



**Master professionnel de sciences de gestion,
Mention finances de marché
Spécialité Actuariat**

**Modélisation de la dépendance entre branches d'assurance
avec les modèles de chocs communs multiplicatifs
lognormaux**

Par : **Guillaume Galataud**

Tuteur : **Hervé Pinaud**

NON CONFIDENTIEL

Mémoire soutenu le 15 Décembre 2010

JURY

Président : Michel FROMENTEAU

Membre : Safia LEKEHAL

Pierre PETAUTON

Vincent RUOL

François WEISS

Résumé

La prise en compte de la corrélation, et plus largement, de la dépendance est un sujet de plus en plus crucial dans le calcul des provisions techniques et des fonds propres exigibles pour l'activité d'assurance. Et l'arrivée de Solvency II confirme cette tendance.

Les modèles à chocs communs permettent de modéliser la dépendance entre branches : ils consistent à appliquer un même choc à différentes branches, ce qui a pour effet de créer de la dépendance entre ces dernières. Parmi les modèles à chocs communs, ceux portant sur les lois de Poisson modélisant le nombre de sinistres ont déjà fait l'objet de nombreuses études et mémoires ; voici un exemple d'un tel modèle :

$$\begin{cases} N_1 = N_{10} + N_0 \\ N_2 = N_{20} + N_0 \end{cases}$$

(avec : N_1, N_2 qui modélisent le nombre de sinistres de deux branches d'assurance, N_0 qui est le choc commun; N_{10}, N_{20}, N_0 sont des variables de Poisson)

Le présent mémoire se concentrera sur une autre famille de chocs communs : les chocs communs multiplicatifs log-normaux avec ou sans shift, c'est-à-dire, dans le cas de deux branches, sur les modèles suivants :

$$\begin{cases} X_1 = X_{10} \times \beta \\ X_2 = X_{20} \times \beta \end{cases}$$

(avec : X_1, X_2 qui modélisent la sinistralité de deux branches d'assurance, β qui est le choc commun; X_{10}, X_{20} sont des lognormales et β est une lognormale avec ou sans shift ; une lognormale avec shift est en fait la somme d'une lognormale et d'une constante)

Les structures de dépendance induites par ces modèles seront étudiées.

Pour les chocs log-normaux sans shift, il sera démontré l'équivalence avec une copule normale, lors d'une mise en parallèle avec les modèles basés sur des copules. Cette équivalence sera vérifiée numériquement.

Les structures de dépendance dans le cas de choc log-normal avec shift seront étudiées. Il apparaîtra que le recours au shift donne lieu à des formes variées de structures de dépendance qui peuvent être approchées par celles des copules HRT, de Gumbel ou de Clayton.

Enfin, des applications numériques seront abordées pour avoir un aperçu de l'impact de ces modèles sur des données réelles. Le présent mémoire s'inscrivant dans une démarche exploratoire, il n'aborde pas les tests de validation et de robustesse qui seraient requis pour une utilisation certifiée dans un modèle interne.

Dans un premier temps, quelques applications portant sur deux branches d'assurance (ou Line of Business) seront réalisées. Puis, un cas concret de filiale d'assurance avec plusieurs LoBs sera étudié.

Abstract

Taking into account the correlation and, more broadly, the dependence is a topic increasingly important in the calculation of technical provisions and capital requirements for insurance business. And the arrival of Solvency II confirms this trend.

Models with common shocks are used to model the dependence between the branches: they apply the same shock at different LoBs (Lines of Business), which has the effect of creating dependency between them. Among the models with common shocks, those using Poisson distributions to model the number of claims have already been the subject of numerous studies and thesis, here is an example of such a model:

$$\begin{cases} N_1 = N_{10} + N_0 \\ N_2 = N_{20} + N_0 \end{cases}$$

(with: N_1, N_2 modelling the number of claims of two LoBs, N_0 being the common shock, N_{10}, N_{20} , N_0 being Poisson distributions)

This paper will focus on another family of common shocks: common shocks multiplicative lognormal with or without shift, that is to say, in the case of two branches, the following models:

$$\begin{cases} X_1 = X_{10} \times \beta \\ X_2 = X_{20} \times \beta \end{cases}$$

(with : X_1, X_2 modelling the claims amount of two LoBs, β being the common shock; X_{10}, X_{20} being log-normal distribution and β being a lognormal distribution which is either shifted or not; a shifted lognormal distribution equals to a lognormal distribution added to a constant)

The dependence structures induced by these models will be studied.

For lognormal shocks without shift, it will be demonstrated the equivalence with a normal copula. This equivalence will be verified numerically.

The structures of dependency in the case of shocks with shifted lognormal distributions will be explored. It appears that using the shift gives rise to various forms of dependence structures that can be approximated by those of HRT copula, Gumbel or Clayton.

Finally, numerical applications are discussed for an overview of the impact of these models on real data. Since this document is about an exploratory approach, it does not address the validation tests and robustness tests that would be required for use in a certified internal model.

First, some applications on two classes of insurance (or Line of Business) will be performed. Then, a concrete case of insurance subsidiary with several LoBs will be studied.

Remerciements

Je remercie ma famille, et surtout mes parents, sans qui rien de tout cela n'aurait été possible.

Merci aux professeurs qui m'ont formé tout au long de ma vie d'étudiant, et plus particulièrement aux professeurs du CNAM et à leur faculté à partager leurs connaissances pointues, leur enthousiasme et leur goût pour l'actuariat.

Un grand merci à Hervé Pinaud, mon tuteur au sein de Groupama, qui m'a guidé tout au long de ce mémoire et qui m'a apporté de très précieux conseils. Enfin, je remercie également mes supérieurs René Cado, Marc Raymond et Fanny Pouget qui ont apporté le cadre nécessaire à la constitution du présent mémoire.

SOMMAIRE

I.	LA NOTION DE DEPENDANCE ET LE MONDE DE L'ASSURANCE	8
I.1.	INTRODUCTION	8
I.2.	LE CADRE REGLEMENTAIRE ET LA NOTION DE DEPENDANCE.....	8
I.3.	CALCUL DES BESOINS EN FONDS PROPRES DANS SOLVENCY II.....	9
I.4.	CADRAGE DU MEMOIRE DANS LE CONTEXTE SOLVENCY II	13
II.	LES CONCEPTS MATHÉMATIQUES DE LA DEPENDANCE ET DE MODELE A CHOC COMMUN	14
II.1.	RAPPELS SUR LES DISTRIBUTIONS MULTIVARIEES	14
II.2.	DEFINITION DE LA DEPENDANCE	14
II.3.	CAS NOTABLES DE DEPENDANCE	15
II.3.1.	<i>Dépendance parfaites</i>	15
II.3.2.	<i>Dépendances positives (resp. négatives) par quadrant</i>	16
II.4.	QUANTIFICATION DE LA DEPENDANCE	17
II.4.1.	<i>Mesure de dépendance</i>	17
II.4.2.	<i>Les principaux indicateurs de dépendance</i>	17
II.4.2.1.	La notion de concordance	17
II.4.2.2.	Tau de Kendall	18
II.4.2.3.	Rho de Spearman	18
II.4.3.	<i>Les indicateurs de dépendance de queue</i>	18
II.5.	LES DIFFERENTS TYPES DE MODELISATION DE LA DEPENDANCE	19
II.5.1.	<i>Les approches</i>	19
II.5.1.1.	L'approche « Top-down ».....	19
II.5.1.2.	L'approche « Bottom-up »	20
II.5.2.	<i>Les chocs communs</i>	20
II.5.2.1.	Introduction au modèle de chocs communs	20
II.5.2.2.	Définition d'un modèle de chocs communs	21
II.5.2.3.	Les différents types de chocs communs	21
II.5.2.4.	Une illustration par un exemple de chocs communs	22
II.5.3.	<i>Les copules</i>	24
II.5.3.1.	Présentation générale des copules	24
II.5.3.2.	Définition des copules.....	24
II.5.3.3.	Théorème de Sklar	25
II.5.3.4.	Théorème d'invariance.....	25
II.5.3.5.	Densité d'une copule.....	25
II.5.3.6.	Les copules usuelles.....	26
II.6.	CONCLUSION.....	26
III.	METHODES D'ANALYSE DE LA DEPENDANCE INDUITE PAR UN MODELE DE CHOCS COMMUNS	27
III.1.	PRESENTATION DES AXES D'ANALYSE	27
III.2.	METHODE DE SIMULATION DE MONTE-CARLO	27
III.3.	DETERMINATION DES MARGINALES OBTENUES SUITE A L'APPLICATION DU CHOC	28
III.4.	DESCRIPTION QUALITATIVE DE LA STRUCTURE DE DEPENDANCE INDUITE PAR LE CHOC COMMUN	29
III.4.1.	<i>Visualisation suite à la « normalisation » des données</i>	29
III.4.2.	<i>Visualisation de la structure de dépendance à l'aide de copule empirique</i>	31
III.4.2.1.	Formule théorique de la copule empirique naturelle	31
III.4.2.2.	Construction de la copule empirique de Deheuvels	31
III.5.	EVALUATION NUMERIQUE DES INDICATEURS DE DEPENDANCE	35
III.5.1.	<i>Les estimateurs classiques</i>	35
III.5.2.	<i>Les estimateurs basés sur les copules empiriques</i>	36
III.6.	MISE EN PARALLELE DU CHOC COMMUN AVEC LES COPULES	37
III.6.1.	<i>Fit d'une copule empirique avec les copules usuelles</i>	37
III.6.2.	<i>Mesure de l'adéquation avec la copule usuelle</i>	38
III.6.2.1.	Sélection d'une copule à partir du test d'adéquation du chi-deux	38
III.6.2.2.	Distance entre les classes	40
III.6.2.3.	Choix de la mesure d'adéquation	41

IV.	ETUDE DES CHOCS COMMUNS ET DE LA DEPENDANCE INDUITE.....	42
IV.1.	CHOC MULTIPLICATIF LOG-NORMAL	42
IV.1.1.	Rappel du choc.....	42
IV.1.2.	Détermination des marginales obtenues suite à l'application du choc.....	42
IV.1.3.	Analyse qualitative de la structure de dépendance obtenue.....	42
IV.1.4.	Analyse quantitative de la structure de dépendance obtenue.....	44
IV.1.5.	Vérification numérique de l'analyse qualitative :	46
IV.1.6.	Evaluation des indicateurs de dépendance et leur sensibilité aux paramètres du modèle.....	47
IV.1.7.	Utilisation concrète du modèle et ses limites	50
IV.1.8.	Extension du modèle.....	50
IV.1.9.	Conclusion sur le choc multiplicatif log-normal.....	51
IV.2.	CHOC MULTIPLICATIF LOG-NORMAL SHIFTE	52
IV.2.1.	Présentation du modèle.....	52
IV.2.2.	Détermination des marginales obtenues suite à l'application du choc.....	53
IV.2.3.	Analyse qualitative de la structure de dépendance obtenue.....	53
IV.2.4.	Mise en parallèle avec les copules	55
IV.2.5.	Analyse quantitative de la structure de dépendance obtenue.....	57
IV.2.6.	Extension du modèle.....	60
IV.2.7.	Conclusion sur le choc log-normal shifté.....	61
V.	APPLICATION A UN CAS CONCRET	63
V.1.	OBJECTIFS DE L'APPLICATION A UN CAS CONCRET	63
V.2.	PRESENTATION DU PORTEFEUILLE	63
V.2.1.	Structure des bases de données.....	63
V.2.2.	Traitement des données et préparation des variables à modéliser.....	64
V.2.3.	Statistiques exploratoires	64
V.3.	EXPLORATION DE LA DEPENDANCE OBSERVEE.....	67
V.3.1.	Prise en compte de la saisonnalité.....	67
V.3.2.	Mesures de la dépendance	68
V.3.2.1.	Corrélation linéaire.....	68
V.3.2.2.	Indicateurs de dépendance	70
V.3.2.3.	Conclusion	70
V.4.	METHODOLOGIE POUR LA MODELISATION MULTIVARIEE DE LA SINISTRALITE	70
V.4.1.	Calibrage des modèles de chocs communs	70
V.4.1.1.	Méthode du maximum de vraisemblance	71
V.4.1.2.	Adaptation de la méthode du maximum de vraisemblance	71
V.4.1.3.	Méthode basée sur les indicateurs de dépendance.....	72
V.4.1.4.	Choix de la méthode et mise en application.....	72
V.4.2.	Adéquation et comparaison des différents modèles de chocs communs.....	73
V.4.2.1.	Examen graphique.....	73
V.4.2.2.	Valeurs des log-vraisemblances maximales et critère d'Akaike	73
V.4.2.3.	Distances entre les classes.....	74
V.4.2.4.	Probabilité simultanée d'excès de quantiles.....	74
V.4.2.5.	Choix de la méthode.....	74
V.5.	APPLICATION DANS LE CAS DE 2 BRANCHES.....	75
V.5.1.	Modélisation de la dépendance entre la RC Auto et le Dommage Auto.....	75
V.5.2.	Modélisation de la dépendance entre Accident&Health et Worker Compensation	77
V.5.3.	Modélisation de la dépendance entre la somme des deux branches automobiles et « Fire and other Property Damage ».....	80
V.6.	APPLICATION A L'ENSEMBLE DU PORTEFEUILLE POUR CALCULER LE SCR	83
V.6.1.	Approche utilisée.....	83
V.6.2.	Résultats	84
V.6.2.1.	1 ^{er} modèle hiérarchique	84
V.6.2.2.	2 ^{ème} modèle hiérarchique.....	85
V.6.2.3.	3 ^{ème} modèle hiérarchique.....	86
V.6.3.	Conclusion	87
VI.	CONCLUSION.....	89
VI.1.	RESUME	89
VI.2.	SYNTHESE SUR LA COMPARAISON AVEC LES COPULES	91
VII.	BIBLIOGRAPHIE.....	92

VIII.	ANNEXES.....	93
VIII.1.	LES COPULES DE REFERENCE	93
VIII.1.1.	<i>Copule d'indépendance.....</i>	93
VIII.1.2.	<i>Copule normale.....</i>	93
VIII.1.3.	<i>Copule de Gumbel.....</i>	94
VIII.1.4.	<i>Copule de Clayton.....</i>	95
VIII.1.5.	<i>Copule HRT.....</i>	96
VIII.2.	RESULTATS OBTENUS POUR LE CHOC COMMUN MULTIPLICATIF LOG-NORMAL	97
VIII.3.	RESULTATS OBTENUS POUR LE CHOC COMMUN MULTIPPLICATIF LOG-NORMAL SHIFTE	98
VIII.3.1.	<i>Cas de cor = .3.....</i>	98
VIII.3.2.	<i>Cas de cor = .5.....</i>	98
VIII.3.3.	<i>Cas de cor = .7.....</i>	101
VIII.4.	PROGRAMME « R » UTILISE	105

I. LA NOTION DE DEPENDANCE ET LE MONDE DE L'ASSURANCE

I.1. Introduction

La prise en compte de la corrélation, et plus largement, de la dépendance est un sujet de plus en plus crucial dans le calcul des provisions technique et des fonds propres exigibles pour l'activité d'assurance.

L'arrivée de « Solvency II », en introduisant des critères de Value-at-Risk et en incitant les compagnies d'assurance à utiliser un modèle interne de détermination du besoin en capital, incite fortement à l'utilisation de méthodes stochastiques (et des dépendances stochastiques).

L'assureur ne peut se contenter de modéliser individuellement ses différentes branches, puis de sommer les besoins en fonds propres de chaque branche afin d'obtenir le besoin en capital de l'entreprise dans sa globalité.

En effet, les différentes branches de l'entreprise ne sont pas indépendantes :

- Un facteur particulier à un niveau national au international peut avoir un impact sur des branches à priori indépendantes. Par exemple, l'influence du climat peut être source de dépendance : dans le cas d'une année particulièrement pluvieuse, le seul point commun entre deux accidents dans deux régions différentes sera la dégradation des conditions de circulation par la pluie.
- Un événement unique peut induire des sinistres dans plusieurs branches d'activités différentes. Par exemple, pour un accident de voiture, il faut dans certains cas réparer la voiture (branche automobile), indemniser et réparer les dommages causés au tiers (responsabilité civile), et indemniser le conducteur en cas d'invalidité ou prendre en charge des frais de santé.

Afin de mieux maîtriser ses risques et pour répondre à une réglementation européenne de plus en plus exigeante, l'assureur est amené à modéliser les **dépendances** entre branches d'assurance, et ce, de différentes manières.

I.2. Le cadre réglementaire et la notion de dépendance.

Après Solvency I qui prévoyait une marge de solvabilité déterminée en fonction de pourcentages appliqués sur les primes et les sinistres, la réglementation des assurances passe à des règles plus complexes intégrant la notion de risque et de dépendance entre risques. Ces règles sont applicables soit par l'application de formules standards, soit par la prise en compte de modèles internes.

Dans ce chapitre, nous décrirons globalement Solvency II, puis nous nous concentrerons sur la partie dédiée à la dépendance.

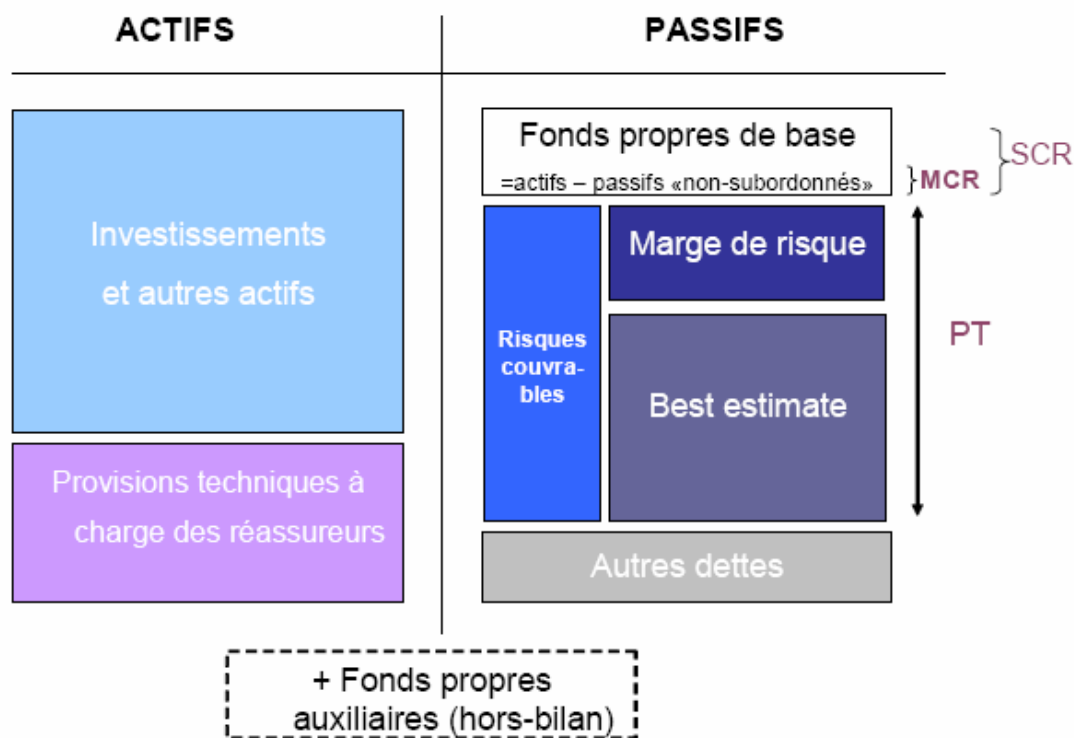
Solvency II prévoit trois piliers :

1) **Le premier pilier** a pour objectif de définir des seuils quantitatifs aussi bien pour les provisions techniques que pour les fonds propres. Ces seuils deviendront des seuils réglementaires. Deux niveaux de fonds propres, MCR et SCR, seront définis :

- le MCR représente le niveau minimum de fonds propres en-dessous duquel l'intervention de l'autorité de contrôle sera automatique
- le SCR représente le capital cible nécessaire pour absorber le choc provoqué par un scénario catastrophe se produisant en moyenne une fois tous les 200 ans.

Le schéma ci-dessous illustre ces notions en présentant le bilan d'une compagnie d'assurance :

Bilan Solvabilité II



2) Le **deuxième pilier** a pour vocation de fixer des normes qualitatives de suivi des risques en interne aux sociétés et de décrire 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 et/ou de réduire leur exposition aux risques.

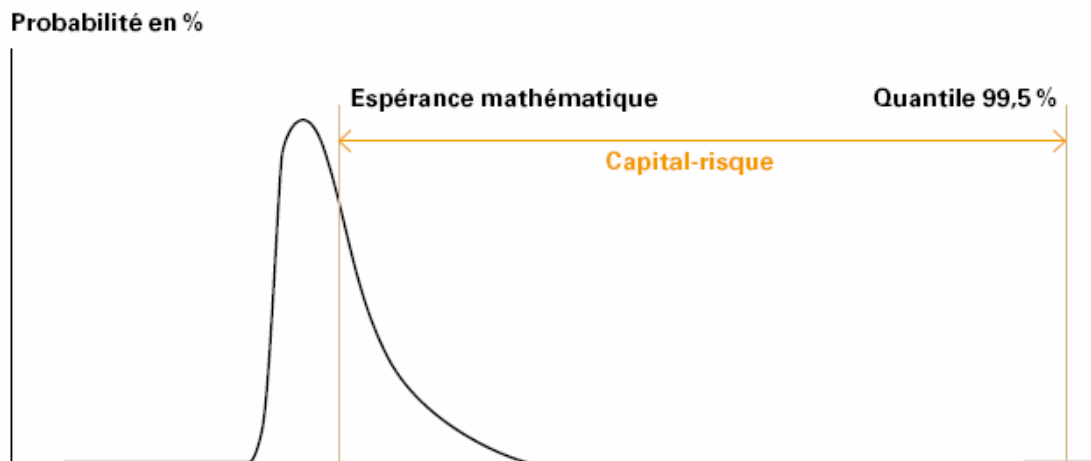
3) Le **troisième pilier** a pour objectif de définir l'ensemble des informations détaillées que les autorités de contrôle jugeront nécessaires pour exercer leur pouvoir de surveillance ; il inclut également tout l'aspect de communication en direction des marchés.

Le présent mémoire s'inscrit dans le cadre du premier pilier de Solvency II, et plus particulièrement du calcul du SCR.

I.3. Calcul des besoins en fonds propres dans Solvency II

Les assureurs et réassureurs seront contraints de mesurer leurs risques et de s'assurer qu'ils ont suffisamment de fonds propres pour les couvrir. Il est sûr que le niveau de probabilité de 99,5% sera retenu sur un horizon d'un an, au vu des différents QIS (Quantitative Impact Study).

Mathématiquement, la notion de VaR à (Value At Risk) à 99,5% et à un horizon d'un an est utilisée. Le schéma ci-dessous montre le capital-risque correspondant :



Ainsi, une compagnie d'assurance devra avoir un niveau de fonds propre suffisant pour faire face au cas le plus extrême qui peut se présenter en moyenne une fois tous les 200 ans.

La suite de l'explication sur Solvency II s'appuie sur le QIS4.

Les compagnies auront le choix entre trois options. Elles pourront soit adopter une approche standard, soit un modèle interne, ou encore la solution intermédiaire du modèle interne partiel.

