille particulier de risque devra investir pour sa couverture.

Ainsi, parmi toutes ces interprétations de la mesure de risque, nous allons dans cette section nous intéresser à la troisième interprétation de la mesure de risque c'est-à-dire la conception de cette dernière comme capital économique pour la couverture des risques d'une compagnie d'assurance et qui n'est autre que le « Solvency Capital Requirement » dans le cadre de Solvabilité 2.

### **VIII.1. Propriétés de la mesure de risque**

Notons par  $\rho$  la mesure du risque.

#### **P 1 : Invariance ou objectivité**

L'une des premières propriétés de  $\rho$  provient de Denuit et al (2006) qui stipule que la mesure de risque doit être basée sur la fonction de répartition du risque. En clair  $\rho(X) = \rho(F_X)$ , avec  $F_X$ , la fonction de répartition de  $X$ . Ceci sous-entend que  $X$  est égale en distribution à  $F_X$  et l'on notera par  $X =_d F_X$ .

#### **P 2 : Non négativité**

Cette propriété impose à  $\rho$  une valeur toujours positive ce qui implique

$$\rho(X) > 0 \quad \forall X$$

#### **P 3 : Homogénéité positive**

$$\rho(aX) = a\rho(X) \quad \forall a > 0$$

#### **P 4 :**

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

#### **P 5 :**

$$\rho(X + b) \leq \rho(X) + b \quad \forall b \in \mathbb{R}$$

**P 6 : Convexité**

$$\rho(\lambda X + (1 - \lambda)Y) \leq \lambda\rho(X) + (1 - \lambda)\rho(Y) \quad \forall \lambda \in [0,1]$$

**P 7 : Invariance par translation**

$$\rho(X + b) = \rho(X) + b \quad \forall b \in \mathbb{R}$$

**P 8 : Monotonie**

$$X \leq Y \xrightarrow{\text{implique}} \rho(X) \leq \rho(Y)$$

NB : Une mesure de risque qui satisfait les propriétés 3, 4,7 et 8 est appelé « coherent risk measure », voir Artzner et al. (1999)—*coherent risk measures*.

**P 9 : Expectation boundedness (expression anglo-saxonne)**

$$\rho(X) > E(-X) \quad \forall X \text{ non constant}$$

$$\rho(X) = E(-X) \quad \forall X \text{ constant}$$

**P 10 : Additivité comonotonique**

Si X et Y sont deux variables aléatoires comonotoniques, alors :

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

**P 11 : Additivité indépendante**

Si X et Y sont deux variables aléatoires indépendantes, alors :

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

**P 12 : Condition du régulateur**

Soit  $(X, Y)$  un couple de variables aléatoire,  $\varepsilon$  un nombre compris entre 0 et 1. Dans la perception du régulateur, le capital de solvabilité  $\rho_{SCR}$  doit vérifier la condition suivante :

$$\begin{aligned} E([X + Y - \rho_{SCR}(X, Y)]_+) + \varepsilon \times \rho_{SCR}(X + Y) \\ \leq E([X - \rho_{SCR}(X)]_+) + E([Y - \rho_{SCR}(X)]_+) + \varepsilon \times \rho_{SCR}(X) + \varepsilon \times \rho_{SCR}(Y) \end{aligned}$$

$\varepsilon$  peut représenter le coût de capital (COC) en solvabilité 2.

### VIII.2. Des résultats importants

Ces résultats dérivent de la propriété P 12 concernant les exigences du régulateur. En effet, ce dernier exige seulement suffisant pour la couverture du risque en question. Pour mettre en relief cette idée, considérons un portefeuille dont les pertes futures sont notées par  $X$ .

Introduisons l'expression anglo-saxonne "shortfall risk" noté par  $\rho^{SR}(X)$ .

La mesure de risque induite par la perception du régulateur peut être formulée comme :

$$\rho^{SR}([X - \rho_{SCR}(X)]_+)$$

Soit  $c(X, d)$  la fonction de coût définie par :

$$c(X, d) = \rho^{SR}([X - d]_+) + \varepsilon \times d$$

Le capital requis optimal est obtenu en minimisant la fonction de coût  $c(X, d)$ . L'un des résultats important est obtenu en s'inspirant de **Dhaene et al. (2008a)** qui utilise l'espérance de  $X$  comme "shortfall risk" c'est à dire que  $\rho^{SR}(X) = E(X)$ .

En minimisant la fonction de coût pour  $\rho^{SR}(X) = E(X)$  et  $\rho_{SCR}(X)$ , nous obtenons que :

$$\rho_{SCR}(X) = VaR_{1-\varepsilon}(X)$$

### VIII.3. Types de mesures de risques

Dans cette sous-section, nous allons nous intéresser à quelques mesure de risque utilisées tout au long du mémoire, soit pour le calcul des provisions techniques par la méthode GAM ou GAM mixed, GLM et GLM mixed, soit pour le calcul du capital agrégé, soit pour l'allocation du capital agrégé entre les lignes d'affaires.

#### VIII.3.1. Mesures de risques basées sur les moments

Il existe deux types de mesures de risques basées sur les moments. La première mesure est représentée par la variance  $\sigma^2$  ou l'écart type  $\sigma$ .

$$\sigma^2 = E([X - E(X)]^2)$$

$$\sigma = \sqrt{\sigma^2}$$

La deuxième mesure dérive des principes de calcul de prime rencontrés en théorie du risque.

Cette mesure s'écrit de la manière suivante :  $\rho^{EV}(X) = \beta E(X) \quad \forall \beta > 1$

### VIII.3.2. Value at Risk (VaR)

Etant donné un niveau de confiance ou un niveau de probabilité  $1 - \alpha$  ;  $\forall \alpha \in [0, 1]$ , la value at Risk (VaR) pour un risk  $X$  n'est rien d'autre que le quantile associé pour un niveau de confiance de  $1 - \alpha$ .

En clair c'est l'inverse de la fonction de répartition pour un niveau de probabilité égale à  $1 - \alpha$ .

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

Pratiquement, le calcul se fera en utilisant le package « actuar » sous R.

### VIII.3.3. Tail Value at Risk (TVAR) et Conditional Expected Loss (CTE)

Etant donné un niveau de confiance ou un niveau de probabilité  $1 - \alpha$  ;  $\forall \alpha \in [0, 1]$ , la Tail Value at Risk (TVAR) pour un risk  $X$  se définit comme suit :

$$TVaR_{1-\alpha}(X) = \frac{1}{\alpha} \int_{1-\alpha}^1 VaR_{\xi}(X) d\xi$$

En clair, la TVAR peut être interprété comme la moyenne arithmétique ou une mixture des  $VaR_s$  de  $X$  de  $1 - \alpha$  à  $1$ .

Une définition alternative de la TVAR est la Conditional Expected Loss (CTE) qui s'écrit comme suit :

$$CTE_{1-\alpha}(X) = E(X|X > VaR_{1-\alpha}(X))$$

Cette dernière peut être vue comme la moyenne des pertes dans  $100\alpha\%$  des cas.

Lorsque la variable d'intérêt  $X$  est continue, la TVAR et la CTE sont égales c'est-à-dire :

$$TVaR_{1-\alpha}(X) = CTE_{1-\alpha}(X)$$

Par contre, lorsque la variable n'est pas continue, il est conseillé au lecteur de se pencher vers Denuit et al (2005).

Pratiquement, les calculs se feront en utilisant encore une fois le package « actuar » développé par Vincent Goulet et son équipe.

**VIII.3.4. Conditional VaR**

Cette grandeur peut être calculée comme la différence entre la CTE et la VaR.

$$CVaR_{1-\alpha}(X) = CTE_{1-\alpha}(X) - VaR_{1-\alpha}(X)$$

Il existe d'autres mesures de risque dans la littérature mais nous allons nous en arrêter là pour aborder l'un des chapitres les plus importants de ce mémoire à savoir le chapitre 3 concernant les provisions techniques d'une compagnie d'assurance non vie.

*Chapitre 3 : Les provisions techniques en assurance non vie*

Les provisions techniques constituent un élément important pour les entreprises d'assurance car elles représentent les engagements auxquels doit faire face l'assureur. En effet un sous-provisionnement peut avoir des conséquences à long terme sur la pérennité de l'entreprise d'assurance et par extension sur l'économie en ce sens que l'assurance est un moyen de collecter l'épargne des ménages et de financer l'économie. C'est pourquoi les autorités de contrôle accordent une attention particulière à ce poste du passif, le rôle de ces derniers étant de défendre les intérêts de l'assuré.

Ainsi un rôle central de l'actuaire non vie est le calcul de ces provisions techniques essentiellement pour les sinistres à payer. Ces derniers sont destinés à permettre le règlement intégral des engagements pris envers les assurés et bénéficiaires de contrats.

Les provisions techniques mesurent ainsi les engagements restant à honorer par l'assureur et représentent en assurance non vie 75% du bilan total.

Dans cette section, nous essayerons de calculer les provisions de différentes façons et nous déduirons les différences observées.

Le formalisme général se définit comme suit :

Soit :

- $Y_{i,j,l}$  la charge de sinistre non cumulé pour la ligne  $l$  ( $l=1 \dots L$ ) correspondant à l'année de survenance  $i$  ( $i=1 \dots I$ ) et l'année de développement  $j$  ( $j=1 \dots J$ ).
- $C_{i,j,l}$  la charge de sinistre cumulé pour la ligne  $l$  ( $l=1 \dots L$ ) correspondant à l'année de survenance  $i$  ( $i=1 \dots I$ ) et l'année de développement  $j$  ( $j=1 \dots J$ ).
- $D_{inf,l} = \{i + j > I\}$  : le domaine de définition des couples  $(i, j)$  appartenant au triangle inférieur pour la ligne  $l$ .
- $D_l = \{i + j < I\}$  : le domaine de définition des couples  $(i, j)$  appartenant au triangle supérieur pour la ligne  $l$ .

S'agissant des branches à développement long, le coût réel d'un sinistre est connu plusieurs années après sa survenance. C'est pourquoi l'assureur est dans l'obligation de constituer des provisions pour faire face à ses engagements futurs de sorte à ne pas être surpris dans l'avenir. Nous pouvons matérialiser ce cas de figure à partir du schéma ci-dessous :

### Chapitre 3 : Les provisions techniques en assurance non vie

Année	1	2	3	4
1	$Y_{1,1}$	$Y_{1,2}$	$Y_{1,3}$	$Y_{1,4}$
2	$Y_{2,1}$	$Y_{2,2}$	$Y_{2,3}$	$Y_{2,4}$
3	$Y_{3,1}$	$Y_{3,2}$	$Y_{3,3}$	$Y_{3,4}$
4	$Y_{4,1}$	$Y_{4,2}$	$Y_{4,3}$	$Y_{4,4}$

Les valeurs  $Y_{i,j,l}$  à prédire correspondent aux valeurs se trouvant dans le triangle inférieur du tableau.

Les provisions à constituer sont égales dans le cas des montants non cumulés à :

$$Provision = \sum_{j>I-i} Y_{i,j,l}$$

Dans le cas de montants cumulés, nous avons pour la branche l fixé :

Année	1	2	3	4
1	$C_{1,1}$	$C_{1,2}$	$C_{1,3}$	$C_{1,4}$
2	$C_{2,1}$	$C_{2,2}$	$C_{2,3}$	$C_{2,4}$
3	$C_{3,1}$	$C_{3,2}$	$C_{3,3}$	$C_{3,4}$
4	$C_{4,1}$	$C_{4,2}$	$C_{4,3}$	$C_{4,4}$

Les valeurs  $C_{i,j} = \sum_{k=1}^j Y_{i,k}$  à prédire correspondent aux valeurs se trouvant dans le triangle inférieur du tableau. Soit  $R_i$ , la réserve de l'année  $i$ , nous avons :

$$R_i = \sum_{j>I-i} (C_{i,j} - C_{i,j-i})$$

Les provisions à constituer sont égales dans le cas des montants cumulés à :

$$Provision = \sum_{i=1}^I R_i$$

**Méthodes basiques de calcul des provisions techniques**

**I.1. Méthode de Chain Ladder (CL)**

La méthode de Chain Ladder est basée sur l'hypothèse qu'il existe des facteurs de développements  $f_1, \dots, f_j$  avec  $f_j = 1$  tel que pour tout  $i$  ( $i=1\dots I$ ) et  $j$  ( $j=1\dots J$ ) :

$$E(C_{i,j}|C_{i,1}, \dots, C_{i,j-1}) = E(C_{i,j}|C_{i,j-1}) = f_{j-1} * C_{i,j-1} \quad \text{I.1-1}$$

Les valeurs des charges  $C_{i,j}$  sont supposées être indépendantes. A partir de l'équation I.1-1, nous pouvons prévoir la charge ultime comme suit :

$$E(C_{i,j}|D_I) = f_{j-1} * E(C_{i,j-1}|C_{i,j-1}) = \dots = C_{i,j-1} * \prod_{j=J-1}^{j-1} f_j \quad \text{I.1-2}$$

Les  $f_j$  sont appelés les facteurs de Chain Ladder et sont estimés à partir des données connus :

$$\hat{f}_j = \frac{\sum_{i=1}^{I-1} C_{i,j+1}}{\sum_{i=1}^{I-1} C_{i,j}}$$

Ce qui nous amène à écrire que :

$$\hat{C}_{i,j}^{CL} = C_{i,j-1} * \prod_{j=J-1}^{j-1} \hat{f}_j$$

La reserve de Chain Ladder facilement par la connaissance de la charge ultime prédite par :

$$\hat{R}_i^{CL} = \hat{C}_{i,j}^{CL} - C_{i,j-1}$$

**I.2. Méthode de Bornhuetter-Ferguson (BF)**

Cette méthode vient de Bornhuetter and Ferguson (1972). La méthode BF est une méthode très populaire de calcul des provisions du fait de sa robustesse. Aussi cette méthode fait intervenir des idées bayésiennes en intégrant dans le modèle des avis d'expert, les primes et le business plan stratégique.

La prédiction des charges futures se définit comme suit :

$$\hat{C}_{i,j}^{BF} = C_{i,j-1} + (1 - \beta_{j-1})\vartheta_i \quad \text{I.2-1}$$

Et

$$\hat{R}_i^{BF} = \hat{C}_{i,j}^{BF} - C_{i,j-1} = (1 - \beta_{j-1})\vartheta_i \quad \text{I.2-1}$$

### Chapitre 3 : Les provisions techniques en assurance non vie

$\vartheta_i$  représente une information exogène. C'est une variable aléatoire estimant la charge moyenne à l'ultime pour chaque année de survénance  $i$  en tenant pas compte des observations observées.

Les coefficients betas sont équivalent à :

$$\beta_{J-i} = \prod_{j=J-i}^{J-1} \frac{1}{\hat{f}_j}$$

#### I.3. Combinaison de la méthode Chain Ladder (CL) et de la méthode Bornhuetter-Ferguson (BF)

Cette combinaison est connue sous le nom de la méthode de Benktander-Hovinen (BH) proposé par Benktander (1979) et étudié en détail dans Mack (2000).

Rappelons que nous sommes en présence de deux situations extrêmes : une première méthode BF basée sur l'estimation de  $\vartheta_i$  obtenu en ignorant les données et une deuxième méthode CL basée sur l'estimation de la charge ultime à partir des observations.

A partir de ces deux situations, Benktander (1979) proposa une méthode incluant ces deux situations ci-dessus en introduisant une autre expression de l'information exogène  $\gamma_i$ :

$$\gamma_i = c_i \hat{C}_{i,J}^{CL} + (1 - c_i) \vartheta_i$$

Ou  $c_i$  est un facteur de crédibilité.

En posant  $c_i = \beta_{J-i}$ , nous obtenons :

$$\hat{C}_{i,j}^{BH} = C_{i,J-i} + (1 - \beta_{J-i}) \gamma_i = C_{i,J-i} + (1 - \beta_{J-i}) (\beta_{J-i} \hat{C}_{i,J}^{CL} + (1 - \beta_{J-i}) \vartheta_i)$$

Et

$$\hat{R}_i^{BH} = \hat{C}_{i,j}^{BH} - C_{i,J-i} = (1 - \beta_{J-i}) (\beta_{J-i} \hat{C}_{i,J}^{CL} + (1 - \beta_{J-i}) \vartheta_i)$$

$\forall i = 1 \dots I$

#### I.4. Application sur la branche Automobile Corporelle

Comme données, nous disposons d'un triangle de liquidation montant cumulé de 18 années de survénance et 18 années de développement.

L'implémentation de ces trois méthodes a été faite à partir du logiciel R qui est un outil puissant en ce sens qu'il possède un langage de programmation matriciel très développé.

