

ESTIMATING LOSS RESERVES IN CAR INSURANCE USING COPULA THEORY

ESTIMATION DES PROVISIONS TECHNIQUES EN ASSURANCE AUTOMOBILE SELON LA THEORIE DES COPULES

* **Lamia CHOUCHAOUI**

Ecole Nationale Supérieure de Statistique et d'Economie Appliquée ENSSEA
Laboratoire de Modélisation des Processus Stochastiques LAMOPS
anya.lamia@hotmail.com

Ghania SAIDI

Ecole Nationale Supérieure de Statistique et d'Economie Appliquée ENSSEA
Laboratoire de Modélisation des Processus Stochastiques LAMOPS
ghsaidi@yahoo.fr

Reçu le : 11/03/2018 **Accepté le :** 22/09/2018 **Publication en ligne le :** 31/12/2018

ABSTRACT: Outstanding claims reserves in general insurance are a type of technical reserve or accounting provision in the financial statements of an insurer. Classical methods in claims reserving can be used to evaluate technical reserves like Chain Ladder Method, GLM model. These methods, however, have been shown to suffer from several drawbacks. For example they give no estimate of the reserve distributions tail which is of great importance for risk calculation. In this paper, an alternative method based on dependence theory by using Archimedean copulas (Joe Copula), is proposed to estimate the reserves. The copula estimated claims reserves will be compared to the evaluation by the generalized linear model on a real data set from a public insurance company having significant car portfolio.

Keywords : provision, GLM model, dependence, correlation coefficient, copula, simulation.

JEL Classification : C13, C14, C15,

RESUME : La provision pour sinistres à payer (PSAP), constitue la majeure partie des provisions techniques en assurance non-vie, elle correspond à la valeur estimative des dépenses nécessaires au règlement de l'ensemble des sinistres survenus non payés. Pour permettre l'évaluation actuarielle de ces provisions, plusieurs méthodes classiques basées sur l'hypothèse d'indépendance entre les branches ou les garanties peuvent être utilisées, elles s'appuient généralement sur l'étude des triangles de liquidation. : Le monde de l'assurance a longtemps été basé sur cette notion d'indépendance. Cependant, des phénomènes de grande ampleur sont venus remettre en question cette hypothèse. Les dépendances entre les risques en assurance non-vie peuvent être une des causes de la sous-tarification ou de provisions insuffisantes pour couvrir la charge de sinistre afférent à une année de survenance. Dans cet article, on s'intéresse à l'étude des dépendances entre deux sous-branches de l'assurance automobile ainsi que leurs implications en introduisant les

* *Auteur Correspondant*

copules en particulier les copules archimédiennes (Joe copula) qui permettent de décrire ces dépendances. Les conséquences de ces dépendances en termes de provisionnement pour sinistres restant à payer sont mises en évidence à l'aide des simulations.

Mots clés : Provision, modèle GLM, dépendance, coefficient de corrélation, copule, simulation.

1. INTRODUCTION

Ces dernières années, le monde des assurances a connu des changements importants, la compétition s'est accrue, la sinistralité a augmenté avec une plus grande concentration des risques assurés. Pour cela, les assureurs sont fortement incités à développer une gestion optimale de leurs fonds propres qui doit satisfaire des intérêts divergents.

Le risque de provisionnement constitue l'un des principaux risques, auxquels est soumise la compagnie d'assurance. L'actuariat met en pratique plusieurs méthodes de calcul des provisions à savoir les méthodes déterministes (méthode Chain-Ladder, méthode London-Chain, méthode Bornhuetter-Ferguson) et les méthodes stochastiques (modèle de Mack, modèle linéaire généralisé GLM, méthode de Bootstrap). Chaque méthode donne un résultat possible, et le montant de ces provisions techniques ne demeure qu'une estimation dont la précision peut être plus ou moins grande.

Les techniques utilisées proposent une vision à long terme, c'est-à-dire, l'étude du développement des sinistres jusqu'aux sinistres ultimes, mais aujourd'hui, la nouvelle réglementation Solvabilité II impose également une vision à un an.

Afin de mieux maîtriser les risques et pour répondre à une réglementation européenne de plus en plus exigeante, l'assureur ne peut pas se contenter de modéliser individuellement les différentes branches, puis de sommer les besoins en fonds propres de chaque branche afin d'obtenir les besoins en capital de l'entreprise dans sa globalité, il est amené à prendre en compte les dépendances entre les branches d'assurance et même entre les sous-branches, et ce, de différentes manières. Parmi les travaux dans cette optique, On peut citer le travail de Freez and Valdez (2008)[†] et Géraldine Krauth[‡] (2007).