Globalement, le calcul du SCR consiste :

- à calculer les VaR à 99.5% et à un an pour des modules de risques
- puis à agréger les modules de risques pour obtenir le résultat final du SCR

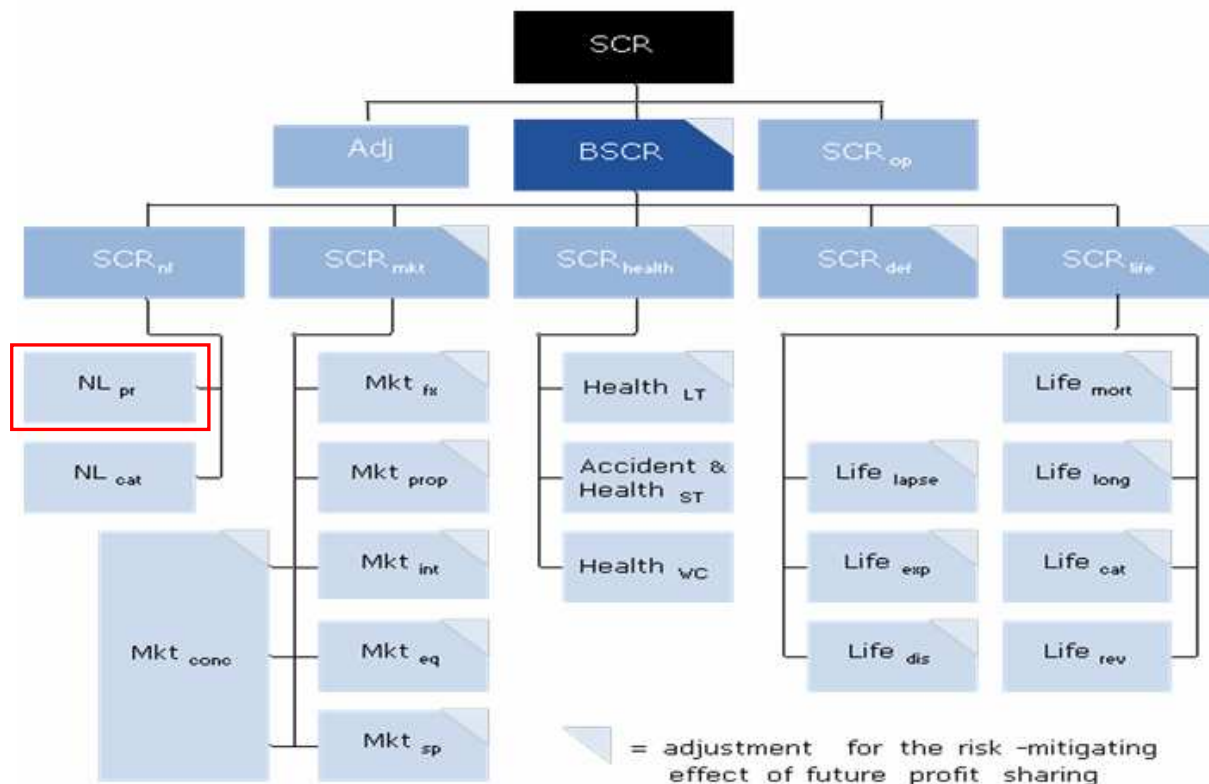
Le calcul du SCR est réalisé par l'agrégation de trois modules (définis dans le QIS4) :

- un module d'ajustement
- un BSCR (Basic SCR)
- un module pour les risques opérationnels : SCROp

Le BSCR est également obtenu par l'agrégation des cinq modules :

- SCRnl : pour la partie non-vie (Non Life)
- SCRmkt : pour les risques de marché (market)
- SCRhealth : pour les risques santé
- SCRdef : pour le risque de contre partie (default)
- SCRlife : pour la partie vie

Voici la structure des modules de risques pour le calcul du SCR (telle qu'elle est définie dans le QIS4) :



Le SCR_{nl} est à son tour décomposé en deux sous-modules :

- NL_{pr} comportant le risque de prime qui fera l'objet d'applications numériques
- NL_{cat} pour les risques catastrophes

Le risque de prime traite de la nouvelle année de souscription et du risque que les primes ne soient pas suffisantes pour faire face à la charge de sinistre.

NL_{pr} est constitué des différentes LoB (lines of business, ou sous-branches d'assurance) suivantes :

Numéro de LoB	
1	Motor, third-party liability
2	Motor, other classes
3	Marine, aviation, transport (MAT)
4	Fire and other property damage
5	Third-party liability
6	Credit and suretyship
7	Legal expenses
8	Assistance
9	Miscellaneous
10	Non-proportional reinsurance - property
11	Non-proportional reinsurance - casualty
12	Non-proportional reinsurance - MAT

Un élément décisif dans le calcul du SCR est la méthode d'agrégation. En effet, elle peut induire des effets de diversifications importants. Il est utile de distinguer deux niveaux d'agrégation : une agrégation entre modules de risques, et une agrégation entre branches au sein d'un même module de risques.

Il existe plusieurs **méthodes d'agrégation** :

- le recours à matrice de corrélation linéaire qui apparaît souvent dans la formule standard, par exemple :

$$(B)SCR = \sqrt{\sum_{1 \leq r, c \leq n} CorrSCR_{r,c} \cdot SCR_r \cdot SRC_c}$$

où $(CorrSCR_{r,c})_{1 \leq r, c \leq n}$ est la matrice de corrélation n*n

- le recours à des méthodes statistiques plus complexes, avec notamment les copules et **les modèles de chocs communs (qui font l'objet du présent mémoire)**

Exemple numérique :

Voici un exemple numérique simple basé sur la méthode d'agrégation standard avec matrice de corrélation, qui se place dans un environnement gaussien :
Supposons deux modules de risques ayant les SCR suivants :

SCR module de risque 1	100
SCR module de risque 2	150

La matrice de corrélation est la suivante :

$$\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$$

Le tableau ci-dessous présente l'évolution du SCR total :

ρ	SCR Modules 1+2	Gain de diversification	%de SCR par rapport à la situation d'indépendance
1.00	250.00	0.00%	38.68%
0.75	234.52	6.19%	30.09%
0.50	217.94	12.82%	20.89%
0.25	200.00	20.00%	10.94%
0.00	180.28	27.89%	0.00%
-0.25	158.11	36.75%	-12.29%
-0.50	132.29	47.08%	-26.62%

Ce simple exemple illustre à quel point la prise en compte de la dépendance peut avoir un impact fort sur le SCR global : pour des valeurs du paramètre ρ passant de 1 à -.5, le SCR global varie de +38% à -26% par rapport à la situation d'indépendance. Le gain de diversification représente la diminution de SCR par rapport au cas de dépendance totale.

Exemple d'agrégation de Solvency II :

L'exemple présenté ci-dessus peut être mis en parallèle avec l'agrégation des NL_{pr} et NL_{cat} pour constituer le SCR_{nl} :

$$SCR_{nl} = \sqrt{\sum_{r^*c} CorrNL^{pr^*cat} \cdot NL_{pr} \cdot NL_{cat}}$$

avec la matrice de corrélation $CorrNL^{pr^*cat}$ définie ainsi dans le QIS4 :

$CorrNL^{pr^*cat}$	NL_{pr}	NL_{cat}
NL_{pr}	1	0
NL_{cat}	0	1

Enfin, pour illustrer le cas de matrice de corrélation de grande taille, voici celle du module NL_{pr} dans le QIS4 :

CorrLob _{pr}	1	2	3	4	5	6	7	8	9	10	11	12
1 : M(3rd party)	1											
2 : M(other)	.5	1										
3 : MAT	.5	.25	1									
4 : Fire	.25	.25	.25	1								
5 : 3rd party liab	.5	.25	.25	.25	1							
6 : credit	.25	.25	.25	.25	.5	1						
7 : legal exp.	.5	.5	.25	.25	.5	.5	1					
8 : assistance	.25	.5	.5	.5	.25	.25	.25	1				
9 : misc.	.5	.5	.5	.5	.5	.5	.5	.5	1			
10 : reins. (prop)	.25	.25	.25	.5	.25	.25	.25	.5	.25	1		
11 : reins. (cas)	.25	.25	.25	.25	.5	.5	.5	.25	.25	.25	1	
12 : reins. (MAT)	.25	.25	.5	.5	.25	.25	.25	.25	.5	.25	.25	1

Note : Dans le cas du module NL_{pr}, la matrice de corrélation ne s'applique pas directement aux SCR qui pourraient être calculés pour chaque branche, mais à leurs volatilités.

I.4. Cadrage du mémoire dans le contexte Solvency II

Avec Solvency II, la notion de dépendance à l'intérieur de modules de risques et entre différents modules de risques prend un rôle clef et peut avoir des impacts quantitatifs significatifs.

Le présent mémoire va se concentrer sur une de ces méthodes d'agrégation, à savoir les modèles de chocs communs.

Et plus précisément sur les modèles de chocs communs multiplicatifs log-normaux.

Etant donné que le mémoire se concentre davantage sur des problématiques d'assurance que sur des sujets financiers, les applications numériques porteront principalement sur le risque de primes de la partie non-vie d'une compagnie d'assurance du module de risque NL_{pr}.

Le risque de prime traite de la nouvelle année de souscription et du risque que les primes ne soient pas suffisantes pour faire face à la charge de sinistre.

Egalement, les modèles de chocs communs peuvent constituer une possibilité pour l'agrégation de tout module de risque.

II. LES CONCEPTS MATHÉMATIQUES DE LA DÉPENDANCE ET DE MODÈLE À CHOC COMMUN

II.1. Rappels sur les distributions multivariées

Un bref rappel sur les distributions multivariées peut s'avérer nécessaire avant d'aborder les différents types de modélisation de dépendance.

Fonctions de répartition conjointe

Soit un vecteur aléatoire $X=(X_1, X_2, \dots, X_d)$ de dimension d . La **fonction de répartition conjointe** de X est la fonction de d variables définie par :

$$F(x_1, x_2, \dots, x_d) = P(X_1 \leq x_1, X_2 \leq x_2, \dots, X_d \leq x_d)$$

Lois marginales

Les **fonctions de répartition marginales** sont les fonctions de répartition F_1, F_2, \dots, F_d des variables aléatoires X_1, X_2, \dots, X_d et sont définies par :

$$F_i(x_i) \rightarrow P(X_i \leq x_i) = F(\infty, \infty, \dots, x_i, \dots, \infty, \infty)$$

Cette notation doit se comprendre comme la limite de F lorsque chacune des composantes x_j ($i \neq j$) tend vers l'infini.

II.2. Définition de la dépendance

Définition :

Les v.a.r. X et Y sont **dépendantes** si et seulement si il existe (x, y) tel que :

$$F_{X,Y}(x, y) \neq F_X(x) \cdot F_Y(y)$$

Au contraire, X et Y sont **indépendantes** ssi $F_{X,Y}(x, y) = F_X(x) \cdot F_Y(y)$ pour tout $(x, y) \in \mathbb{R}^2$.

Nous pouvons donc noter que la dépendance entre les v.a.r. X et Y est totalement présente en la distribution du couple (X, Y) .

En dimension d , nous obtenons :

Les composantes du vecteur aléatoire X sont 2 à 2 indépendantes si pour tout (x_1, x_2, \dots, x_d) dans \mathbb{R}^d on a :

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

ou encore si X a pour densité :

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

Remarques :

La notion de dépendance est beaucoup plus large que celle de la corrélation linéaire.

Le coefficient de corrélation linéaire, également connu sous le nom de coefficient de corrélation de Pearson, est le plus couramment utilisé dans ce cadre.

Pour rappel, voici la définition du coefficient de corrélation de Pearson :

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_X \cdot \sigma_Y}$$

avec : $\text{Cov}(X, Y) = E[(X - E[X]) \times (Y - E[Y])]$

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 ne suivent pas toujours de telles lois.

Aussi, il est indispensable de s'intéresser à la notion de dépendance et aux moyens disponibles pour la mesurer.

Enfin, il est bon de noter que non corrélation et indépendance ne sont pas équivalents :

« X, Y indépendantes implique X, Y non corrélées »,

est vraie mais la réciproque est fautive : deux variables non corrélées peuvent être dépendantes (et l'être fortement), comme l'illustre le contre-exemple suivant :

Ex : X suit une loi symétrique (donc telle que $E[X^3] = 0$), par exemple une loi Normale $N(0, 1)$ et $Y = X^2$. Alors :

$$\rho(X, Y) = E(XY) = E(X^3) = 0$$

Ainsi, X et Y sont liées fonctionnellement mais ne sont pas corrélées. En fait, le coefficient de corrélation ne capture que les liaisons linéaires entre X et Y

II.3. Cas notables de dépendance

Il existe quelques cas de dépendances particuliers qui sont à noter.

II.3.1. Dépendance parfaites

La notion de dépendance parfaite inclut la comonotonie et l'antimonotonie.

Le couple de risques (X, Y) est **comonotone** s'il existe une v.a.r. Z et des fonctions croissantes positives a et b telles que :

$$(X, Y) \text{ a même distribution que } (a(Z), b(Z))$$

Le couple de risques (X, Y) est **antimonotone** s'il existe une v.a.r. Z et une fonction croissante positive a et une fonction décroissante b telles que :

$$(X, Y) \text{ a même distribution que } (a(Z), b(Z))$$

Ex : la comonotonie s'introduit naturellement dans les contrats comportant un partage de risque entre assureur et assurés.

Soit pour un montant (individuel ou cumulé) de sinistre X, I(X) la charge sinistre de l'assureur et X-I(X) celle de l'assuré.

I étant supposé croissante, (X, I(X)) sont comonotones.

Ex : pour un traité de réassurance en stop-loss

$$X_a = \begin{cases} X(\text{pour } : X \leq d) \\ d(\text{pour } : X > d) \end{cases}$$

$$X_r = \begin{cases} 0(\text{pour } : X \leq d) \\ X - d(\text{pour } : X > d) \end{cases}$$

X_a et X_r représentant les sinistralités vues respectivement par l'assureur et le réassureur sont comonotones.

La notion de la dépendance parfaite ouvre la porte à des propriétés intéressantes. A titre d'exemple, voici une des propriétés d'un couple de risques comonotones :

Proposition : Soit X, Y deux risques de fonction de répartition F_X et F_Y et de quantiles q_X et q_Y

Soit $S = X + Y$ la charge globale

Si (X, Y) est comonotone, on a :

$$q_S(p) = q_X(p) + q_Y(p), \text{ pour } p \in]0, 1[$$

II.3.2. Dépendances positives (resp. négatives) par quadrant

Définition :

Soient X et Y deux risques de fonction de répartition jointe $F_{X,Y} = F$ et marginales F_X et F_Y .

Soit \bar{F} la fonction de queue (classiquement nommée fonction de survie) bivariées associée, définie par : $\bar{F}(x, y) = P(X \geq x, Y \geq y)$ pour tout $(x, y) \in R_+^2$

Alors, X et Y sont **dépendants en quadrant positif** si :

$$P(X \geq x, Y \geq y) \geq P(X \geq x) \cdot P(Y \geq y) \text{ pour tout } (x, y) \in R_+^2$$

soit, à l'aide des fonctions de queue jointe et marginales :

$$\bar{F}(x, y) \geq \bar{F}(x) \cdot \bar{F}(y) \text{ pour tout } (x, y) \in R_+^2$$

Par ailleurs, X et Y sont **dépendants en quadrant négatif** si :

$$P(X \geq x, Y \geq y) \leq P(X \geq x) \cdot P(Y \geq y) \text{ pour tout } (x, y) \in R_+^2$$

soit, à l'aide des fonctions de queue jointe et marginales :

$$\bar{F}(x, y) \leq \bar{F}(x) \cdot \bar{F}(y) \text{ pour tout } (x, y) \in R_+^2$$

Remarque :

Si X et Y sont dépendants en quadrant positif, on a : $\text{Cov}(X, Y) \geq 0$, si cette covariance existe

Démonstration :

$$\text{Cov}(X, Y) = E(XY) - E(X) \cdot E(Y)$$

$$= \int_0^{\infty} \int_0^{\infty} \bar{F}(x, y) \cdot dx \cdot dy - \int_0^{\infty} \bar{F}(x) \cdot dx \times \int_0^{\infty} \bar{F}(y) \cdot dy$$

$$= \int_0^{\infty} \int_0^{\infty} [\overline{F}(x, y) - \overline{F}(x) \cdot \overline{F}(y)] dx \cdot dy \geq 0$$

II.4. Quantification de la dépendance

Nous allons maintenant chercher à quantifier cette notion de dépendance. Pour ce faire, dans un premier temps, nous regarderons les propriétés que doit respecter une mesure de dépendance.

Puis, nous étudierons les principaux indicateurs existants.

II.4.1. Mesure de dépendance

L'objet de ce paragraphe est de définir ce qu'est une mesure de dépendance et de présenter les mesures les plus couramment utilisées. Une **mesure de dépendance** est une application qui associe à deux variables aléatoires un réel permettant de quantifier la force de la dépendance qui lie les deux variables aléatoires.

Pour être jugée satisfaisante, une mesure de dépendance doit satisfaire les cinq conditions suivantes :

- (a) Symétrie : $\delta(X, Y) = \delta(Y, X)$ pour tout (X, Y)
- (b) Normée : $|\delta(X, Y)| \leq 1$ pour tout (X, Y) (cette propriété étant inspirée par le coefficient de corrélation linéaire)
- (c) $\delta(X, Y) = 1$ ssi (X, Y) est comonotone
 $\delta(X, Y) = -1$ ssi (X, Y) est antimonotone
- (d) si α est une fonction réelle strictement croissante
 $\delta[\alpha(X), \alpha(Y)] = \delta(X, Y)$ pour tout (X, Y)
- (e) X et Y indépendantes implique $\delta(X, Y) = 0$

Ainsi, par exemple, nous pouvons voir que le coefficient de corrélation linéaire de Pearson ne vérifie que (a), (b) et (e). De ce fait, il ne peut être considéré comme une mesure de dépendance.

II.4.2. Les principaux indicateurs de dépendance

II.4.2.1. La notion de concordance

La notion de concordance sera utilisée pour définir le Tau de Kendall dans le chapitre suivant.

Définition :

Les couples (X,Y) et (X',Y') sont en **concordance** si on a :
 $(X < X' \text{ et } Y < Y')$ ou $(X' < X \text{ et } Y' < Y)$

Ils sont en **discordance** si
 $(X < X' \text{ et } Y' > Y)$ ou $(X' > X \text{ et } Y < Y')$

Les probabilités de concordance puis de discordance s'écrivent :

$$\pi_c = P(X < X', Y < Y') + P(X' < X, Y' < Y)$$

$$\pi_d = P(X < X', Y' > Y) + P(X' > X, Y < Y')$$

Nous pouvons d'ailleurs remarquer que $\pi_c + \pi_d = 1$

II.4.2.2. Tau de Kendall

Définition :

La mesure Tau de Kendall $\tau(X,Y)$ est définie par :

$$\tau(X,Y) = \pi_c - \pi_d$$

Cette mesure peut également s'écrire sous la forme développée suivante :

$$\tau(X,Y) = P[(X-X')(Y-Y') > 0] - P[(X-X')(Y-Y') < 0]$$

où (X',Y') est un vecteur aléatoire identique à (X,Y)

Nous pouvons noter que cette mesure est utilisable pour tout couple (X,Y) sans restriction.

τ vérifie les propriétés (a), (b), (c), (d) et (e) d'une mesure de dépendance.

De plus, τ est indépendante des marginales. Plus précisément, étant donné deux fonctions de répartition F_X et F_Y et τ dans $[-1,1]$, on peut trouver une loi jointe produisant ces marginales et cette valeur de τ .

II.4.2.3. Rho de Spearman

Définition :

Le ρ_s de Spearman entre deux variables aléatoires X, Y est défini par :

$$\rho_s(X,Y) = 3 (P[(X-X')(Y-Y^*) > 0] - P[(X-X')(Y-Y^*) < 0])$$

où (X',Y') et (X^*,Y^*) sont des vecteurs aléatoires identiques à (X,Y)

Cette mesure peut également être exprimée ainsi :

$$\rho_s = \rho(F_X(X), F_Y(Y))$$

Il est donc égal au coefficient de corrélation linéaire de Pearson appliqué aux fonctions de répartition de X et Y .

Tout comme le Tau de Kendall, le ρ_s de Spearman les propriétés (a), (b), (c), (d) et (e) d'une mesure de dépendance.

II.4.3. Les indicateurs de dépendance de queue

La **Tail Dependence** ou **dépendance de queue** décrit la probabilité limite qu'une variable aléatoire excède un certain seuil sachant qu'une autre variable aléatoire a déjà excédé ce seuil.

Cette notion est particulièrement utile pour toute considération relative à la VaR (Value at Risk) qui s'intéresse aux quantiles (par exemple à 99.5%)

Définition :

On dit que (X, Y) est **upper tail dependent** si :

$$\lambda_U = \lim_{u \rightarrow 1} P(X > F_X^{-1}(u) / Y > F_Y^{-1}(u)) > 0$$

De manière similaire, on dit que (X, Y) est **lower tail dependent** si :

$$\lambda_L = \lim_{u \rightarrow 0} P(X < F_X^{-1}(u) / Y < F_Y^{-1}(u)) > 0$$

où F_X^{-1} et F_Y^{-1} correspondent aux inverses généralisées des fonctions de répartition de X et Y respectivement.

λ_U et λ_L sont appelés les **tail dependent coefficients**.

Egalement, afin d'étudier la dépendance de queue, il est intéressant d'étudier les tail concentration functions dont la définition est la suivante :

Définition :

La **upper tail concentration function** est la fonction R définie par :

$$\begin{aligned} R(z) &= P(X > F_X^{-1}(z) / Y > F_Y^{-1}(z)) \\ &= P(U > z / V > z) \end{aligned}$$

La **lower tail concentration function** est la fonction R définie par :

$$\begin{aligned} L(z) &= P(X < F_X^{-1}(z) / Y < F_Y^{-1}(z)) \\ &= P(U < z / V < z) \end{aligned}$$

avec U, V respectivement les fonctions quantiles de X et Y

II.5. Les différents types de modélisation de la dépendance

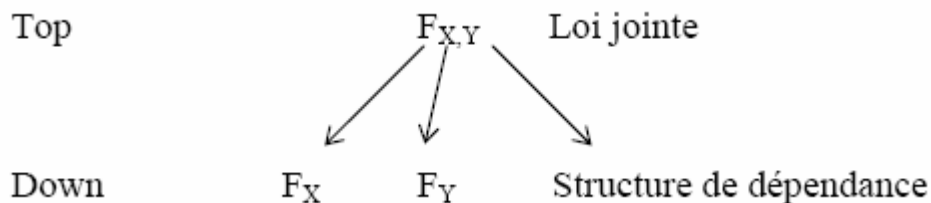
II.5.1. Les approches

Traditionnellement, deux approches existent pour prendre en compte la dépendance entre deux variables de risques :

- l'approche « Top-down »
- l'approche « Bottom-up »

II.5.1.1. L'approche « Top-down »

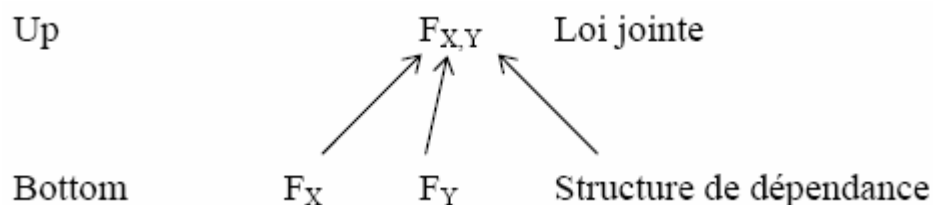
L'approche « Top-down » est l'approche standard : elle consiste en la donnée directe de la loi jointe $F_{X,Y}$; les deux lois marginales F_X , F_Y et la structure de dépendance s'en déduisent. Il s'agit donc de la construction de modèles bivariés, comme le montre le schéma ci-dessous :



Les modèles à chocs communs sont traditionnellement classés comme étant une approche « Top-Down ».

II.5.1.2. L'approche « Bottom-up »

Dans cette approche, on part des marginales et de la structure de dépendance pour obtenir la loi bivariee :



Les modèles à base de copules sont traditionnellement classés comme étant une approche « Bottom-Up ».