Abstraction faite de la théorie, nous nous sommes rendu compte lors de l'implémentation que les méthodes Chain Ladder, Bornhuetter-Ferguson et Benktander-Hovinen converge lorsque

### Chapitre 3 : Les provisions techniques en assurance non vie

les estimateurs bayesiens  $\vartheta_i$  et  $\gamma_i$  sont égaux à la charge ultime  $\hat{C}_{i,J}^{CL}$  obtenu par la méthode de Chain Ladder. En clair, lorsque  $\vartheta_i = \gamma_i = \hat{C}_{i,J}^{CL}$ , les trois méthodes sont équivalentes.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
1602821,71	9125208,62	16091012,7	26317541,3	38100908,8	46081223	54421552,7	57861936,6	61579100,3	63673022,7	66165956,3	67386646,5	69043836,7	69268105,3	69673644,9	69986363,2	7001
2470164,84	9823118,76	18223059,3	28912581,5	38761591	54563770,4	60948409,4	65538851	68127972,3	71007392,4	72732000,6	73740465,6	74335005,8	75078939,4	75410944,6	75791029,7	7584
3012740,91	11920882,95	23806271,4	37516643,7	53668556,3	60658150,3	68404543,6	73940671,9	78036492,9	81049853,2	82039666,4	83133114,7	83842260,6	84296161,4	84555610,9	84573887,7	
1927342,88	9545715,39	21491394,9	39226039,3	50762091	60181717,8	67520924,3	72683434,7	77265367,8	79636426,6	80489180,3	81560557,7	82415653,6	82806575,2	83147562		
2656241,47	13235198,7	44687618,1	72085689,3	93312360,9	110043112	120151895	130254694	135222185	138159578	140282058	141880790	142618497	143118755			
3217218,05	26415183,01	59366431,9	95764802,2	129813663	153135967	169291515	179577763	185243927	187844946	191234542	193237686	194073600				
6291739,57	30845973,23	75907414	116594359	147487980	171224667	188862313	197071124	202945819	209761692	212213692	212890132					
4626342,28	26742493,1	66321713,8	111707570	148134084	182814558	195080369	205068768	212355846	216509877	218528444						
4425464,44	30483676,85	72039020,8	120330356	159902609	179341642	193423241	203498041	209919870	213021889							
6538559,1	37715246,25	86212723,1	147661711	182392309	206478776	221163462	230721010	235171135								
6830197,27	37899736,56	102435800	163359109	206729202	240585432	257826523	268026075									
6401024,18	39444122,81	96763066,3	157322868	208178374	235954366	251059279										
4737304,19	34430393,64	96437490,1	165906646	211242297	236293383											
3751062,68	33334345,92	103314430	166244940	203683989												
5604852,59	42700142,17	107940426	167337468													
6669473,66	52278750,39	114762425														
6485076,96	41061878,07															
9427989,87																

**Figure 7 : Triangle de liquidation de la branche Automobile Corporel**

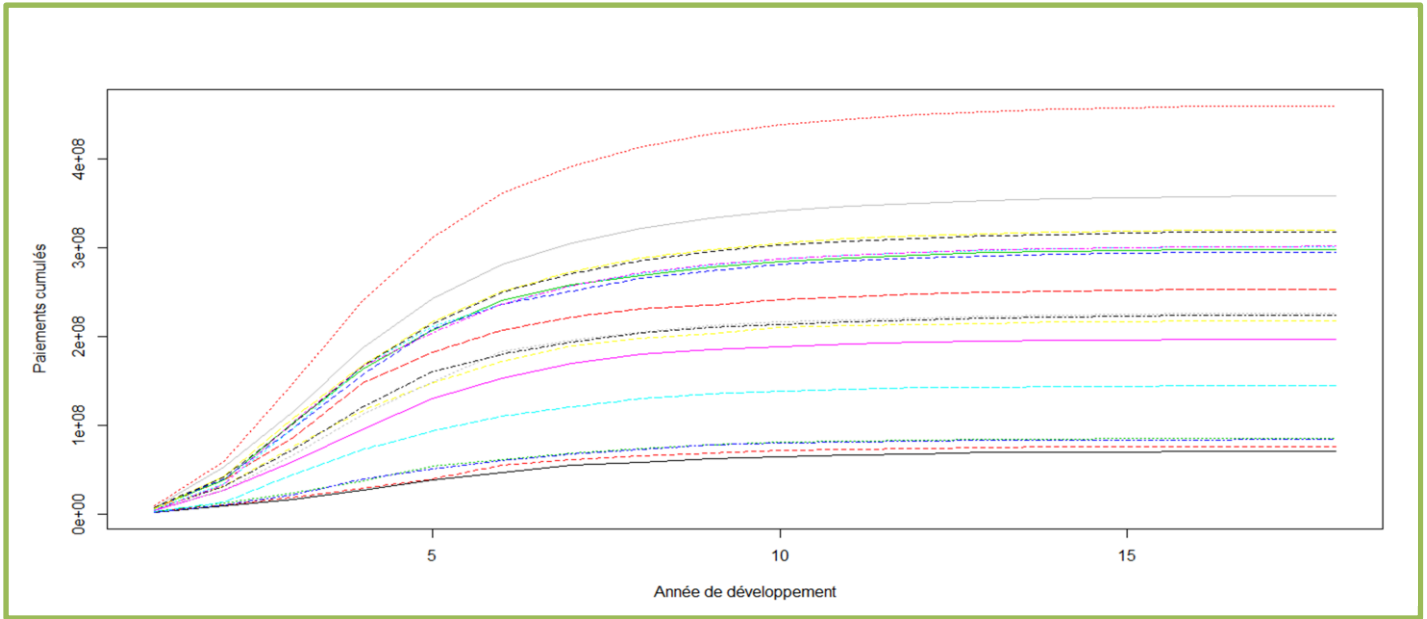
#### I.4.1. Méthode de Chain Ladder

Les facteurs de développement sont donnés à partir de la capture d'écran ci-dessous :

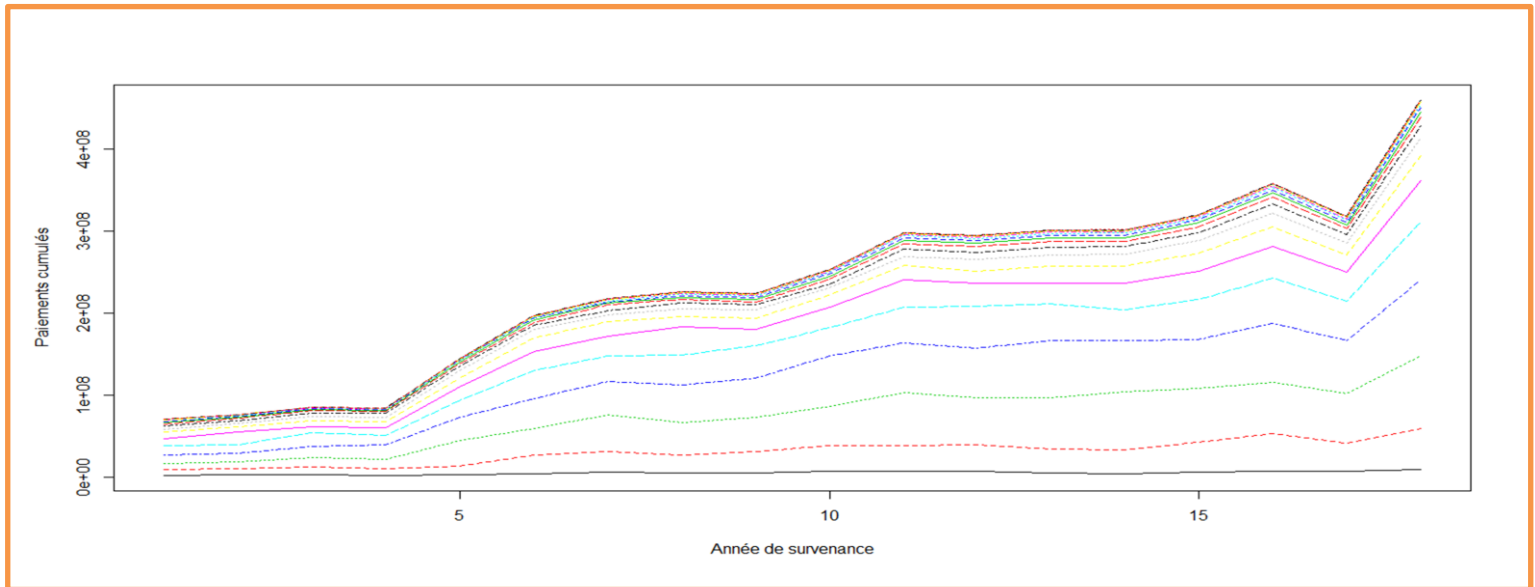
```
> L
[1] 6.304428 2.479705 1.630905 1.292087 1.161147 1.086470 1.054566 1.035059
[9] 1.024351 1.015313 1.010261 1.008409 1.005115 1.004296 1.003096 1.000531
[17] 1.000402 1.000000
```

A partir de ces facteurs de développement, nous déduisons les provisions techniques, voir ci-dessus pour l'aspect théorique :

$$\text{Provision} = 14002562 \text{ DHRS}$$



**Figure 8: Evolution des paiements cumulés en fonction de l'année de développement**



**Figure 9: Evolution des paiements cumulés en fonction de l'année de survenance**

**NB :**

Certains chiffres clés ont été supprimés dans la valeur des provisions techniques par soucis de confidentialité et ceci sera le cas tout le long du rapport, le plus important étant la méthodologie et non la valeur exacte des provisions qui est dans notre cas confidentiel.

**I.4.2. Méthode de Bornhuetter-Ferguson (BF)**

$$\hat{C}_{i,j}^{BF} = C_{i,j-i} + (1 - \beta_{j-i})\vartheta_i$$

Et

$$\hat{R}_i^{BF} = \hat{C}_{i,j}^{BF} - C_{i,j-i} = (1 - \beta_{j-i})\vartheta_i$$

Les coefficients  $\beta_j$  pour chaque année de développement  $j$  sont donnés par :

```

> beta
[1] 0.02052797 0.12941713 0.32091635 0.52338399 0.67625745 0.78523399
[7] 0.85313304 0.89968539 0.93122764 0.95390352 0.96851082 0.97844887
[13] 0.98667655 0.99172339 0.99598381 0.99906787 0.99959841 1.00000000
    
```

Si nous faisons l’hypothèse que  $\vartheta_i = \hat{C}_{i,j}^{CL}$ , alors nous obtenons une provision identique à celle calculée par la méthode de Chain Ladder.

Dans notre cas, nous allons supposer qu’en fonction du business plan stratégique de l’entreprise, nous aurons :  $\vartheta_i = \hat{C}_{i,j}^{CL} + 1000$  d’où la provision obtenue est :

**Provision=14002601 Dhrs**

**I.4.3. Méthode de Benktander-Hovinen (BH)**

$$\hat{C}_{i,j}^{BH} = C_{i,j-i} + (1 - \beta_{j-i})\gamma_i = C_{i,j-i} + (1 - \beta_{j-i})(\beta_{j-i}\hat{C}_{i,j}^{CL} + (1 - \beta_{j-i})\vartheta_i)$$

Et

$$\hat{R}_i^{BH} = \hat{C}_{i,j}^{BH} - C_{i,j-i} = (1 - \beta_{j-i})(\beta_{j-i}\hat{C}_{i,j}^{CL} + (1 - \beta_{j-i})\vartheta_i)$$

Connaissant déjà la valeur des  $\beta_j$  et de  $\vartheta_i$ , voir la section I.4.2, la provision obtenue est :

**Provision= 14002588 Dhrs**

	Provisions (Dhrs)
Chain Ladder (CL)	14002562
Bornhuetter-Ferguson (BF)	14002601
Benktander-Hovinen (BH)	14002588

**Tableau 4 : récapitulatif des provisions techniques pour les trois méthodes déterministes**

**I. Modèles non linéaires stochastiques de calcul des provisions**

Comme énoncé plus haut, le calcul des provisions techniques se fait à partir d'un triangle de liquidation dans le cas des branches à développement long. Ceci est conforme à notre situation vu que nous disposons de trois branches à développement long qui est l'accident de travail, l'automobile corporelle et l'automobile matérielle.

Dans ce triangle de liquidation, les charges de sinistres ou les règlements de sinistres dépendent de l'année de survenance  $i$  et de l'année de développement  $j$ .

Pour mettre en place le modèle et pour faciliter les notations, notons par :

- $year_i$  : l'année de survenance  $i$
- $lag_j$  : l'année de développement  $j$

Pour calculer les provisions, il suffira de prédire les charges futures qui correspondent dans notre cas aux valeurs du triangle inférieur. Comme nous disposons d'un triangle dont les montants sont non cumulés, les provisions se déduisent en faisant la somme des valeurs prédites par le modèle correspondant au triangle inférieur.

Nous remarquons que l'élément clé du calcul des provisions dépend de l'estimation des charges futures. Une estimation non optimale du modèle conduit soit à un sur-provisionnement ou à un sous-provisionnement.

Par contre une bonne estimation du modèle conduit à la « best estimate » des provisions techniques qui est celle recommandée par **Solvency II**.

Nous essayerons dans ce mémoire de déterminer la meilleure estimation des provisions techniques en utilisant toutes sortes de modèle contenue dans la famille des modèles non linéaires.

## II.1. Modèles de Calcul des provisions techniques par la méthode GAM (Generalized Additive Model)

Nous proposerons dans cette sous-section, l'utilisation des modèles GAMs dans le calcul des provisions techniques. En effet ces modèles répondent à la problématique de « best estimate » recommandé par Solvabilité 2 du fait de la robustesse du modèle et de la qualité de l'ajustement. Les modèles GAMs fournit un paramétrage non paramétrique au modèle indépendamment de la forme paramétrique de la variable expliquée par l'introduction de fonctions jointes aux variables explicatives. L'utilisation d'un GAM requiert :

- ◆ Le choix d'une distribution pour la variable dépendante  $Y_{ij}$  appartenant à la famille exponentielle et dans le même le choix d'une fonction lien.

Si  $Y_{ij} \sim$  famille exponentielle, alors :

$$f_{\theta}(y_{ij}) = \exp\left(\frac{y_{ij}\theta - b(\theta)}{a(\phi)} + c(y_{ij}, \phi)\right)$$

Avec b, c et a des fonctions arbitraires,  $\phi$  est le paramètre de dispersion et  $\theta$  est connu sous le nom de 'paramètre canonique'.

- ◆ Le choix d'une fonction de lissage à appliquer aux variables explicatives. Cette fonction peut être polynomial, spline cubique etc.

### Remarque

Après le choix de la fonction de lissage, des paramétrages de cette fonction doivent être effectués. Ceux-ci concernent le choix de la dimension, le type de régression (penalized regression) etc.

#### II.1.1. Modèle GAM

Dans le cas du calcul des provisions techniques, les variables explicatives du modèle retenues sont :

- L'année de survenance notée  $year_i$
- L'année de développement noté  $lag_j$

Soit  $\eta_{ij} = g(\widehat{\mu}_{ij})$  le prédicteur du modèle GAM, nous obtenons les modèles ci-dessous :

1.  $\eta_{ij} = c + s_\lambda(year_i) + s_\nu(lag_j)$
2.  $\eta_{ij} = s_\lambda(year_i) + s_\nu(lag_j)$
3.  $\eta_{ij} = c + \beta_{year_i} year_i + \beta_{lag_j} lag_j + s_{\lambda,\nu}(year_i, lag_j)$
4.  $\eta_{ij} = c + s_\lambda(year_i) + s_\nu(lag_j) + s_{\lambda,\nu}(year_i, lag_j)$

$$\widehat{\mu}_{ij} = g^{-1}(\widehat{\eta}_{ij})$$

Ces modèles ont été présentés pour dire que dans un contexte de « best estimate », plusieurs modèles doivent être testés pour choisir celle qui explique le mieux notre situation.

Le modèle étant spécifié, il reste à le calibrer. Dans un modèle GLM, il suffira juste d'estimer les paramètres du modèle. Par contre dans un modèle GAM (cas du modèle 3 par exemple), il s'agira d'estimer les coefficients des variables  $year_i$  et  $lag_j$  et dans le même temps, d'estimer les meilleures fonctions de lissage  $s_\lambda$  et  $s_\nu$  et les paramètres  $\lambda, \nu$  de ces fonctions.

### **Calibrage de la fonction de lissage**

Soit deux fonctions  $h$  et  $g$  de classe  $C^2$  c'est-à-dire que ces deux fonctions sont deux fois différentiable, de dérivée seconde  $h''$  et  $g''$ .

Le calibrage des deux fonctions de lissage se fait en minimisant la quantité :

$$Q = \sum \sum (y_{ij} - \eta_{ij})^2 + \lambda \int (h''(t))^2 dt + \nu \int (g''(t))^2 dt$$

On peut montrer qu'il existe une solution explicite et unique à ce problème de minimisation et que le minimum est un spline cubic naturel  $s_\lambda$  et  $s_\nu$  avec des nœuds uniquement aux valeurs  $year_i$  et  $lag_j$ .

**Choix des paramètres de la fonction de lissage**

Les deux fonctions de lissage  $s_\lambda$  et  $s_\nu$  n'ont qu'un seul paramètre de lissage chacune. Chaque paramètre gouverne le niveau de lissage ainsi que le compromis entre le biais et la variance :

- Quand  $\lambda, \nu$  tendent vers 0, il n'y a pas de pénalité de lissage et le modèle fournit un ajustement parfait : les valeurs ajustées sont les données elles-mêmes.
- Quand  $\lambda, \nu$  sont grands (tendent vers l'infini), l'ajustement est un lissage parfait : les valeurs ajustées tombent le long d'une ligne droite forçant ainsi la liaison à être linéaire. Il en résulte une faible variance mais un biais important.

Les paramètres  $\lambda, \nu$  doivent être compris entre ces deux cas extrêmes pour produire le niveau désiré de lissage qui permet d'avoir un bon ajustement du modèle avec un équilibre biais - variance. Le choix de ces paramètres de lissage peut se faire à l'aide de la méthode par validation croisée générale (GCV). Mais une valeur particulière pour  $\lambda$  et  $\nu$  peut également être choisie en fixant soi-même le nombre de degrés de liberté.

La méthode d'estimation des paramètres  $\lambda$  et  $\nu$  est connue sous le nom de Generalized Cross Validation (GCV). Wahba a proposé cette dernière pour résoudre les problèmes qui nécessitent la sélection de paramètres minimisant le risque de prédiction.

On peut montrer que  $\hat{\eta} = f(A(\lambda, \nu), Y)$  ou  $A(\lambda, \nu)$  est une matrice qui dépend de  $\lambda, \nu$  et de la série de données ; c'est l'analogie avec une matrice de projection dans une régression linéaire. La fonction GCV à minimiser est alors :

$$GCV(\lambda, \nu) = \frac{\sum_{i=1}^n \sum_{j=1}^n (y_{ij} - \hat{\eta}_{ij})^2}{(n - \text{tr}(A(\lambda, \nu)))^2}$$

*Ou  $\text{tr}(A(\lambda, \nu))$  est la trace de la matrice  $A(\lambda, \nu)$ , soit la somme des éléments diagonaux de cette matrice.*

Malheureusement, la méthode GCV est très coûteuse pour les problèmes à paramètres de lissage multiples. Simon Wood a développé une approche améliorée pour résoudre ce problème dans le cas d'une classe de modèles de spline de lissage incluant les modèles GAMs avec le problème d'optimisation présenté ci-dessus.

### Chapitre 3 : Les provisions techniques en assurance non vie

Le logiciel R utilise cette fonction GCV « améliorée » comme critère de sélection des paramètres de lissage  $\lambda$  et  $\nu$  par le biais du package « mgcv ».

En pratique, l'utilisation de la méthode GCV nous fournit un GCV score qui peut être utilisé comme critère de sélection du meilleur modèle. On choisira un GCV score le plus petit.

Il existe aussi plusieurs critères de sélection de modèle comme le critère AIC (**Akaike's information criteria**, Akaike, 1973). **On choisira aussi le 'AIC' le plus petit.**

Pour les intéressés, je les invite à consulter l'ouvrage de GRACE WAHBA intitulé « Spline Models for Observational Data ».

#### II.1.2. Modèle GAM Mixed (GAMM)

Le Generalized Additive Model Mixed (GAMM) est une extension du model GAM. Sa particularité est qu'elle induit des effets aléatoires entre les facteurs (ou variables) et tient aussi compte de la corrélation existante entre les résidus du modèle.

Elle se définit comme suit :

$$\eta_{ij} = f(s_\lambda, s_\nu) + b_{year_i} + b_{lag_j} + \varepsilon$$

Avec  $b_{year_i} \sim N(0, \sigma_{year_i})$  et  $b_{lag_j} \sim N(0, \sigma_{lag_j})$  et  $\varepsilon \sim N(0, \sigma_\varepsilon)$

Cependant il peut arriver que  $\varepsilon$  suive un processus autorégressif d'ordre 1 c'est-à-dire que :

$$\forall t \mid \varepsilon_t = \phi \varepsilon_{t-1} + \zeta_t$$

Avec  $\zeta_t$  un bruit blanc ou (*white noise*) et si

$f(s_\lambda, s_\nu) = s_\lambda(year_i) + s_\nu(lag_j) + s_{\lambda,\nu}(year_i, lag_j)$ , dans ce cas, le modèle devient :

$$\eta_{ij} = s_\lambda(year_i) + s_\nu(lag_j) + s_{\lambda,\nu}(year_i, lag_j) + b_{year_i} + b_{lag_j} + \phi \varepsilon_{t-1} + \zeta_t$$

D'où :  $\mu_{ij} = g^{-1}(\eta_{ij})$

$$= g^{-1}(s_\lambda(year_i) + s_\nu(lag_j) + s_{\lambda,\nu}(year_i, lag_j) + b_{year_i} + b_{lag_j} + \phi \varepsilon_{t-1} + \zeta_t)$$

Le modèle finalement retenu est :

$$\mu_{ij} = g^{-1}(s_\lambda(year_i) + s_\nu(lag_j) + \phi \varepsilon_{t-1} + \zeta_t)$$

### *Chapitre 3 : Les provisions techniques en assurance non vie*

Dans notre cas, c'est ce modèle qui s'ajuste le mieux à nos données parce qu'il fournit un GCV score et un AIC plus faible. Aussi les résidus étant autocorrolés dans notre cas, nous avons induit un processus autorégressif d'ordre 1 dans le modèle. Au vu de ces arguments, nous allons l'utiliser pour prédire les charges futures et introduire la méthode bootstrap pour dégager les provisions techniques à partir des quantiles, voir la section suivante.