Dans un portefeuille d'assurance automobile, différents types de réclamations peuvent survenir pour chaque police en vigueur. En cas de collision entre deux véhicules, l'assuré peut déposer une réclamation pour dommages corporels et matériels à lui-même et à l'autrui. Traditionnellement, ces types de risques sont considérés comme indépendants afin d'en faciliter la modélisation stochastique. Dans la pratique, nous observons toutefois une dépendance entre les montants de ces réclamations dont il est important de tenir compte.

Dans ce contexte, notre étude porte sur l'estimation des provisions pour sinistres à payer en assurance automobile en considérant la dépendance entre les sinistres matériels et corporels. Cette dépendance sera illustrée à travers la théorie des copules, ensuite nous

[†] **Freez and Valdez**, « *Hierarchical Insurance Claims Modeling* », Journal of the American Statistical Association, 2008.

[‡] **Krauth G.**, « *Provisionnement et corrélation entre branches* », CNAM, Paris, 2007.

comparons les résultats obtenus avec ceux calculés avec le modèle linéaire généralisé GLM.

2. LES METHODES CLASSIQUES D'ESTIMATION DES PROVISIONS TECHNIQUES

On distingue deux méthodes d'estimation des provisions techniques à savoir les méthodes déterministes et les méthodes stochastiques.

2.1. Les méthodes déterministes

Elles permettent de mesurer l'incertitude associée à la prédiction du montant des provisions techniques sans aucune hypothèse probabiliste. Elles sont toutes caractérisées par une hypothèse commune « les cadences de développement sont constantes par année d'accident », donc les tendances sont identiques sur la période historique des données et sur les projections futures.

Les méthodes de provisionnement les plus classiques sont la méthode Chain-Ladder (CL), la méthode London-Chain (LC) et la méthode Bornhuetter-Ferguson (BF).

2.2. Les méthodes stochastiques

La finalité première des approches stochastiques est de mesurer l'incertitude présente dans le triangle de liquidation afin d'améliorer les résultats des méthodes déterministes.

Le but des méthodes stochastiques est d'explicitier les hypothèses utilisées dans les méthodes déterministes afin de les valider au moins partiellement, et obtenir des intervalles de confiances et des estimations de la variabilité de la provision prévue par les modèles déterministes.

Parmi les méthodes stochastiques, on distingue le modèle de Mack qui est une méthode complémentaire à la méthode de Chain-Ladder (elle estime la volatilité des provisions techniques alors que la méthode de Chain-Ladder estime leur espérance), le modèle linéaire généralisé GLM qui correspond à une généralisation du modèle linéaire classique du type « variable à expliquer /variables explicatives », et la méthode de Bootstrap qui consiste à fabriquer de l'information et à fournir des réponses là où les autres méthodes ne sont pas applicables. Le choix entre l'une ou l'autre de ces catégories de méthodes dépend des données disponibles, ainsi que de la probabilité d'interpréter aisément les résultats qui en découlent.

3. ETUDE DE LA DEPENDANCE DANS LE DOMAINE DE L'ASSURANCE

Les études de la dépendance n'ont pas cessé d'intéresser les statisticiens. Lorsque l'on évoque la dépendance, le premier mot qui vient généralement à l'esprit est celui du coefficient de corrélation linéaire. Ce dernier connu également sous le nom du coefficient de corrélation de Pearson, était et demeure en effet le plus couramment utilisé. Ce coefficient est tout à fait approprié lorsque nous étudions des distributions Normales ou de Student Multivariées, mais celui-ci perd de son intérêt si le modèle est différent.

Malheureusement, les distributions dans le domaine de l'assurance suivent très rarement de telles lois, car elles n'attribuent pas d'importance aux queues des distributions marginales qui peuvent présenter un très grand risque en cas de modélisation des risques.

Cependant, Une alternative de construction de fonctions de répartition bivariées et multivariées est basée sur les copules qui permettent un large choix de marginales et de types de dépendance. La modélisation de l'assurance qui en résulte est étonnamment simple et générale. Les lois marginales étant données, cette théorie montre qu'il suffit de les joindre par une copule ayant les propriétés souhaitées. La distribution résultante a pour marginales les distributions marginales d'origine.

3.1. Dépendance et corrélation

Les notions de dépendance et de corrélation sont souvent confondues; ce sont pourtant deux concepts différents qui sont toutefois associés :

- Si X et Y sont indépendants, alors X et Y sont non corrélés;
- Si X et Y sont non corrélés, cela n'implique pas que X et Y sont indépendants. Par exemple, considérons une variable symétrique X avec $Y=X^2$, Y dépend de X , mais pourtant $r(X, Y) = 0$.