II.5.2. Les chocs communs

II.5.2.1. Introduction au modèle de chocs communs

Un choc commun permet de faire apparaître dans le modèle un type de dépendance qui peut être :

- Un évènement/accident commun qui touche plusieurs types de garanties :
 - o Un accident de voiture peut causer un sinistre pour l'assurance RC automobile, l'assurance dommage automobile et l'assurance incapacité/invalidité
 - o Un sinistre incendie peut mobiliser les garanties risques direct et perte d'exploitation d'un contrat Risques industriels
 - o Un évènement naturel (Tempêtes de 1999, inondations, tremblements de terre) impactera les contrats Dommages auto et Multirisques habitation
- Un facteur commun impactant plusieurs branches :
 - o L'influence du temps sur l'assurance santé et sur l'assurance agricole. Dans ce cas, les sinistres semblent apparaître indépendamment, alors que le temps est le facteur commun
 - o L'influence des contrôles de police. En France, le développement des radars a diminué le nombre d'accidents ainsi que le nombre de blessés graves sur la route. La fréquence des contrôles peut varier suivant le temps et crée une corrélation temporelle.
- Une modification touchant par exemple le montant des indemnités dans le cadre de sinistres corporels comme le coût de la Tierce Personne et qui affectera donc

toutes les branches ministérielles impactées par des sinistres corporels comme la RC Corporelle Automobile, les Dommages Corporels ou les sinistres RC Générale corporels (comme les sinistres RC Médicale)

- Etc ...

II.5.2.2. Définition d'un modèle de chocs communs

Un modèle basé sur un choc commun suppose que les variables aléatoires modélisant la sinistralité des branches de l'assureur peuvent être corrélées par le biais d'un même paramètre (ou de plusieurs paramètres).

Ainsi, le modèle avec k chocs et n branches s'écrira :

$$X_i = X_{i0} \oplus \delta_{i1}C_1 \oplus \dots \oplus \delta_{ij}C_j \oplus \dots \oplus \delta_{ik}C_k, \text{ pour } j=1, \dots, k \text{ et } i=1, \dots, n$$

Avec :

- Les X_{i0} sont indépendants deux à deux pour tout $i=1, \dots, n$
- Les X_{i0} et les C_j sont indépendants deux à deux pour tout $i=1, \dots, n$ et tout $j=1, \dots, k$
- Les C_j sont indépendants deux à deux pour tout $j=1, \dots, k$

Dans ce modèle :

- Il y a n branches et X_i est la sinistralité de la i -ième branche
- Il y a k chocs et C_j est le j -ième choc du modèle
- $\delta_{ij} = \{0,1\}$ indique si le j -ième choc impacte la i -ième branche. De ce fait, le j -ième choc, impacte $\sum_{i=1}^n \delta_{ij}$ branches

Pour illustrer cette définition générale, prenons le cas d'un modèle avec un seul choc et deux branches. Le choc commun sera multiplicatif.

Le modèle est alors le suivant :

- la branche 1 a pour sinistralité : $X_1 = \beta \times X_{10}$
- la branche 2 a pour sinistralité : $X_2 = \beta \times X_{20}$
- le choc commun est : β

Avec :

- X_{10} , X_{20} et β trois variables aléatoires indépendantes

Dans cet exemple, les deux branches sont liées par le choc commun β .

II.5.2.3. Les différents types de chocs communs

Les modèles à choc commun peuvent être découpés en plusieurs familles, suivant la manière dont le choc impacte les marginales. Voici les trois familles généralement identifiées :

- les modèles à choc additifs : la sinistralité d'une branche est la somme d'une composante indépendante et d'un choc stochastique
- les modèles à choc multiplicatif : la sinistralité d'une branche est le produit d'une composante indépendante et d'un choc stochastique
- les modèles à choc sur les paramètres : la sinistralité d'une branche est une variable aléatoire dont un des paramètres (par exemple, l'espérance) est déformé par le choc commun

Par ailleurs, le choc est généralement représenté par une variable aléatoire qui peut être :

- une variable continue, ayant pour loi :
 - o une loi normale
 - o une loi lognormale
 - o une loi lognormale décalée (« shifted log-normal » en anglais ; si X est une loi log-normale et c une constante, X+c est une loi lognormale shiftée)
 - o une loi gamma
 - o tout autre type de loi : loi de weibull, loi de pareto, ...
- une variable discrète, ayant pour loi :
 - o multinomiale
 - o poisson
 - o multinomiale négative, etc...

Comme vu précédemment, le choc s'applique à la sinistralité entre deux ou plusieurs branches. Il peut donc être appliqué à toute variable, en particulier au :

- nombre de sinistre (ou de la fréquence moyenne)
- coût d'un sinistre
- ou à un niveau plus global, celui de la sinistralité totale, si l'on modélise par exemple la loi S/P

De plus, il est possible d'appliquer un choc commun au sein d'une même branche. Dans ce cas, le choc ne sera pas commun à différentes branches, mais il sera commun aux sinistralités individuelles des polices d'assurance d'une même branche.

De surcroît, il est possible de réaliser un choc commun entre la sinistralité attritionnelle et les sinistres graves.

Assez souvent, les ouvrages traitant des modèles de chocs communs se concentrent sur le choc commun additif portant sur le nombre de sinistres à l'aide d'une loi de Poisson.

Le mémoire se concentrera sur des modèles de chocs appliqués au niveau de la sinistralité totale entre deux ou plusieurs branches. Il s'agira donc de chocs modélisés à l'aide de lois continues et non discrètes.

Il explorera les modèles suivants :

- choc multiplicatif log-normal
- choc multiplicatif log-normal shifté

A l'aide de ces deux cas, nous pourrons étudier un spectre de possibilités assez large, et ainsi identifier des exemples de :

- Chocs symétriques : Un choc symétrique est un choc qui impacte autant la sinistralité faible que la sinistralité forte, alors qu'un choc asymétrique les impacte asymétriquement
- Chocs asymétriques : un choc asymétrique est un choc qui ne remplit pas les conditions d'un choc symétrique

II.5.2.4. Une illustration par un exemple de chocs communs

Considérons le cas d'un choc commun multiplicatif sur deux branches d'assurance :

- la branche 1 a pour sinistralité : $X_1 = \beta \times X_{10}$
- la branche 2 a pour sinistralité : $X_2 = \beta \times X_{20}$
- le choc commun est : β

De plus, les variables aléatoires suivent des lois log-normale :

- $X_{10} \sim \text{LN}(\mu_1, \sigma_1)$
- $X_{20} \sim \text{LN}(\mu_2, \sigma_2)$
- $\beta \sim \text{LN}(\mu_c, \sigma_c)$

$\text{LN}(\mu, \sigma)$ désigne la loi log-normale, et les paramètres précédemment cités sont choisis tels que :

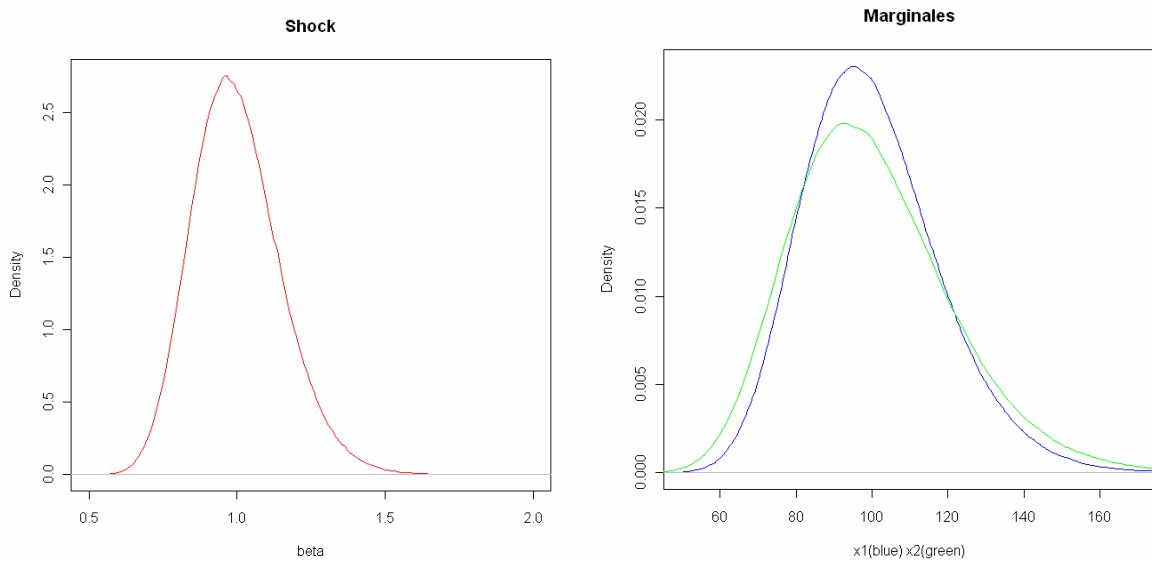
- $E[X_{10}] = 100$ et $\text{Var}(X_{10}) = 10^2$
- $E[X_{20}] = 100$ et $\text{Var}(X_{20}) = 15^2$
- $E[\beta] = 1$ et $\text{Var}(\beta) = 0.2^2$

Nous pouvons remarquer si nous appliquons la fonction logarithme sur ce modèle, nous retombons sur un modèle additif gaussien.

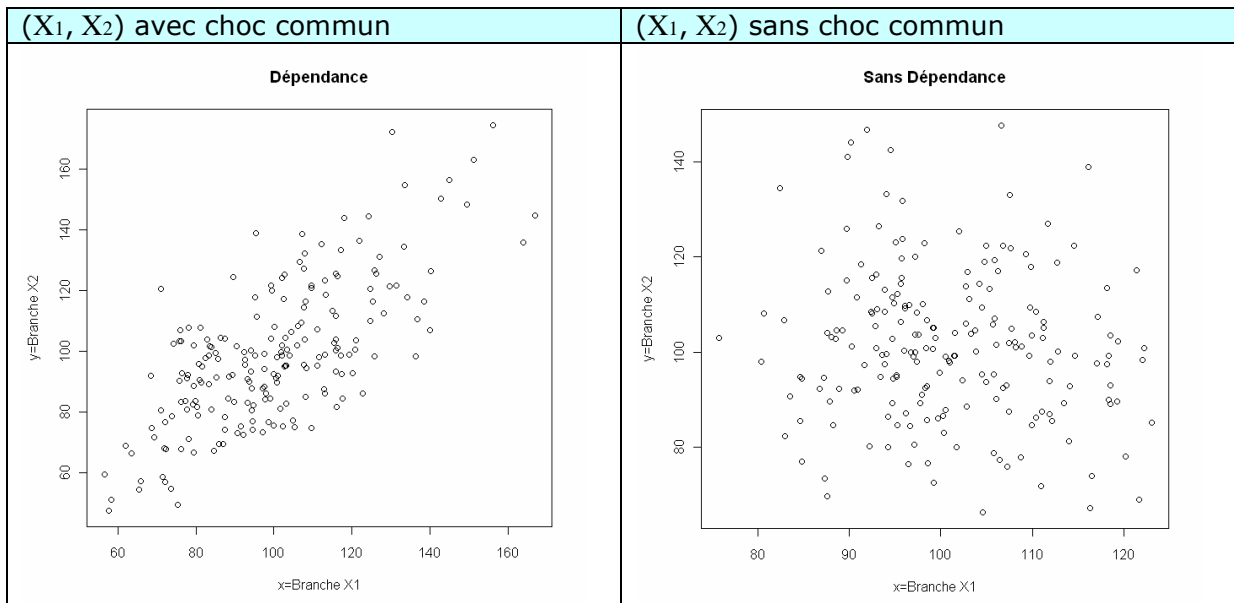
D'après la propriété de multiplicativité des lois log-normales, nous pouvons facilement déterminer les marginales obtenues : elles suivent les lois suivantes :

- $X_1 \sim \text{LN}(\mu_1 + \mu_c, \sqrt{\sigma_1^2 + \sigma_c^2})$
- $X_2 \sim \text{LN}(\mu_2 + \mu_c, \sqrt{\sigma_2^2 + \sigma_c^2})$

Voici les graphes des fonctions de densité du choc commun et des marginales :



Regardons maintenant le couple (X_1, X_2) . Le graphe suivant représente 200 réalisations du couple (X_1, X_2) . Il est comparé au couple qui aurait été obtenu sans choc commun, c'est-à-dire dans le cas où β serait une constante égale à 1.



Clairement, nous voyons que le choc commun tend à aligner les points suivant la première bissectrice. En effet, le choc β tend à rapprocher X_1 et X_2 l'un de l'autre. Au contraire, en l'absence de choc, les points ont d'avantage tendance à occuper tout le cadran et non à se concentrer autour de la première bissectrice.

Avec cette analyse graphique, nous pouvons conclure que le choc commun tend donc bien à lier la sinistralité des deux branches, et à y introduire une dépendance.

II.5.3. Les copules

II.5.3.1. Présentation générale des copules

Le concept des copules a été introduit en 1959 par Abe Sklar. Les copules sont devenues un élément essentiel dans la modélisation des distributions multivariées en finance et en assurance. Si la distribution normale multivariée a souvent servi de modèle, c'est pour la simplicité relative des calculs et le fait que la dépendance est entièrement déterminée par le coefficient de corrélation. Malheureusement la loi normale s'ajuste mal à certains problèmes en assurance.

Dans la théorie des copules, si on considère un vecteur aléatoire $X=(X_1, X_2, \dots, X_d)$, on considère que toute l'information à propos de la structure de dépendance de X est contenue dans une fonction appelée la copule. Ce concept a l'avantage d'être un outil puissant et flexible dans la mesure où il permet de modéliser la dépendance sans tenir compte de l'effet du comportement des marginales, c'est-à-dire des fonctions de répartition F_1, F_2, \dots, F_d des variables aléatoires X_1, X_2, \dots, X_d prises individuellement.

Pour résumer : une copule est une fonction qui lie des marginales pour générer la loi jointe ; et l'idée de base est de séparer la structure de dépendance (représentée par la copule) des marginales.

II.5.3.2. Définition des copules

Une **copule C de dimension d** est une fonction de répartition multivariée définie sur $[0;1]^d$ et ayant des lois marginales uniformes.

Ainsi,

$$C : [0;1]^d \rightarrow [0;1]$$

$$(u_1, u_2, \dots, u_d) \mapsto P(U_1 \leq u_1, U_2 \leq u_2, \dots, U_d \leq u_d)$$

Puisque C est une fonction de répartition, il en découle les propriétés suivantes :

- $C(u_1, u_2, \dots, u_d)$ est croissante en chaque composante u_i
- $C(1, \dots, 1, u_i, 1, \dots, 1) = u_i$ pour $1 \leq u_i \leq d$ et $u_i \in [0;1]$
- la probabilité $P(U_1 \in [a_1, b_1], \dots, U_d \in [a_d, b_d])$ ne doit pas être négative pour tous $(a_1, a_2, \dots, a_d), (b_1, b_2, \dots, b_d)$ dans $[0;1]^d$ tels que $a_i \leq b_i$:

$$\sum_{i_1=1}^2 \dots \sum_{i_d=1}^2 (-1)^{i_1 + \dots + i_d} \times C(u_{1,i_1}, \dots, u_{1,i_d}) \geq 0$$

$$\text{où } u_{j,1} = a_j \text{ et } u_{j,2} = b_j$$

Réciproquement, une fonction vérifiant ces trois propriétés est une copule.

II.5.3.3. Théorème de Sklar

Théorème de Sklar :

Soit F une fonction de répartition de dimension d , et dont les marginales F_1, F_2, \dots, F_d sont continues. Alors, il existe une copule unique telle que :

$$F(x_1, x_2, \dots, x_d) = C(F_1(x_1), F_2(x_2), \dots, F_d(x_d))$$

Remarque :

Lorsque les marginales ne sont pas continues, il est toujours possible de définir une copule, mais celle-ci n'est plus unique et perd beaucoup d'intérêt. On peut en effet toujours poser :

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

avec F_i^{-1} qui est l'inverse de la fonction de répartition de la i -ième marginale :

$$F_i^{-1}(u) = \inf\{t \mid F_i(t) \geq u\}$$

Ce résultat est important en pratique pour les applications à l'assurance et à la finance dans le sens où il indique que dans le cadre d'une analyse d'une problématique multivariée, celle-ci peut être faite en 2 étapes :

- l'identification des distributions marginales
- l'analyse de la structure de dépendance entre les composantes à l'aide de la copule

II.5.3.4. Théorème d'invariance

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

Théorème : Soient deux variables aléatoires continues X_1 et X_2 ayant comme marginales F_1 et F_2 et ayant une copule, associée à la distribution F du vecteur aléatoire $X = (X_1, X_2)$, que nous noterons $C_{\langle X_1, X_2 \rangle}$. Si h_1 et h_2 sont deux fonctions strictement croissantes sur $\text{Im}(X_1)$ et $\text{Im}(X_2)$ respectivement, alors :

$$C_{\langle h_1(X_1), h_2(X_2) \rangle} = C_{\langle X_1, X_2 \rangle}$$

La copule est invariante par transformations strictement croissantes des variables aléatoires.

II.5.3.5. Densité d'une copule

Si elle existe, la densité d'une copule se définit par :

$$c(u_1, u_2, \dots, u_d) = \frac{\partial C(u_1, u_2, \dots, u_d)}{\partial u_1 \partial u_2 \dots \partial u_d}$$

A partir du théorème de Sklar, on peut donc en déduire la densité multivariée des variables X_1, X_2, \dots, X_d :

$$F(x_1, x_2, \dots, x_d) = C(F_1(x_1), F_2(x_2), \dots, F_d(x_d))$$

On peut donc en déduire la densité f de la variable aléatoire X :

$$f(x_1, x_2, \dots, x_d) = c(F_1(x_1), F_2(x_2), \dots, F_d(x_d)) \times \prod_{i=1}^d f_i(x_i)$$

II.5.3.6. Les copules usuelles

Il existe un certain nombre de copules usuelles qui apparaissent dans la littérature actuarielle. Le lecteur est invité à se reporter à l'annexe : '[VIII.1. Les copules de référence](#)' pour y trouver un descriptif des copules suivantes :

- Copule d'indépendance
- Copule gaussienne
- Copule de Gumbel
- Copule de Clayton
- Copule HRT

II.6. Conclusion

Dans cette première partie, nous avons vu l'intérêt de modéliser la dépendance, ainsi que les différentes méthodes de modélisation : à savoir une approche à base de copules ou une approche avec les chocs communs.

Dans la littérature actuarielle, les modèles avec copules ont été analysés à de nombreuses reprises ; de plus, les copules présentent un moyen efficace de visualiser la dépendance existante entre deux branches.

Aussi, le présent mémoire qui est centré sur les chocs communs fera souvent une mise en parallèle avec les copules, dans le but de s'appuyer dès que possible sur la théorie existante des copules, et dans l'objectif de fournir un élément tangible de comparaison.

Enfin, cette première partie a été l'occasion d'introduire les différentes mesures de dépendance, en particulier le Tau de Kendall, le Rho de Spearman et les indicateurs de dépendance : upper tail concentration et lower tail concentration.

III. METHODES D'ANALYSE DE LA DEPENDANCE INDUITE PAR UN MODELE DE CHOCS COMMUNS

Cette troisième partie a pour but d'introduire la méthode et les axes d'analyse communs qui seront utilisés lors de l'étude des différents types de chocs communs en quatrième partie.

III.1. Présentation des axes d'analyse

Lors de l'exploration des différents modèles de chocs communs, il sera question de :

1. déterminer les marginales obtenues suite à l'application du choc
2. déterminer qualitativement la structure de dépendance obtenue
3. évaluer des indicateurs de dépendance et leur sensibilité aux paramètres du modèle
4. mettre en parallèle la dépendance induite par le choc commun avec les modèles à base de copules

Cette étape exploratrice sera en général réalisée à l'aide d'une approche théorique mathématique et également avec le support d'outils de simulation numérique.

En effet, une approche théorique peut donner lieu à des formules peu facilement exploitables ou interprétables, et de fait, nécessiter un recours à la simulation numérique.

Par ailleurs, coupler approche théorique avec méthode de simulation numérique permet de faire une vérification significative des résultats obtenus.

Aussi, avant de nous pencher sur les 4 axes d'analyse listés plus haut, nous présenterons dans un premier temps la méthode de simulation de Monte-Carlo permettant de faire des simulations numériques.

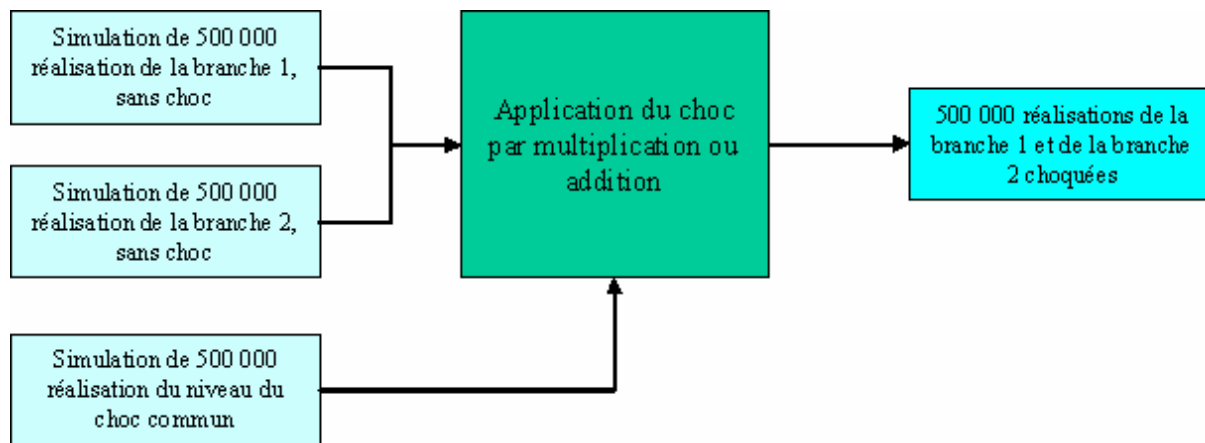
III.2. Méthode de simulation de Monte-Carlo

Le logiciel de simulation utilisé est le logiciel « R » qui permet de concevoir et d'utiliser des fonctions statistiques tout en les imbriquant dans des programmes. Il permet donc de réaliser des simulations de Monte-Carlo et grâce à une programmation algorithmique, il rend possible l'estimation de copules empiriques ou des valeurs d'indicateurs de dépendance.

Les temps de calcul sont tout à fait convenables : dans l'exemple du choc multiplicatif log-normal, le programme tourne en moins de 10 secondes pour une simulation avec 500 000 échantillons.

Le logiciel « R » permet de générer facilement des réalisations de variables aléatoires suivant les lois les plus courantes.

Regardons maintenant le processus de simulation de 500 000 échantillons d'un modèle à chocs communs (qui peut s'appliquer tant à un modèle à choc commun multiplicatif qu'additif) :



III.3. Détermination des marginales obtenues suite à l'application du choc

Nous utiliserons des lois statistiques courantes pour modéliser le choc et les marginales non choquées.

Après, application du choc, il faut tout d'abord déterminer les marginales obtenues suite à l'application du choc ; en effet, il n'y a aucune raison pour que les marginales obtenues après le choc suivent le même type de loi qu'elles suivaient avant le choc.

Approche théorique :

La plupart des chocs permettent une obtention assez aisée des lois des marginales :

- une loi normale déformée additivement par un choc normal donne une loi normale
- il en est de même pour les lois gamma et inverse gaussien
- une loi log-normale déformée multiplicativement par un choc log-normal donne une loi log-normale

Toutefois, certains chocs nécessiteront une étude approfondie :

- La somme de deux lois aléatoire peut être définie à l'aide du produit de convolution pour l'estimation des fonctions de densité
- Le produit de deux variables aléatoires peut être défini par la formule suivante :

$$f(x) = \int_{-\infty}^{+\infty} f_1(u) f_2(x/u) \cdot du/u \text{ où } f \text{ est la fonction de densité du produit de deux}$$

lois qui ont pour fonction de densité f_1 et f_2

Il est bon de remarquer que des travaux ont déjà été menés pour certains de ces cas : par exemple, des études ont déjà été réalisées concernant la somme de deux lois log-normales (ce qui sera utile pour le produit d'une log-normale avec une shifted log-normale, dans le cas d'un choc de type shifted log-normal)

Enfin, dans le cas où la marginale (via sa fonction de répartition ou de densité) ne pourrait être exprimée plus simplement que sous forme d'intégrale complexe, il est toujours possible de la visualiser via un calcul numérique d'intégrale.

Approche par la simulation numérique :

Grâce à la simulation de Monte-Carlo, il est possible d'estimer la fonction de répartition empirique.