De ce fait, nous sommes sur la voie du best estimate des PT ce qui est conforme à la vision de Solvabilité II.

#### **II.1.3. Cas pratique**

Dans cette sous-section, nous allons appliquer successivement les modèles GAM, GAM Mixed dans le calcul des provisions techniques.

Pour simplifier, nous allons tester deux lois à savoir :

- la loi log-normale et Gamma pour les distributions continues
- la loi de poisson pour les distributions discrètes

en utilisant la fonction lien log.

Nous allons aussi distinguer deux cas :

- cas ou les données sont « as if »
- cas ou les données ne sont pas « as if »

Enfin, nous allons utiliser le package « mgcv » sous R pour obtenir les résultats qui seront présentés ci-dessous.

#### **1. Modèle GAM**

Cette étude de cas portera sur les trois plus importantes branches de la RMA WATANYA. Il s'agit de :

- la branche 'Accident de Travail (AT) '
- la branche 'Automobile Corporelle' et
- la branche 'Automobile Matérielle '

### Chapitre 3 : Les provisions techniques en assurance non vie

#### ✓ loi log-normale

Nous supposons que les règlements incrémentaux suivent une distribution log-normale. Pour illustrer la méthodologie de travail, nous allons tester par la méthode GAM deux modèles à savoir :

$$a. \eta_{ij} = c + s_{\lambda}(\text{year}_i) + s_v(\text{lag}_j)$$

$$b. \eta_{ij} = c + s_{\lambda}(\text{year}_i) + s_v(\text{lag}_j) + s_{\lambda,v}(\text{year}_i, \text{lag}_j)$$

L'outil informatique utilisé est le logiciel R en occurrence le package « mgcv ». En utilisant la commande 'gam', nous obtenons pour le modèle a, l'output suivant pour la branche AT :

```
Parametric coefficients:
      Estimate Std. Error t value Pr(>|t|)
(Intercept) 15.81862    0.07205   219.5  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Approximate significance of smooth terms:
      edf Ref.df      F p-value
s(year)   5     5  0.504  0.773
s(lag)    6     6 94.398 <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R-sq.(adj) = 0.805  Deviance explained = 81.8%
GCV score = 0.95478  Scale est. = 0.88778  n = 171
```

Pour le modèle a, nous avons un GCV score =0.95478 et un AIC= 478.4802

```
> AIC(ModeleAT.gam)
[1] 478.4802
```

*Pour le modèle b, nous obtenons toujours pour la branche AT, l'output suivant :*

```
Parametric coefficients:
      Estimate Std. Error t value Pr(>|t|)
(Intercept) 15.81862    0.06502   243.3  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Approximate significance of smooth terms:
      edf Ref.df      F p-value
s(year)  5.000  5.000  3.495  0.0051 **
s(lag)   6.000  6.000 11.075 3.15e-10 ***
s(year,lag) 6.935  7.677  4.479 8.62e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

R-sq.(adj) = 0.841  Deviance explained = 85.8%
GCV score = 0.81288  Scale est. = 0.72287  n = 171
```

### Chapitre 3 : Les provisions techniques en assurance non vie

Pour le modèle b, nous avons un GCV score =0.81288 et un AIC= 449.5849

```
> AIC(ModeleAT.gam)
[1] 449.5849
```

*Au vu de ces résultats, le modèle b est le plus meilleur car il fournit un GCV score et un AIC plus faible. C'est cette approche qui a été utilisée tout le long de la modélisation GAM et le modèle retenu pour les trois branches est le suivant :*

$$\eta_{ij} = c + s_{\lambda}(\text{year}_i) + s_v(\text{lag}_j) + s_{\lambda,v}(\text{year}_i, \text{lag}_j)$$

*J'ai préféré commencer par ces étapes préliminaires pour montrer que le plus important dans un modèle GAM réside dans ces inputs. Plus le modèle est bien calibrer (bonne spécification du modèle, bonne calibration de la fonction de lissage), plus les outputs (les prédictions) seront bonnes.*

#### **NB**

*Nous rappelons aussi que nous avons essayé d'introduire l'année calendaire comme variable explicative du modèle, mais ceci a donné un résultat raisonnable pour la branche « Accident de travail » et « Automobile matériel ». Cependant pour la branche « Automobile Corporel », nous avons obtenu des résultats qui s'écartent de la réalité, du coup cette variable a été écartée.*

#### **Résultats du calcul des provisions techniques (link=log)**

	AT	Automobile Corporelle	Automobile Matérielle
Provisions	8491683	14859011	1323171

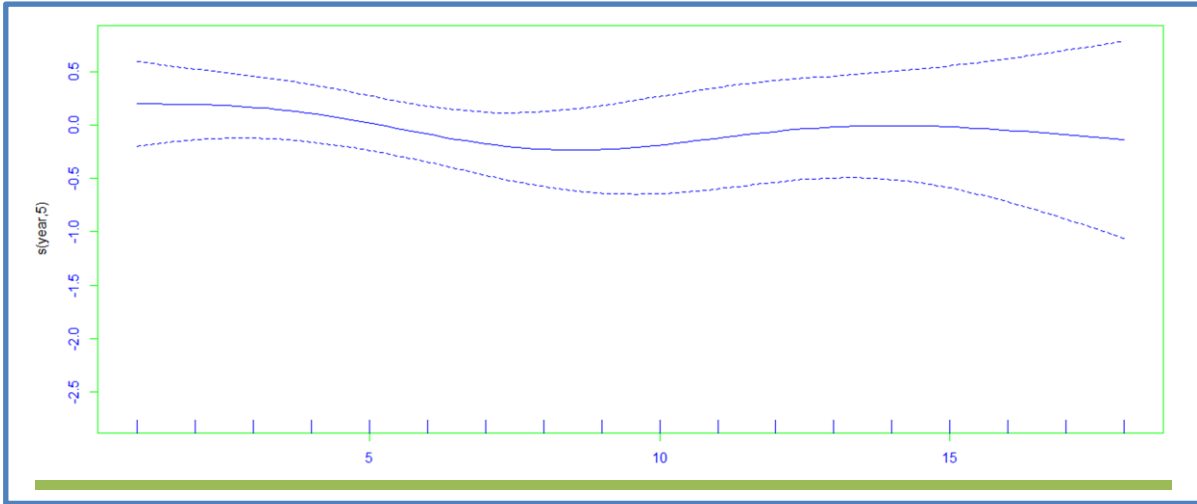
**Tableau 5 : Provisions techniques sans introduction du Bootstrap : Données « as if »**

	AT	Automobile Corporelle	Automobile Matérielle
Provisions	8991033	15922377	1368256

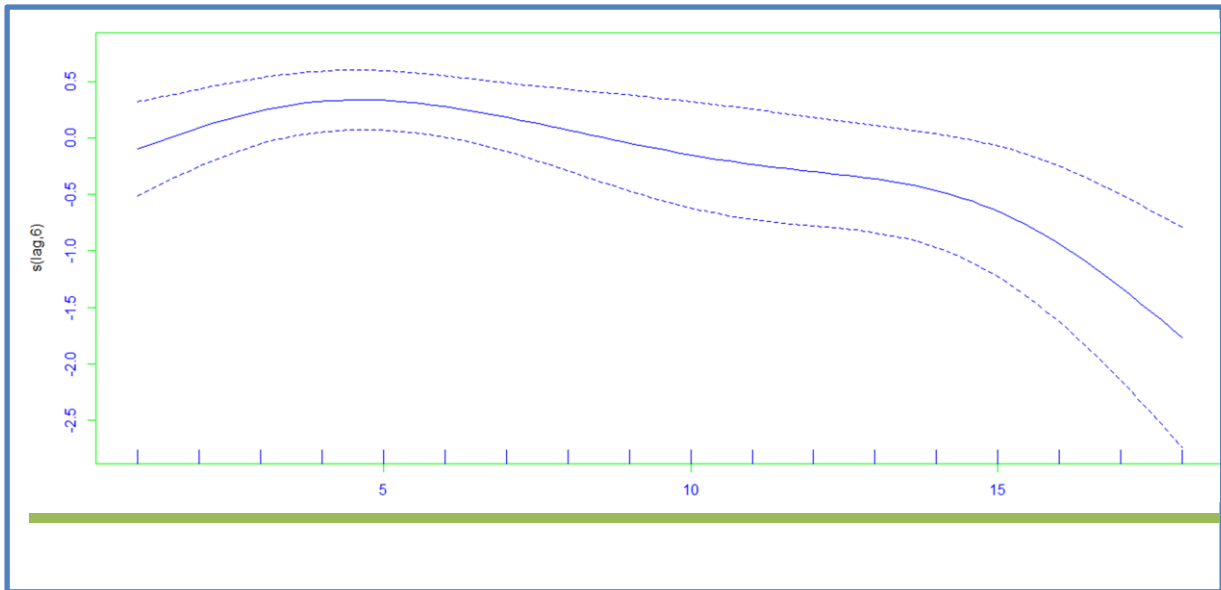
**Tableau 6 : Provisions techniques sans introduction du Bootstrap : Données non « as if »**

Visualisation de la structure de lissage et diagnostic du modèle GAM ‘Données non as if’

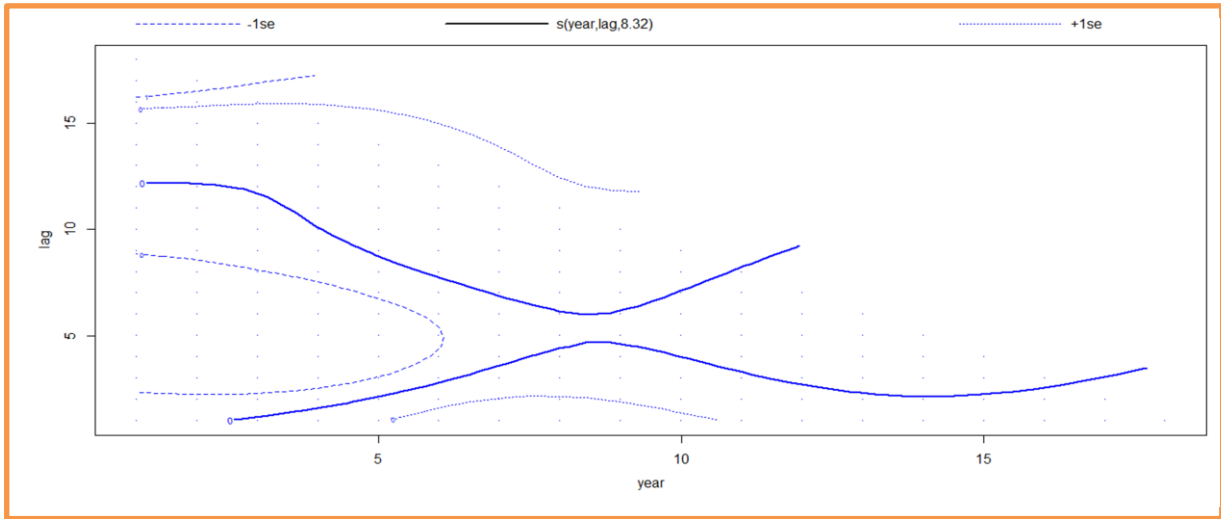
Branche Accident de Travail



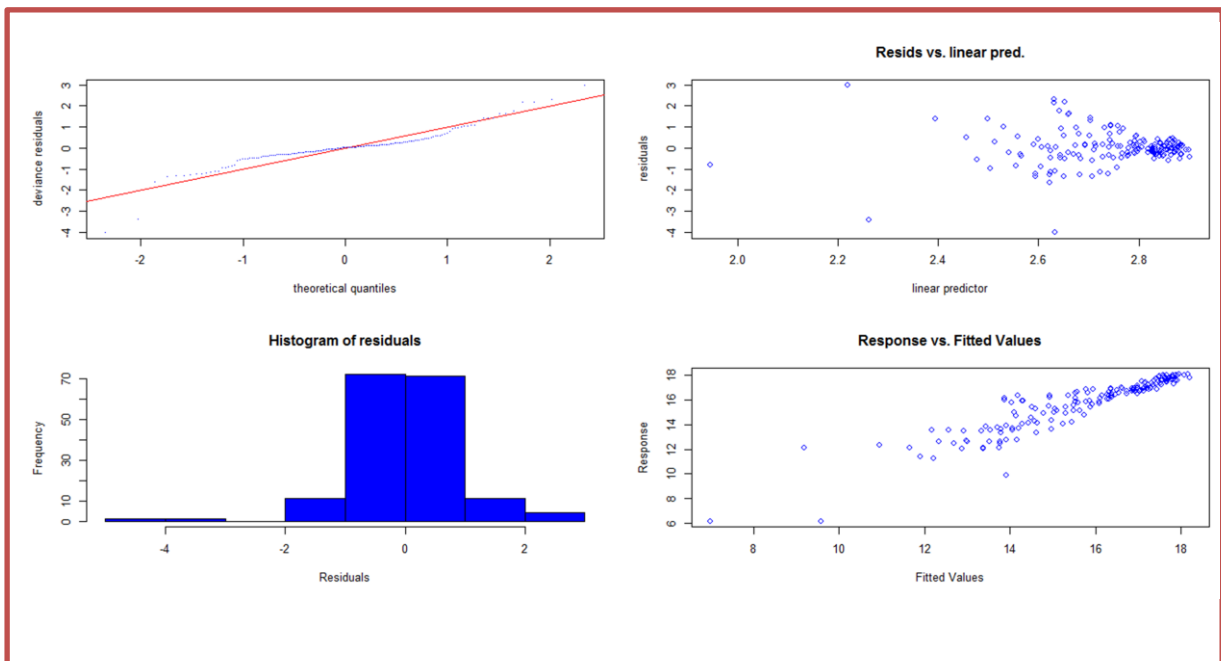
**Figure 10: Structure de lissage de la variable « année de survenance » de la branche Accident de travail pour la loi log normale**



**Figure 11 : Structure de lissage de la variable « année de développement » de la branche Accident de travail pour la loi log normale**

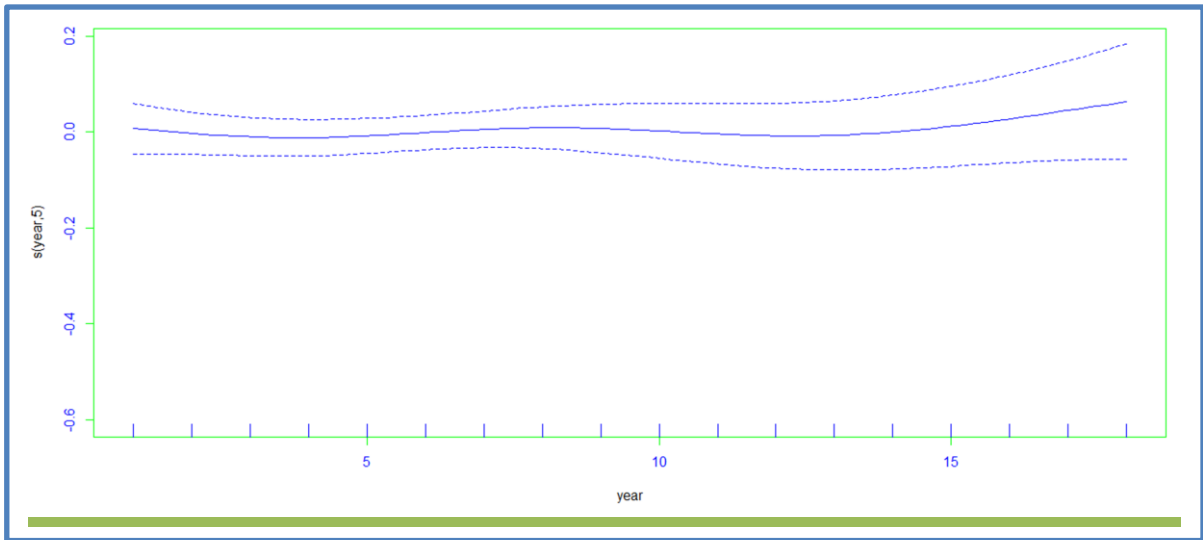


**Figure 12: Structure de lissage simultané des deux variables « année de survenance et année de développement » de la branche Accident de travail pour la loi log normale**