Le coefficient de corrélation linéaire mesure le degré de dépendance entre deux variables, la présence de forte corrélation implique à son tour une dépendance en moyenne, mais c'est une mesure de dépendance particulière qui perd sa pertinence en dehors de l'univers gaussien[§].

3.2. Théorie des copules

Les concepts de copule et de structure de dépendance sont récents. Il s'agit d'un outil statistique de modélisation mis en avant pour la première fois par Abe Sklar en 1959, comme solution d'un problème de probabilité énoncé par Maurice Fréchet dans le cadre des espaces métriques aléatoires (travaux réalisés avec Berthold Schweizer). Ce concept est resté longtemps très peu utilisé en statistique; on peut toutefois citer les travaux de Kimeldorf et Sampson sur la dépendance (1975), Galambos (1978) ainsi que les recherches de Paul Deheuvels à la fin des années 70, qui lui donnèrent le nom de fonction de dépendance.

L'étude systématique des copules ont débuté au milieu des années 1980 avec Christian Genest et son équipe. L'article fondateur est celui de Genest et Mac Kay (1986). Depuis, de nombreux développements statistiques ont été menés par C. Genest et ses co-auteurs R.J. MacKay, L.P. Rivest, P. Capéraà, A.L. Fougères, K. Ghoudi, B. Rémillard. Les copules sont devenues un outil de base dans la modélisation des distributions multivariées dans les domaines de finance et des assurances.

3.2.1. Définition de la copule

Une copule est définie comme une fonction de répartition un peu particulière, qui lie les fonctions de répartition multivariées des lois de probabilité dans $\mathbb{R}^d$ pour $d \geq 2$, aux fonctions de répartition marginales de leurs coordonnées. La caractéristique des copules étant une fonction de répartition multivariée ayant des lois marginales uniformes sur $[0,1]$. Cette fonction dans le cas bivarié est définie de $[0,1]^2$ vers $[0,1]$ par la relation :

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

[§] Fréchet, Rapport pour la 22^e session de l'IIS à Londres, Bulletin de l'IIS, 1934.

vérifie les trois propriétés suivantes:

1. $C(u, v)$ est une fonction croissante pour chacune de ses composantes u et v ;
2. $C(u, 1) = u$ et $C(1, v) = v$;
3. Pour tout $a_1 \leq a_2$ et $b_1 \leq b_2$, nous avons :

$$C(a_2, b_2) - C(a_1, b_2) - C(a_2, b_1) + C(a_1, b_1) \geq 0.$$

Dans ce qui suit, nous énonçons le **théorème de Sklar** qui met en évidence comment il est possible de relier le principe de la copule à la fonction de répartition bivariable et aux lois marginales.

Théorème de Sklar :

Soit F une fonction de distribution en deux dimensions ayant des marginales F_X et F_Y . Alors, il existe une copule C telle que pour tout $(x, y) \in [-\infty, +\infty]^2$, nous avons :

$$F(x, y) = C(F_X(x), F_Y(y)).$$

Pour $(u, v) \in I^2$, nous avons :

$$C(u, v) = C(F_X(x), F_Y(y)) = F(F_X^{-1}(u), F_Y^{-1}(v)).$$

Cette dernière formule est un point de départ pour construire une copule bivariable à partir d'une fonction de distribution conjointe et des lois marginales.

3.2.2. Familles de copules

Il existe une vaste gamme de familles de copules dans lesquelles on peut sélectionner un modèle approprié, selon les variables aléatoires des données que nous essayons de modéliser.

3.2.2.1. Copules archimédiennes

Une classe particulièrement importante de copule est celle des copules archimédiennes, définies par :

$$C(u_1, u_2, \dots, u_d) = \varphi^{-1}(\varphi(u_1) + \varphi(u_2) + \dots + \varphi(u_d)),$$

Où φ (appelé générateur de la copule archimédienne) est au moins d-2 fois continuellement dérivable, dont la dérivée d'ordre d-2 est décroissante convexe, et telle que $\varphi(1) = 0$.

Elles ont le grand avantage de décrire des structures de dépendance très diverses dont notamment les dépendances dites asymétriques où les coefficients de queue inférieure et de queue supérieure sont différents.