Pour rappel, La fonction de distribution empirique $F_n(x)$ basée sur l'échantillon (x_1, x_2, \dots, x_n) est une fonction en escalier définie ainsi :

$$F_n(x) = \frac{\text{Nombre_éléments_dans_échantillon} \leq x}{n} = \frac{1}{n} \cdot \sum_{i=1}^n I(x_i \leq x)$$

avec :

- $I(A)$ est la fonction indicatrice de l'événement A .

D'après :

- la loi forte des grands nombres qui affirme que :
 - o $F_n(x) \rightarrow F(x)$ presque sûrement pour x fixé
 - o $F_n(x)$ est un estimateur non biaisé de la fonction de répartition $F(x)$
- le théorème de la limite centrale qui dit que :
 - o $\sqrt{n} \cdot (F_n(x) - F(x))$ converge vers une loi normale $N(0, F(x)(1-F(x)))$ pour un x fixé

Et étant donné :

- le nombre élevé d'échantillons simulés (ici n) : le programme écrit sous « R » permet de générer facilement plus de 500 000 échantillons en un temps tout à fait correct inférieur à 10 secondes

Nous pouvons conclure que le recours à la fonction de répartition empirique est un indicateur fiable.

III.4. Description qualitative de la structure de dépendance induite par le choc commun

Une fois les marginales identifiées, il convient alors d'étudier la structure de dépendance induite par le choc commun.

III.4.1. Visualisation suite à la « normalisation » des données

Reprenons l'exemple du chapitre [« II.5.2.4. Une illustration par un exemple de chocs communs »](#)

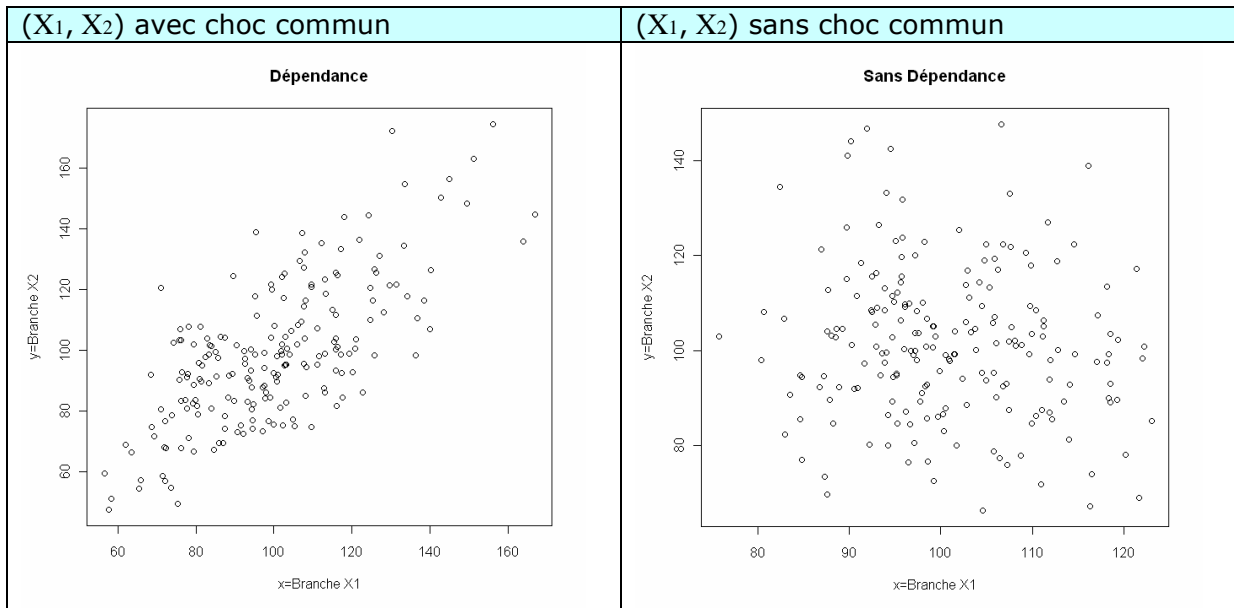
Pour rappel, dans cet exemple, il y a deux branches et :

- la branche 1 a pour sinistralité : $X_1 = \beta \times X_{10}$
- la branche 2 a pour sinistralité : $X_2 = \beta \times X_{20}$
- le choc commun est : β

avec :

- $E[X_{10}] = 100$ et $\text{Var}(X_{10}) = 10^2$
- $E[X_{20}] = 100$ et $\text{Var}(X_{20}) = 15^2$
- $E[\beta] = 1$ et $\text{Var}(\beta) = 0.2^2$

Le couple (X_1, X_2) représente la sinistralité de deux branches. Le graphe suivant représente 200 réalisations du couple (X_1, X_2) . Il est comparé au couple qui aurait été obtenu sans choc commun, c'est-à-dire dans le cas où β serait une constante égale à 1.

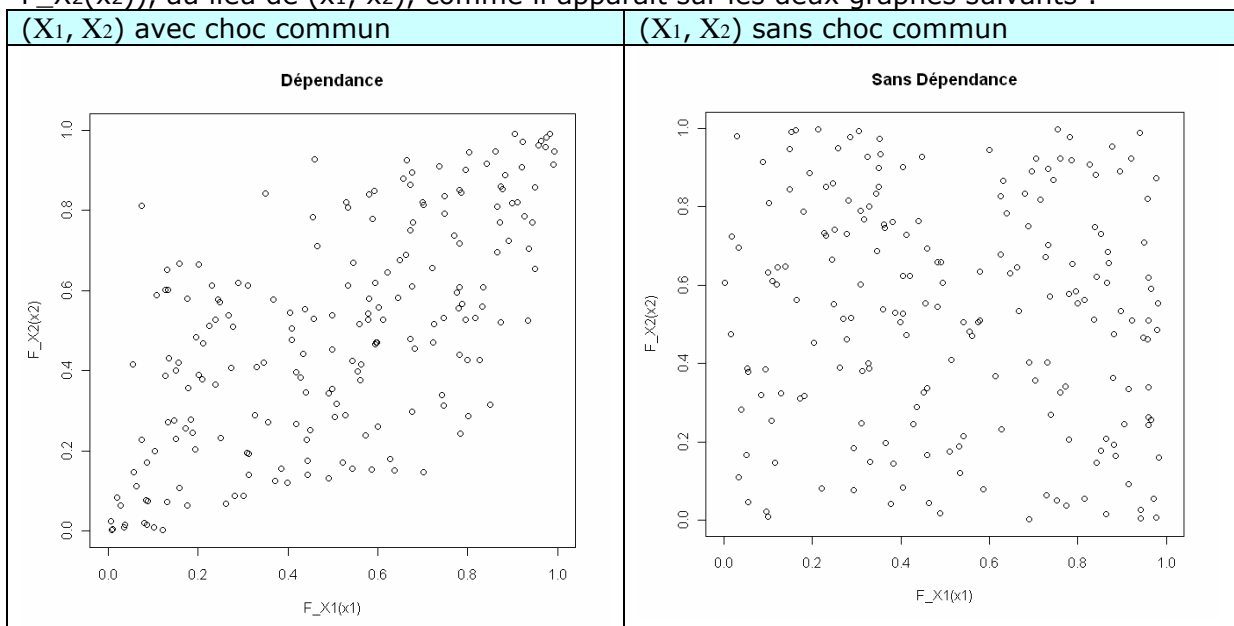


Afin de mieux visualiser la dépendance induite par le choc commun, nous allons normaliser les graphes. Cette « normalisation » sera indispensable pour pouvoir ensuite comparer différents types de dépendances induites par différents chocs communs. Aussi, nous allons afficher chaque point non pas en fonction du montant de la sinistralité, mais en fonction de son quantile (par rapport à la fonction de répartition de la marginale). Ainsi, nous obtiendrons des graphes tenant sur un carré $[0,1] \times [0,1]$.

Pour réaliser ce type de graphe « normalisé », il faut avoir les fonctions de répartition de marginales, qui sont notées :

- F_{X_1} pour la branche 1
- F_{X_2} pour la branche 2

Puis, il faudra afficher les points de coordonnées des 200 échantillons ($F_{X_1}(x_1)$, $F_{X_2}(x_2)$), au lieu de (x_1, x_2) , comme il apparaît sur les deux graphes suivants :



Les remarques émises précédemment se retrouvent sur ce type de graphe. Dans le cas d'indépendance, les points représentant les 200 simulations occupent uniformément le plan $[0,1] \times [0,1]$.

Au contraire, en présence de choc, ils délaissent le cadran gauche supérieur et le cadran droit inférieure pour se concentrer sur la première bissectrice.

Autrement dit, quand la branche 1 enregistre une forte ou une faible sinistralité, la dépendance induite par le choc commun tend à faire en sorte que la branche 2 enregistre également une forte ou une faible sinistralité (respectivement).

III.4.2. Visualisation de la structure de dépendance à l'aide de copule empirique

Les copules permettent d'explicitier la structure de dépendance entre deux variables aléatoires, et ce indépendamment de leurs marginales.

De ce fait, il est intéressant d'essayer de rapprocher les résultats obtenus suite à l'application d'un choc commun avec la théorie des copules. Aussi, nous allons essayer de déterminer la copule empirique correspondante au choc commun, suivant les deux approches suivantes :

- via la formule théorique de la copule empirique
- via la simulation numérique de Monte-Carlo

III.4.2.1. Formule théorique de la copule empirique naturelle

En nous appuyant sur le théorème de Sarkl (cf [chapitre II.5.3.3. Théorème de Sklar](#)), nous pouvons expliciter une copule empirique avec la relation suivante

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

avec F qui est la fonction de répartition jointe et F_i^{-1} qui est l'inverse de la fonction de répartition de la i -ième marginale :

$$F_i^{-1}(u) = \inf \{t \mid F_i(t) \geq u\}$$

Dans le cas d'un rapprochement d'un modèle à choc commun avec la théorie des copules, cette formule ne sera pas toujours facilement utilisable. En effet, elle nécessite de connaître et d'avoir pu expliciter la fonction de répartition jointe, ainsi que l'inverse des fonctions de répartition des marginales.

III.4.2.2. Construction de la copule empirique de Deheuvels

Dans le cas où la formule théorique explicitant la copule empirique naturelle n'est pas facilement utilisable, nous utiliserons la copule empirique de Deheuvels.

Définition de la copule empirique de Deheuvels :

Considérons le cas d'une distribution multivariée de dimension n et pour laquelle nous avons j observations.

Soit $\{r_1^j, r_2^j, \dots, r_n^j\}$ la statistique de rang associée au j -ième échantillon multivarié $\{x_1^j, x_2^j, \dots, x_n^j\}$ avec $1 \leq j \leq J$; autrement dit :

$$\forall i \in [1; n], r_i^j \text{ est le rang de } x_i^j \text{ parmi } (x_i^j)_{j=1, \dots, J}$$

Alors, toute copule \hat{C} définie sur le treillis $\left\{ \left(\frac{t_1}{T}, \frac{t_2}{T}, \dots, \frac{t_n}{T} \right) \mid 1 \leq k \leq n, 0 \leq t_k \leq T \right\}$ par :

$$\hat{C} \left(\frac{t_1}{T}, \frac{t_2}{T}, \dots, \frac{t_n}{T} \right) = \frac{1}{J} \cdot \sum_{j=1}^J \prod_{k=1}^n \mathbf{1}_{\{r_k^j / J \leq t_k / T\}}$$

est une copule empirique.

Dans le cas d'une distribution bivariable, la définition peut être simplifiée comme suit :

$$\hat{C} \left(\frac{t_1}{T}, \frac{t_2}{T} \right) = \frac{1}{J} \cdot \sum_{j=1}^J \mathbf{1}_{\{r_1^j / J \leq t_1 / T\}} \cdot \mathbf{1}_{\{r_2^j / J \leq t_2 / T\}}$$

Remarque :

Nous pouvons remarquer que la définition est particulièrement adaptée, dans le cas de données issues d'une simulation numérique.

En effet, si l'on note \hat{C}_J la copule empirique d'ordre J , où J représente la taille de l'échantillon utilisé pour sa construction, alors \hat{C}_J converge asymptotiquement vers C (voir Faivre [2002])

Exemple :

Reprenons l'exemple introduit dans le chapitre « [II.5.2.4. Une illustration par un exemple de chocs communs](#) » puis repris au chapitre « [III.4.1. Visualisation suite à la « normalisation » des données](#) »

Pour rappel, dans cet exemple, il y a deux branches et :

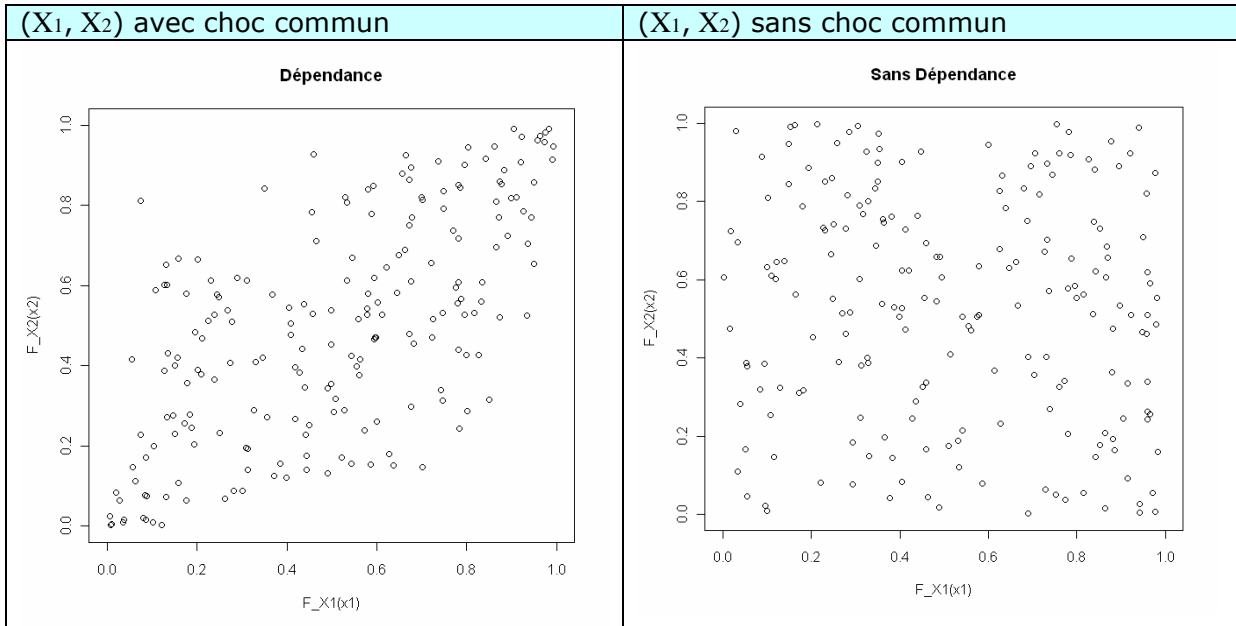
- la branche 1 a pour sinistralité : $X_1 = \beta \times X_{10}$
- la branche 2 a pour sinistralité : $X_2 = \beta \times X_{20}$
- le choc commun est : β

avec :

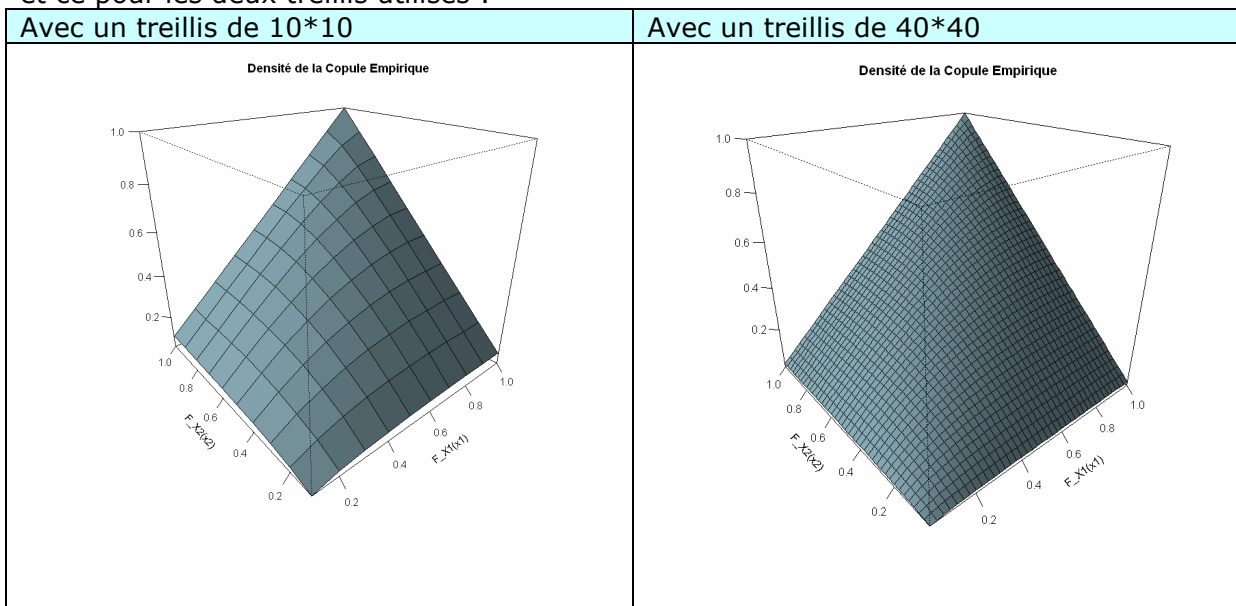
- $E[X_{10}] = 100$ et $\text{Var}(X_{10}) = 10^2$
- $E[X_{20}] = 100$ et $\text{Var}(X_{20}) = 15^2$
- $E[\beta] = 1$ et $\text{Var}(\beta) = 0.2^2$

Ci-dessous, les graphes normalisés affichant les 200 échantillons ($F_{X_1}(x_1)$, $F_{X_2}(x_2)$) avec :

- F_{X_1} étant la fonction de répartition de la marginale de la branche 1
- F_{X_2} étant la fonction de répartition de la marginale de la branche 2



Pour l'exemple considéré, nous obtenons les copules empiriques de Deheuvels suivantes, et ce pour les deux treillis utilisés :



Dans les deux cas, les graphes ne semblent guère facilement interprétables.

Il est probable que les fonctions de densité des copules nous apportent des résultats plus faciles à analyser. Pour ce faire, considérons la définition de la densité d'une copule :

$$c(u_1, u_2, \dots, u_d) = \frac{\partial C(u_1, u_2, \dots, u_d)}{\partial u_1 \partial u_2 \dots \partial u_d}$$

Dans le cas bivarié, d'après la définition d'une dérivée partielle, nous obtenons que :

$$c(u, v) = \lim_{g \rightarrow 0} \lim_{h \rightarrow 0} \left[\frac{C(u+g, v+h) - C(u+g, v) - C(u, v+h) + C(u, v)}{g \cdot h} \right]$$

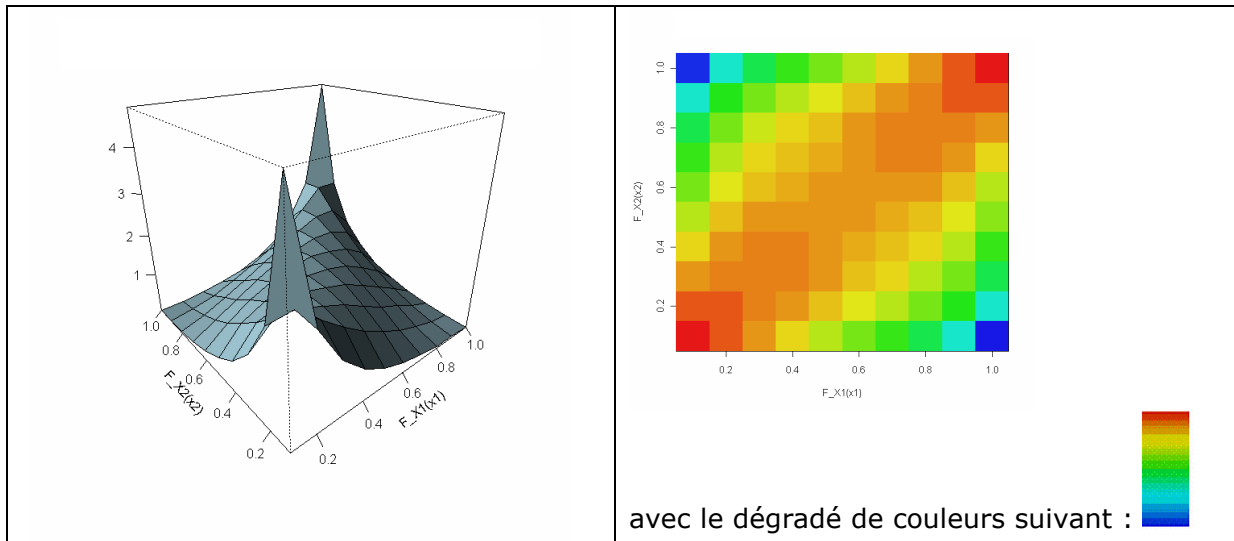
Le terme entre crochets (sans prendre les limites) correspond à la probabilité que les quantiles d'un échantillon appartienne au carré $[u;u+g]*[v;v+h]$, le tout divisé par l'aire de ce carré.

Etudions d'abord ce terme entre crochets sans faire tendre g et h vers zéro.

Pour ce faire, nous découpons alors les deux intervalles $[0,1]$ en dix tranches chacun, ce qui revient à découper le plan $[0,1]*[0,1]$ en cent parcelles.

Puis, pour chaque parcelle, nous nous intéressons à la probabilité pour qu'une réalisation (x_1, x_2) appartienne à la parcelle considérée, divisée par l'aire de la parcelle.

En simulant cette fois-ci 500 000 réalisations du couple (x_1, x_2) , nous obtenons les deux graphes suivants :



Le graphe de gauche est une représentation tridimensionnelle et celui de droite affiche les valeurs via un dégradé de couleurs.

Même avec le treillis de $10*10$ qui donne un graphe de dégradé de couleurs pouvant être considéré comme assez grossier, nous pouvons doré et déjà affiner nos conclusions sur cet exemple.

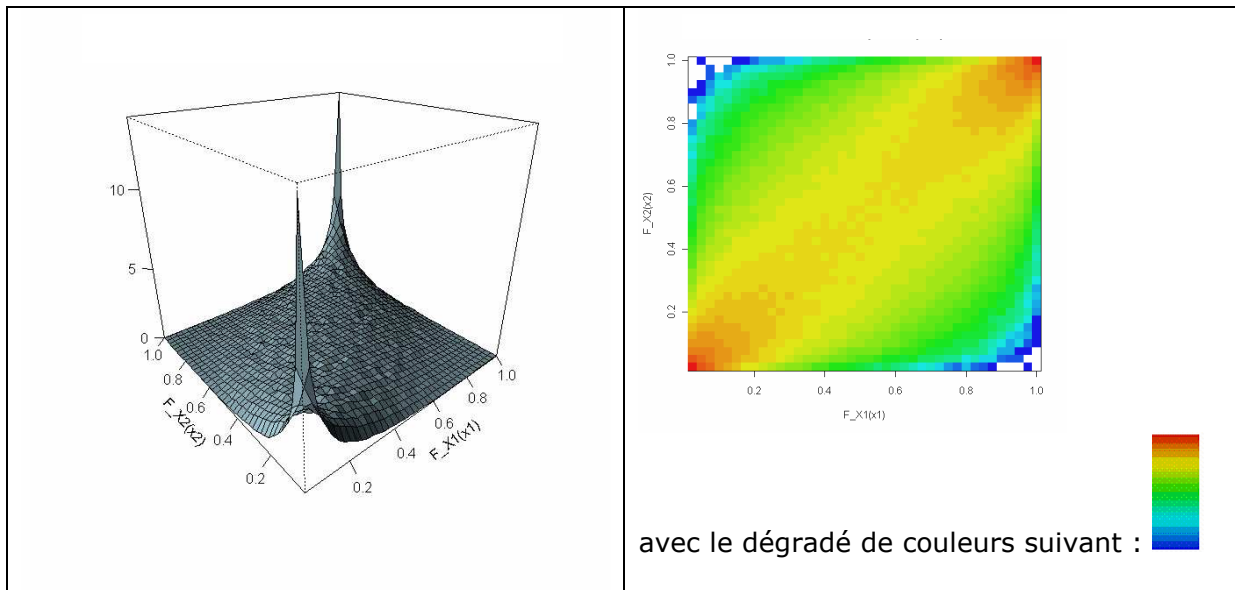
En effet, nous constatons des valeurs extrêmes quand les deux branches (X_1 et X_2) ont simultanément des quantiles ou très faibles ou très élevés. Par exemple, nous voyons qu'il y a de fortes valeurs pour les carrés $[0,0;0,1]*[0,0;0,1]$ et $[0,9;1,0]*[0,9;1,0]$.