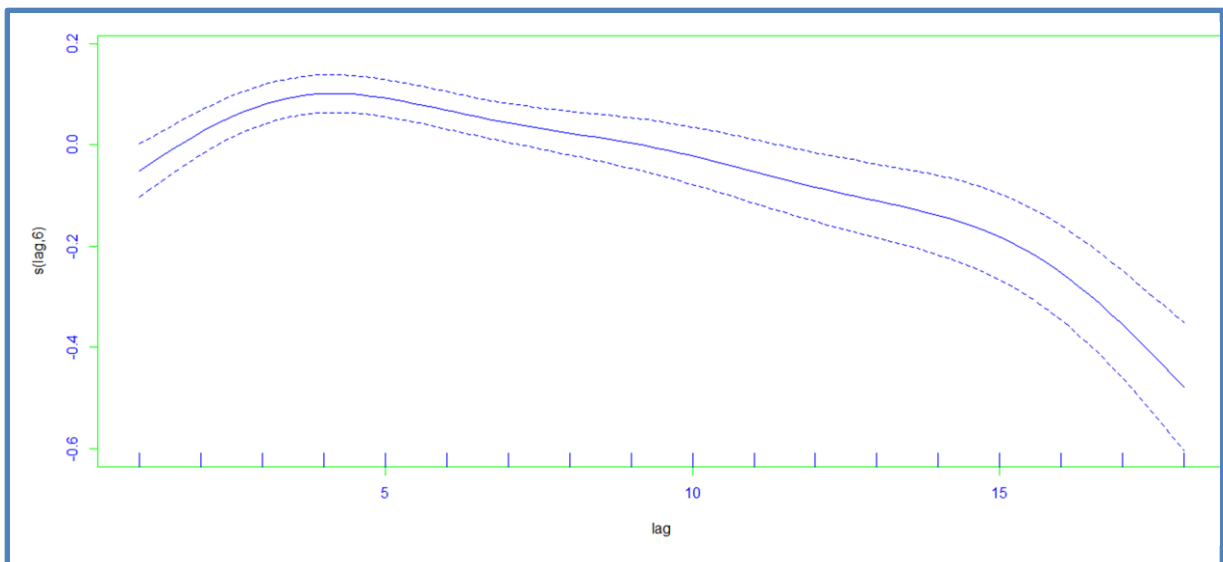


**Figure 13 : Diagnostics du modèle GAM de la branche de la branche Accident de travail pour la loi log normale**

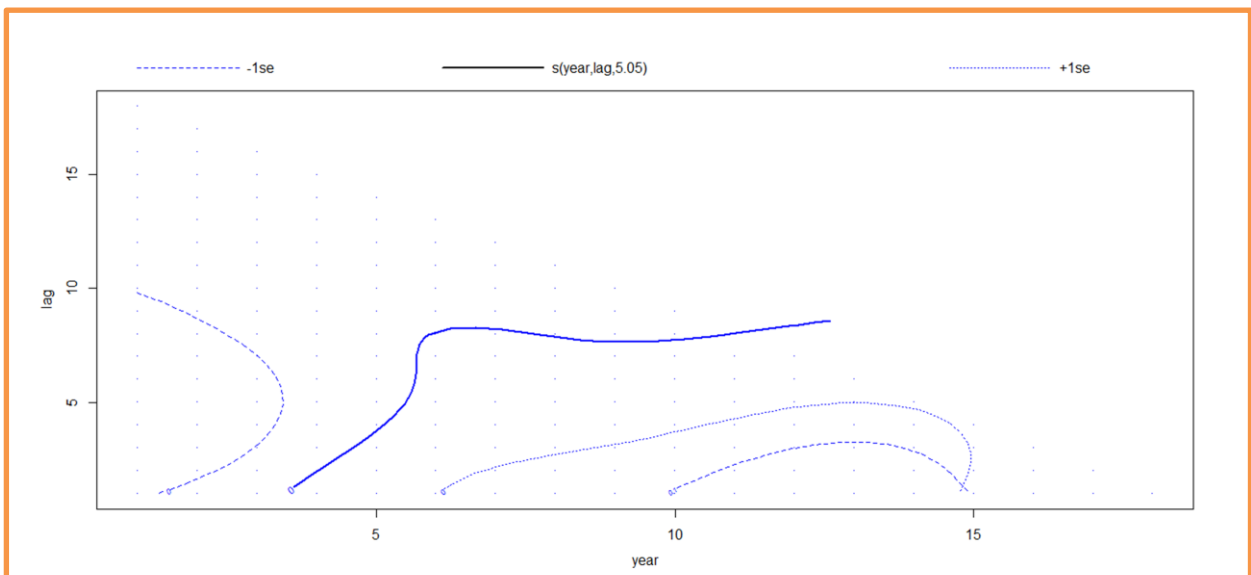
**Branche « Automobile Corporelle »**



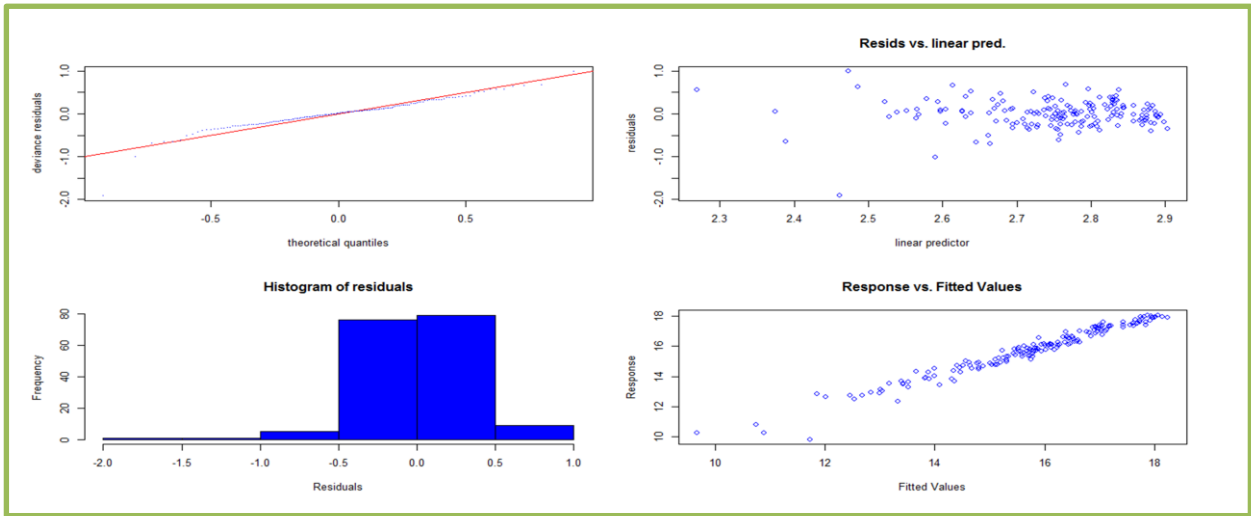
**Figure 14 : Structure de lissage de la variable « année de survenance » de la branche Automobile Corporel pour la loi log normale**



**Figure 15 : Structure de lissage de la variable « année de développement » de la branche Automobile Corporel pour la loi log normale**

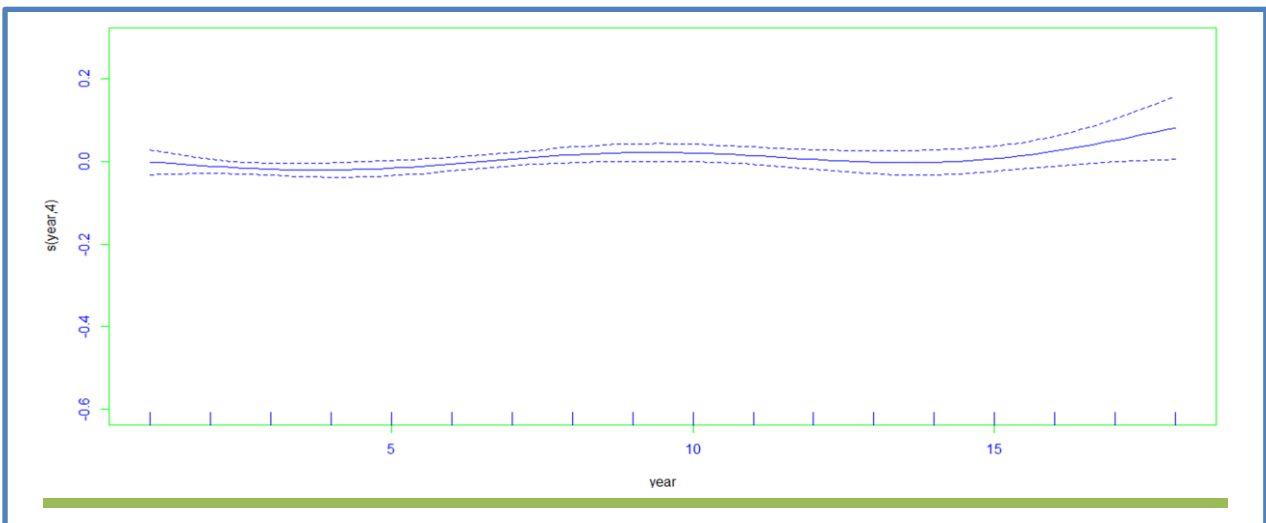


**Figure 16 : Structure de lissage simultané des 2 variables « année de survenance et année de développement » de la branche Automobile Corporel pour la loi log normale**

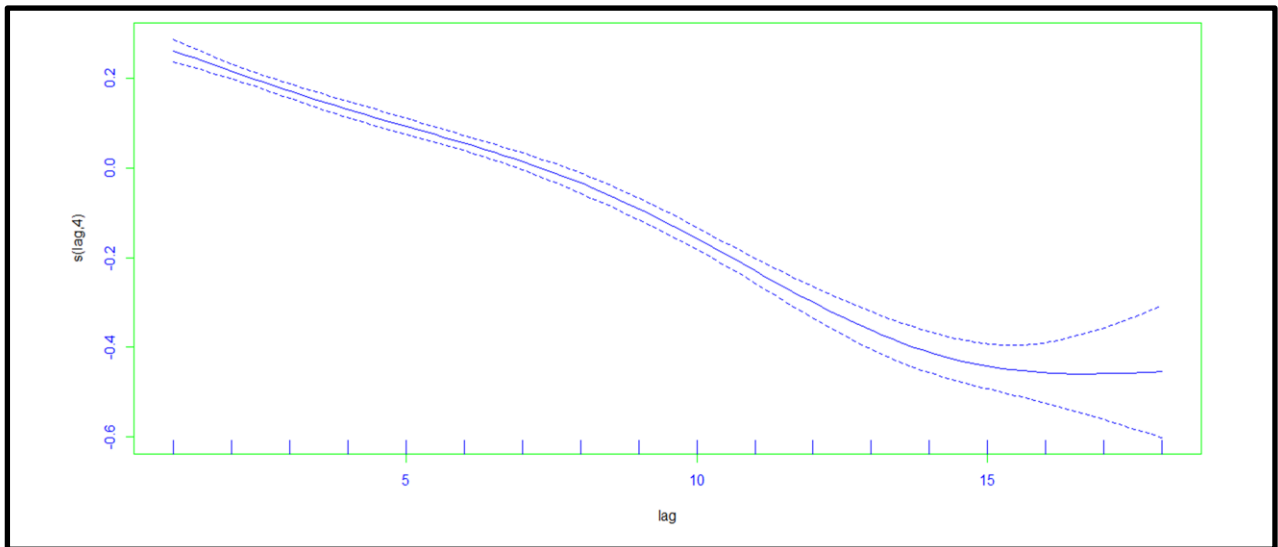


**Figure 17 : Diagnostics du modèle GAM de la branche Automobile Corporel pour la loi log normale**

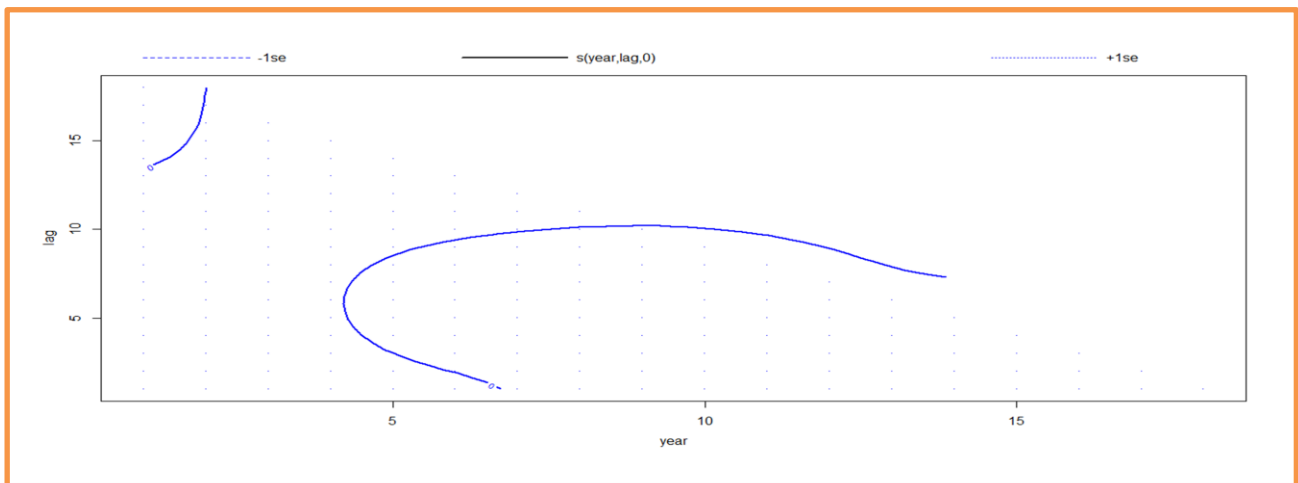
**Branche « Automobile Matériel »**



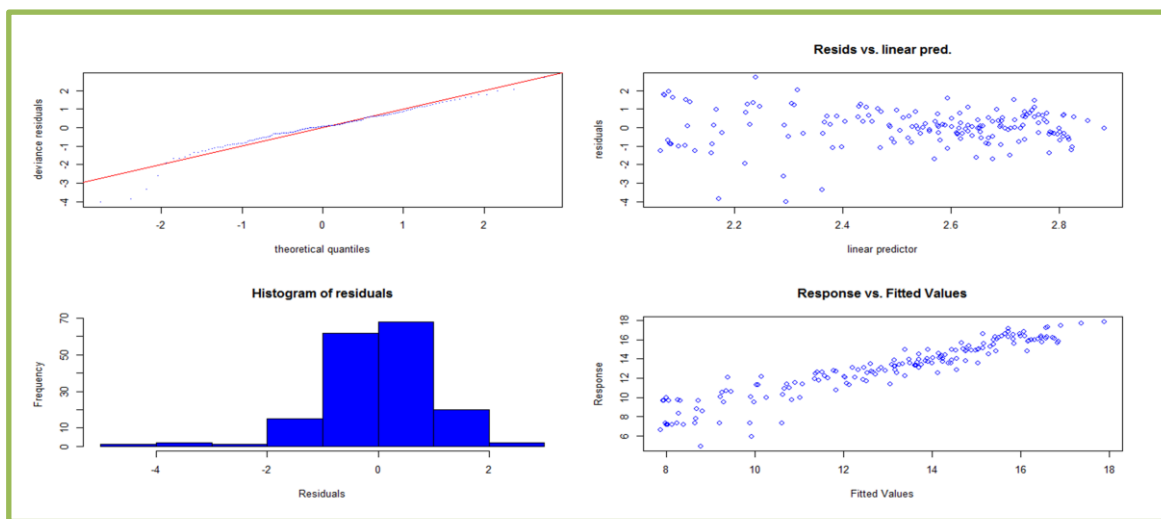
**Figure 18 : Structure de lissage de la variable « année de survenance » de la branche Automobile Matériel pour la loi log normale**



**Figure 19 : Structure de lissage de la variable « année de développement » de la branche Automobile Matériel pour la loi log normale**



**Figure 20 : Structure de lissage simultané des 2 variables « année de survenance et année de développement » de la branche Automobile Matériel pour la loi log normale**



**Figure 21 : Diagnostics du modèle GAM de la branche Auto Matériel pour la loi log normale**

**Chapitre 3 : Les provisions techniques en assurance non vie**

**Bootstrap-GAM : Données « as if »**

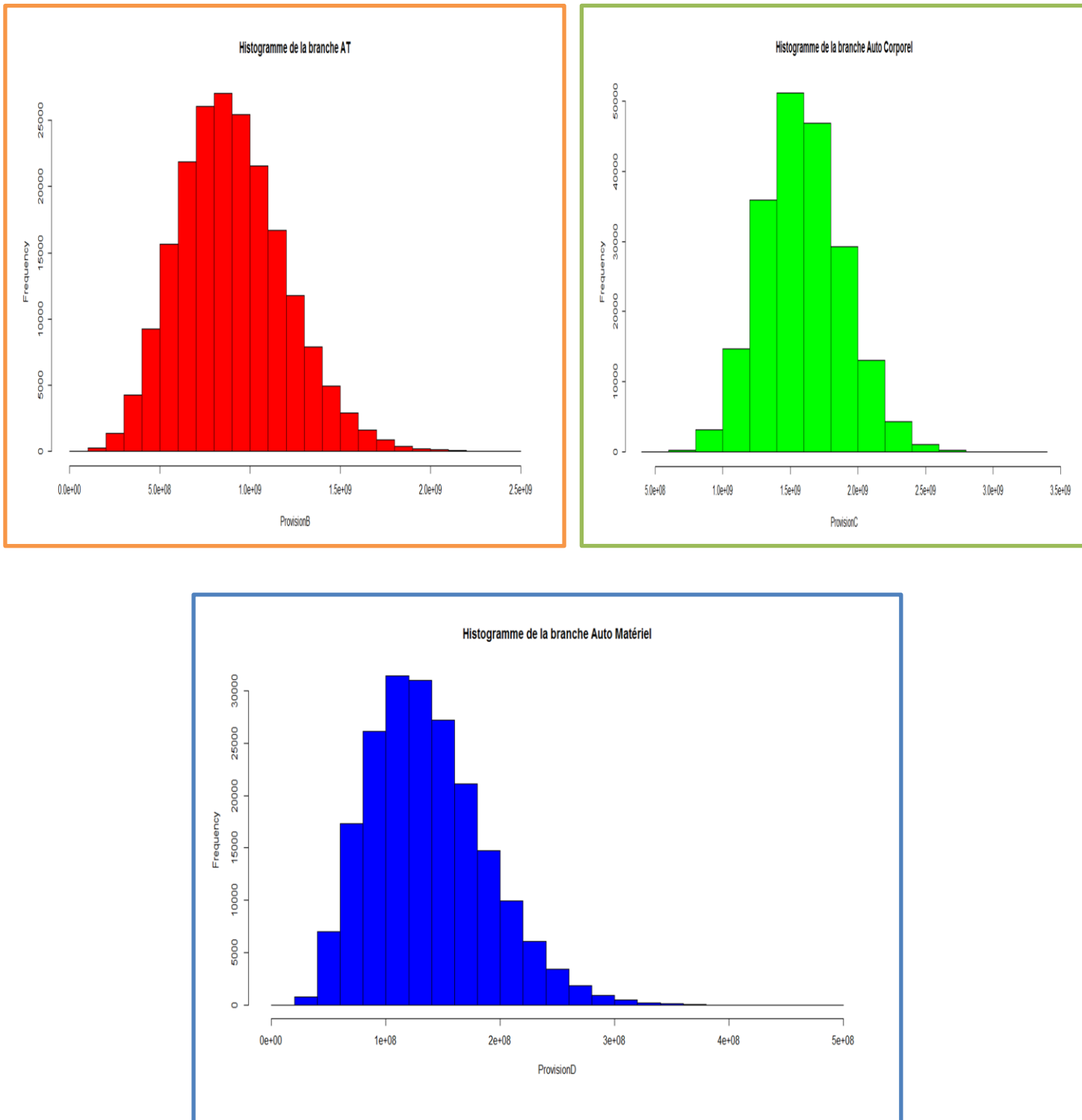
*Pour le calcul des provisions par la méthode « Bootstrap », c'est l'approche par quantile qui est privilégiée. Nous allons à partir de la distribution des provisions techniques des trois branches, déterminer la Value at Risk (VaR) et la Conditional Tail Expectation (CTE), voir tableau ci-dessous :*

	AT	Automobile Corporelle	Automobile Matérielle
<b>VaR à 50 %</b>	8014567	14583270	1230638
<b>VaR à 75 %</b>	10155703	16629811	1605979
<b>VaR à 99 %</b>	18133375	22929062	3082746
<b>CTE à 50 %</b>	10860321	17157827	1739005
<b>CTE à 75 %</b>	12720763	18770267	2075111
<b>CTE à 99 %</b>	20536378	24524521	3554865