Cette notion de copule, définie par GENEST et MACKAY (1986), regroupe un certain nombre de copules dont on présente quelques exemples dans le cas bivarié dans le Table suivant :

Table N°1 : Les exemples des copules archimédiennes **

Nom de la copule	Générateur	Copule bivarié $C(u, v)$
Indépendance	$-\ln(u)$	uv
Clayton ($\theta > 0$)	$u^{-\theta} - 1$	$(u^{-\theta} + v^{-\theta} - 1)^{-\frac{1}{\theta}}$
Franck ($\theta \neq 0$)	$-\ln\left(\frac{e^{-\theta u} - 1}{e^{-\theta} - 1}\right)$	$-\frac{1}{\theta} \ln \left[1 + \frac{(e^{-\theta u} - 1)(e^{-\theta v} - 1)}{e^{-\theta} - 1} \right]$
Gumbel ($\theta \geq 1$)	$(-\ln(u))^{\theta}$	$\exp\left(-\left[(-\ln(u))^{\theta} + (-\ln(v))^{\theta}\right]^{\frac{1}{\theta}}\right)$
Joe ($\theta \geq 1$)	$1 - (1 - e^{-u})^{\frac{1}{\theta}}$	$1 - \left[(1 - u)^{\theta} + (1 - v)^{\theta} - (1 - u)^{\theta}(1 - v)^{\theta}\right]^{\frac{1}{\theta}}$

3.2.2.2. Copules elliptiques

Les courbes de niveau des densités de telles distributions forment des ellipses d'excentricité constante. Elles ont la forme suivante :

$$C_{\rho}(u, v) = \frac{1}{\sqrt{1 - \rho^2}} \int_{-\infty}^{\varphi_{g_1}^{-1}(u)} \int_{-\infty}^{\varphi_{g_2}^{-1}(v)} g\left(\frac{x^2 - 2\rho xy + y^2}{\sqrt{1 - \rho^2}}\right) dx dy$$

$$= H_{\rho}\left(\varphi_{g_1}^{-1}(u), \varphi_{g_2}^{-1}(v)\right)$$

Avec ρ est le coefficient de corrélation, H_{ρ} est la distribution conjointe des variables X, Y , et $\varphi_{g_1}^{-1}, \varphi_{g_2}^{-1}$ sont les fonctions quantiles.

Dans cette famille se trouve :

- Les copules gaussiennes sous-jacentes à la distribution Normale multivariée, la copule gaussienne de paramètre $\rho \in]-1, 1[$ dans le cas bivarié est donnée par :

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

** **Planchet F.**, « Modèles financiers en assurance et analyses dynamiques : Introduction à la théorie des copules », ISFA, 2017.

- Les copules Student construites à partir de la distribution de Student centrée réduite, il s'agit de la **t-copule** de corrélation $\rho \in]-1,1[$ et de degré de liberté ν dont la fonction de densité est donnée par :

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

3.2.3. Choix de la copule optimale

Pour déterminer les relations entre les variables, il faut choisir la copule qui résume le mieux la structure de dépendance, soit la copule la plus adaptée à la série de données étudiée. Ce choix doit être fait par référence à certains critères tels que la flexibilité, la simplicité analytique et la diversité des formes de dépendance, et non pas par le coefficient de corrélation.

Cependant, pour choisir la meilleure copule et donc pour estimer le(s) paramètre(s) de cette copule, différentes méthodes sont disponibles. Parmi ces méthodes, la méthode du maximum de vraisemblance demeure la plus souvent utilisée^{††}. Cette méthode est plus simple à mettre en œuvre et donne de meilleurs résultats pour les échantillons de grande taille. Elle présente aussi l'avantage d'être utilisables pour tous types de copules, alors que la méthode IFM repose sur la détermination de la matrice de Godambe qui peut s'avérer très complexe en raison des multiples calculs de dérivées, ce qui justifie notre choix pour cette méthode.

La fonction log vraisemblance $L(\theta)$ peut s'écrire :

$$\begin{aligned} L(\theta) &= \sum_{j=1}^n \ln \left(C_\theta \left(F_1(x_{i_1}), \dots, F_d(x_{i_d}) \right) \right) \prod_{j=1}^d f_j(x_j) \\ &= \sum_{j=1}^n \ln \left(C_\theta \left(F_1(x_{i_1}), \dots, F_d(x_{i_d}) \right) \right) + \sum_{i=1}^n \sum_{j=1}^d \ln f_j(x_j) \end{aligned}$$

avec C_θ est la densité de la copule, $f(x_1, x_2, \dots, x_d)$ est la fonction de densité jointe.

L'estimateur de θ est : $\hat{\theta}_n^{MV} = \arg \max L(\theta)$

4. PRESENTATION ET DESCRIPTION DES DONNEES

Notre étude consiste à estimer les provisions techniques pour sinistres à payer pour deux garanties de la branche automobile à savoir la responsabilité civile matérielle et la responsabilité civile corporelle. Pour cela, nous allons appliquer dans un premier temps la méthode GLM pour le cas d'indépendance et la théorie des copules pour le cas de dépendance, ensuite nous comparons les résultats obtenus à travers ces deux approches.