Ce qui tend à dire que la structure de dépendance favorise l'apparition simultanée de sinistralité très haute ou très basse pour les deux branches.

De plus, nous pouvons remarquer que par simple lecture graphique nous pouvons obtenir la valeur de la lower tail concentration à 10% (c'est l'altitude du point $(0.1 ; 0.1)$ divisé par 10, ce qui donne à peu près .47) et celle de la upper tail concentration à 90% (c'est l'altitude du point $(1.0 ; 1.0)$ divisé par 10, ce qui donne à peu près .47).

Dans le cas d'indépendance, nous aurions obtenu des valeurs de .1 ($.1*.1/.1=.1$) ; donc le choc multiplie par 4.7 la probabilité d'avoir des sinistralités simultanées extrêmement fortes à un quantile de 90% ou extrêmement faibles à un quantile de 10%.

Afin d'obtenir une résolution de meilleure qualité, il est possible d'affiner le treillis : au lieu de couper le carré $[0,1]^2$ en $10*10$ tranches, découpons le maintenant en $40*40$ tranches. Nous obtenons alors les graphes suivants pour les densités empiriques des copules :



Affiner la résolution du treillis permet d'obtenir des résultats plus précis. Toutefois, nous pouvons également remarquer que les graphes obtenus sont moins lisses, ce qui indique qu'il faudrait simuler plus que 500 000 échantillons pour que la densité empirique ainsi estimée semble converger sur un graphe découpé en 40*40 petits carrés.

Egalement, d'après la définition de la fonction de densité, augmenter la résolution du treillis permet de tendre (en n'oubliant pas de diviser par $g.h$, ici $(1/40)^2$) vers la valeur de la fonction de densité de la copule.

Conclusion :

Les tracés ainsi obtenus de la densité empirique de la copule empirique permet une visualisation intéressante et efficace de la structure de dépendance induite par un modèle de chocs communs.

Et même sans prendre un treillis infiniment fin, et donc sans tendre totalement vers la définition mathématique de la fonction de densité de la copule empirique, les graphes obtenus sont appropriés pour des analyses.

Dans la suite du document, nous utiliserons un treillis de 20*20 ou de 40*40.

III.5. Evaluation numérique des indicateurs de dépendance

III.5.1. Les estimateurs classiques

Même si le coefficient de Pearson n'est pas un indicateur de dépendance dans le sens strict du terme, il est un indicateur très répandu qui permet de donner un premier résultat.

Estimateur du coefficient de Pearson :

Notons $(x_i, y_i)_{1 \leq i \leq n}$ les n échantillons de (X, Y) .

Alors le coefficient de Pearson est estimé ainsi :

$$\rho(X, Y) = \frac{n \cdot \sum_{i=1}^n x_i \cdot y_i - \sum_{i=1}^n x_i \cdot \sum_{i=1}^n y_i}{\sqrt{n \cdot \sum_{i=1}^n x_i^2 - \left(\sum_{i=1}^n x_i\right)^2} \cdot \sqrt{n \cdot \sum_{i=1}^n y_i^2 - \left(\sum_{i=1}^n y_i\right)^2}}$$

Estimateur du Tau de Kendall :

En notant $(x_i, y_i)_{1 \leq i \leq n}$ les n échantillons de (X,Y), nous pouvons construire un estimateur du Tau de Kendall de la manière suivante :

$$\hat{\tau}(X, Y) = \frac{2}{n \cdot (n-1)} \cdot \sum_{i=2}^n \sum_{j=1}^{i-1} \text{sign}[(x_i - x_j) \cdot (y_i - y_j)]$$

avec $\text{sign}(z) = \begin{cases} 1 & \text{si } z \geq 0 \\ -1 & \text{si } z < 0 \end{cases}$

Estimateur du Rho de Spearman :

En notant $(x_i, y_i)_{1 \leq i \leq n}$ les n échantillons de (X,Y), on peut construire un estimateur du Rho de Spearman de la manière suivante :

$$\hat{\rho}_S = \frac{\sum_{i=1}^n (R_i - \bar{R}) \cdot (S_i - \bar{S})}{\sqrt{\sum_{i=1}^n (R_i - \bar{R})^2} \cdot \sqrt{\sum_{i=1}^n (S_i - \bar{S})^2}}$$

avec :

- R_i est le rang de x_i
- S_i est le rang de y_i
- et la notation : $\bar{Z} = \frac{1}{n} \cdot \sum_{i=1}^n Z_i$

III.5.2. Les estimateurs basés sur les copules empiriques

Soient X et Y deux variables aléatoires continues de copule C.

Le Tau de Kendall du couple (X, Y) est donné par :

$$\tau(X, Y) = 4 \iint_{[0;1]^2} C(u, v) \cdot dC(u, v) - 1$$

Et le Rho de Spearman est donné par :

$$\rho(X, Y) = 12 \iint_{[0;1]^2} u \cdot v \cdot dC(u, v) - 3 = 12 \iint_{[0;1]^2} C(u, v) \cdot du \cdot dv - 3$$

Et concernant les indicateurs de dépendance de queue :

- pour la upper tail concentration function :

$$R(z) = P(X > F_X^{-1}(z) / Y > F_Y^{-1}(z)) = P(U > z / V > z) = \frac{C^*(z, z)}{1 - z}$$

- pour la lower tail concentration function :

$$L(z) = P(X < F_X^{-1}(z) / Y < F_Y^{-1}(z))$$

$$= \frac{C(z, z)}{z}$$

avec :

- U, V respectivement les fonctions quantiles de X et Y
- C* est la copule de survie définie par :

$$C^*(u, v) = 1 - u - v + C(u, v)$$

III.6. Mise en parallèle du choc commun avec les copules

III.6.1. Fit d'une copule empirique avec les copules usuelles

Avant de faire un fit statistique, un simple examen visuel des fonctions de densité des copules permet d'exclure rapidement les rapprochements entre deux structures de dépendance totalement différentes.

Nous avons considéré les copules usuelles suivantes :

- copule gaussienne
- copule de Gumbel
- copule de Clayton
- copule HRT
- copule de Student

La copule de Student a pour paramètres une matrice symétrique définie positive et un nombre de degrés de liberté ; ce dernier paramètre requiert une procédure de maximisation de vraisemblance compliquée, aussi, nous ne retiendrons pas cette copule.

Nous nous sommes limités à ces copules pour deux raisons principales. En premier lieu, les quatre copules retenus permettent de couvrir un spectre de formes de structure de dépendance satisfaisant. De plus, elles proposent des méthodes de fit efficaces et relativement simples à mettre en œuvre.

Pour le fit, nous avons utilisé les méthodes simples et usuelles listées dans le tableau suivant :

Copule	Méthode pour fitter la copule
copule Gaussienne	La matrice de covariance de la copule Gaussienne est la suivante : $\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$ et ρ a pour valeur la corrélation linéaire des composantes normales
Copule de Gumbel	Pour la copule de Gumbel, il existe une relation entre son paramètre θ et le taux de Kendall : $\tau(\theta) = 1 - \frac{1}{\theta}$ De ce fait, il suffit de prendre le paramètre θ de la façon suivante : $\theta = 1 - \frac{1}{\tau_{\text{induit par le choc commun}}}$

copule de Clayton et copule HRT	<p>Pour les copules de Clayton et copules HRT, il existe également une relation entre leur paramètre et le tau de Kendall :</p> $\tau(\theta) = \frac{\theta}{\theta + 2}$ <p>De ce fait, il suffit de prendre le paramètre θ de la façon suivante :</p> $\theta = \frac{2 \cdot \tau_{\text{induit par le choc commun}}}{1 - \tau_{\text{induit par le choc commun}}}$
---------------------------------	---

III.6.2. Mesure de l'adéquation avec la copule usuelle

Une fois le fit réalisé, nous obtenons la copule usuelle qui s'approche le plus de la copule générée, mais ce, uniquement à l'intérieur d'une famille de copule usuelle.

Il convient alors de vérifier qu'il y a une bonne adéquation avec la copule usuelle. Pour ce faire, il existe plusieurs méthodes ; nous nous concentrerons sur les deux méthodes suivantes, puis choisirons celle qui nous semble la plus adaptée :

- test d'adéquation du chi-deux
- distance entre les classes

III.6.2.1. Sélection d'une copule à partir du test d'adéquation du chi-deux

Le test du chi-deux utilisé dans le cadre de l'ajustement d'une distribution paramétrique à une distribution empirique constitue à cet égard un outil intéressant pour le choix d'une copule. Habituellement mis en oeuvre pour ajuster une distribution univariée, il offre l'avantage de facilement s'adapter au cas des distributions multivariées.

Voici les étapes à suivre pour sa mise en oeuvre :

1. Fit avec les copules usuelles

On estime le paramètre de la copule bivariée à l'aide d'une des méthodes explicitées dans le chapitre précédent. On note $C(u,v)$ la copule paramétrique obtenue.

2. Calcul de la copule empirique

On définit un treillis d'ordre T adapté au nombre J d'observations et aux contraintes de temps de calcul puis on calcule la copule empirique bivariée sur ce treillis (il s'agit de la copule empirique de Deheuvels) :

$$\hat{C}\left(\frac{t_1}{T}, \frac{t_2}{T}\right) = \frac{1}{J} \cdot \sum_{j=1}^J \mathbf{1}_{\{r_1^j/J \leq t_1/T\}} \cdot \mathbf{1}_{\{r_2^j/J \leq t_2/T\}}$$

avec : $\{r_1^j, r_2^j, \dots, r_n^j\}$ la statistique de rang associée au j -ième échantillon multivarié $\{x_1^j, x_2^j, \dots, x_n^j\}$ avec $1 \leq j \leq J$

3. Choix d'un nombre d'intervalles pour découper les uniformes univariées

On se donne un nombre k afin de découper chacune des uniformes (i.e. intervalles $[0 ; 1]$ des marginales) en k intervalles et de construire un tableau de contingence bivarié des effectifs. Dans les notations, nous avons donc $T=k$. Sans a priori sur la distribution empirique bivariée, on découpe chaque uniforme sur des intervalles identiques. On note $(b_0, b_1], (b_1, b_2], \dots, (b_{k-1}, b_k]$ et $(c_0, c_1], (c_1, c_2], \dots, (c_{k-1}, c_k]$ les k intervalles pour découper respectivement u et v .

4. Calcul des effectifs et construction du tableau de contingence bivarié

L'objet de cette étape est de calculer les effectifs empiriques et théoriques sur chacun des k^2 intervalles bidimensionnels $(b_{i-1}, b_i] * (c_{i-1}, c_i]$, pour $i, j = 1 \dots k$

Pour ce faire, il convient de calculer les probabilités bivariées $p_{i,j}$ d'appartenance à un intervalle. La formule de calcul est la suivante :

$$p_{i,j} = C(b_i, c_j) - C(b_{i-1}, c_j) - C(b_i, c_{j-1}) + C(b_{i-1}, c_{j-1})$$

où $i, j = 1 \dots k$ et $C(u, v) = \hat{C}(u, v)$ ou $C(u, v)$ respectivement pour le calcul des effectifs empiriques et théoriques.

Les effectifs $f_{i,j}$ sont calculés en multipliant les probabilités bivariées par le nombre total d'observations. On obtient alors un tableau de contingence bivarié donnant les effectifs empiriques issus de la copule empirique et les effectifs théoriques issus de la copule paramétrique. Il a la forme suivante :

u / v	$(c_0, c_1]$,	$(c_1, c_2]$...	$(c_{k-1}, c_k]$
$(b_0, b_1]$	$f_{1,1}$	$f_{1,2}$...	$f_{1,k}$
$(b_1, b_2]$	$f_{2,1}$	$f_{2,2}$...	$f_{2,k}$
...
$(b_{k-1}, b_k]$	$f_{k,1}$	$f_{k,2}$...	$f_{k,k}$

5. Application du critère de Cochran et regroupement en classes

Dans la pratique, les effectifs théoriques de certaines cases peuvent être très faibles voir proches de zéro. Il faut alors procéder à un regroupement des k^2 intervalles initiaux en n classes permettant de respecter le critère de Cochran qui recommande d'avoir des effectifs théoriques au moins supérieurs à 1% du nombre total d'observations dans chaque classe, et supérieurs à 5% du nombre total d'observations dans au moins 80% des classes. Ce regroupement est aussi à effectuer sur le tableau bivarié des effectifs empiriques. Selon les zones de dépendance recherchées, on pourra notamment isoler les cellules des queues de distribution ou retenir des directions particulières comme la diagonale principale par exemple.

6. Calcul de la statistique bivariée du chi-deux et sélection de la copule

Après avoir défini le regroupement en n classes vérifiant le critère de Cochran, on calcule la statistique bivariée du chi-deux observée comme suit :

$$\chi_{obs}^2 = \sum_{j=1}^n \frac{(O_j - E_j)^2}{E_j}$$

où O_j et E_j représentent respectivement les effectifs observés et attendus dans chaque classe j , $j = 1, \dots, n$. Cette statistique suit une loi du chi-deux à $(n-r-1)$ degrés de liberté avec r le nombre de paramètres estimés de la copule. On se donne ensuite un seuil critique α pour la zone de rejet et on calcule la p-valeur $\hat{\alpha} = P(\chi_{n-r-1}^2 > \chi_{obs}^2)$. On rejette l'adéquation de la copule paramétrique si $\hat{\alpha} < \alpha$.

Point faible de la méthode :

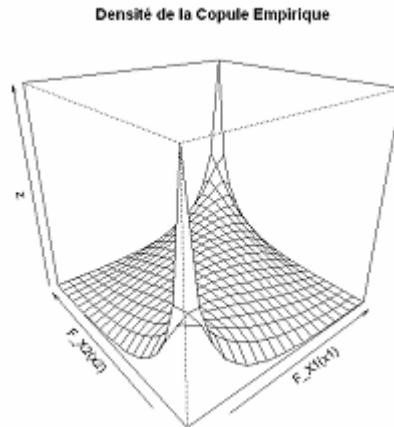
Pour que le test d'adéquation du chi-deux fonctionne correctement, il faut que le critère de Cochran soit respecté, ou du moins, que l'on ne s'en s'écarte pas trop.

Or la condition « d'avoir des effectifs théoriques au moins supérieurs à 1% du nombre total d'observations dans chaque classe, et supérieurs à 5% du nombre total d'observations dans au moins 80% des classes » est contraignante et difficile à mettre en œuvre.

En effet, elle limite le nombre de classes à 25 et incite à un choix des classes 'pertinent' (sinon, il est impossible que 80% des classes contiennent au moins 5% du nombre total d'observations). Par exemple, le choix le plus simple qui serait de découper $[0,1]^2$ en :

$$\left(\left[\frac{i-1}{5}; \frac{i}{5} \right] \times \left[\frac{j-1}{5}; \frac{j}{5} \right] \right)_{1 \leq i, j \leq 5}$$

sera très insatisfaisant vis-à-vis du critère de Cochran (et du coup rendra inutilisable le test d'adéquation du chi-deux) dans le cas d'une copule empirique ayant, par exemple, la forme suivante :



En effet, les classes $[0 ;.2] \times [0 ;.2]$ et $[.8 ;1] \times [.8 ;1]$ capteraient la plupart des échantillons, ce qui empêcherait la plupart des classes restantes de remplir le critère de Cochran ; en particulier, nous pouvons voir que les classes $[0 ;.2] \times [.8 ;1]$ ou $[.8 ;1] \times [0 ;.2]$ auraient un nombre quasi-nul d'échantillon.

Pour utiliser le test du chi-deux, il faudrait alors définir un treillis 'intelligent' (et si possible 'naturel') qui s'adapterait à chaque cas, ce qui rend la procédure relativement complexe ; de plus, il pourrait y avoir des treillis différents pour des cas différents, ce qui biaiserait la comparaison.

III.6.2.2. Distance entre les classes

Cette deuxième méthode reprend les étapes 1, 2, 3 et 4 de la méthode du chi-deux, avec le treillis suivant :

$$\left(\left[\frac{i-1}{k}; \frac{i}{k} \right] \times \left[\frac{j-1}{k}; \frac{j}{k} \right] \right)_{1 \leq i, j \leq k}$$

Nous obtenons alors :

- $p_{i,j}^{empirique}$ pour $1 \leq i, j \leq k$
- $p_{i,j}^{copule\ testées}$ pour $1 \leq i, j \leq k$

qui sont les probabilités bivariées d'appartenance à un intervalle, exprimées en pourcentage, respectivement pour la copule empirique et pour la copule testée.

Nous pouvons alors définir la notion de distance, de la manière suivante :

$$D_{copule\ testée} = \sum_{1 \leq i, j \leq k} (p_{i,j}^{empirique} - p_{i,j}^{copule\ testées})^2$$

Par ailleurs, pour avoir un ordre de grandeur des distances ainsi mesurée, il peut être très utile de comparer la distance obtenue avec celle que l'on obtiendrait dans le cas d'une copule d'indépendance.

Egalement, dans le cas où la distance est très faible, nous pourrions nous intéresser à une deuxième distance qui est la suivante :

$$D_{\text{copule testée}}^* = \text{Max}_{1 \leq i, j \leq k} \left(\frac{p_{i,j}^{\text{empirique}} - p_{i,j}^{\text{copule testées}}}{p_{i,j}^{\text{copule testées}}} \right)$$

Cette distance est plus facilement interprétable en tant que telle : il s'agit du maximum en valeur absolue de l'écart relatif entre les probabilités d'appartenance à une case entre la copule empirique et la copule testée.

Toutefois, si la probabilité théorique est très faible pour une case donnée (par exemple, pour la case : $[0 ; 0.1] \times [0,9 ; 1]$ dans le cas étudié), il risque de n'y avoir aucun (ou très peu d') échantillon simulé par Monte-Carlo, ce qui viendrait à fausser la pertinence de cette deuxième mesure.

Aussi, dans le mémoire, nous utiliserons seulement la première mesure de distance.

III.6.2.3. Choix de la mesure d'adéquation

Dans la suite du mémoire, nous utiliserons la deuxième méthode, i.e. la distance entre classes pour sa simplicité de mise en œuvre et sa pertinence. En effet, elle n'est pas limitée dans le nombre de classes (alors que le test d'adéquation du chi-deux le limite à 25) , ce qui permet de tester l'adéquation de la copule en un plus grand nombre d'intervalles dans le carré $[0 ; 1]^2$.

De plus, elle ne nécessite pas la mise au point d'un algorithme très complexe de découpage de classes ; en effet le simple treillis suivant :

$$\left(\left[\frac{i-1}{k}, \frac{i}{k} \right] \times \left[\frac{j-1}{k}, \frac{j}{k} \right] \right)_{1 \leq i, j \leq k}$$

, en plus d'être intuitif, donne des résultats convenables (contrairement à la méthode du chi-deux).

IV. ETUDE DES CHOCS COMMUNS ET DE LA DEPENDANCE INDUITE

IV.1. Choc multiplicatif log-normal

IV.1.1. Rappel du choc

Considérons le cas d'un choc commun multiplicatif sur deux branches d'assurance :

- la branche 1 a pour sinistralité : $X_1 = \beta \times X_{10}$
- la branche 2 a pour sinistralité : $X_2 = \beta \times X_{20}$
- le choc commun est : β

avec :

- $X_{10} \sim \text{LN}(\mu_{10}, \sigma_{10})$
- $X_{20} \sim \text{LN}(\mu_{20}, \sigma_{20})$
- $\beta \sim \text{LN}(1, \sigma_c)$

$\text{LN}(\mu, \sigma)$ désigne la loi log-normale

L'espérance de β est prise égale à 1 pour que l'espérance de X_1 reste inchangée par rapport à X_{10} , et de même pour X_2 par rapport à X_{20} . Faire autrement aurait ajouté un degré de liberté supplémentaire inutile.

IV.1.2. Détermination des marginales obtenues suite à l'application du choc

D'après la propriété de multiplicativité des lois log-normales, nous pouvons facilement déterminer les marginales obtenues : elles suivent les lois log-normales suivantes :

- $X_1 \sim \text{LN}(\mu_{10} + \mu_c, \sqrt{\sigma_{10}^2 + \sigma_c^2})$
- $X_2 \sim \text{LN}(\mu_{20} + \mu_c, \sqrt{\sigma_{20}^2 + \sigma_c^2})$

IV.1.3. Analyse qualitative de la structure de dépendance obtenue

Afin de visualiser la dépendance obtenue, nous allons faire des simulations de Monte-Carlo et estimer numériquement la fonction de densité empirique de la copule induite par le choc commun. Le lecteur est invité à se reporter à la partie [III.4.2. Visualisation de la structure de dépendance à l'aide de copule empirique](#) pour de plus amples informations.

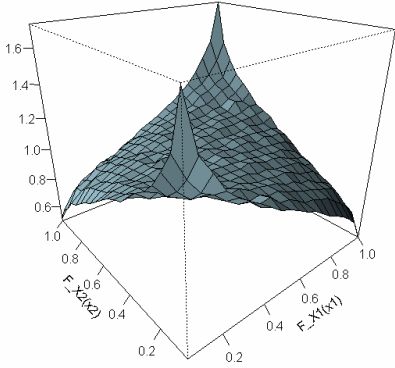
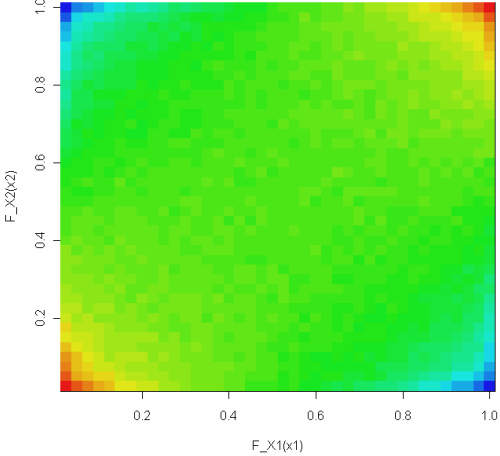
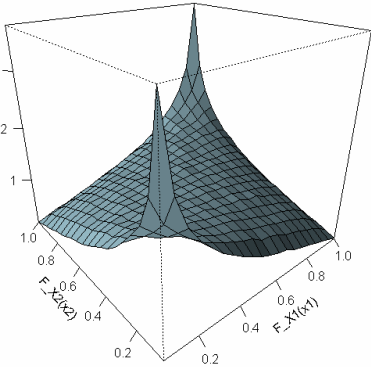
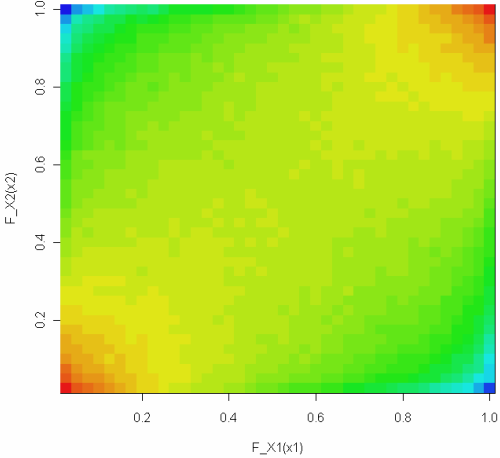
Pour ce faire, nous prenons les valeurs suivantes pour les paramètres du modèle :

- $E[X_{10}] = 100$ et $\text{Var}(X_{10}) = 10^2$
- $E[X_{20}] = 100$ et $\text{Var}(X_{20}) = 15^2$
- $E[\beta] = 1$ et $\text{Var}(\beta)$ prendra plusieurs valeurs

Les paramètres μ et σ sont déterminés à l'aide des relations suivantes (méthode des moments) :

$$\mu = \ln(E[X]) - \frac{1}{2} \ln\left(1 + \frac{\text{Var}(X)}{E[X]^2}\right) \text{ et } \sigma^2 = \ln\left(1 + \frac{\text{Var}(X)}{E[X]^2}\right)$$

Nous pouvons constater que ce modèle a un unique paramètre : la variance du choc appliqué aux deux branches $\text{Var}(\beta)$. Aussi, pour comprendre l'influence de ce paramètre sur la structure de dépendance, nous allons le faire varier.