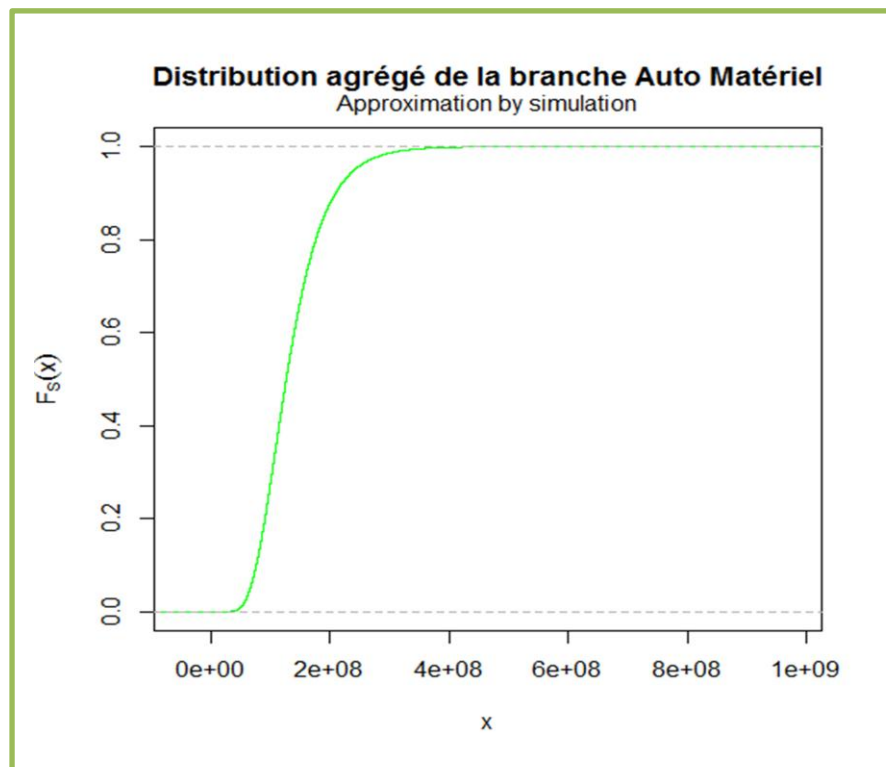
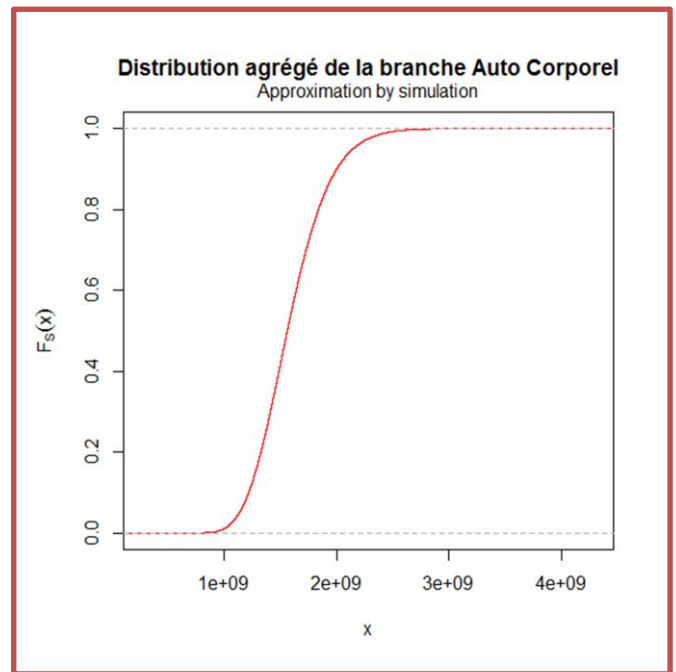
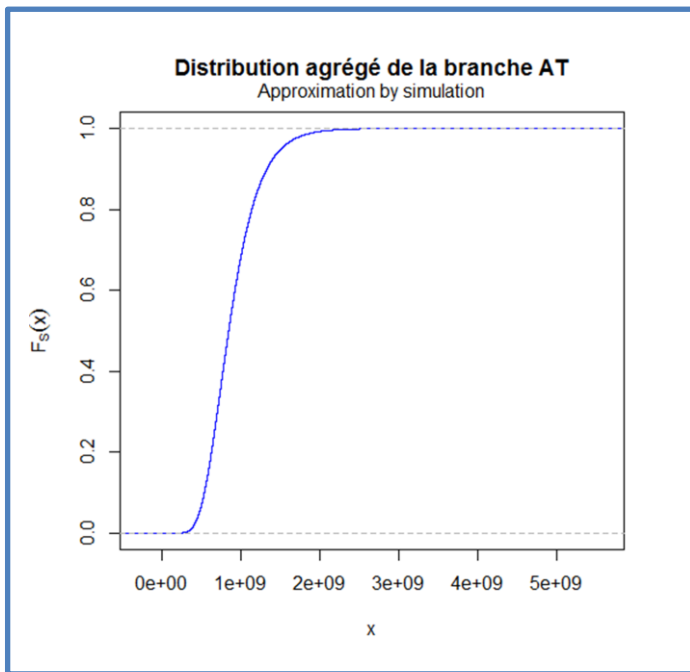
**Tableau 7 : Provisions techniques avec introduction du Bootstrap dans un modèle GAM pour la loi log normale: Données « as if » pour la loi log normale**

	AT	Automobile Corporelle	Automobile Matérielle
<b>VaR à 50 %</b>	8487669	15642230	1271186
<b>VaR à 75 %</b>	10760164	17818729	1654404
<b>VaR à 99 %</b>	19250774	24498142	3153636
<b>CTE à 50 %</b>	11511924	18375172	1789469
<b>CTE à 75 %</b>	13488088	20083683	2131552
<b>CTE à 99 %</b>	21849438	26201713	3633921

**Tableau 8 : Provisions techniques avec introduction du Bootstrap dans un modèle GAM pour la loi log normale: Données non « as if »**



**Figure 22 : Histogramme des trois branches dans un modèle GAM-bootstrap pour la loi log normale**



**Figure 23 : Distribution agrégée des trois branches dans un modèle GAM-Bootstrap pour la loi log normale**

✓ loi de Gamma (link=log)

	AT	Automobile Corporelle	Automobile Matérielle
<b>Provisions</b>	6555494	15004915	997417

**Tableau 9 : Provisions techniques sans introduction du Bootstrap dans un modèle GAM pour la loi Gamma : Données non « as if »**

	AT	Automobile Corporelle	Automobile Matérielle
<b>Provisions</b>	6168497	13862642	947785

**Tableau 10 : Provisions techniques sans introduction du Bootstrap dans un modèle GAM pour la loi Gamma : Données « as if »**

	AT	Automobile Corporelle	Automobile Matérielle
<b>VaR à 50 %</b>	6316978	14842258	961734
<b>VaR à 75 %</b>	7836644	16746787	1193213
<b>VaR à 99 %</b>	12507077	22078522	1906219

<b>CTE à 50 %</b>	8226570	17175060	1252863
<b>CTE à 75 %</b>	9424277	18604653	1435301
<b>CTE à 99 %</b>	13660915	23315368	2082457

**Tableau 11 : Provisions techniques avec introduction du Bootstrap dans un modèle GAM pour la loi Gamma : Données non « as if »**

**GAM-Bootstrap ‘Données as if’**

	AT	Automobile Corporelle	Automobile Matérielle
<b>VaR à 50 %</b>	5958617	13717022	912167
<b>VaR à 75 %</b>	7383119	15504559	1136321
<b>VaR à 99 %</b>	11774821	20501008	1828361
<b>CTE à 50 %</b>	7751637	15906074	1194374
<b>CTE à 75 %</b>	8877240	17247391	1371364
<b>CTE à 99 %</b>	12860256	21649603	2000684

**Tableau 12 : Provisions techniques avec introduction du Bootstrap dans un modèle GAM pour la loi Gamma : Données « as if »**

✓ **loi de poisson(link=log)**

	AT	Automobile Corporelle	Automobile Matérielle
<b>Provisions</b>	8175239	16792898	1173286

**Tableau 13 : Provisions techniques sans introduction du Bootstrap dans un modèle GAM pour la loi de poisson : Données non « as if »**

	AT	Automobile Corporelle	Automobile Matérielle
<b>Provisions</b>	7936112	22625522	947784

**Tableau 14 : Provisions techniques sans introduction du Bootstrap dans un modèle GAM pour la loi de poisson : Données « as if »**

**2. Prise en compte de la dépendance : GAM-Copule-Bootstrap**

Dans cette section, nous allons introduire les copules dans le calcul des provisions techniques. La prise en compte de la dépendance se fera sur deux branches à savoir l'accident de travail (AT) et l'automobile corporelle (AC). Cette décision s'est faite en étudiant la corrélation entre les branches et par la représentation graphique de l'une des branches en fonction de l'autre. Pour simplifier le problème, nous traiterons dans la suite du mémoire seulement le cas où les **données ne sont pas as if** et la variable dépendante suit une loi de gamma.

**Etude de la corrélation entre les branches Accident de travail et Automobile Corporel**

Notation :

increLossB : Montants incrémentaux de l'accident de travail (AT)

increLossC : Montants incrémentaux de l'automobile corporelle (AC)

increLossD : Montants incrémentaux de l'automobile matérielle (AM)

```
> cor
      increLossB increLossC
increLossB 1.0000000 0.5488887
increLossC 0.5488887 1.0000000
```

Taux de Kendal

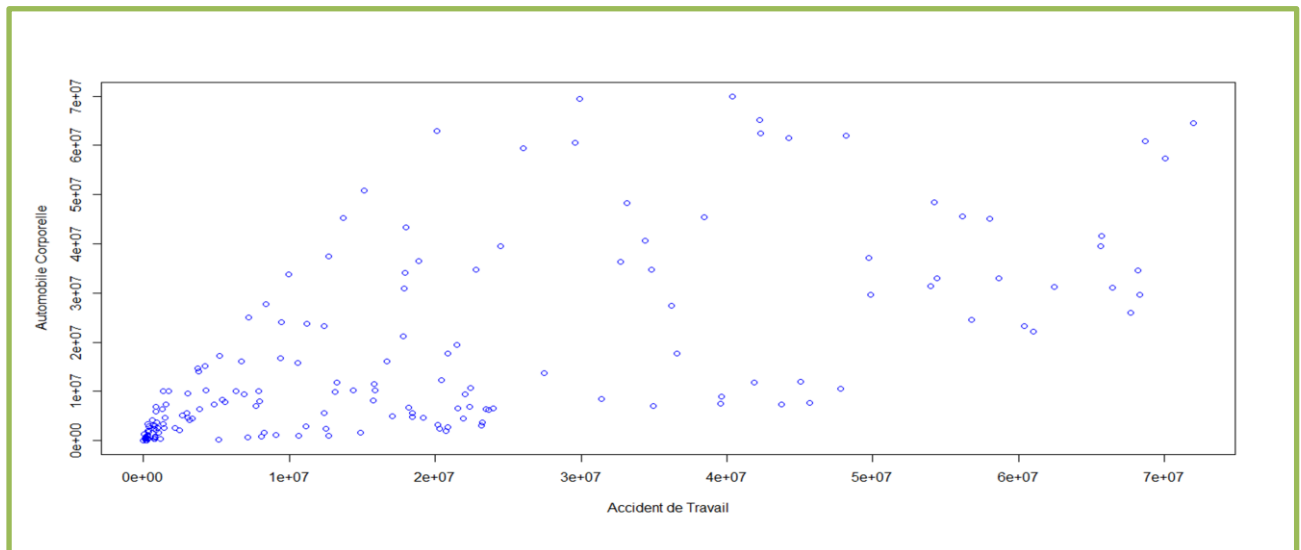
```
> cor1
      increLossB increLossC
increLossB 1.0000000 0.6522404
increLossC 0.6522404 1.0000000
```

Taux de pearson

```
> cor2
      increLossB increLossC
increLossB 1.0000000 0.7433382
increLossC 0.7433382 1.0000000
```

Taux de spearman

La valeur du coefficient de corrélation de Kendall et de spearman, nous fait penser qu'il existe une corrélation non linéaire entre les branches AT et AC.



**Figure 24 : Représentation graphique de la dépendance entre la branche Automobile Corporelle et Accident de travail**

### Chapitre 3 : Les provisions techniques en assurance non vie

Au vu des résultats (étude de la corrélation, représentation graphique) ci-dessus, l'existence d'une corrélation entre les branches AT et AC peut être soulevée.

La problématique est : **Comment introduire la copule dans un modèle GAM ?**

Ainsi pour répondre à cette question, j'ai introduit une notion appelée « complément des provisions techniques ».

En effet, Cette approche à laquelle j'ai pensé sert en quelques sortes à corriger les provisions techniques en cas d'erreurs de prédiction en prenant en compte d'autres types d'informations à savoir :

- ✚ L'information contenue dans les marginales dans le cas où nous disposons d'une bonne connaissance de la distribution de la variable aléatoire dépendante.
- ✚ L'information obtenue selon une approche non paramétrique dans le cas où nous avons moins d'information sur la distribution de la variable aléatoire.
- ✚ La prise en compte de la dépendance entre les branches par l'introduction des copules.

Dans le souci d'obtenir une expression close du complément et du fait que le complément peut être soit positif ou négatif, nous avons supposé qu'il suivait une loi normale d'espérance nulle et d'écart type  $\sigma$ , voir ci-dessous:

$$C_{ij} \sim N(0, \sigma) \xrightarrow{\text{implique}} C_{ij} = \sigma \Phi^{-1}(u) \quad \forall u \in [0; 1], \quad \sigma = \sigma(Y_{ij})$$

Nous obtenons donc une expression close du complément ce qui nous facilitera sa mise en œuvre.

Le modèle GAM-Copule devient alors :

$$\mu_{ij} = g^{-1}(s_{\lambda}(year_i) + s_v(lag_j) + s_{\lambda,v}(year_i, lag_j)) + \sigma_{ij} \Phi^{-1}(u_{ij})$$

$$u_{ij} = \frac{\text{Rang}(\widehat{Y}_{ij})}{N+1}, \quad N \text{ étant le nombre d'observations prédites.}$$

Résultats sur des données non « as if »

	AT	Automobile Corporelle	Automobile Matérielle
Provisions	14148795	18461603	2130401

**Tableau 15 : Provisions techniques sans introduction du Bootstrap dans un modèle GAM-Copule pour la loi de Gamma**

	AT	AC	AM
VaR à 50 %	13529785	18178396	2008197
VaR à 75 %	17222447	20972163	2640774

**Tableau 16 : Provisions techniques dans un modèle GAM-Copule-Bootstrap pour la loi de Gamma**

**3. Modèle GAM Mixed**

Cette section concerne le cas pratique pour le modèle GAM Mixed. Le modèle retenu lors de l'implémentation sous R est :

$$\mu_{ij} = g^{-1}(s_{\lambda}(year_i) + s_{\nu}(lag_j) + \phi\varepsilon_{t-1} + \varsigma_t)$$

Nous allons utiliser ce modèle pour calculer les provisions techniques pour les trois branches AT, Auto Corporel et Auto Matériel sous l'hypothèse que les règlements de sinistres non cumulés ont une distribution Gamma.

Ce qui nous a motivé à passer de la méthode GAM à la méthode GAM Mixed, c'est la valeur de la statistique du test de Durbin Watson qui prouve qu'il existe une autocorrélation entre les erreurs. Ce test s'est fait sous R en utilisant le package lmtest ou car, voir ci-dessous :

**Branche B (Package car)**

```

> testb1
lag Autocorrelation D-W Statistic p-value
1      0.1796719      1.640426      0
Alternative hypothesis: rho != 0
    
```

Il existe une autocorrélation positive car la statistique du test =1.64 < 2

**Chapitre 3 : Les provisions techniques en assurance non vie**

**Branche C**

```

lag Autocorrelation D-W Statistic p-value
1 -0.02206492 2.042061 0
Alternative hypothesis: rho != 0
    
```

Il existe une autocorrélation négative égale à -0.022

**Branche D**

```

> testd1
lag Autocorrelation D-W Statistic p-value
1 0.2205699 1.555012 0
Alternative hypothesis: rho != 0
    
```

Existence d'une autocorrélation positive

	AT	AC	AM
<b>Provisions</b>	<b>5761794</b>	<b>13286457</b>	<b>763855</b>

**Tableau 17 : Provisions techniques dans un modèle GAM Mixed pour la loi de Gamma**

	AT	Automobile Corporelle	Automobile Matérielle
<b>VaR à 50 %</b>	5706679	13192070	744661
<b>VaR à 75 %</b>	6361978	14562858	895222
<b>VaR à 99 %</b>	8170001	18301474	1344495
<b>VaR à 100 %</b>	11616259	24990175	2304440
<b>CTE à 50 %</b>	6506538	14860352	932304
<b>CTE à 75 %</b>	6994641	15876477	1048950
<b>CTE à 99 %</b>	8586454	19147190	1453542

**Tableau 18 : Provisions techniques dans un modèle GAM Mixed-Bootstrap pour la loi de Gamma**

**Conclusion du modèle GAM**

Le modèle GAM a été introduit dans le calcul des provisions techniques parce que nous étions confrontés à un problème dans lequel notre variable dépendante ne suivait aucune loi. Ceci, la prédiction des charges futures devraient se faire par un modèle de régression non paramétrique. Après avoir opté pour le modèle GAM, le choix du meilleur modèle était encore un souci. Pour contourner ce problème, plusieurs modèles ont été testé et le meilleur modèle fût celui possédant un GCV score faible et un AIC faible. Ensuite, un test de Durbin Watson nous a permis de constater qu'il y avait une autocorrélation entre les erreurs du modèle. Ce nouveau problème nous a poussé à introduire le modèle GAM Mixed qui en tient compte dans la prédiction des charges futures. Une nouvelle, une panoplie de modèle GAM Mixed fût testé et le meilleur a été choisi sur la base d'un AIC plus faible. Les meilleurs AIC obtenus pour les trois branches sont :

```
> AIC(ModeleAT.gam$lme)
[1] 449.3646
```

```
> AIC(ModeleCorp.gam$lme)
[1] 156.6221
```

```
> AIC(ModeleMat.gam$lme)
[1] 435.6481
```

**Branche AT : 449,3646**

**Branche AC : 156,6221**

**Branche AM : 435,6481**

Pour simplifier la rédaction, toutes ces étapes n'ont pas été présentées dans ce rapport. C'est au praticien d'en tenir compte lors de l'estimation de son modèle.

Cette analyse nous permet de dire que dans cette étude, le meilleur modèle GAM pour le calcul des provisions techniques est le Generalized Additive Model Mixed ou GAM Mixed.

Ce modèle nous fournit même des provisions techniques inférieur au modèle de Chain Ladder (modèle de référence), voir tableau ci-dessous :

	<b>GAM Mixed</b>	<b>Chain Ladder</b>
<b>Accident de Travail</b>	<b>5761794</b>	<b>7080085</b>
<b>Auto Corporelle</b>	<b>13286457</b>	14002562
<b>Auto Matérielle</b>	<b>763855</b>	1302433

**Ce résultat est intéressant et montre que le modèle GAM Mixed ne cherche ni à surestimer ou à sous-estimer la valeur des provisions. Ainsi, il peut nous servir de base pour faire du bootstrap qui nous permettra de converger vers une valeur stable et donc vers le best estimate qui pour ma part est le quantile à un seuil de 100 % (AT : 11616259, AutoC : 24990175, AutoM : 2304440)**

**II.2. Modèle de Calcul des provisions techniques par la méthode GLM**

**II.2.1. Approche simple de la méthode GLM**

**Rappel du modèle**

$$\widehat{\mu}_{ij} = g^{-1}(\widehat{\eta}_{ij}) = g^{-1}(\beta_0 + \beta_1 year + \beta_2 lag)$$

g étant la fonction lien et  $\widehat{\mu}_{ij}$  les valeurs prédites du modèle GLM.

$$\widehat{\mu}_{ij} = E(X_{ij}|D_I), \quad D_I = \{i + j < I\}$$

Le terme « Données as if » signifie que les données ont été actualisé en tenant compte de l'inflation.