L'étude a été menée sur un échantillon de l'agence d'assurance CAAT. Les données portent sur les montants de règlements sinistres pour la responsabilité civile matérielle et la

^{††} C. Kharoubi-Rakotomalala, « Les fonctions copules en finance », Sorbonne, 2008.

responsabilité civile corporelle. Elles sont présentées dans des triangles de règlements par année de survenance et par année de développement sur la période allant de 2004 à 2016.

Dans le Table suivant, nous présentons un récapitulatif des résultats de l'analyse statistique des données recueillies obtenus en utilisant le logiciel XL STAT :

Table N°2 : Analyse descriptive des données

Statistiques	CORPOREL	MATERIEL
Nb. d'observations	91	91
Minimum	2 216 892,730	2 475 045,100
Maximum	466 185 241,280	2 768 632 133,780
1er Quartile	26 914 384,650	24 859 858,195
Médiane	105 193 751,090	98 053 819,270
3ème Quartile	187 923 624,100	587 579 966,696
Moyenne	130 241 115,743	469 401 917,443
Ecart-type (n-1)	119 972 629,712	710 878 663,298

Nous constatons que les sinistres matériels ont en moyenne un montant de règlement plus important. Cela s'explique par une fréquence plus élevée pour ce type de sinistres comparé au corporel.

5. ESTIMATION DES PROVISIONS TECHNIQUES

Nous commençons par estimer les provisions techniques dans le cas indépendant en utilisant le modèle linéaire généralisé GLM, ensuite dans le cas dépendant en utilisant la théorie des copules.

5.1. Estimation des provisions techniques par le modèle GLM

Dans le cadre de l'hypothèse d'indépendance, nous estimons les provisions techniques pour les deux garanties responsabilité civile matérielle et responsabilité civile corporelle en utilisant le modèle classique stochastique « le modèle GLM ^{††} ». Dans ce modèle, les régresseurs sont des variables indicatrices (une variable indicatrice par année de survenance, et une variable indicatrice par année de développement). Pour modéliser les montants de sinistres, il est possible d'utiliser des lois discrètes ou des lois continues. Les lois les plus utilisées pour modéliser des données discrètes sont la loi de Poisson et la loi Binomiale Négative.

Les paramètres du modèle GLM sont estimés par la méthode du maximum de vraisemblance, appliquée aux données du triangle supérieur (les règlements connus) ^{§§}.

Le calcul de la provision se fait selon la formule :

^{††} Développé initialement par **Nelder et Wedderburn** en 1972.

^{§§} **C. Partrat**, « *Evaluation stochastique de la provision pour sinistres* », Institut des Actuaire, 2004.

$$\mu_{ij} = e^{u+\alpha_i+\beta_j},$$

Avec μ_{ij} : la provision des sinistres à payer estimé pour l'année de survenance i et pour l'année de développement j ;

u : la constante ;

α_i : le paramètre estimé de la variable explicative année de survenance i ;

β_j : le paramètre estimé de la variable explicative année de développement j .

Pour l'année 0, nous avons $\alpha_0 = \beta_0 = 0$.

Dans le Table suivant, nous présentons les estimateurs des paramètres α_i et β_j pour les deux sous-branches obtenus en utilisant le logiciel R :

Table N°3 : Estimation des paramètres u , α_i et β_j du modèle GLM

Dommages Matériels u = 20,4889		Dommages Corporels u= 18,8550271	
α_i	β_j	α_i	β_j
0	0	0	0
0,1383	-0,4436	-0,4048009	0,8336862
0,2453	-1,8659	-0,1406768	0,4499782
0,3542	-2,6093	0,0353261	-0,0008184
0,5692	-2,988	0,0303858	-0,5372343
0,5961	-3,1113	-0,0334859	-1,1535355
0,6599	-3,6511	-0,2584407	-1,6343122
0,9712	-4,1794	0,0474541	-2,1091092
1,2866	-4,2722	0,0646102	-2,199275
1,2453	-4,8529	0,0901959	-2,3743944
1,2786	-5,1044	0,1045232	-2,9250007
1,2008	-5,7302	-0,0669069	-2,690282
1,1086	-4,8808	-0,3728338	-4,2434099

Pour

l'évaluation des provisions techniques, nous avons calculé la moyenne et l'écart type pour les sinistres corporels plus matériels qui vont nous permettre par la suite de faire la comparaison avec le cas dépendant. Les résultats obtenus en utilisant le logiciel XL STAT sont présentés dans le Table suivant :

Table N°4 : Estimation des provisions techniques par le modèle GLM

(Matériel + Corporel) cas indépendant		
Moyenne	Ecart type	
50 683 141,2	135	265 930

5.2. Estimation des provisions techniques par la théorie des copules

Tout d’abord, nous commençons par vérifier s’il existe une corrélation entre les données supérieures des deux triangles. Une fois nous confirmons ceci, nous passons à l’estimation des provisions techniques en utilisant la théorie des copules.