Intensité du choc	Structure de dépendance : copule empirique liée au choc commun	
$\text{Var}(\beta)$	Représentation tridimensionnelle	Dégradé de couleur
0.05 ²	<p style="text-align: center;">Densité de la Copule Empirique</p> 	<p style="text-align: center;">Densité de la Copule Empirique</p> 
0.1 ²	<p style="text-align: center;">Densité de la Copule Empirique</p> 	<p style="text-align: center;">Densité de la Copule Empirique</p> 



avec le dégradé de couleurs suivant :

Nous pouvons donc remarquer que plus l'intensité du choc augmente (plus la variance de β augmente), et plus la fonction de densité empirique de la copule empirique a tendance à se tasser sur la première diagonale du carré $[0,1]^2$.

De même, quand le choc augmente en intensité, la probabilité de sinistralité extrême (extrêmement forte ou extrêmement faible) augmente significativement (l'échelle maximale pour l'axe d'altitude augmente significativement). En terme d'indicateur, cela revient à une forte augmentation des upper tail concentration et des lower tail concentration.

Au contraire, la probabilité d'avoir une sinistralité extrêmement forte sur une branche tout en ayant une sinistralité extrêmement faible sur l'autre branche (et vice-versa) tend vers 0.

IV.1.4. Analyse quantitative de la structure de dépendance obtenue

Une première approche serait d'essayer d'explicitier la copule induite par le choc commun.

Par exemple, à l'aide de la formule :

$$C(u_1, u_2) = F(F_{X_1}^{-1}(u_1), F_{X_2}^{-1}(u_2))$$

Il faudrait alors expliciter : $F_{X_1, X_2}(x, y) = \Pr(\beta X_{10} < x, \beta Y_{10} < y)$

Pour ce faire, il serait possible d'intégrer $\Pr(\beta X_{10} < x, \beta Y_{10} < y \mid \beta = \beta_0)$ suivant β_0 .

Néanmoins, le calcul est complexe et ne permet pas d'arriver facilement à une solution exploitable.

Une deuxième approche est envisageable en s'appuyant sur le théorème d'invariance des copules (cf chapitre [II.5.3.4. Théorème d'invariance](#)) et sur une des propriétés des indicateurs de mesure de concordance que satisfont les indicateurs de dépendance (cf. chapitre [II.4.1. Mesure de dépendance](#)).

En effet, si α est une fonction réelle strictement croissante sur $\text{Im}(X_1)$ et sur $\text{Im}(X_2)$, alors :

- pour une mesure de dépendance δ , nous avons: $\delta[\alpha(X_1), \alpha(X_2)] = \delta(X_1, X_2)$
- $C_{\alpha(X_1), \alpha(X_2)} = C_{X_1, X_2}$

Autrement dit, étudier la structure de dépendance entre $\alpha(X_1), \alpha(X_2)$ revient à étudier la structure de dépendance entre X_1, X_2 .

Or, en prenant la fonction $\alpha: x \mapsto \exp(x)$ (définie sur : $R \mapsto R$, et qui est par définition strictement croissante sur $\text{Im}(X_1)$ et sur $\text{Im}(X_2)$), nous pouvons nous ramener à l'étude d'un modèle de choc commun additif normal ($\beta + X_{10}, \beta + X_{20}$) (au lieu d'un modèle de choc commun multiplicatif log-normal) :

- la branche 1 a pour sinistralité : $X_1 = \beta + X_{10}$
- la branche 2 a pour sinistralité : $X_2 = \beta + X_{20}$
- le choc commun est : β

Et X_{10}, X_{20} et β sont des lois normales.

Considérons alors le théorème suivant :

Théorème : La distribution d'une variable multidimensionnelle (X_1, X_2, \dots, X_n) est une distribution normale, si et seulement si toute combinaison linéaire de X_1, X_2, \dots, X_n (i.e. tout $L_1 X_1 + L_2 X_2 + \dots + L_n X_n$ avec L_1, L_2, \dots, L_n des réels) est une distribution normale unidimensionnelle.

D'après ce théorème, il est alors facile de vérifier que $(\beta + X_{10}, \beta + X_{20})$ est une distribution normale bi-dimensionnelle.

De plus, le coefficient de covariance de la distribution normale bi-dimensionnelle s'obtient en calculant la covariance : $\text{cov}(\beta + X_{10}, \beta + X_{20})$

Reprenons maintenant la définition d'une copule normale :

Avec R une matrice symétrique définie positive avec $\text{diag}(R) = (1, 1, \dots, 1)$ et ϕ_R la fonction de répartition de la loi normale multivariée centrée ayant R comme matrice de corrélation, la copule Normale est définie par :

$$C_R^{\text{Normale}}(u, v) = \phi_R(\phi^{-1}(u), \phi^{-1}(v))$$

où ϕ^{-1} désigne l'inverse de la fonction de répartition de la loi normale centrée réduite ϕ

Nous remarquons que nous sommes très proches de cette définition, si nous explicitons la copule induite par le choc commun avec la formule : $C(u, v) = F_{X_1, X_2}(F_{X_1}^{-1}(u), F_{X_2}^{-1}(v))$

Il suffit d'utiliser le théorème d'invariance des copules (cf chapitre [II.5.3.4. Théorème d'invariance](#)) pour « normaliser » la structure de dépendance, avec les fonctions strictement monotones suivantes :

- $\alpha_1 = x \mapsto E[\beta + X_{10}] + x \cdot \sqrt{\text{Var}(\beta + X_{10})}$
- $\alpha_2 = x \mapsto E[\beta + X_{20}] + x \cdot \sqrt{\text{Var}(\beta + X_{20})}$

Aussi, nous obtenons que la structure de dépendance d'un choc commun additif normal ($\beta + X_{10}, \beta + X_{20}$) est semblable à celle d'une copule gaussienne.

Conclusion : Etudier quantitativement la structure de dépendance induite par un choc commun additif normal ou par un choc commun multiplicatif lognormal revient à étudier une copule normale.

IV.1.5. Vérification numérique de l'analyse qualitative :

Nous avons alors cherché à confirmer numériquement cette conclusion.

Dans ce but, nous avons comparé :

- la fonction de répartition de la copule empirique induite par un choc commun multiplicatif log-normale, simulée avec la méthode de Monte-Carlo avec un million de points
- avec la fonction de répartition de la copule gaussienne ayant pour coefficient de covariance : $\text{cov}(\beta X_{10}, \beta X_{20})$

Nous avons pris :

- $E[X_{10}] = 100$ et $\text{Var}(X_{10}) = 10^2$
- $E[X_{20}] = 100$ et $\text{Var}(X_{20}) = 15^2$
- $E[\beta] = 1$ et $\text{Var}(\beta) = 0.2$

Et nous avons comparé les deux copules sur les 100 points suivants :

- $(i/10, j/10)$ pour $1 \leq i, j \leq 10$

Pour calculer la fonction de répartition d'une copule normale, nous avons besoin de la fonction de répartition d'une distribution normale bivariée.

Cependant, il n'existe pas d'expression analytique pour cette dernière.

Aussi, nous avons utilisé la fonction « pmvnom » du package « mvtnorm » de R.

Voici les valeurs prises par la copule normale gaussienne $C^{Normale}(u, v)$:

C(u,v)	u=0,1	u=0,2	u=0,3	u=0,4	u=0,5	u=0,6	u=0,7	u=0,8	u=0,9	u=1
v=0,1	0,0475	0,0698	0,0827	0,0905	0,0951	0,0977	0,0991	0,0997	0,1000	0,1000
v=0,2	0,0698	0,1142	0,1447	0,1658	0,1802	0,1896	0,1953	0,1984	0,1997	0,2000
v=0,3	0,0827	0,1447	0,1922	0,2283	0,2553	0,2745	0,2875	0,2953	0,2991	0,3000
v=0,4	0,0905	0,1658	0,2283	0,2795	0,3203	0,3517	0,3745	0,3896	0,3977	0,4000
v=0,5	0,0951	0,1802	0,2553	0,3203	0,3753	0,4203	0,4553	0,4802	0,4951	0,5000
v=0,6	0,0977	0,1896	0,2745	0,3517	0,4203	0,4795	0,5283	0,5658	0,5905	0,6000
v=0,7	0,0991	0,1953	0,2875	0,3745	0,4553	0,5283	0,5922	0,6447	0,6827	0,7000
v=0,8	0,0997	0,1984	0,2953	0,3896	0,4802	0,5658	0,6447	0,7142	0,7698	0,8000
v=0,9	0,1000	0,1997	0,2991	0,3977	0,4951	0,5905	0,6827	0,7698	0,8475	0,9000
v=1	0,1000	0,2000	0,3000	0,4000	0,5000	0,6000	0,7000	0,8000	0,9000	1,0000

Le tableau suivant représente l'écart relatif entre la valeur théorique et la valeur

générée :
$$\frac{C^{Empirique}(u, v) - C^{Normale}(u, v)}{C^{Normale}(u, v)}$$

écart	u=0,1	u=0,2	u=0,3	u=0,4	u=0,5	u=0,6	u=0,7	u=0,8	u=0,9	u=1
v=0,1	1,2630%	0,7787%	0,6794%	0,5579%	0,3660%	0,1719%	0,0715%	0,0034%	0,0053%	-0,0010%
v=0,2	0,9119%	0,7369%	0,6794%	0,6052%	0,4629%	0,2811%	0,1519%	0,0412%	0,0077%	-0,0005%
v=0,3	0,6213%	0,6510%	0,5303%	0,5806%	0,4494%	0,3112%	0,1585%	0,0402%	0,0207%	-0,0003%
v=0,4	0,3291%	0,3609%	0,3896%	0,4715%	0,4433%	0,3494%	0,2230%	0,1011%	0,0387%	-0,0003%
v=0,5	0,1756%	0,3036%	0,3295%	0,3649%	0,4021%	0,3481%	0,2425%	0,1327%	0,0570%	-0,0002%
v=0,6	0,0983%	0,1682%	0,2391%	0,2635%	0,2845%	0,2456%	0,1894%	0,1268%	0,0689%	-0,0002%
v=0,7	0,0574%	0,0833%	0,1557%	0,2089%	0,2348%	0,1879%	0,1763%	0,1211%	0,0656%	-0,0001%
v=0,8	-0,0036%	0,0125%	0,0409%	0,1065%	0,1339%	0,1342%	0,1072%	0,0937%	0,0678%	-0,0001%
v=0,9	0,0013%	0,0092%	0,0264%	0,0490%	0,0568%	0,0862%	0,0813%	0,0631%	0,0417%	-0,0001%
v=1	-0,0010%	-0,0005%	-0,0003%	-0,0003%	-0,0002%	-0,0002%	-0,0001%	-0,0001%	-0,0001%	0,0000%

Le plus gros écart relatif est de 1.3%, et il porte sur C(0.1,0.1), ce qui est logique dans le sens où C(0.1,0.1) est la plus petite valeur.

Nous pouvons donc conclure que nous retrouvons numériquement le fait qu'un choc commun multiplicatif log-normale induit une structure de dépendance semblable à celle d'une copule normale.

Enfin, nous pouvons aussi recourir à la distance entre copules telle qu'elle est définie dans la partie [III.6.2.2. Distance entre les classes](#), nous avons défini deux distances :

$$D_{\text{copule testée}} = \sum_{1 \leq i, j \leq k} (p_{i,j}^{\text{empirique}} - p_{i,j}^{\text{copule testées}})^2$$

En comparant la distance avec la copule normale avec celle de la copule d'indépendance, nous obtenons :

Copule	Distance
Indépendance	68.017
Normale	0.015
De Gumbel	5.035
De Clayton	18.362
HRT	18.382

Ce qui confirme également l'adéquation avec la copule normale.

IV.1.6. Evaluation des indicateurs de dépendance et leur sensibilité aux paramètres du modèle

Intéressons-nous dans un premier temps au calcul du coefficient linéaire de Pearson :

$$\begin{aligned} Cov[\beta X_{10}, \beta X_{20}] &= E_{\beta} [Cov[\beta X_{10}, \beta X_{20} | \beta]] + Cov_{\beta} [E[\beta X_{10} | \beta], E[\beta X_{20} | \beta]] \\ &= E_{\beta} [\beta^2 Cov[X_{10}, X_{20} | \beta]] + Cov_{\beta} [\beta E[X_{10}], \beta E[X_{20}]] \\ &= E_{\beta} [\beta^2 \cdot 0] + E[X_{10}] \cdot E[X_{20}] \cdot Cov_{\beta} [\beta, \beta] \\ &= E[X_{10}] \cdot E[X_{20}] \cdot Cov_{\beta} [\beta, \beta] \end{aligned}$$

D'où :

$$Cor[X_1, X_2] = Cor[\beta X_{10}, \beta X_{20}] = \frac{E[X_{10}] \cdot E[X_{20}] \cdot Var[\beta]}{\sqrt{Var[X_1] \cdot Var[X_2]}}$$

Or,

$$\begin{aligned} Var[X_1] &= Var[\beta X_{10}] = E_{\beta} [Var[\beta X_{10} | \beta]] + Var_{\beta} [E[\beta X_{10} | \beta]] \\ &= E_{\beta} [\beta^2 \cdot Var[X_{10}]] + Var_{\beta} [\beta \cdot E[X_{10}]] \end{aligned}$$

$$\begin{aligned}
 &= \text{Var}[X_{10}] \cdot E_{\beta}[\beta^2] + E[X_{10}]^2 \cdot \text{Var}_{\beta}[\beta] \\
 &= \text{Var}[X_{10}] \cdot (1+b) + E[X_{10}]^2 \cdot b
 \end{aligned}$$

avec $b = \text{Var}_{\beta}[\beta]$

Donc, finalement :

$$\text{Cor}[X_1, X_2] = \text{Cor}[\beta X_{10}, \beta X_{20}] = \frac{E[X_{10}] \cdot E[X_{20}] \cdot b}{\sqrt{(\text{Var}[X_{10}] \cdot (1+b) + E[X_{10}]^2 \cdot b) \cdot (\text{Var}[X_{20}] \cdot (1+b) + E[X_{20}]^2 \cdot b)}}$$

avec $b = \text{Var}_{\beta}[\beta]$

Enfin, l'expression de la corrélation en fonction des marginales et du choc β nous donne :

$$\text{Cor}[X_1, X_2] = \frac{E[X_1] \cdot E[X_2] \cdot \text{Var}[\beta]}{\sqrt{\text{Var}[X_1] \cdot \text{Var}[X_2]}}$$

Concernant le Tau de Kendall et le Rho de Spearman :

Nous pouvons nous appuyer sur une propriété des copules elliptiques (dont les copules normales font partie) et qui nous donne :

$$\begin{aligned}
 - \quad \tau_{Kendall} &= \frac{2}{\pi} \cdot \text{Arc sin}(\rho) \\
 - \quad \rho_{Spearman} &= \frac{6}{\pi} \cdot \text{Arc sin}\left(\frac{\rho}{2}\right)
 \end{aligned}$$

avec $\rho = \text{Cor}(X_1, X_2)$; en toute rigueur, X_1 et X_2 doivent correspondre au logarithme des observations.

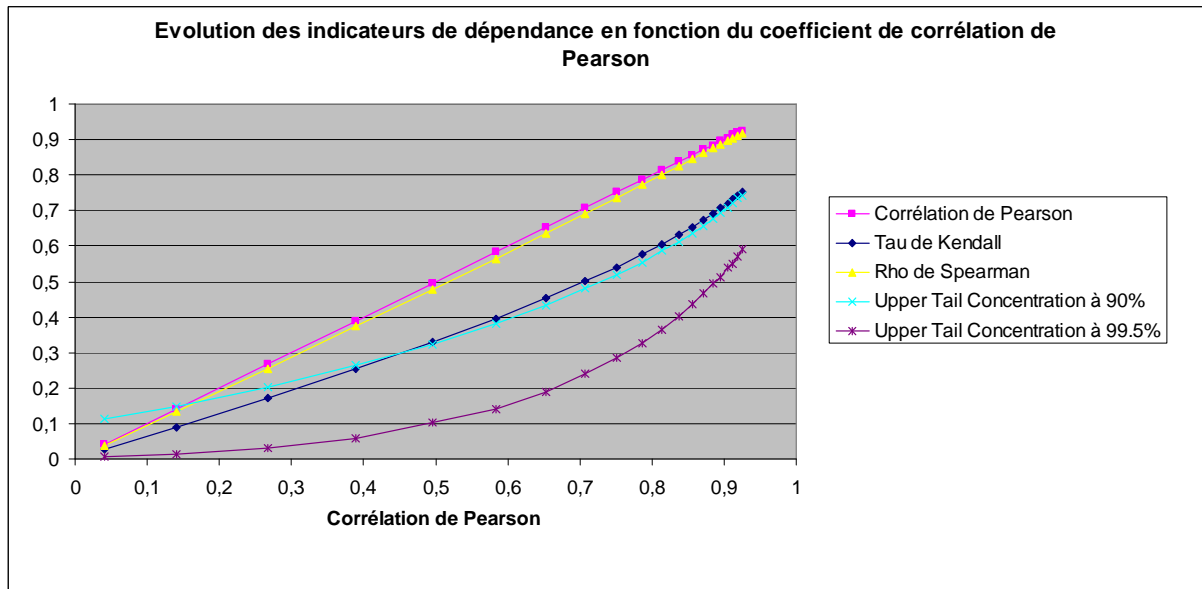
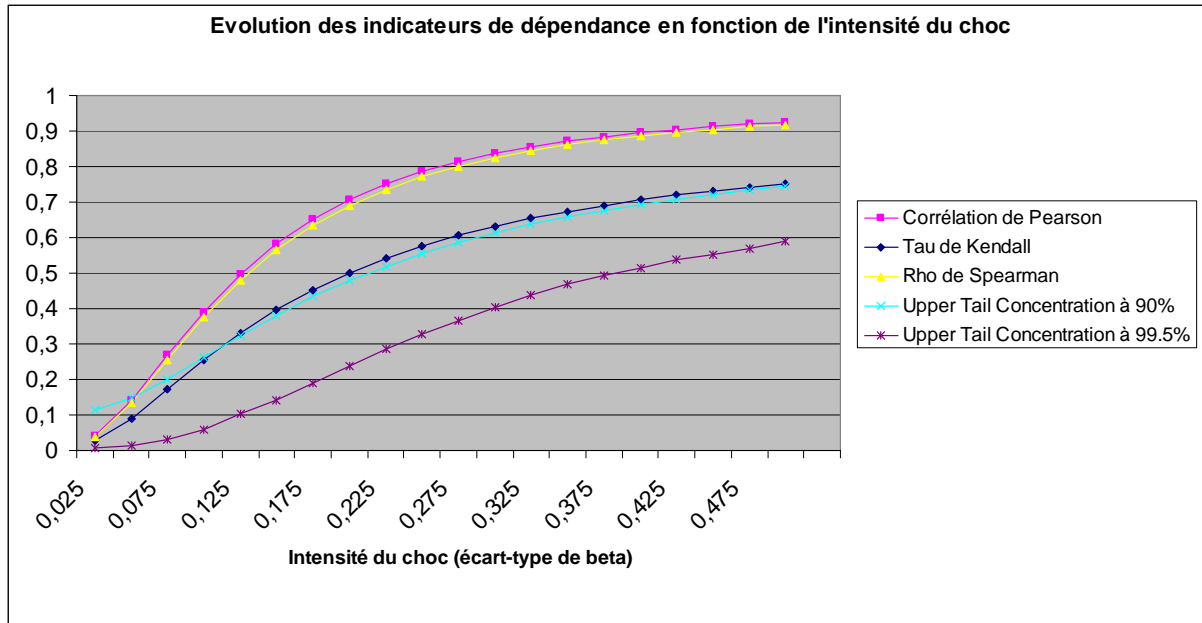
Concernant la dépendance de queue :

Nous pouvons donner l'expression analytique en s'appuyant sur fonction de répartition de la copule gaussienne : cf [III.5.2. Les estimateurs basés sur les copules empiriques](#)
Toutefois, cette expression n'est guère exploitable.

Résultats numériques :

Regardons l'évolution des indicateurs de dépendance en fonction de :

- l'intensité du choc
- du coefficient de corrélation de Pearson



Logiquement tous les facteurs de dépendance augmentent avec l'intensité du choc.

Nous pouvons faire quelques remarques sur les données obtenues :

- les simulations de Monte-Carlo ayant été réalisées avec des échantillons d'un million de points, les droites obtenues sont lisses et les valeurs mesurées sont très proches des valeurs théoriques attendues (cf annexe [VIII.2. Résultats obtenus pour le choc commun multiplicatif log-normale](#))
- assez logiquement, les indicateurs de corrélation et de dépendance augmentent dans le même sens que l'intensité du choc
- le choc étant symétrique, la lower tail concentration à 10% (qui n'est pas représentée sur le graphe) est égale à la upper tail concentration à 90% ; il en est de même pour la lower tail concentration à .5% et la upper tail concentration à 99.5% (aussi, par souci d'affichage, seules les valeurs à 90% et 99.5% sont affichées)

- en cas d'indépendance, les tail concentrations à 10%/90% sont égales à .1 ($.1 \cdot .1 / .1 = .1$) et les tail concentrations à .5%/99.5% sont égales à .005, ce que nous retrouvons en prolongeant les courbes du graphique au point d'abscisse 0. Nous remarquons donc le fort impact du choc sur les tail concentrations : par exemple, pour $\rho = .5$, la upper tail concentration vaut .1, alors qu'elle vaut .005 en cas d'indépendance.

IV.1.7. Utilisation concrète du modèle et ses limites

Le modèle spécifie que :

- la branche 1 a pour sinistralité : $X_1 = \beta \times X_{10}$
- la branche 2 a pour sinistralité : $X_2 = \beta \times X_{20}$

En pratique, nous partons de X_1 et de X_2 . Puis, nous fixons la variance de β ; les lois X_{10} et X_{20} sont déduites à l'aide des relations suivantes :

$$\begin{cases} E[X_{10}] = E[X_1] / E[\beta] \\ E[X_{20}] = E[X_2] / E[\beta] \end{cases} \quad \text{et} \quad \begin{cases} Var(X_{10}) = \frac{Var(X_1) - Var(\beta) \cdot E[X_{10}]^2}{Var(\beta) + E[\beta]^2} \\ Var(X_{20}) = \frac{Var(X_2) - Var(\beta) \cdot E[X_{20}]^2}{Var(\beta) + E[\beta]^2} \end{cases} \quad (1)$$

En effet, nous avons vu que :

$$\begin{aligned} Var[X_1] &= Var[\beta X_{10}] = E_\beta [Var[\beta X_{10} | \beta]] + Var_\beta [E[\beta X_{10} | \beta]] \\ &= Var[X_{10}] \cdot (E[\beta]^2 + b) + E[X_{10}]^2 \cdot b \end{aligned}$$

A partir de (1), comme une variance est toujours positive, nous avons que la variance de β est bornée par les inégalités suivantes :

$$\begin{cases} Var(\beta) \leq Var(X_1) / E[X_{10}]^2 \\ Var(\beta) \leq Var(X_2) / E[X_{20}]^2 \end{cases} \quad \text{et donc :} \quad \begin{cases} Var(\beta) \leq CV_1^2 \\ Var(\beta) \leq CV_2^2 \end{cases} \quad (\text{l'espérance du choc vaut 1})$$

(CV_1 et CV_2 représentent les coefficients de variation respectifs de X_1 et X_2)

Or, nous avons :

$$Cor[X_1, X_2] = Cor[\beta X_{10}, \beta X_{20}] = \frac{E[X_{10}] \cdot E[X_{20}] \cdot Var[\beta]}{\sqrt{Var[X_1] \cdot Var[X_2]}} = \frac{Var(\beta)}{CV_1 \cdot CV_2}$$

De ce fait, la corrélation linéaire que peut induire le choc est limitée par :

$$Cor[X_1, X_2] \leq \text{Min} \left(\frac{CV_1}{CV_2}; \frac{CV_2}{CV_1} \right)$$

Et de même le Tau de Kendall et le Rho de Spearman qui peuvent être engendrés grâce au choc peuvent être fortement limités dans le cas où les deux marginales ont des coefficients de variations très différents.