**Application : Calcul des provisions pour les branches AT, Auto Corp et Auto Mat**

**Hypothèse 1** :  $X_{ij}$  suivent une distribution log normale (lien=log)

	AT	Auto Corporelle	Auto Matérielle
Data « as if »	7250942	16164088	778236
Data non « as if »	7574855	17427594	810842

**Tableau 19 : Provisions techniques dans un modèle GLM pour la loi log normal**

	AT	Automobile Corporelle	Automobile Matérielle
VaR à 50 %	7442897	17293001	765082
VaR à 75 %	8456744	18831907	966776
VaR à 99 %	11566194	23206203	1713175
CTE à 50 %	8716366	19186383	1032568
CTE à 75 %	9512023	20350219	1207024
CTE à 99 %	12356372	24239665	1940047

**Tableau 20 : Provisions techniques dans un modèle GLM-Bootstrap pour la loi log normal**

**NB**

Pour la méthode GLM-Bootstrap, nous ne traiterons que le cas où les données ne sont pas « as if ». Ce qu'il faut retenir, c'est que le fait de rendre les données « as if » est très bénéfique pour l'assureur en ce sens que nous observons une baisse des provisions.

**Chapitre 3 : Les provisions techniques en assurance non vie**

**Hypothèse 2** :  $X_{ij}$  suivent une distribution Gamma (lien=log)

	AT	Auto Corporelle	Auto Matérielle
Data « as if »	5377323	14180808	759472
Data non « as if »	5718846	15268670	798354

**Tableau 21 : Provisions techniques dans un modèle GLM pour la loi de Gamma**

	AT	Automobile Corporelle	Automobile Matérielle
VaR à 50 %	5682770	15165971	782118
VaR à 75 %	6220336	16682551	924873
VaR à 99 %	7677670	20801847	1343966
CTE à 50 %	6334926	17009441	959176
CTE à 75 %	6731475	18130762	1068880
CTE à 99 %	8006476	21741750	1443339

**Tableau 22 : Provisions techniques dans un modèle GLM-Bootstrap pour la loi de Gamma**

**Hypothèse 3** :  $X_{ij}$  suivent une distribution poisson (lien=log)

	AT	Auto Corporelle	Auto Matérielle
Data « as if »	8178769	22544634	901112
Data non « as if »	8434051	23792070	943096

**Tableau 23 : Provisions techniques dans un modèle GLM pour la loi de poisson**

### II.2.2. Approche complexe de la méthode GLM

Une approche complexe d'utiliser le modèle GLM dans le calcul des provisions techniques est de le combiner avec d'autres méthodes stochastiques. Dans cette section, nous étudierons un nouveau modèle de provisionnement qui s'inscrit dans un cadre GLM et qui prend compte aussi d'autres méthodes stochastiques en occurrence le modèle de Bornhuetter-Ferguson stochastique et le modèle bayésien. Ce modèle a été présenté dans l'article du suisse *Mario V. Wuthrich*. Nous montrerons que pour des choix optimaux des paramètres du modèle, ce modèle converge vers la méthode de Chain Ladder.

#### Hypothèses du modèle

- ◆ Nous supposons qu'il existe des facteurs de développement  $\beta_0, \dots, \beta_j$  avec  $\beta_0 > 0$ ,  $\beta_j = 1$  et  $\beta_j > \beta_{j-1}$ , ( $j \geq 1$ ).

Nous définissons aussi la quantité  $\gamma_0 = \beta_0$  et  $\gamma_j = \beta_j - \beta_{j-1}$

- ◆ Conditionnellement, étant donné le paramètre  $\Theta_i$ , nous avons que les montants incrémentaux  $X_{i,j}$  sont indépendants avec :

$$Z_{i,j} = \frac{X_{i,j}}{\gamma_j \times \mu_0^i} \sim dF_{i,j}^{\Theta_i}(x) = a \left( x, \frac{\sigma^2}{\gamma_j (\mu_0^i)^2} \right) \exp \left\{ \frac{x\Theta_i - b(\Theta_i)}{\sigma^2 \gamma_j^{-1} (\mu_0^i)^{-2}} \right\}$$

Ceci montre que  $Z_{i,j}$  appartient à la famille exponentielle.

- ◆ Les vecteurs aléatoires  $(\Theta_i, (X_{i,0}, \dots, X_{i,j}))$  sont indépendants et  $\Theta_i$  est indépendant et identiquement distribué dont la densité s'écrit :

$$u_{\mu, \tau^2}(\theta) = d(\mu, \tau^2) \exp \left\{ \frac{\mu\theta - b(\theta)}{\tau^2} \right\}, \mu = 1, \tau^2 > 0$$

D'où  $\Theta_i$  appartient aussi à la famille exponentielle.

#### **NB**

$\mu_0^i = E(C_{i,j})$  Joue le rôle de la charge ultime moyenne à priori pour l'année de survenance  $i$ . C'est une information exogène, qui peut être estimée en ne tenant pas compte des données de l'entreprise.

Le but étant de calculer les provisions techniques, la prédiction des charges futures est alors une étape cruciale.

### Chapitre 3 : Les provisions techniques en assurance non vie

En s'inspirant du modèle stochastique de Bornhuetter-Ferguson, Mario montre que l'ajout d'une approche bayésienne à ce modèle revient à multiplier ce modèle par un estimateur bayésien  $\widehat{\mu}(\Theta_i)$  qui est en fait un facteur de crédibilité.

En clair, ce modèle s'écrit :

$$E(X_{i,j}) = \underbrace{\gamma_j \times \mu_0^i}_{\text{facteur de crédibilité}} \times \widehat{\mu}(\Theta_i)$$

#### Modèle stochastique de Bornhuetter-Ferguson

Quant à l'estimateur bayésien, il se définit comme suit :

$$\widehat{\mu}(\Theta_i) = \alpha_i^{\min(I-i,J)} \times \bar{Z}_i + (1 - \alpha_i^{\min(I-i,J)})$$

Comme  $I=J$  dans notre cas alors  $\min(I-i,J) = I-i$  d'où nous obtenons :

$$\widehat{\mu}(\Theta_i) = \alpha_i^{I-i} \times \bar{Z}_i + (1 - \alpha_i^{I-i})$$

Avec :

$$\bar{Z}_i = \frac{C_{\min(I-i,J)}}{\beta_{\min(I-i,J)} \times \mu_0^i} = \frac{C_{I-i}}{\beta_{I-i} \times \mu_0^i}$$

$$\alpha_i^{I-i} = \frac{\beta_{I-i}}{\beta_{I-i} + k_i}$$

$$\text{Avec } k_i = \frac{\sigma^2}{\tau^2(\mu_0^i)^2}$$

Le but de cette section est de montrer que ce nouveau modèle converge vers la méthode de Chain Ladder si nous posons :

$$\alpha_i^{I-i} = \beta_{I-i}$$

$$\mu_0^i = \hat{C}_{i,J}^{CL}$$

**Démonstration**

Si  $\mu_0^i = \hat{C}_{i,J}^{CL}$  alors

$$\bar{Z}_i = \frac{C_{\min(I-i,J)}}{\beta_{\min(I-i,J)} \times \mu_0^i} = \frac{C_{I-i}}{\beta_{I-i} \times \hat{C}_{i,J}^{CL}}$$

Or  $\hat{C}_{i,J}^{CL} = \frac{C_{I-i}}{\beta_{I-i}}$  ce qui implique  $\bar{Z}_i = \frac{C_{I-i}}{\beta_{I-i}} \times \frac{\beta_{I-i}}{C_{I-i}} = 1 \quad \forall i$

Comme  $\alpha_i^{I-i} = \beta_{I-i}$  implique :

$$\mu(\bar{\Theta}_i) = \alpha_i^{I-i} \times \bar{Z}_i + (1 - \alpha_i^{I-i}) = \beta_{I-i} + (1 - \beta_{I-i}) = 1$$

On déduit que :

$$E(X_{i,j}) = \gamma_j \times \mu_0^i \times \mu(\bar{\Theta}_i) = \gamma_j \times \hat{C}_{i,J}^{CL} = E(X_{i,j}^{CL})$$

Ce qui nous permet de conclure que le modèle de Mario et celle de la méthode de Chain Ladder converge lorsque :

$\alpha_i^{I-i} = \beta_{I-i}$  Ce qui veut dire que le facteur de crédibilité coïncide avec le facteur de développement  $\beta$ .

Nous rappelons au lecteur que  $\beta_{J-i} = \prod_{j=J-i}^{J-1} \frac{1}{\hat{f}_j}$

$$\mu_0^i = \hat{C}_{i,J}^{CL}$$

**Application sous R**

- Facteurs de développement  $\hat{f}_j$

```
> L
[1] 6.304428 2.479705 1.630905 1.292087 1.161147 1.086470 1.054566 1.035059
[9] 1.024351 1.015313 1.010261 1.008409 1.005115 1.004296 1.003096 1.000531
[17] 1.000402 1.000000
```

### Chapitre 3 : Les provisions techniques en assurance non vie

- Facteurs de développement  $\hat{\beta}_j$

```
> beta
[1] 0.02052797 0.12941713 0.32091635 0.52338399 0.67625745 0.78523399
[7] 0.85313304 0.89968539 0.93122764 0.95390352 0.96851082 0.97844887
[13] 0.98667655 0.99172339 0.99598381 0.99906787 0.99959841 1.00000000
```

- Charge ultime de Chain Ladder  $\mu_0^i = \hat{C}_{i,J}^{CL}$

```
> u0
[1] 70042620 75870783 84652795 83482845 144313178 196694246 217579209
[8] 225633456 223315969 252538825 297910890 294279167 300920979 301192968
[15] 319722177 357608529 317283180 459275212
```

- Coefficients  $\hat{\gamma}_j = \hat{\beta}_j - \hat{\beta}_{j-1}$  avec  $\sum_{j=1}^J \hat{\gamma}_j = 1$

```
> gammi
[1] 0.0205279746 0.1088891539 0.1914992254 0.2024676359 0.1528734615
[6] 0.1089765429 0.0678990430 0.0465523510 0.0315422564 0.0226758737
[11] 0.0146073050 0.0099380485 0.0082276782 0.0050468425 0.0042604210
[16] 0.0030840602 0.0005305387 0.0004015878
> sum(gammi)
[1] 1
```

- Coefficients  $\bar{Z}_i$  et  $\mu(\bar{\Theta}_i)$

Nous avons montré, voir ci-dessus que  $\bar{Z}_i = 1 \forall i$  et  $\mu(\bar{\Theta}_i) = 1 \forall i$

```
> zi
[1] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
```

```
> cred
[1] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
```

- Provisions

```
> prov
[1] 14002562
```

*Cette provision est identique à la provision de Chain Ladder ce qui confirme le résultat théorique.*

**Chapitre 4 : Ratio de sinistralité d'une compagnie d'assurance non vie**

**I. Introduction**

Le ratio S/ P se définit comme étant le rapport des charges de sinistres payées par les primes encaissées au cours de l'exercice.

C'est un moyen de mesurer l'impact de la sinistralité sur le portefeuille en question et il constitue de ce fait un indicateur de rentabilité pour les entreprises d'assurance.

**II. Présentation du modèle**

Soit :

- $LR_{i,j,l}$  le ratio de perte pour la ligne  $l$  ( $l=1 \dots L$ ) correspondant à l'année de survenance  $i$  ( $i=1 \dots T$ ) et l'année de développement  $j$  ( $j=1 \dots K$ ).
- $LR_i$  le ratio agrégé pour l'année de survenance  $i$ .
- $S$  la charge de sinistre et  $P$  la prime encaissé

Nous avons donc :

$$LR_i = \frac{S_i}{P_i} = \frac{\sum_{l=1}^L \sum_{j=1}^K S_{i,j,l}}{\sum_{l=1}^L \sum_{j=1}^K P_{i,j,l}}$$

Après les calculs, nous obtenons :

$$LR_i = \sum_{l=1}^L \sum_{j=1}^K LR_{i,j,l} \times \omega_{i,j,l}$$

Avec  $\omega_{i,j,l} = \frac{P_{i,j,l}}{\sum_{l=1}^L \sum_{j=1}^K P_{i,j,l}}$  et  $LR_{i,j,l} = \frac{S_{i,j,l}}{P_{i,j,l}}$

En observant l'expression de  $LR_i$  , seule la variable  $LR_{i,j,l}$  est aléatoire. Il s'agira de prédire les valeurs futures de  $LR_{i,j,l}$  et de déduire le ratio de sinistralité par année de survenance sous l'hypothèse que les primes sont constantes au cours de chaque année de survenance  $i$ .

### III. Application sur trois branches d'activités d'un assureur non vie

*Cette application a été effectuée sur un portefeuille d'un assureur non vie composé de trois branches B, C et D. Par souci de confidentialité, nous ne pouvons pas donner plus de détail. Ces trois branches étudiées suivaient tous une distribution normale ce qui nous a permis d'établir le modèle ci-dessous :*

$$LR_{i,j,l} = g^{-1} \left( \sum_{k=1}^2 \beta_k x_k \right) + \sigma(LR_{i,j,l}) \times \Phi^{-1}(u), \quad \forall u \sim U(0,1)$$

Les  $x_k$  représentent l'année de survenance et l'année de développement.

*Nous allons calculer le ratio de sinistralité agrégé en combinant simultanément l'approche bayésienne et la théorie des copules. Pour ne pas alourdir le mémoire, beaucoup de choses seront supprimé dans les chapitres suivants.*

La méthode de calcul peut être énoncée comme suit :

#### Etape 1

Estimer les paramètres initiaux du modèle par la méthode GLM.

#### Etape 2

Utiliser la méthode bayésienne MCMC pour déterminer  $\beta_k^{moy}$  et  $\sigma^{moy}$  en supposant que :

$$\beta_i \sim Normal(\mu, \sigma) \quad \forall i = 1,2 \quad \text{et} \quad \varphi \sim Gamma(\alpha, \gamma)$$

#### Etape 3

Utiliser la théorie des copules pour calculer  $U = C(u_1, \dots, u_L)$ ,  $u_i \sim U(0,1)$  qui joint les trois branches B, C et D.

#### Etape 4

Calculer

$$LR_{i,j,l} \approx g^{-1} \left( \sum_{k=1}^2 \beta_k^{moy} x_k \right) + \sigma^{moy} \times \Phi^{-1}(U), \quad \forall U \sim C(u_1, \dots, u_L)$$

**Résultats**

<b>Année de survenance</b>	<b>Ratio de perte</b>
1	70,61424 %
2	60,56165 %
3	80,49338 %
4	65,46457 %
5	50,41838 %
6	70,31078 %
7	70,25398 %
8	70,22811 %
9	71,34026 %
10	80,61344 %

**Tableau 24 : Résultats des ratios de sinistralité agrégés par année de survenance**

*Chapitre 5 : Capital agrégé pour des lignes d'affaires*

*1. Introduction : Problème d'agrégation des risques*

Considérons un portefeuille d'assurance soumis à  $n$  risques provenant des  $K$  lignes de business, avec  $K$  un entier connu.

Le problème de l'assureur est de connaître le coût total auquel il est soumis de sorte à prendre des dispositions (réassurance, augmentation du chargement de sécurité, etc...) pour réduire les conséquences liées à ce coût. Ce qui intéresse l'assureur est le comportement de la variable aléatoire  $S = \sum_{j=1}^K \sum_{i=1}^n X_{ij}$ , à une date fixée  $T$ .

Il s'agira pour l'assureur de déterminer la distribution exacte ou approximative (par simulation) de  $S$ , ce qui lui permettra par exemple de demander à l'assuré une prime pure égale à  $\pi = E(S) \times p$ , avec  $p$  la fréquence d'apparition du risque qui couvrira exactement le risque.

Pour atteindre cet objectif, deux hypothèses peuvent être énoncées :

- ✚ Supposer que les variables aléatoires sont indépendantes et identiquement distribuées (iid). Cette partie est traitée dans l'ouvrage intitulé « Loss Models : From data to decisions », dont les auteurs sont Panjer, Klugman et Willmot.
- ✚ Supposer qu'il y a dépendance entre les variables et que ceux-ci peuvent avoir des comportements différents. Cette hypothèse introduit la notion de copule qui est abordée dans l'ouvrage de Roger Nelsen intitulé « An introduction to copulas, 2<sup>nd</sup> édition ».

En pratique, il s'agira de déterminer la fonction de répartition de la variable  $S$  qui peut être obtenue de manière directe (par une formule fermée) ou indirecte (par simulation dans le cas où  $F_S(s)$  n'est pas explicite).

Plus loin, nous nous placerons dans un modèle individuel et nous présenterons les méthodes d'agrégation pour la détermination du capital agrégé selon les hypothèses énoncées mais avant cela, nous allons présenter brièvement les deux grandes approches d'agrégation du risque à savoir l'approche individuelle et l'approche collective.

### I.1. Modèle collectif

#### Notation

$X_{jl}$ : Le montant de la charge relatif au risque  $l$  pour la branche  $j$ .

$Y_{ij}$ : La charge individuelle pour la branche  $j$ .

$N(t)$ : Une variable aléatoire représentant le nombre de sinistre observé sur le portefeuille à la date  $t$ . Pour  $t$  fixé, on a :  $N(t) = N$

$K$ : Le nombre de ligne d'affaire que possède la compagnie.

Le modèle collectif peut être formulé comme suit :

$$X_{jl}(t) = \sum_{j=1}^K \sum_{i=1}^{N(t)} Y_{ij} \xrightarrow{t \text{ fixé}} X_{jl} = \sum_{j=1}^K \sum_{i=1}^N Y_{ij}$$

#### Interprétation de cette formule

En prenant comme hypothèse, le risque de sinistralité lié à la branche  $j$ , cette expression suppose que la détermination de la distribution de la charge de sinistre doit tenir compte du nombre de sinistre et du coût observé.

Pour pouvoir utiliser plus facilement ce modèle, deux hypothèses (très fortes) sont faites à savoir :

- ◆ Indépendance entre la fréquence et le coût des sinistres c'est à dire que les variables aléatoires  $N$  et  $Y_{ij}$  sont supposées indépendantes.
- ◆ Indépendance et stationnarité des montants de sinistres c'est-à-dire que les  $Y_{ij}$  sont indépendantes et surtout ne varient pas avec le temps.

En pratique, l'approche par simulation est privilégiée et surtout la simulation de Monte Carlo. Abstraction faite de la branche et du risque, considérons le modèle collectif simplifié suivant :

$$X = \sum_{i=1}^N Y_i, N \text{ v. a}$$

Notre but sera de déterminer la distribution de  $X$  par simulation c'est-à-dire sous l'hypothèse qu'il n'existe pas d'expression analytique de la distribution de  $X$ .

En revanche, nous supposons connu la distribution des variables aléatoires  $Y_i$  et  $N$ . A partir de cette hypothèse, nous allons résumer ce problème par les étapes ci-dessous :

**Etape 1**

Simuler  $m$  valeurs entières  $N_1, \dots, N_m$  selon la loi de  $N$

**Etape 2**

Simuler  $N_i$  valeurs selon la loi de  $Y_i$  et calculer  $\forall i, X_i = \sum_{i=1}^{N_i} Y_i$

**Etape 3**

A partir des  $m$  valeurs  $X_1, \dots, X_m$ , déterminer la distribution de la variable aléatoire  $X$ .

**NB :**

Je rappelle que mon but est d'agréger les risques provenant de plusieurs lignes d'affaires. En considérant le risque de sinistralité comme seul risque, il s'agira de déterminer la distribution de la charge de sinistre pour chaque ligne d'affaire à partir d'un **modèle collectif** et de tout simplement utiliser le **modèle individuel** pour obtenir la charge totale.