5.2.1. Etude de la corrélation

L’étude de la corrélation est faite en se basant sur les coefficients de corrélation de Pearson, Spearman et Kendall dont les valeurs sont données dans les matrices de corrélation suivante :

Matrice de corrélation (Spearman)		
Variables	Corporel	Matériel
Corporel	1	0,846
Matériel	0,846	1

Matrice de corrélation (Kendall)		
Variables	Corporel	Matériel
Corporel	1	0,666
Matériel	0,666	1

Matrice de corrélation (Pearson)		
Variables	Corporel	Matériel
Corporel	1	0,528
Matériel	0,528	1

L’analyse des valeurs des trois matrices de corrélation nous a permis de constater qu’il existe une forte corrélation entre la responsabilité civile matérielle et la responsabilité civile corporelle. Cette corrélation est confirmée également à travers les nuages de points représentés dans les figures suivantes :

Figure N° 1 : Nuage de points des règlements corporels en fonction
des règlements matériels

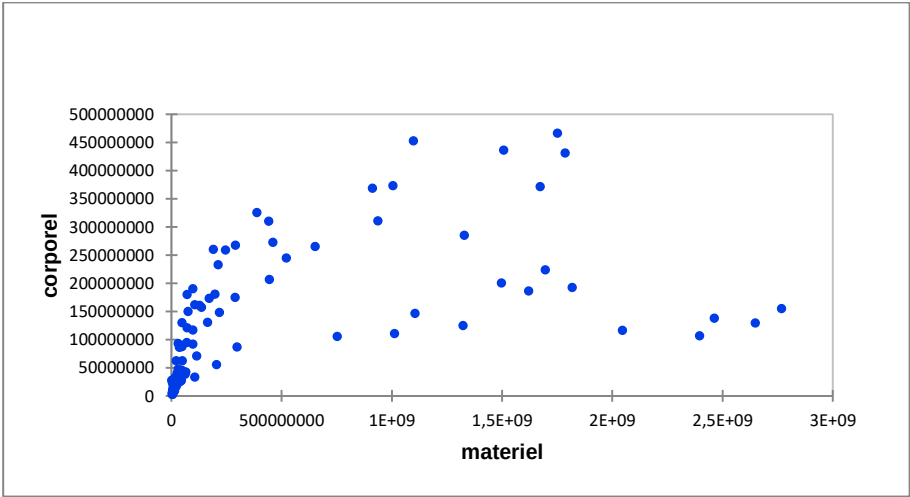
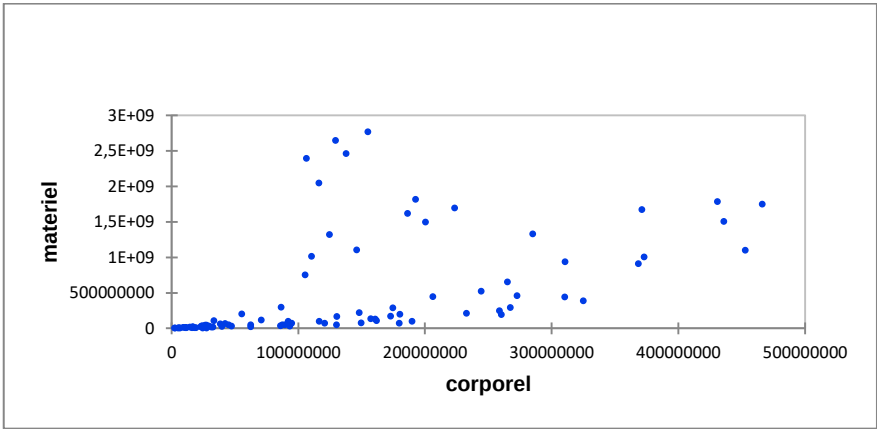


Figure N° 2 : Nuage de points des règlements matériels en fonction des règlements corporels



D’après les figures 1 et 2, il s’avère que les règlements de sinistres matériels et corporels sont fortement corrélés, il y a une forte concentration autour des petites valeurs, ce qui renforce l’hypothèse de dépendance.

5.2.2. **Choix des fonctions marginales**

Afin d’estimer les paramètres de la structure de dépendance, nous estimons d’abord les paramètres des fonctions marginales. Pour cela, il faut tester plusieurs lois utilisées dans le domaine de l’assurance pour modéliser les provisions techniques pour sinistres à payer.