Enfin, par défaut, le modèle décrit jusqu'ici ne sait pas générer des dépendances négatives.

IV.1.8. Extension du modèle

La version présentée du modèle semble ne pas permettre des dépendances à caractère négatif.

Toutefois, en prenant le modèle suivant :

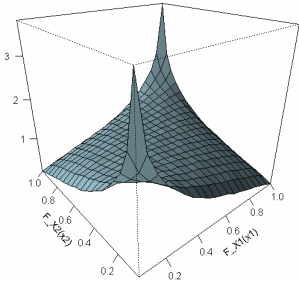
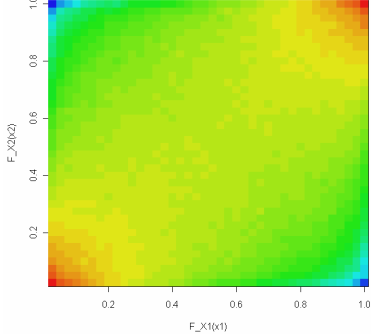
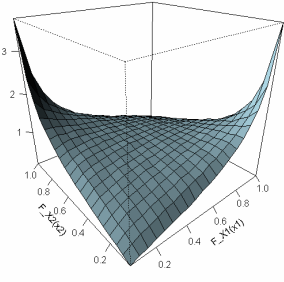
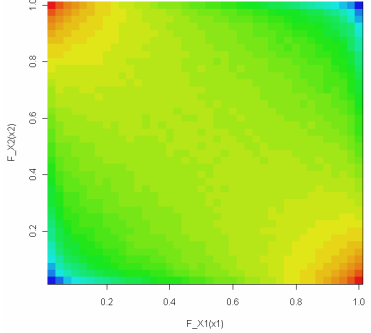
$$\begin{cases} X_1 = X_{10} \times \beta \\ X_2 = X_{20} / \beta \end{cases}$$

Nous obtenons les corrélations opposées à celles obtenues avec le modèle où le choc commun β est multiplié sur les deux branches.

En repassant par le modèle additif normal, cette propriété apparaît facilement.

A titre d'exemple, considérons le cas suivant :

- $E[X_{10}] = 100$ et $Var(X_{10}) = 10^2$
- $E[X_{20}] = 100$ et $Var(X_{20}) = 15^2$
- Et $Var(\beta) = .1$

Modèle utilisé	Structure de dépendance : copule empirique liée au choc commun	
	Représentation tridimensionnelle	Dégradé de couleur
$\begin{cases} X_1 = X_{10} \times \beta \\ X_2 = X_{20} \times \beta \end{cases} \quad (1)$		
$\begin{cases} X_1 = X_{10} \times \beta \\ X_2 = X_{20} / \beta \end{cases} \quad (2)$		

Dans ce tableau comparatif, nous voyons clairement que le modèle (2) présente une structure de dépendance 'opposée' par rapport à celle du modèle (1). Plus précisément, la fonction de densité de la copule empirique a connu une rotation de 90 degrés par rapport au centre du carré $[0 ; 1]^2$ pour passer du modèle (1) au modèle (2).

En bref, à l'aide du modèle (2), il est tout à fait possible de générer des structures de dépendance induisant des corrélations/dépendances négatives.

IV.1.9. Conclusion sur le choc multiplicatif log-normal

Le modèle de choc multiplicatif log-normal permet de reproduire des structures de dépendance symétriques.

Il présente un seul degré de liberté qu'il est simple de représenter directement par ρ , la corrélation linéaire de Pearson qu'il induit. La portée du modèle peut être limitée dans le cas où les marginales ont des coefficients de variations très différent, et ce par la relation suivante :

$$\rho = Cor[X_1, X_2] \leq \text{Min} \left(\frac{CV_1}{CV_2}; \frac{CV_2}{CV_1} \right) \quad (1)$$

Le Tau de Kendall et le Rho de Spearman peuvent être directement explicités en fonction du paramètre ρ , ce qui lui confère une possibilité de calibrage simple.

Les indicateurs de dépendance (Tau de Kendall, Rho de Spearman, et tail concentration) croissent dans le même sens que le paramètre ρ .

Enfin, une légère modification du modèle permet de générer des dépendances négatives.

IV.2. Choc multiplicatif log-normal shifté

IV.2.1. Présentation du modèle

Considérons le cas d'un choc commun multiplicatif sur deux branches d'assurance :

- la branche 1 a pour sinistralité : $X_1 = \beta \times X_{10}$
- la branche 2 a pour sinistralité : $X_2 = \beta \times X_{20}$
- le choc commun est : β

avec :

- $X_{10} \sim \text{LN}(\mu_{10}, \sigma_{10})$
- $X_{20} \sim \text{LN}(\mu_{20}, \sigma_{20})$
- $\beta \sim \text{LN}(1-C, \sigma_c) + C$

$\text{LN}(\mu, \sigma)$ désigne la loi log-normale

C désigne une constante positive strictement inférieure à 1. En prenant $C=0$, nous retombons sur un choc multiplicatif log-normal non-shifté.

L'espérance de β est prise égale à 1 pour que l'espérance de X_1 reste inchangée par rapport à X_{10} , et de même pour X_2 par rapport à X_{20} .

En reprenant les calculs réalisés pour les chocs commun multiplicatifs log-normaux non shiftés, nous obtenons :

$$Cor[X_1, X_2] = Cor[\beta X_{10}, \beta X_{20}] = \frac{E[X_{10}] \cdot E[X_{20}] \cdot Var[\beta]}{\sqrt{Var[X_1] \cdot Var[X_2]}} = \frac{Var(\beta)}{CV_1 \cdot CV_2}$$

Et également une limitation sur la corrélation linéaire que peut induire le choc est limitée par :

$$Cor[X_1, X_2] \leq \text{Min} \left(\frac{CV_1}{CV_2}; \frac{CV_2}{CV_1} \right)$$

Enfin, il est bon de noter que le modèle shifté présente deux paramètres (deux degrés de liberté) pour la structure de dépendance : la variance de β et le shift C .

IV.2.2. Détermination des marginales obtenues suite à l'application du choc

Intéressons-nous aux marginales obtenues : X_1 et X_2 .

En développant le calcul, nous nous apercevons qu'il s'agit de la somme de :

- une log-normale multipliée par une log-normale : qui donne une log-normale
- une log-normale multipliée par une constante : qui reste une log-normale.

Malheureusement, il n'existe pas de distribution usuelle caractérisant la somme de deux log-normales. Toutefois, il est généralement accepté qu'elle peut être raisonnablement approchée par une loi log-normale sur la queue droite, alors qu'au voisinage de 0, il n'y a pas de ressemblance avec les distributions log-normales.

Etant donné que nous sommes principalement intéressés par les valeurs extrêmement fortes de sinistralités (pour calculer des VaR à 99.5% et des SCR), il apparaît donc convenable d'approcher les marginales X_1 et X_2 par des log-normales.

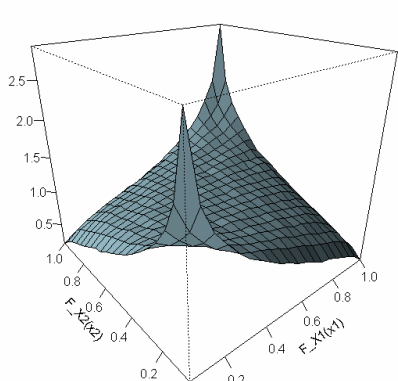
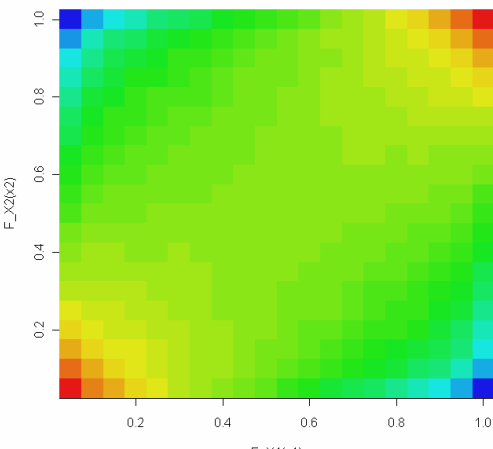
IV.2.3. Analyse qualitative de la structure de dépendance obtenue

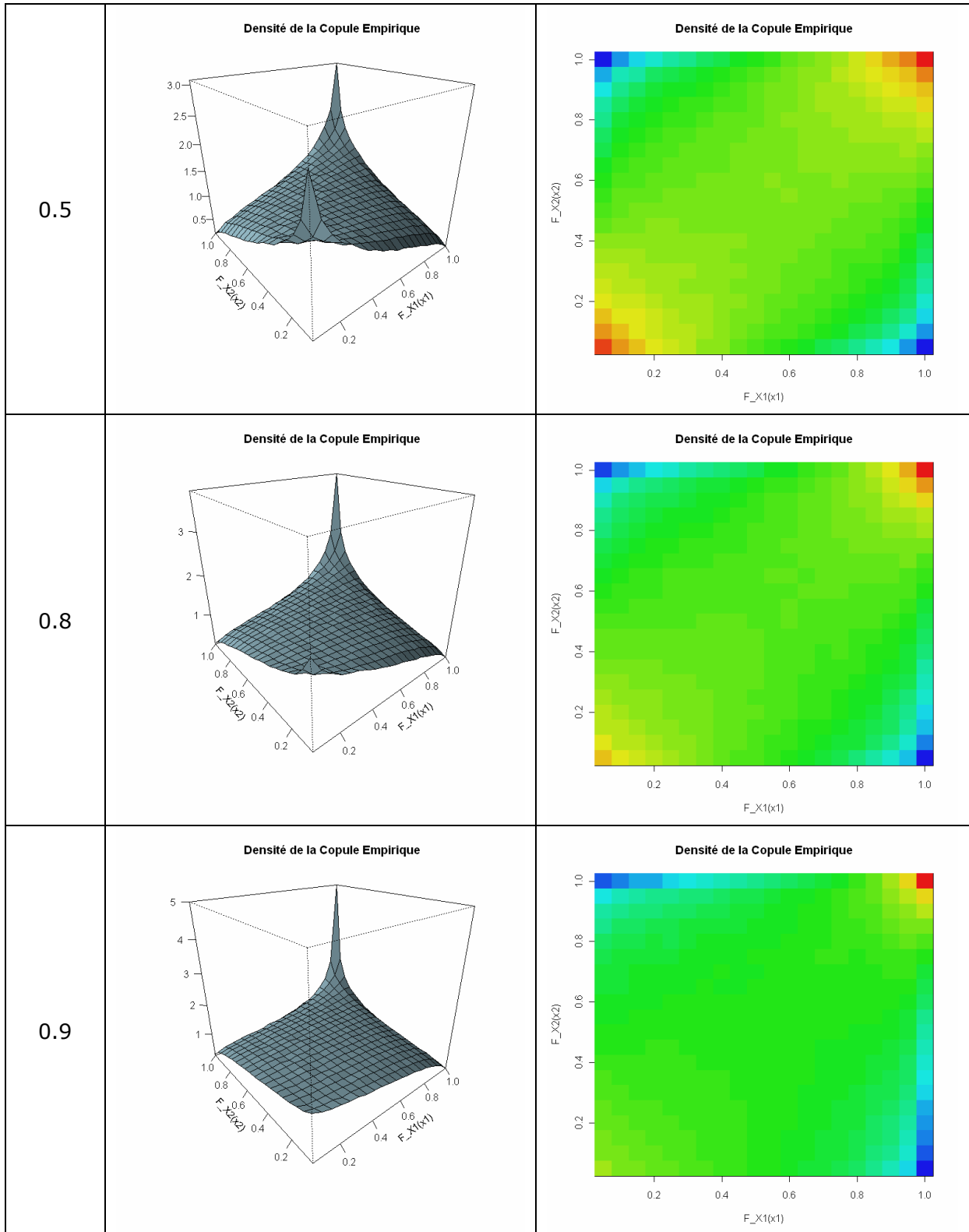
En prenant un shift (variable C) nul, nous retombons dans le cas précédent. Donc, pour appréhender la dépendance induite par un choc multiplicatif, nous allons faire varier le shift de 0 à 1 pour une valeur de corrélation linéaire donnée (ce qui revient à fixer la variance de β).

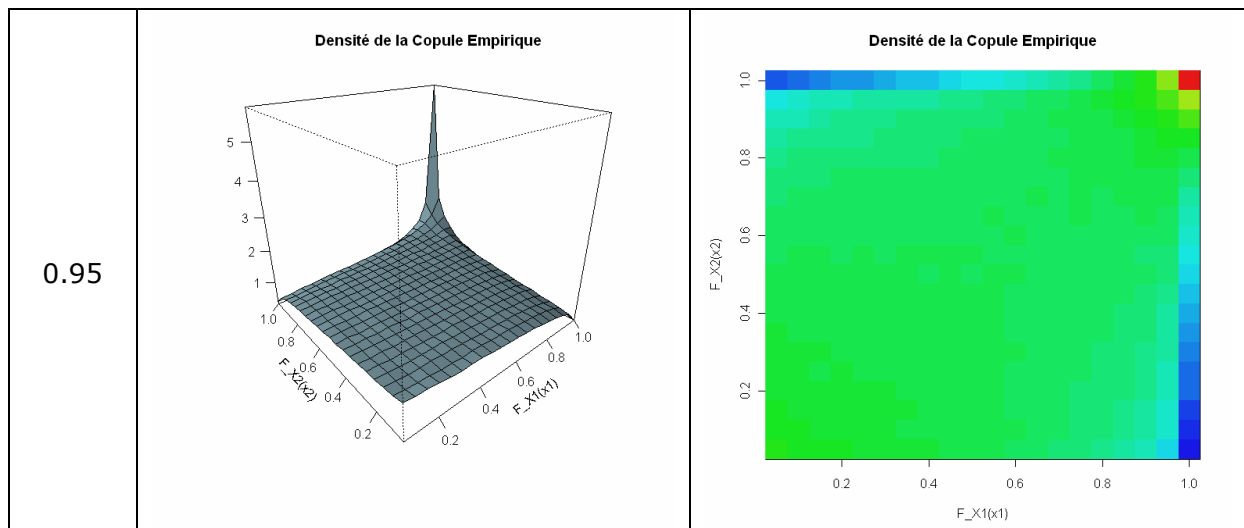
Donc, avec les hypothèses suivantes :

- $E[X_{10}] = 100$ et $Var(X_{10}) = 10^2$
- $E[X_{20}] = 100$ et $Var(X_{20}) = 15^2$
- $Var(\beta) = 0.06$ et donc $Cor(X_1, X_2) = .3$

Nous obtenons les résultats suivants :

Valeur du shift	Structure de dépendance : copule empirique liée au choc commun	
C	Représentation tridimensionnelle	Dégradé de couleur
0	<p style="text-align: center;">Densité de la Copule Empirique</p> 	<p style="text-align: center;">Densité de la Copule Empirique</p> 





Pour un shift nul, nous retrouvons une structure de dépendance symétrique : la survenance simultanée de sinistralité extrême est aussi probable pour des valeurs extrêmement fortes que pour des valeurs extrêmement faibles.

Nous pouvons constater deux effets quand la valeur du shift augmente :

- la structure de dépendance devient de plus en plus asymétrique en faveur des queues de distribution supérieures : la concentration pour les quantiles supérieurs $[\cdot 9 ; 1]^2$ augmente fortement alors que celle pour les quantiles inférieurs $[0 ; \cdot 1]^2$ diminue
- pour des valeurs de shift supérieures à $\cdot 9$, en dehors des quantiles supérieurs ($[\cdot 9 ; 1]^2$), la structure de dépendance tend à s'aplanir et donc à se rapprocher de celle du cas d'indépendance ; ce qui correspondrait à un cas d'indépendance sauf pour le cas où une des deux branches connaissent une sinistralité élevée.

Nous observons le même phénomène pour d'autres valeurs de $\text{Var}(\beta)$: cf [VIII.3. Résultats obtenus pour le choc commun multiplicatif log-normal shifté](#)

Pour conclure, l'introduction du shift permet de rendre la structure de dépendance asymétrique, et l'asymétrie ainsi produite est d'autant plus forte que la valeur du shift est élevée.

IV.2.4. Mise en parallèle avec les copules

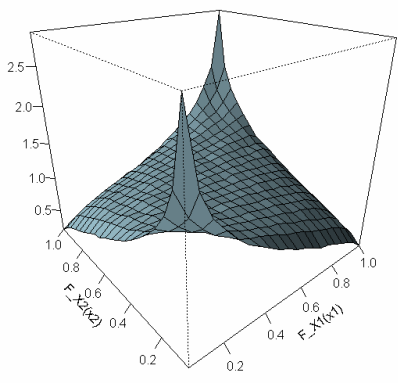
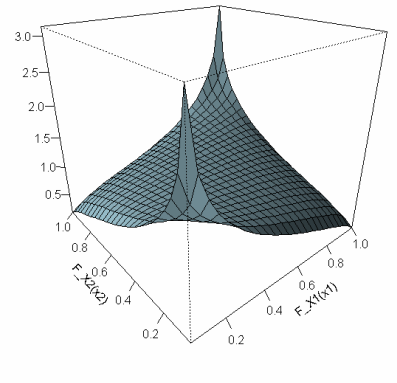
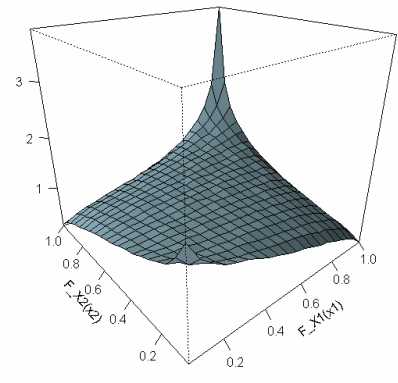
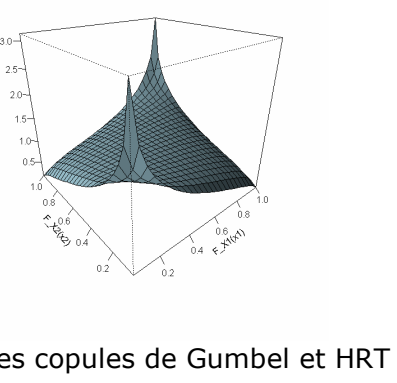
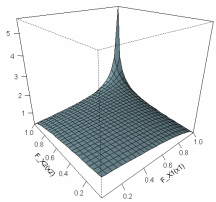
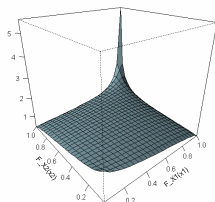
Nous avons testé la possibilité de rapprochement entre la structure de dépendance induite par le choc commun avec celle produite par une copule. Pour ce faire, nous avons considéré les copules usuelles suivantes :

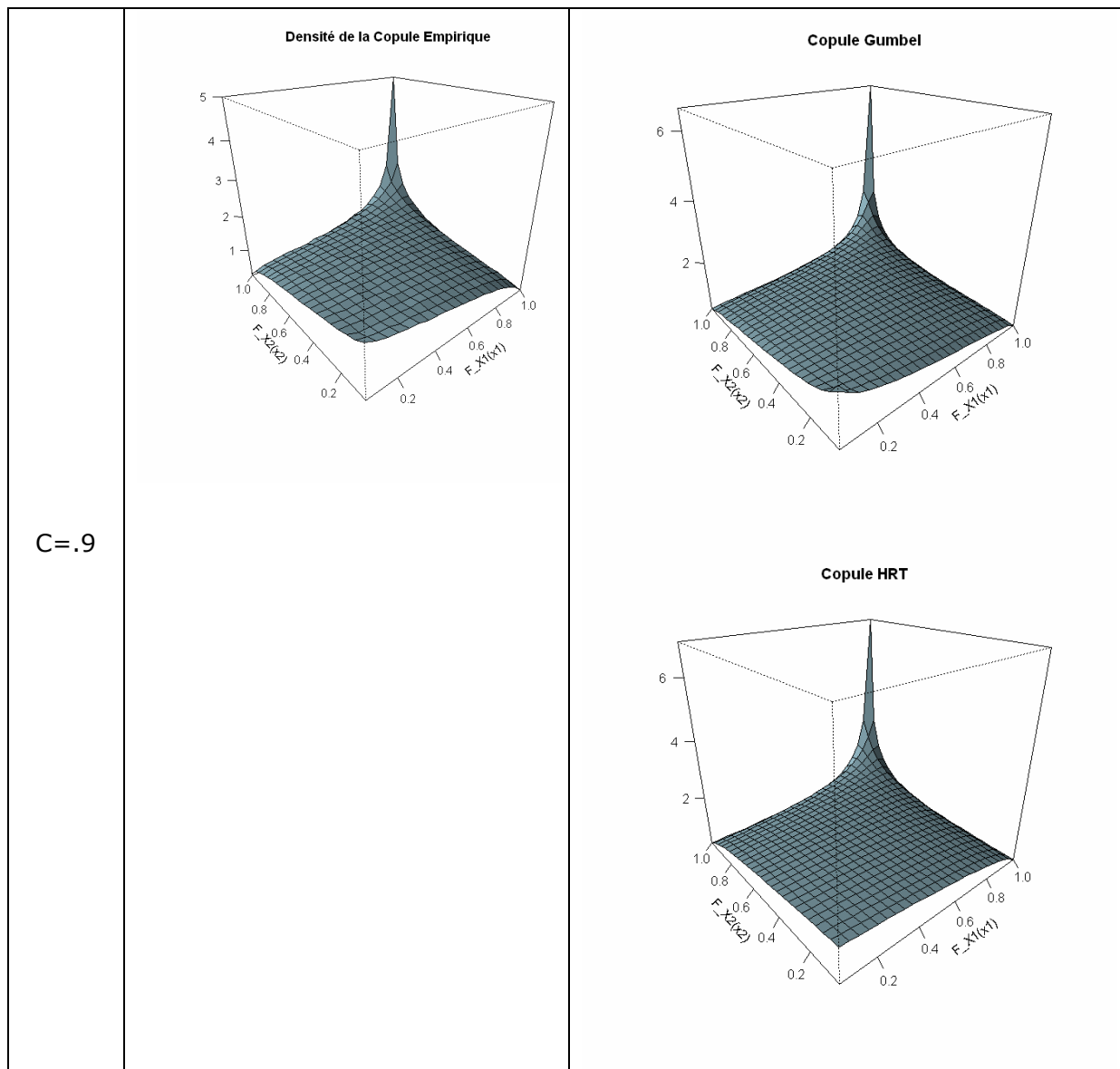
- copule Normale
- copule Gumbel
- copule Clayton
- copule HRT
- copule de Student

Toutefois, la copule de Student n'a pas été retenue. En effet, elle a pour paramètres une matrice symétrique définie positive et un nombre de degrés de liberté ; ce dernier paramètre requiert une procédure de maximisation de vraisemblance compliquée.

Pour fitter ces copules, nous avons utilisé les méthodes simples et usuelles listées dans le tableau figurant dans le chapitre [III.6.1. Fit d'une copule empirique avec les copules usuelles](#).