**Remarque**

Ces deux hypothèses peuvent impliquer plusieurs commentaires. En effet, il est évident qu'avec l'inflation, par exemple, les coûts de sinistres sont fortement impactés sur une durée longue. L'hypothèse de stationnarité peut tout de même s'appliquer avec actualisation des montants, mais encore faut-il trouver le taux d'actualisation adéquat. De même, l'hypothèse d'indépendance des coûts de sinistres est parfois très forte : il n'est pas rare qu'une jurisprudence motive le montant des indemnisations des affaires futures. La première hypothèse n'est quant à elle vérifiée que si le portefeuille étudié est homogène.

**I.2. Modèle individuel**

La différence entre le modèle individuel et le modèle collectif réside au niveau du nombre de sinistre. En effet dans le modèle individuel, le nombre de sinistre  $N$  est connu et fixe. Il se définit pour le risque de sinistralité  $l$  comme suit :

$$Z = \sum_{j=1}^K X_{jl}, K \text{ un entier}$$

## Chapitre 5 : Capital agrégé pour des lignes d'affaires

Dans notre cas, nous allons utiliser ce modèle pour déterminer la distribution de la charge totale des lignes d'affaire en utilisant les copules qui matérialisent la structure de dépendance entre les  $X_{jl}$ .

En considérant le modèle individuel simplifié ci-dessous :

$$X = \sum_{i=1}^n Y_i, n \text{ entier}$$

La simulation est simpliste et se résume par les étapes suivantes :

### **Etape 1**

Soit  $M$  le nombre d'itération, simuler  $M$  valeurs selon la loi de  $Y_i, i = 1, \dots, n$

### **Etape 2**

Calculer pour l'itération  $j$ ,  $X_j = \sum_{i=1}^n Y_i, j = 1, \dots, M$ . A partir des  $M$  valeurs  $X_1, \dots, X_M$ , déterminer la distribution de la variable  $X$ .

## **II. Présentation du modèle**

Dans cette section, nous présenterons les modèles d'agrégation de risque et de détermination du capital agrégé. Ce dernier n'est rien autre qu'une mesure de risque appliqué à la variable stochastique égale à la somme des variables aléatoires individuelles. Pour atteindre cet objectif, nous nous placerons dans un modèle individuel présenté ci-dessus.

### **II.1. Hypothèse gaussienne**

L'un des thèmes central soulevé par le dispositif Solvabilité 2 est la prise en compte de la dépendance entre les risques dans le calcul du capital requis.

En effet, ce dispositif préconise aux sociétés d'assurance de disposer de ressources suffisantes pour ne pas être en ruine à un horizon de un an avec une très forte probabilité (99.5%). Ceci sous-entend que l'entreprise doit être capable de contrôler le risque global qu'elle supporte tout le long de l'année avec une très forte probabilité.

Cette exigence de capital peut être déterminée à partir :

- D'un modèle interne :

- D'une approche standard

En ce qui concerne l'approche standard, la CEIOPS a opté pour une approche modulaire. Elle consiste, pour chaque risque identifié, à déterminer les besoins en capitaux nécessaires pour se prémunir de chacun de ces risques pris isolément selon le critère précédemment évoqué, puis en l'agrégation des capitaux ainsi obtenus de manière à refléter l'effet de diversification et les éventuelles dépendances entre ces risques.

La méthode d'agrégation retenue dans ce dispositif standard repose sur le modèle gaussien.

En effet, si :

$$\begin{pmatrix} X \\ Y \end{pmatrix} \approx N \left( \begin{pmatrix} m_X \\ m_Y \end{pmatrix}, \begin{bmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{bmatrix} \right), \text{ alors}$$

$X + Y \approx N(m_X + m_Y, \sigma_X^2 + \sigma_Y^2 + 2\rho\sigma_X\sigma_Y)$ . On déduit la valeur de la Value at Risk égale à :

$$Var_\alpha(X + Y) = m_X + m_Y + \Phi^{-1}(1 - \alpha) \sqrt{\sigma_X^2 + \sigma_Y^2 + 2\rho\sigma_X\sigma_Y}$$

Pour un risque particulier X, le capital de Solvabilité est donné par :

$$SCR_\alpha(X) = Var_\alpha(X) - m_X = \Phi^{-1}(1 - \alpha) \sigma_X$$

Nous déduisons pour deux risques, la formule d'agrégation ci-dessous :

$$SCR_\alpha(X + Y) = \sqrt{SCR_\alpha^2(X) + SCR_\alpha^2(Y) + 2\rho SCR_\alpha(X) SCR_\alpha(Y)},$$

Dont la généralisation pour n risques se définit comme suit :

$$SCR_\alpha \left( \sum_{i=1}^n X_i \right) = \sqrt{\sum_{i=1}^n SCR_\alpha^2(X_i) + 2 \sum_{i \neq j}^n \sum_{j=1}^n \rho_{ij} SCR_\alpha(X_i) SCR_\alpha(X_j)}$$

## **II.2. Hypothèse gaussienne et prise en compte de la dépendance**

Le fil conducteur de cette partie est de calculer le SCR tout en se mettant dans un cadre gaussien selon le dispositif de Solvabilité 2 et en tenant compte de la corrélation linéaire ou non linéaire qui existe entre les risques.

Pratiquement, il s'agira de :

- ◆ Calculer le  $SCR_\alpha(X_i)$  pour tous les risques  $i, i = 1, \dots, n$
- ◆ Calculer selon une approche non paramétrique  $\rho_{ij}, \rho_{ij}^S, \rho_{ij}^K$ , les coefficients de corrélations respectifs de Pearson, Spearman et Kendall pour les couples de risques  $(i, j)$ .

## Chapitre 5 : Capital agrégé pour des lignes d'affaires

- ◆ Calculer  $\widetilde{\rho}_{ij} = \text{Max}(\rho_{ij}, \rho_{ij}^S, \rho_{ij}^K) \forall$  couple de risque  $(i, j)$
- ◆ Calculer le capital agrégé selon la méthode d'agrégation proposé par Solvabilité 2

Pour un portefeuille frappé par n risques, nous obtenons le capital agrégé suivant :

$$SCR_{\alpha} \left( \sum_{i=1}^n X_i \right) = \sqrt{\sum_{i=1}^n SCR_{\alpha}^2(X_i) + 2 \sum_{i \neq j} \sum_{j=1}^n \widetilde{\rho}_{ij} SCR_{\alpha}(X_i) SCR_{\alpha}(X_j)}$$

### II.3. Utilisation des copules pour agréger les risques.

Soit n risques  $X_1, X_2, \dots$  et  $X_n$  de fonctions de répartition  $F_1, F_2, \dots$  et  $F_n$ , C une copule unique et la variable aléatoire agrégé  $Z = \sum_{i=1}^n X_i$ .

L'idée d'utiliser les copules pour agréger les risques est venue à partir du théorème de Sklar.

Dans ce théorème, la fonction de répartition du couple joint  $(X_1, X_2)$  est :

$$F_{X_1, X_2, \dots, X_n}(x_1, x_2, \dots, x_n) = C(F_1(x_1), F_2(x_2), \dots, F_n(x_n)).$$

Sous l'approche par corrélation, une mesure de risque (par exemple Var) est calculée pour chaque risque et le capital total est obtenu en sommant le capital requis pour chaque risque en tenant compte de la corrélation qui existe entre eux.

Cependant il peut arriver que le coefficient de corrélation soit très élevé, alors que l'une des variables ( $X_1$  ou  $\dots$  ou  $X_n$  par exemple) possède des réalisations négatives. Les copules résolvent ce problème en ce sens que la fonction Copula est plane et peut être spécifiée telle que les interactions entre  $X_1, X_2, \dots$  et  $X_n$  diffèrent dans différentes parties de chacune des distributions de  $X_1, X_2, \dots$  et  $X_n$ .

En réalité, agréger un risque, c'est en fait déterminer une distribution de la somme de tous les risques. Dans notre cas, Il s'agira de déterminer la fonction de répartition et la densité de la variable aléatoire Z.

Le problème qui se pose est dans la détermination de la distribution de Z. En effet nous rappelons les cas peuvent se poser :

- Soit il est possible de déterminer une expression explicite de la fonction de répartition de Z. Dans ce cas, la distribution se déduit automatiquement et l'on peut avoir une idée du comportement de l'ensemble des risques.

## Chapitre 5 : Capital agrégé pour des lignes d'affaires

- Soit il est impossible d'obtenir une expression explicite de la fonction de répartition de  $Z$ . Dans ce cas, une alternative s'impose à nous qui est la simulation. Cette dernière est un outil largement utilisée par les actuaires pour déterminer la distribution d'une **variable stochastique** et pour obtenir des résultats provenant d'un modèle complexe. Il s'agira de simuler pour un nombre  $N$  d'itérations chaque marginale de la variable  $Z$  et de faire la somme de celle-ci pour avoir une valeur de  $Z$  correspondant à l'itération en question.

Dans l'hypothèse d'une expression non explicite de la fonction de répartition de  $Z$ , les copules interviennent pour résoudre ce problème tout en tenant en plus de la structure de dépendance qui existe entre les risques.

En pratique les variables stochastiques  $X_1, X_2, \dots, X_n$  sont obtenues par simulation à partir d'une Copula  $C$ .

En effet, ces variables sont transformées en des valeurs uniformes sur  $[0,1]$  en appliquant les fonctions de répartitions respectives  $F_1, F_2, \dots$  et  $F_n$ .

Ensuite, il s'agira de déterminer une Copule  $C$  qui s'ajuste à l'ensemble des données, c'est-à-dire une copula qui matérialise la structure de dépendance entre  $X_1$  et  $X_2$  et d'estimer les paramètres de cette copula. Le théorème de Sklar énoncé plus haut permet de déterminer la distribution jointe du couple  $(X_1, X_2, \dots, X_n)$ .

Enfin, effectuer des simulations pour déterminer les variables  $X_1, X_2, \dots, X_n$  et ensuite la distribution du risque agrégée en respectant les étapes ci-dessous :

### **Etape 1 :**

Une variable uniforme  $u_i, i = 1, \dots, n$  est simulée pour chaque distribution de risque  $i$

### **Etape 2 :**

La fonction de répartition  $F_i$  est ensuite calculée à chaque point  $u_i$ , soit  $F_i(u_i)$

### **Etape 3 :**

Utiliser la fonction Copula pour calculer  $t = C(F_1(u_1), F_2(u_2), \dots, F_n(u_n))$

### **Etape 4 :**

Enfin, utiliser la fonction cumulative inverse  $F_i^{-1}$  pour calculer la charge de sinistre approximative  $\widehat{X}_i$ , soit  $\widehat{X}_i = F_i^{-1}(t)$ .

Pour cette première simulation, la valeur  $Z_1$  de  $Z$  est approximé par  $\widehat{Z}_1 = \sum_{i=1}^n \widehat{X}_i$ .

### **Etape 5 :**

Répéter les étapes 1 à 4, un nombre  $N$  fois,  $N$  étant le nombre de simulation.

## Chapitre 5 : Capital agrégé pour des lignes d'affaires

Dans ce cas, nous obtiendrons N charges agrégées  $\widehat{Z}_1, \widehat{Z}_2, \dots, \widehat{Z}_N$ .

A partir de ces charges, nous pouvons déterminer la distribution empirique de la variable Z et déduire le capital requis en appliquant une mesure de risque à celle-ci.

### III. Application sur trois branches d'activités d'un assureur non vie

Dans cette section, nous allons calculer un capital agrégé pour l'ensemble des trois branches B, C et D sous l'hypothèse que les charges de sinistres de ces dernières suivent un comportement normal.

Le modèle

**Etape 1 : Détermination des  $Y_{i,j,l}$  pour la branche l**

$$Y_{i,j,l} = g^{-1} \left( \sum_{k=1}^2 \beta_k^{moy} x_k \right) + \sigma^{moy}(Y_{i,j,l}) \times \Phi^{-1}(U), \quad \forall U \sim \mathcal{C}(u_B, u_C, u_D)$$

**Etape 2 :** Calcul des  $Z_{i,j} = Y_{i,j,B} + Y_{i,j,C} + Y_{i,j,D}$  et déduire  $Z_i = \sum_{j=1}^{10} Z_{i,j}$  pour l'année de survenance  $i$ .

**Etape 3 :** Déterminer la distribution des  $Z_i$  et déduire le quantile d'ordre  $\alpha$  de cette dernière si l'on considère la Value at Risk comme une mesure de risque de cette variable aléatoire.

**Résultats**

**Valeurs des  $Z_i$**

Année de survenance	Valeurs agrégées pour les trois branches $Z_i$
1	148539
2	156984
3	166403
4	176444
5	187172
6	198857
7	211467
8	224946
9	239358
10	254763

**Tableau 25 : Résultats des capitaux agrégés par année de survenance**

**Valeur de la Value at Risque à 99 %**

$$Z_i \sim \text{Gamma}(33.71422, 5828.203)$$

Et

$$\text{Var}_{0.99}(Z_i) = EC = 283666 \text{ MAD}$$

*Chapitre 6 : Problème d'allocation du capital agrégé*

*1. Introduction*

Le marché de l'assurance au Maroc ou dans le monde est régulé par des lois. La législation contrôle les intérêts des assurés, établit les règles pour déterminer le capital que doit détenir les compagnies. Le niveau de ce capital est déterminé de sorte que la compagnie soit capable de respecter ses engagements financiers avec une très forte probabilité quelque soit les événements aléatoires qui peuvent survenir.

Ainsi, Le calcul de ce capital est très important et intéresse les actionnaires et les investisseurs qui veulent avoir une idée sur les risques auquel sont soumis le capital et le retour sur investissement généré par celui-ci.

Enterprise Risk Management (ERM) est le cadre par excellence pour identifier, mesurer, tarifier et contrôler les risques. Ceci étant, il est donc important pour les compagnies d'accorder une attention particulière à l'ERM.

Une importante composante du cadre ERM est l'exercice de l'allocation de capital, un terme qui se réfère à la subdivision du capital agrégé détenu par la firme entre les lignes d'affaires, le type d'exposition au risque, la localisation territoriale, les produits ou groupes de produits, les entités ou unités organisationnelles.

La plupart des compagnies d'assurance ou organismes financiers possèdent plusieurs lignes d'affaires et trouve un besoin croissant d'allouer un capital à chacune d'elles pour plusieurs raisons :

- Premièrement, la nécessité de redistribuer le capital détenu en fonction des coûts observés entre les lignes d'affaires. En outre, les coûts sont transférés de manière équitable sur les assurés (dans le cas de l'assurance) sous forme de charge.
- Secondairement, l'allocation des frais entre les lignes d'affaire est une activité nécessaire pour le reporting financier.

Enfin, l'allocation de capital fournit un bon dispositif d'estimation et de comparaison de la performance entre les différentes lignes d'affaire déterminé par un indicateur de performance appelé en anglais 'return on allocated capital' pour chacune des lignes. La comparaison de ces lignes via cet indicateur de performance permet de déterminer la ligne la plus profitable pour la compagnie et permet par exemple de rémunérer le meilleur manager ou

## Chapitre 6 : Problème d'allocation du capital agrégé

de prendre des décisions stratégiques concernant celle-ci (expansion ou réduction ou élimination).

Il existe un nombre incalculable de manière d'allouer le capital agrégé entre les différentes unités d'affaire. Cependant il peut exister des dépendances mutuelles entre les performances des différentes unités d'affaires ce qui rend un peu complexe cet exercice.

Nous allons dans la suite de cette partie, exposer la plus part des célèbres méthodes d'allocation de capital rencontrées dans la littérature.

### Problème d'allocation

Considérons un portefeuille constitué de  $n$  pertes individuelles  $X_{1jt}, X_{2jt}, \dots, X_{njt}$  à la date  $t$ , avec  $t = 1, \dots, T$  et pour la ligne d'affaire  $j$ ; avec  $j = 1, \dots, N$

Supposons que :

- $(X_{1jt}, X_{2jt}, \dots, X_{njt})$  est un vecteur aléatoire sur l'espace de probabilité  $(\Omega, \mathfrak{F}, P)$
- Chaque  $X_{ijt}$ ,  $i = 1, \dots, n$ ;  $t = 1, \dots, T$ ;  $j = 1, \dots, N$  a une moyenne finie
- La fonction de répartition  $P(X_{it} \leq x)$  est noté par  $F_{X_{it}}(x)$

Ainsi la somme agrégée de tous les risques est définie par :

$$S = \sum_{t=1}^T \sum_{j=1}^N \sum_{i=1}^n X_{ijt}$$

Que l'on peut simplifier par :

$$S = \sum_{j=1}^N \sum_{i=1}^n X_{ij.} \xrightarrow{\text{Simplification}} S = \sum_{i=1}^n X_{i..}$$

Ou  $S$  est la charge agrégée qui peut être interprétée comme :

- La perte totale d'une entreprise, par exemple une société d'assurance dans lesquelles les pertes individuelles correspondent aux pertes des unités d'affaires.
- La perte d'un portefeuille d'assurance dans lesquelles les pertes individuelles proviennent des différentes polices

Pour alléger les notations, nous noterons  $X_i, K_i, K$  respectivement la perte, le capital de risque de l'unité  $i$  et le capital total requis tel que :

$$\sum_{i=1}^n K_i = K$$

## Chapitre 6 : Problème d'allocation du capital agrégé

Rappelons que le but de l'allocation de capital est de classer les unités d'affaire par niveau de rentabilité.

Dans cette optique, le bénéfice de l'allocation de capital se définit comme suit :

$$R_i = \left( \frac{K_i - X_i}{K_i} \right) = 1 - \frac{X_i}{K_i}, \quad \forall i = 1, \dots, n$$

### II. Modèles d'allocation de capital

#### II.1. Rappels

Pour un niveau de probabilité  $p \in [0,1]$ , nous notons :

- La Value At Risk (VAR) ou quantile de la perte ou de la variable aléatoire  $X$  par  $F_X^{-1}(p)$ . On a :

$$F_X^{-1}(p) = \inf\{x \in \mathbb{R} / F_X(x) \geq p\}$$

Avec  $\inf\{\emptyset\} = +\infty$  par convention.

- La fonction de distribution inverse par  $F_X^{-1+}(p)$  telle que :

$$F_X^{-1+}(p) = \sup\{x \in \mathbb{R} / F_X(x) \leq p\}$$

Avec  $\sup\{\emptyset\} = -\infty$ .

- La fonction de distribution inverse composée d'ordre  $\alpha$  par  $F_X^{-1(\alpha)}(p)$  telle qu'on a :

$$F_X^{-1(\alpha)}(p) = \alpha F_X^{-1}(p) + (1 - \alpha) F_X^{-1+}(p), \forall \alpha \in [0,1]$$

De cette définition, nous déduisons que :

$$\forall X, x \text{ tel que } 0 \leq F_X(x) \leq 1, \exists \alpha_x \in [0,1] / F_X^{-1(\alpha_x)}(F_X(x)) = x.$$

#### II.2. Modèle d'allocation « Haircut »

Une méthode simple d'allocation de capital consiste à attribuer à chaque unité  $i$ , un capital  $K_i = \gamma F_{X_i}^{-1}(p)$ ,  $i = 1, \dots, n$  à chaque ligne d'affaire  $i$  ou le facteur  $\gamma$  est choisi de sorte que l'équation  $\sum_{i=1}^n K_i = K$  soit satisfaite (Full allocation).