Pour la sélection de la loi marginale pour chaque sous-branche, nous utilisons le test d’adéquation de Kolmogorov-Smirnov. Les résultats obtenus en utilisant le logiciel XL STAT sont donnés dans le Table suivant :

Table N°5 : Test d'adéquation de Kolmogorov-Smirnov

Lois	P-Value Corporel	P-Value Matériel
Pareto	0,001	< 0,0001
Normale	0,0437	< 0,0001
Log-normale	0,137	0,145
Gumbel	< 0,0001	< 0,0001
Gamma	< 0,0001	< 0,0001
Fisher	< 0,0001	< 0,0001
Khi-deux	< 0,0001	< 0,0001
Weibull	< 0,0001	< 0,0001

Le test nous a permis de sélectionner la loi Log-normale selon les deux hypothèses suivantes :

H₀ : L'échantillon suit une loi Log-normale ;

H₁ : L'échantillon ne suit pas une loi Log-normale.

Etant donné que la P-Value calculée est supérieure au seuil de signification $\alpha = 0,05$, alors on ne peut pas rejeter l'hypothèse nulle H₀.

Les paramètres estimés de la loi Log-normale (μ, σ) pour les deux sous-branches obtenus à l'aide du logiciel XL STAT sont donnés dans le Table suivant :

Table N°6 : Estimation des paramètres de la loi Log-normale

Paramètres	Matériel	Corporel
μ	16,494	17,121
sigma	1,209	1,218

5.2.3. Sélection de la copule et estimation des paramètres

A partir des triangles de règlements, nous avons créé une matrice de pseudo-observations de la partie inférieure du triangle afin d'uniformiser les données. Pour sélectionner la copule adéquate à notre série de données, nous avons utilisé la fonction « BiCopSelect » du package « CopulaVine » sous le logiciel R sur la matrice uniforme. Cette fonction nous permet de sélectionner la famille des copules à partir des données, et le paramètre de la copule est sélectionné selon la méthode du maximum de vraisemblance.

La sélection est faite parmi les familles suivantes :

0 = Independence Copula

1 = Gaussian Copula

2 = Student t Copula

3 = Clayton Copula

4 = Gumbel Copula

5 = Franck Copula

6 = Joe Copula

7 = BB1 Copula

8 = BB8 Copula

9 = BB7 copula,

10 = BB8 copula.

La copule retenue sur la base de notre série de données est :

Bivariate Copula : Joe (par = 5,97 ; Tau = 0,72)

Joe Copula appartient à la famille des copules archimédiennes, elle se présente sous la forme :

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

Où θ est le paramètre de la copule ; Tau correspond au Tau de Kendall, une statistique qui mesure la corrélation de rang entre les deux variables.

Après la sélection de la copule, nous testons l'adéquation de nos données aux différentes familles de copules avec la fonction « BiCopEst ». Les résultats obtenus en utilisant le logiciel R sont représentés dans le Table suivant :

Table N° 7 : Test d'adéquation pour les différentes familles de copules bivariées

Famille	Paramètre	Tau
0 = Independence Copula	0	0
1= Gaussian Copula	0.79	0.58
2 = Student t Copula (t-Copula)	0.82 ; 2.39	0.62
3 = Clayton Copula	1.02	0.34
4 = Gumbel Copula	3.09	0.68
5 = Franck Copula	8.69	0.63
6 = Joe Copula	5.97	0.72
7 = BB1 Copula	0 ; 8.09	0.68
8 = BB6 Copula	5.96 ; 1	0.70
9 = BB7 Copula	5 ; 0	0.68
10 = BB8 Copula	5.97 ; 1	0.72

Ainsi, selon le critère de maximisation de la vraisemblance, la sélection de Joe Copula est justifiée (elle présente le Tau le plus élevé).

Les graphiques de la densité et celui du contour obtenus en utilisant le logiciel R montrent qu'il y a une forte dépendance entre les deux sous-branches au niveau de la queue inférieure et de la queue supérieure.

Figure N° 3: Joe Copula densité (par =5.97, tau = 0.72)

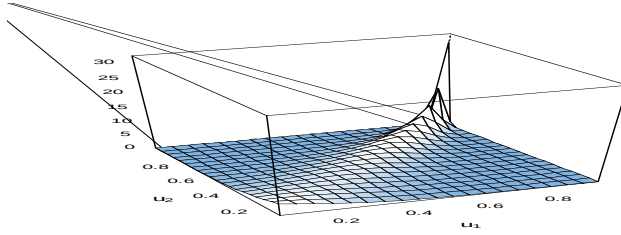
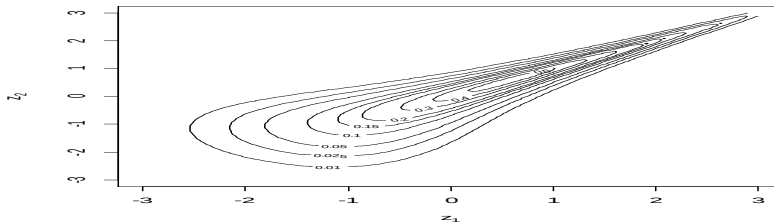