## Chapitre 6 : Problème d'allocation du capital agrégé

En choisissant  $\gamma = \frac{K}{\sum_{j=1}^n F_{X_j}^{-1}(p)}$ , nous obtenons le modèle de « Haircut » ci-dessous :

$$K_i = \frac{K}{\sum_{j=1}^n F_{X_j}^{-1}(p)} \times F_{X_i}^{-1}(p), i = 1, \dots, n$$

Nous vérifions que  $\sum_{i=1}^n K_i = \frac{\sum_{i=1}^n F_{X_i}^{-1}(p)}{\sum_{j=1}^n F_{X_j}^{-1}(p)} \times K = 1 \times K = K = F_S^{-1}(p)$

Cette méthode n'est donc pas influencée par la structure de dépendance existante entre les variables aléatoires  $X_i$  des différentes unités d'affaires. Dans ce cas, nous pouvons dire que cette approche est indépendante du contexte dans lequel nous nous trouvons.

### II.3. Modèle d'allocation quantile

Le modèle d'allocation « Haircut » alloue à chaque unité d'affaire  $i$ , une proportion  $\gamma$  de ces  $p$ -quantiles,  $\gamma$  étant choisi de sorte la propriété « full allocation » soit respectée. Cela signifie qu'une réduction ou augmentation proportionnelle constante est appliquée sur chacune des quantiles  $F_{X_i}^{-1}(p)$ .

Cependant le modèle d'allocation quantile s'appuie sur la fonction de distribution inverse composée d'ordre  $\alpha$ . Nous avons :

$$K_i = F_{X_i}^{-1(\alpha)}(\beta p)$$

Avec  $\alpha$  et  $\beta$  choisi tel que  $\sum_{i=1}^n K_i = K$ .

### II.4. Modèle d'allocation covariance

Ce modèle a été proposé par Overbeck (2000) et est donné par :

$$K_i = \frac{K}{Var(S)} \times Cov[X_i, S], i = 1, \dots, n$$

Avec  $cov[X_i, S]$ , la covariance entre la charge individuelle et la charge agrégée,  $Var(S)$ , la variance de la charge agrégée.

Comme  $\sum_{i=1}^n K_i = \frac{Cov[\sum_{i=1}^n X_i, S]}{Var(S)} \times K = \frac{Var(S)}{Var(S)} \times K = K$ , la propriété « full allocation » est donc vérifiée.

### II.5. Modèle d'allocation CTE

Pour un niveau de probabilité  $p \in [0,1]$  donné, la Conditional Tail Expectation (CTE) se définit comme suit :

$$CTE_p(S) = E(S / S > F_S^{-1}(p))$$

En général, la CTE est une mesure de risque qui ne satisfait pas nécessairement les propriétés de sub-additivité. Cependant elle est connue comme étant une mesure de risque cohérente dans le cas où nous nous limitons aux variables aléatoires possédant des fonctions de répartitions continues.

Le capital alloué à la ligne  $i$  est donné par :

$$K_i = \frac{K}{CTE_p(S)} \times E(X_i / S > F_S^{-1}(p)); i = 1, \dots, n$$

Pour le cas particulier où  $K = CTE_p(S)$ , on a  $K_i = E(X_i / S > F_S^{-1}(p)) = CTE_p(X_i)$ ,

Nous retombons sur le résultat suggéré par **Overbeck (2000) et Denault (2001)**.

### II.6. Modèle d'allocation proportionnel

Cette méthode d'allocation a été discutée plus haut dans des cas spéciaux. Chaque cas a été obtenu en choisissant une mesure de risque  $\rho$  et ensuite en calculant le capital  $K_i = \gamma \rho(X_i)$  pour chaque unité  $i$  dont  $\gamma$  est choisi de sorte que la propriété « full allocation » soit satisfaite. Ainsi,

$$K_i = \frac{K}{\sum_{j=1}^n \rho(X_j)} \times \rho(X_i), i = 1, \dots, n$$

En clair, tous les principes d'allocation discutés dans les sections précédentes proviennent de cette équation en choisissant de manière optimale la mesure du risque  $\rho$ . Ainsi,

$$\text{Haircut allocation} \xrightarrow{\text{correspond}} \rho(X_i) = F_{X_i}^{-1}(p)$$

$$\text{Quantile allocation} \xrightarrow{\text{correspond}} \rho(X_i) = F_{X_i}^{-1}(F_S^c[K])$$

$$\text{Covariance allocation} \xrightarrow{\text{correspond}} \rho(X_i) = \text{cov}[X_i, S]$$

$$\text{TVAR allocation} \xrightarrow{\text{correspond}} \rho(X_i) = E(X_i / S > F_S^{-1}(p))$$

## *Conclusion*

### Conclusion

Ce mémoire s'inscrit dans la vision de Solvabilité 2 qui pousse les assureurs à avoir une autre approche du risque. Nous avons tenté de résoudre ces risques en traitant quelques problématiques des sciences actuarielles. Un accent particulier a été mis sur le calcul des provisions vu leur poids au niveau du passif de l'assureur. Dans le but de converger vers le best estimate des provisions techniques, une panoplie de méthodes déterministes et stochastiques ont été utilisées. Finalement c'est la méthode GAM Mixed que j'ai retenue pour calculer les provisions « best estimate » car elle fournit le meilleur modèle (AIC plus petit) et elle tient aussi compte de l'autocorrélation entre les erreurs dans la prédiction des charges futures. Après la problématique des provisions techniques, celle des ratios de sinistralité fut abordée. Cependant nous avons supprimé certaines étapes et nous avons exposé la dernière étape incluant simultanément la théorie des copules et la théorie bayésienne. Ensuite le troisième problème des sciences actuarielles fut le calcul du capital agrégé pour plusieurs lignes d'affaire. La nouveauté réside dans la prise en compte de la dépendance entre les lignes d'affaires dans le calcul du capital agrégé.

Enfin le dernier problème abordé fut le problème de l'allocation du capital agrégé entre les lignes d'activité. Il serait aussi intéressant d'aborder la notion de mesure de performance ce qui n'a pas été fait dans ce mémoire par manque de temps.

## *Bibliographie*

### Bibliographie

#### ❖ Articles

Lheureux Elise (2006). Une méthode alternative de provisionnement stochastique en Assurance Non Vie : Les Modèles Additifs Généralisés.

Paul Embrechts, Alexander McNeil (2001). Modelling Dependence with Copulas and Applications to Risk Management

Vito Ricci (2005), FITTING DISTRIBUTIONS WITH R

Ivan Kojadinovic, Jun Yan (2010), Modeling Multivariate Distributions with Continuous Margins Using the copula R Package

Jun Yan (2007), Enjoy the Joy of Copulas: With a Package copula

*Mario V. Wuthrich (2008), Using a Bayesian Approach for Claims Reserving*

#### ❖ Cours

**Frédéric PLANCHET (2010).** Dépendance stochastique : Introduction à la théorie des copules, cours ISFA

#### ❖ Mémoires

Marc BAGARRY (2004). Solvency 2, un modèle interne non vie utilisant des copulas elliptiques de dimension n pour évaluer le capital économique.

Cadoux et Loizeau (2006). Dépendances au sein d'un portefeuille d'assurance non vie : cas pratique d'application de la théorie des copules à la détermination du besoin en fonds propres.

#### ❖ Livres

Simon N. Wood (2006). Generalized Additive Models: an introduction with R

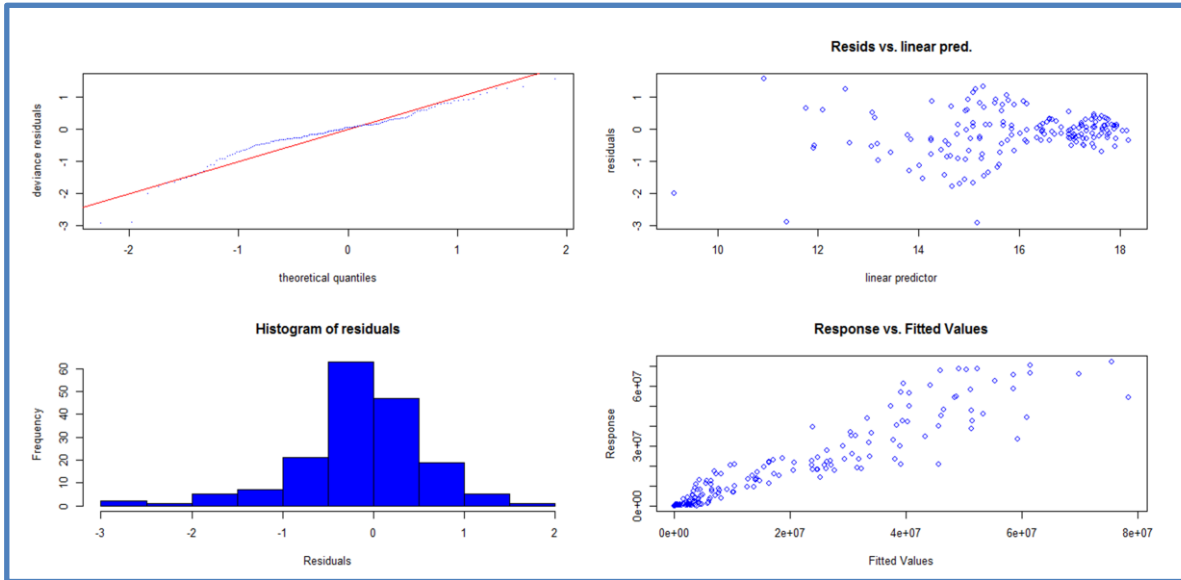
Hadley Wickham (2009), ggplot2: Elegant Graphics for Data Analysis

Julian J. Faraway (2006), Extending the Linear Model with R: Generalized Linear, Mixed Effects and Nonparametric Regression Models

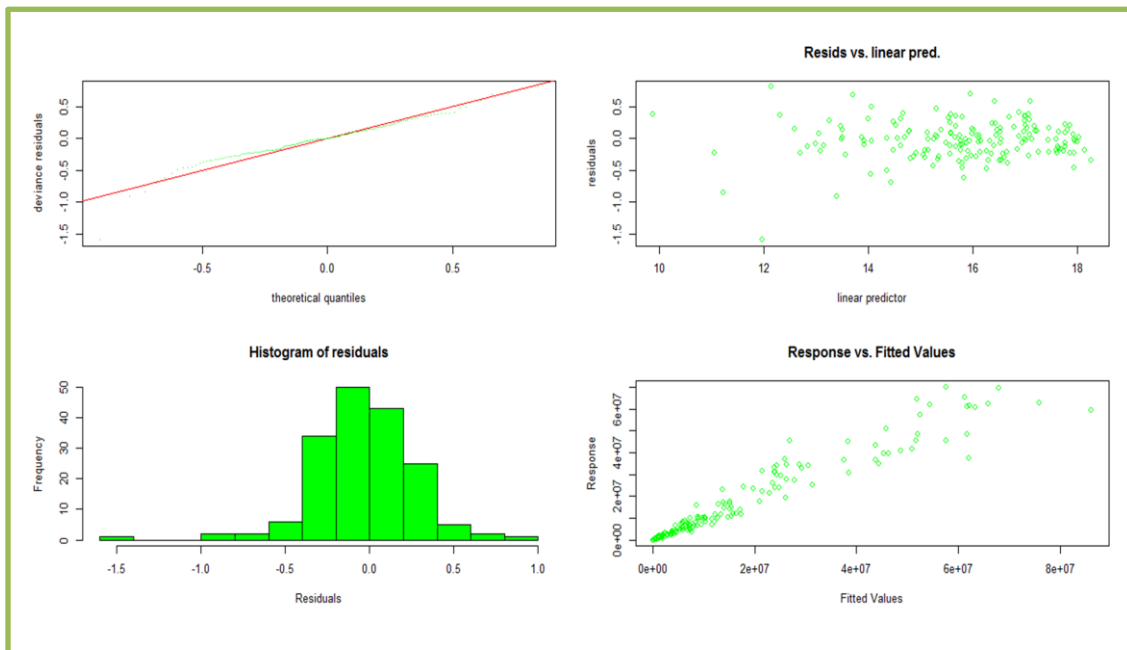
Annexes

**Annexe I: Output du modèle GAM pour la loi de Gamma**

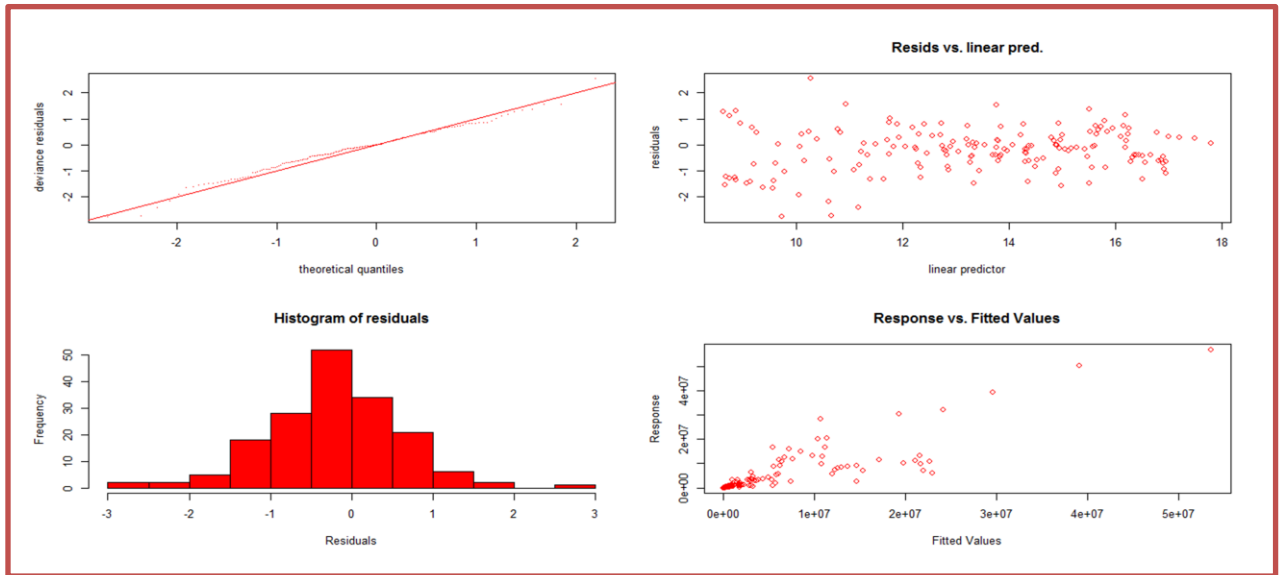
**Branche Accident de travail**



**Branche Automobile Corporel**

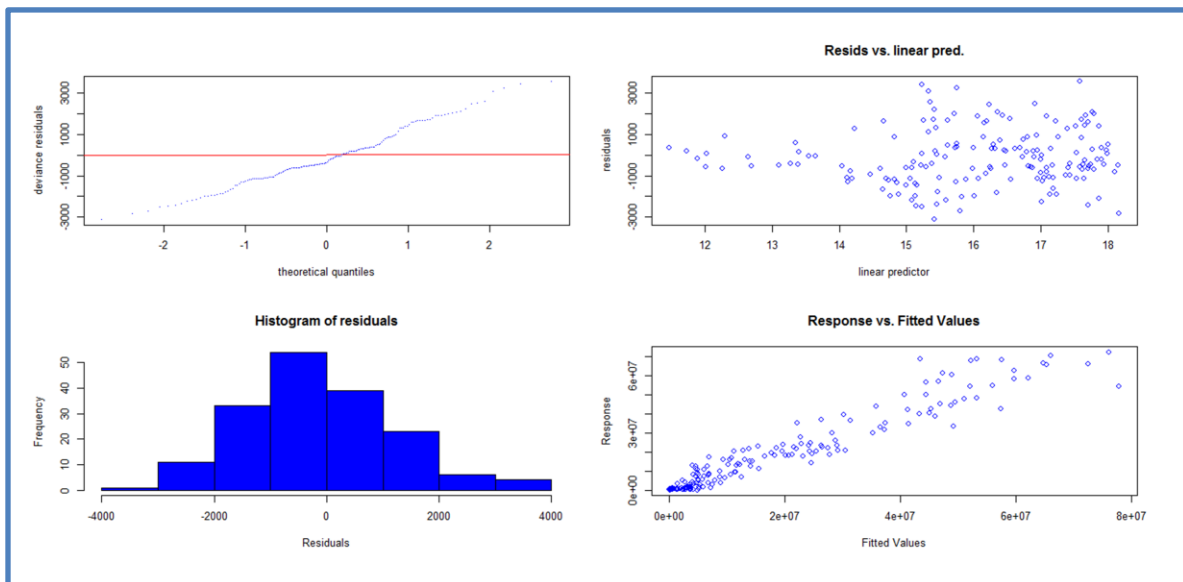


**Branche Automobile Matériel**

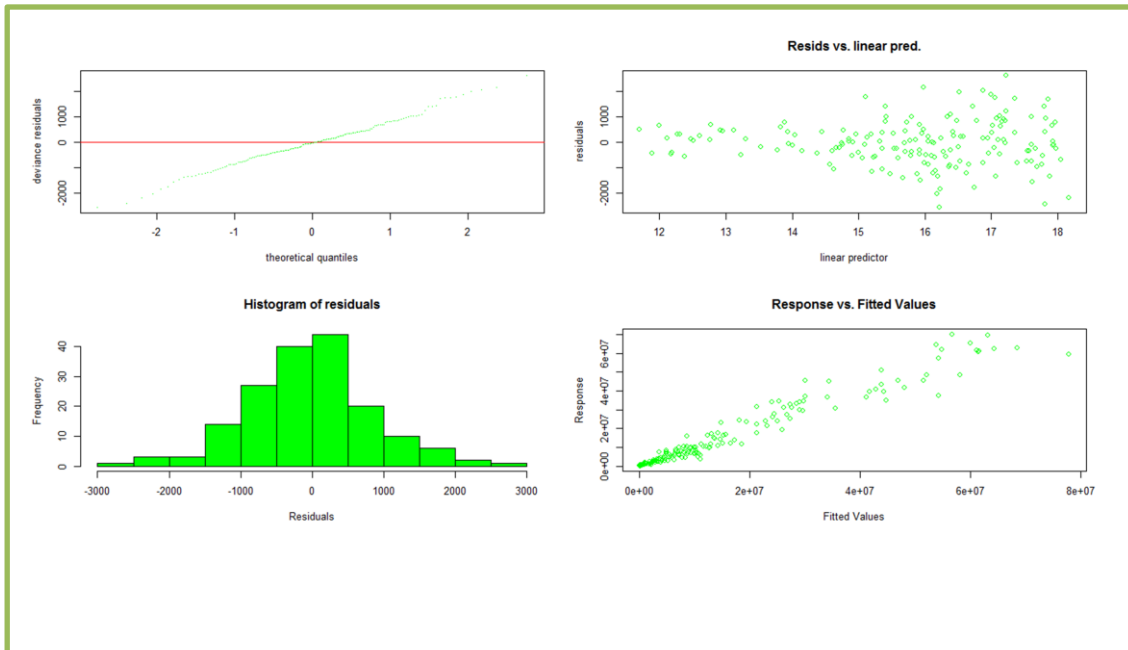


**Annexe II : Output du modèle GAM pour la loi de poisson**

**Branche Accident de travail**



**Branche Automobile Corporel**



**Branche Automobile Matériel**