Figure N° 4: Joe Copula contour (par =5.97, tau = 0.72)



5.2.4. Estimation des provisions techniques

Après avoir estimé les paramètres de la copule sélectionnée sur la base de nos données de la partie inférieure du triangle, nous avons simulé 10000 couples de provisions issues de la copule de Joe. Ensuite, nous avons calculé la moyenne et l'écart-type de ces couples de provisions pour les deux sous-branches confondues. Les résultats obtenus en utilisant le logiciel R sont présentés dans le Table suivant :

Table N°8: Estimation des provisions techniques par la méthode des copules

(Matériel + Corporel) cas dépendant	
Moyenne	Ecart type
53 861 755	234 982 155

En faisant la comparaison entre l'estimation des provisions techniques selon le modèle GLM avec celle trouvée en utilisant la théorie des copules, nous remarquons que la moyenne des provisions techniques estimée est très proche dans les deux cas; par contre l'écart-type est plus élevé dans le cas dépendant.

6. CONCLUSION

Dans notre étude, nous avons tenté de montrer la dépendance entre les deux sous branches de l'assurance automobile, dommages matériels et corporels, à travers l'application de la théorie des copules. Nous avons constaté que cette dépendance existe (**Tau = 0.72**), et que la non prise en considération de la dépendance entre les garanties en

assurance automobile augmenterait sensiblement le risque de sous-estimation des provisions techniques (+73%), pouvant ainsi influencer sur la solvabilité de la compagnie.

La théorie des copules contribue à une meilleure maîtrise du risque lié à l'évaluation des provisions techniques dont la valeur correspondant à notre série de données est égale à **234 982 155** DA comparée aux méthodes classiques qui sous estimeraient ce risque à la valeur égale à **135 265 930** DA. Cette théorie contribue à une meilleure gestion de la solvabilité de la compagnie car elle prend en considération la dépendance à différents niveaux surtout la dépendance de queue, ce qui fait son importance et sa spécificité par rapport aux méthodes classiques.

BIBLIOGRAPHIE

- 1 **Armél K.**, « *Structure de dépendance des générateurs de scénarios économiques –Modélisation et simulation* », EURIA, Bretagne, 2010.
- 2 **Deheuvels P.**, « *La fonction de dépendance empirique et ses propriétés- un test non paramétrique d'indépendance* », Académie Royale de Belgique, Bulletin de la Classe des Sciences, 5^{Eme} Série, 65, 274-292, 1979.
- 3 **Frees E.W & Valdez E.A.**, « *Hierarchical Insurance Claims Modeling* », Journal of the American Statistical Association, 2012.
- 4 **Genest C., Mackay R.J.**, « *The Joy of copulas : Bivariate distributions with uniform marginals* », The American Statistician, 40, 280-283, 1986.
- 5 **Kharoubi-Rakotomolala C.**, « *Les fonctions copules en finance* », Publications de la Sorbonne, 2008.
- 6 **Kimeldorf G et Sampson A.**, « *Uniform Representations of Bivariate Distributions* », FSU Statistics Report M320, Florida, 1974.
- 7 **Krauth G.**, « *Provisionnement et corrélation entre branches* », CNAM, Paris, 2007.
- 8 **Mai J.F. et Scherer M.**, « *Simulating Copulas Stochastic Models, Sampling Algorithms and Applications* », World Scientific Publishing, Singapore, 2017.
- 9 **Nelsen R.B.**, « *An introduction to Copulas* », Springer, Portland, 2007.
- 10 **Partrat C.**, « *Evaluation Stochastique de la provision pour sinistres* », Conférence Institut des Actuaires, Paris, 2004.
- 11 **Roncalli T.**, « *Copules et Aspects Multidimensionnels du Risque* », ENSAI, Paris, 2002.
- 12 **Sauvet C.**, « *Solvency II - Quelle modélisation stochastique des provisions techniques prévoyance et non vie ?* », Université de Lyon, 2006.
- 13 **Sklar A.**, « *Fonctions de répartition à n dimensions et leurs marges* », Publications de l'Institut de Statistique de l'Université de Paris 8, 229-231, 1959.
- 14 **Sklar A.**, « *Random Variables, Joint Distribution Functions, and Copulas* », KYBERNETIKA, Volume 9, Prague, 1973.
- 15 **Vermette R.**, « *Modélisation de la dépendance entre les garanties applicables en assurance automobile* », Université Laval, Québec, 2011.