Regardons visuellement la mise en parallèle de la copule empirique induite par le choc commun avec la copule usuelle (parmi les copules : Normale, de Gumbel, de Clayton, HRT) s'en rapprochant le plus, avec $\text{cor} = .3$:

Paramètres du modèle	Structure de dépendance (Représentation tridimensionnelle) :	
	Copule empirique liée au choc commun	Copule usuelle la plus 'proche'
C=0	<p>Densité de la Copule Empirique</p> 	<p>Copule Gaussienne</p> 
C=.8	<p>Densité de la Copule Empirique</p> 	<p>Entre la copule gaussienne :</p> <p>Copule Gaussienne</p>  <p>et les copules de Gumbel et HRT :</p> <div style="display: flex; justify-content: space-around;"> <div style="text-align: center;"> <p>Copule HRT</p>  </div> <div style="text-align: center;"> <p>Copule Gumbel</p>  </div> </div>



Comme il a été vu pour le choc commun log-normal non shifté, la structure de dépendance produite avec $C=0$ est identique à celle de la copule normale.

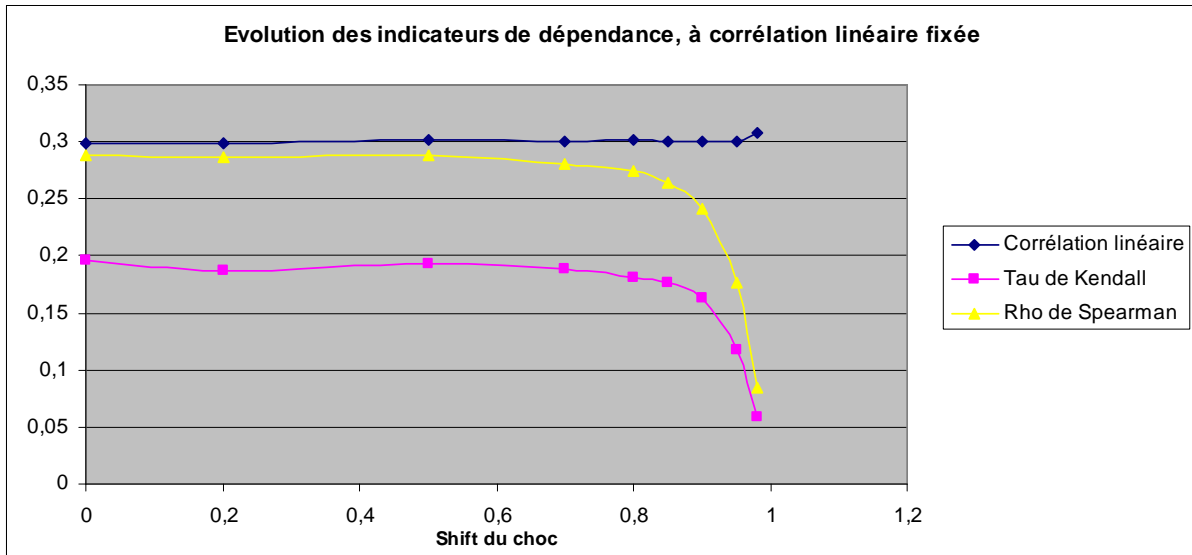
Pour des valeurs de shift importantes, supérieures à .9, la structure de dépendance peut être rapprochée de celle de la copule de Gumbel ou de HRT.

Par contre, dans les situations intermédiaires (shift ayant des valeurs comprises entre .3 et .8), il n'est guère possible de faire de mise en parallèle avec des copules usuelles ; visuellement, la structure de dépendance obtenue semble être à mi-chemin entre la copule normale et la copule HRT ou de Gumbel.

IV.2.5. Analyse quantitative de la structure de dépendance obtenue

Il n'existe pas de formule fermée facilement exploitable pour les indicateurs de dépendance, dans le cas d'un choc commun multiplicatif log-normal shifté. Toutefois, des résultats issus de simulations numériques permettent de faire apparaître certains points intéressants.

Voici le graphe des indicateurs de dépendance, avec une corrélation linéaire fixée à 0.3 :



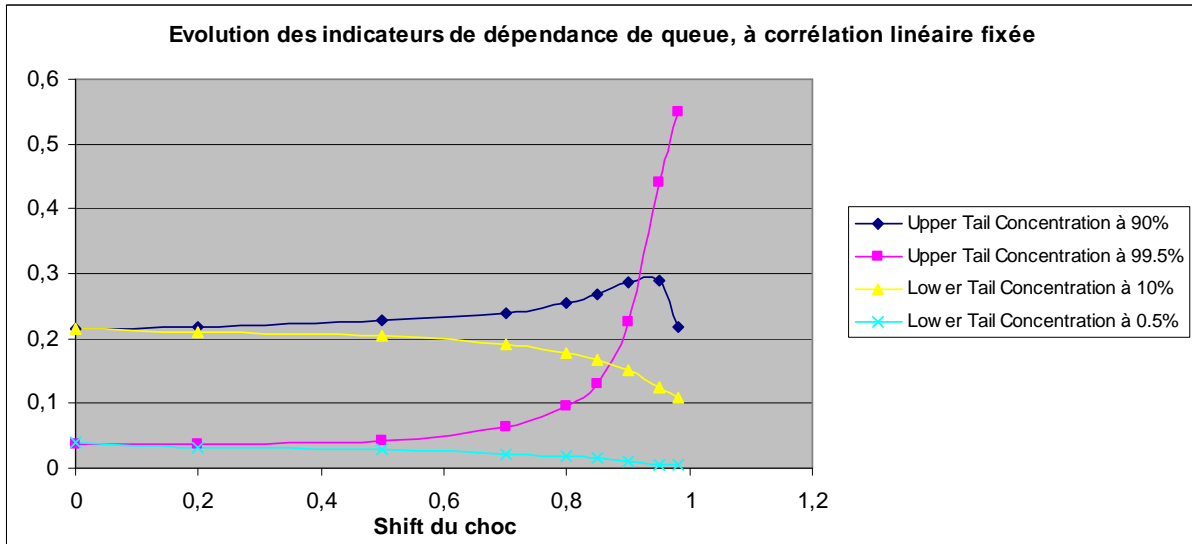
Pour un shift nul, nous retrouvons les relations suivantes :

$$\tau_{Kendall} = \frac{2}{\pi} \cdot \text{Arc sin}(\rho)$$

$$\rho_{Spearman} = \frac{6}{\pi} \cdot \text{Arc sin}\left(\frac{\rho}{2}\right)$$

Il est intéressant de constater que plus le shift est important, plus le Tau de Kendall et le Rho de Spearman ont tendance à décroître.

Voici le graphe des indicateurs de dépendance de queue, toujours pour la corrélation linéaire égale à 0.3 :



Nous retrouvons notre analyse visuelle précédente, à savoir, que plus le shift est important, plus l'asymétrie est forte :

- dans le cas d'un shift nul, nous avons que les upper tail concentration à 90% et 99.5% sont respectivement égaux aux lower tail concentration à 10% et .5%
- puis, l'écart se creuse fortement, notamment pour les tail concentrations à 99.5% et .5%

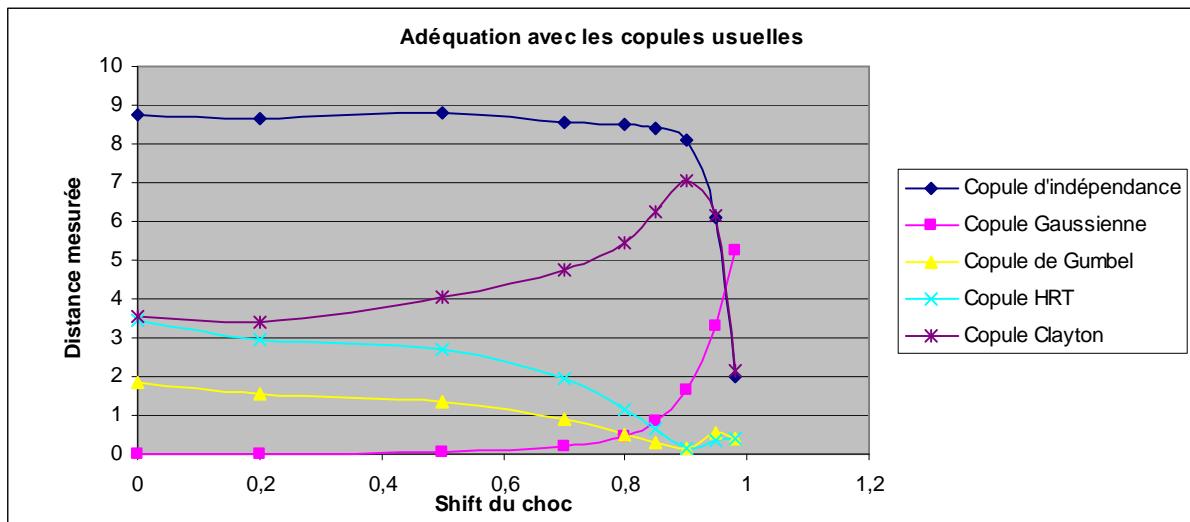
- pour des valeurs extrêmes du shift (.98), la upper tail concentration à 90% tend à diminuer ; ce phénomène est dû au fait que le pic observé dans la structure de dépendance s'agrandit tout en voyant sa base diminuée ; de ce fait, la concentration dans le carré [0.9 ;0.99]² est moins impactée par le pic, d'où la diminution du upper tail concentration à 90%

Il semble donc que le shift a deux effets :

- au niveau de la dépendance de queue : augmentation de l'asymétrie de la structure de dépendance, ce qui est visible via les indicateurs de dépendance de queue
- au niveau de la dépendance générale : diminution de la dépendance de rang.

Enfin, regardons les graphes des distances avec les copules usuelles en fonction de shift, en reprenant la notion de distance introduite dans la partie [III.6.2.2. Distance entre les classes](#) :

$$D_{\text{copule testée}} = \sum_{1 \leq i, j \leq k} (p_{i,j}^{\text{empirique}} - p_{i,j}^{\text{copule testées}})^2$$



Nous retrouvons les remarques visuelles faites précédemment : pour des shifts nuls ou de faible valeur, il y a adéquation avec une copule normale ; pour des valeurs proches de .9, il est possible de faire un rapprochement avec les copules HRT et de Gumbel.

Et de plus, pour des valeurs de shift très proches de 1, la structure de dépendance se rapproche de celle d'un cas d'indépendance.

Regardons le tableau de valeur (correspondant au graphe exposé précédemment) :

Shift	D avec copule d'indépendance	D avec copule gaussienne	D avec copule Gumbel	D avec copule Clayton	D avec copule HRT
0,000	8,769	0,009	1,844	3,527	3,459
0,200	8,637	0,013	1,540	3,406	2,934
0,500	8,794	0,045	1,354	4,063	2,696
0,700	8,573	0,178	0,921	4,739	1,940
0,800	8,523	0,462	0,498	5,471	1,135
0,850	8,418	0,853	0,286	6,248	0,646
0,900	8,111	1,655	0,158	7,043	0,156
0,950	6,088	3,298	0,531	6,135	0,335
0,980	2,020	5,265	0,394	2,162	0,395

Le tableau de valeurs, indique qu'une adéquation mathématique (parfaite) correspond à des distances mesurées <0,01 . Aussi, les rapprochements avec les copules de Gumbel et de HRT (qui présentent des distances de l'ordre de 0.15) ne peuvent être considérés que comme des approximations.

IV.2.6. Extension du modèle

Le modèle présenté était le suivant :

$$\begin{cases} X_1 = X_{10} \times \beta \\ X_2 = X_{20} \times \beta \end{cases}$$

il permettait de créer des niveaux de dépendance très forts pour les sinistralités extrêmement fortes.

En considérant le modèle suivant :

$$\begin{cases} X_1 = X_{10} / \beta \\ X_2 = X_{20} / \beta \end{cases}$$

il est alors possible de créer des niveaux de dépendance très forts pour les sinistralités extrêmement faibles (et non plus celles extrêmement fortes)

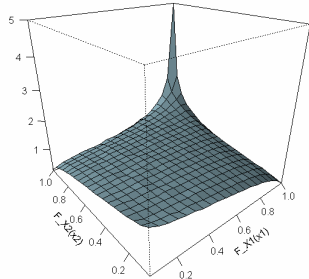
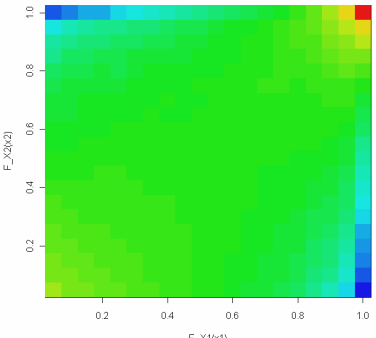
De plus, La version présentée du modèle semble ne pas permettre des dépendances à caractère négatif.

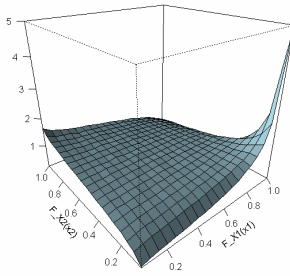
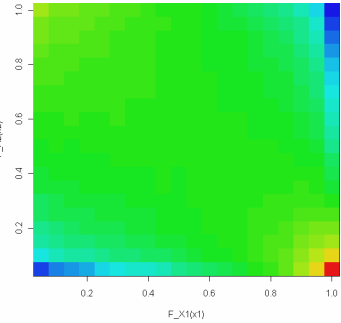
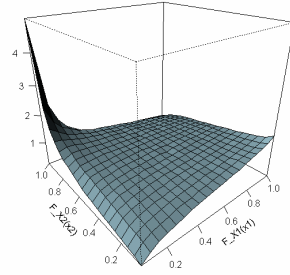
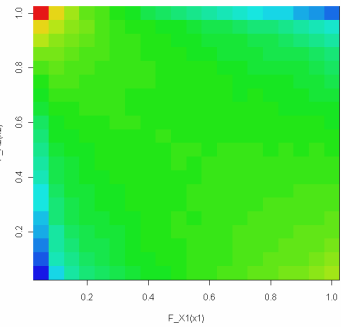
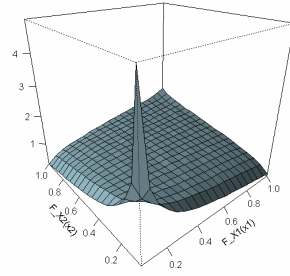
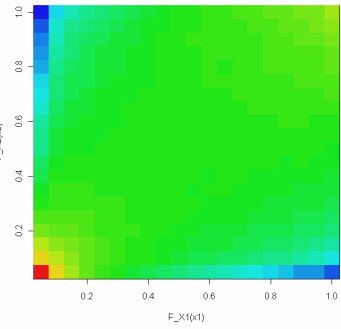
Toutefois, il est possible d'obtenir des indicateurs de dépendance négatifs en prenant le modèle suivant :

$$\begin{cases} X_1 = X_{10} \times \beta \\ X_2 = X_{20} / \beta \end{cases} \text{ ou } \begin{cases} X_1 = X_{10} / \beta \\ X_2 = X_{20} \times \beta \end{cases}$$

A titre d'exemple, considérons le cas suivant :

- $E[X_{10}] = 100$ et $\text{Var}(X_{10}) = 10^2$
- $E[X_{20}] = 100$ et $\text{Var}(X_{20}) = 15^2$
- $\text{Var}(\beta) = 0.06$ et donc que $\text{Cor}(X_1, X_2) = .3$
- $C = .9$ (valeur du shift linéaire)

Modèle utilisé	Structure de dépendance : copule empirique liée au choc commun	
	Représentation tridimensionnelle	Dégradé de couleur
$\begin{cases} X_1 = X_{10} \times \beta \\ X_2 = X_{20} \times \beta \end{cases} \quad (1)$		

$\begin{cases} X_1 = X_{10} \times \beta \\ X_2 = X_{20} / \beta \end{cases} \quad (2)$	<p>Densité de la Copule Empirique</p> 	<p>Densité de la Copule Empirique</p> 
$\begin{cases} X_1 = X_{10} / \beta \\ X_2 = X_{20} \times \beta \end{cases} \quad (3)$	<p>Densité de la Copule Empirique</p> 	<p>Densité de la Copule Empirique</p> 
$\begin{cases} X_1 = X_{10} / \beta \\ X_2 = X_{20} / \beta \end{cases} \quad (4)$	<p>Densité de la Copule Empirique</p> 	<p>Densité de la Copule Empirique</p> 

Dans ce tableau comparatif, nous voyons clairement que le modèle (2) et (3) présentent des structures de dépendance 'opposées' et différentes par rapport à celle du modèle (1). Plus précisément, la fonction de densité de la copule empirique a connu une rotation de 90 degrés par rapport au centre du carré [0 ;1]² pour passer du modèle (1) au modèle (2).

La structure de dépendance du modèle (4) correspond à une rotation d'un angle de 180 degrés. L'asymétrie est inversée par rapport au modèle (1). De ce fait, le rapprochement avec les copules usuelles de Gumbel ou HRT n'a plus lieu ; c'est maintenant la copule de Clayton qui est la plus assimilable au modèle (4).

IV.2.7. Conclusion sur le choc log-normal shifté

Le modèle à choc log-normal shifté est une généralisation du modèle log-normal non shifté : en effet, en prenant un shift de valeur nulle, nous retombons sur ce dernier.

Le modèle de choc commun multiplicatif log-normal shifté permet de jouer sur deux paramètres :

- la valeur du shift : ce paramètre conditionne la forme de la dépendance produite en modifiant son asymétrie ; il est alors possible de jouer sur le rapport entre les upper tail concentration et les lower tail concentration
- la corrélation linéaire qu'il induit : ce paramètre permet d'accentuer la forme de la dépendance produite ; les indicateurs de dépendance linéaire et de rang sont impactés par la modification de ce paramètre.

De plus, en fonction de la variante prise (multiplicatif ou diviseur), il est possible de reproduire une asymétrie portant sur les sinistralités extrêmement fortes ou extrêmement faibles.

Par ailleurs, une légère modification du modèle permet de générer des dépendances négatives.

Enfin, un inconvénient est que la portée du modèle peut être limitée dans le cas où les marginales ont des coefficients de variations très différent, et ce par la relation suivante :

$$\rho = Cor[X_1, X_2] \leq Min\left(\frac{CV_1}{CV_2}; \frac{CV_2}{CV_1}\right)$$

V. APPLICATION A UN CAS CONCRET

V.1. Objectifs de l'application à un cas concret

Les objectifs de l'application au cas concret sont principalement d'appréhender l'impact des modèles de chocs communs sur des données réelles, de vérifier la faisabilité de leur calibration et de présenter un cas réel d'analyse de structure de dépendance. Le présent mémoire s'inscrit dans une démarche exploratoire, aussi, il n'abordera pas les tests de validation et de robustesse qui seraient requis pour une utilisation certifiée dans un modèle interne.

L'application s'intéresse à la dépendance entre LoBs (branches d'assurance) pour le risque de prime. Le risque de prime traite de la nouvelle année de souscription et du risque que les primes ne soient pas suffisantes pour faire face à la charge de sinistre.

V.2. Présentation du portefeuille

V.2.1. Structure des bases de données

Le portefeuille étudié est celui d'une des filiales de Groupama.

Nous avons pris un historique de 8 ans, de 1999 à 2006 inclus. Les branches suivantes ont été considérées :

- Accident
- Maladie
- AutoRisquesdivers
- BateauxRisquesdivers
- Transports
- Incendie
- Autresdommagesauxbiens
- RCAutos
- RCBateaux
- ResponsabilitéCivile
- Crédit
- Cautions
- Pertesfinancières
- Protectionjuridique
- Assistance

Nous avons utilisés trois types de données : les données d'inflation, les nombres de polices et les données relatives aux sinistres.

Le tableau suivant résume les données d'inflation et de sur-inflation :

Année	1999	2000	2001	2002	2003	2004	2005	2006
Inflation économique	1,6%	2,5%	2,7%	2,4%	2,7%	2,2%	1,9%	2,1%
Sur-inflation (uniquement pour RCA et RCD)	3,0%	3,0%	3,0%	3,0%	3,0%	3,0%	3,0%	2,0%

La sur-inflation correspond à l'inflation spécifique à certaines branches (ici RCA et RCD) qu'il faudra ajouter en supplément de l'inflation économique.

Nous disposons des nombres de polices du portefeuille pour chaque année.

Enfin, concernant les données des sinistres, il existe une base de données listant tous les sinistres survenus dans les années allant de 1999 à 2006. Voici ses colonnes susceptibles de nous intéresser dans le cadre de notre étude :

- C_SIN_DNN : Numéro d'identification du sinistre dans le système central
- N_RM_MNST : Numéro de garantie ministérielle
- D_CTB_APR_A : Date à laquelle le sinistre a été ouvert dans le système
- D_AVV_SIN_A : Date à laquelle est survenu le sinistre au format date
- I_RSV_ADAT{1999/.../2006} : Réserve à la fin de la période
- PMT{1999/.../2006} : Paiements de la période nets des recours et comprenant les frais légaux de la contrepartie

Dans cette base, une ligne correspond donc à un sinistre rattaché à une branche ministérielle.

V.2.2. Traitements des données et préparation des variables à modéliser

Les données seront redressées et ramenées dans les conditions de l'année de référence, l'année 2006 ; elles seront alors comparables entre elles. Les données ainsi retraitées seront qualifiées de « as if ».

Dans notre étude, nous modéliserons in fine la charge ultime pour une année de survenance en « as if », et ce par LoB au sens Solvency II.

Dans un premier temps nous chercherons à obtenir les charges ultimes mensuelles par mois de survenance, en « as if ».

Aussi, assez classiquement, les retraitements suivants doivent être réalisés :

- retraitement de l'inflation et de la sur-inflation : il s'agit de ramener les montants en € de la date de référence ;
- retraitement « as if » en fonction de la taille du portefeuille : il s'agit de prendre en compte que si le portefeuille grossit, les montants d'alors doivent être multipliés par un facteur correspondant au développement du portefeuille

De plus, afin d'obtenir les charges ultimes, nous avons utilisé la méthode classique du Chain Ladder.

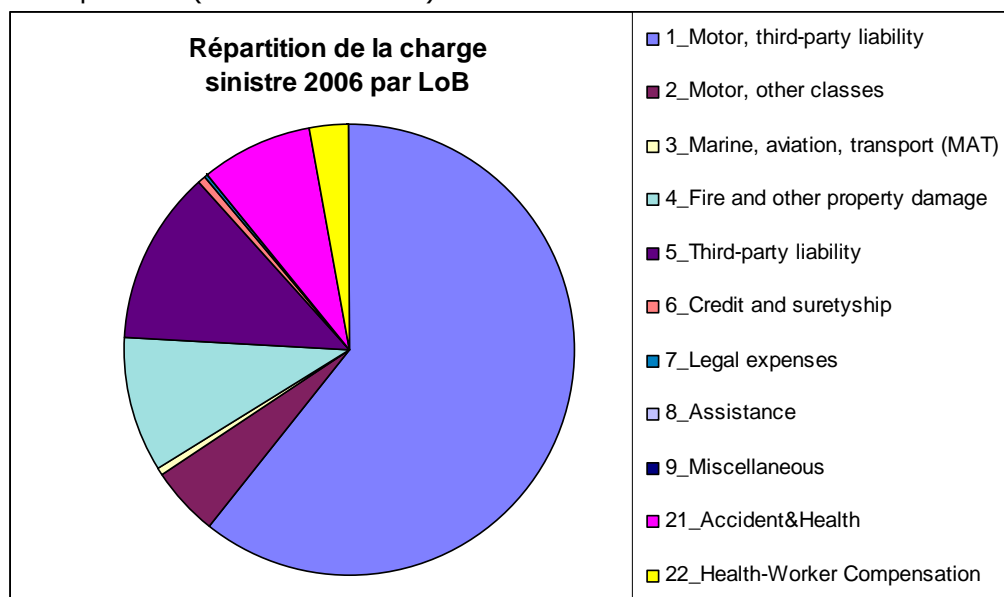
Enfin, nous avons agrégé les sinistres par mois de survenance.

Afin d'obtenir les données par LoB (Line Of Business) au sens de Solvency II, nous avons ensuite utilisé les règles d'affectations résumées dans le tableau suivant :

#LoB	LoB Solvency 2	Branche de l'entité
1	Motor, third party liability	RCAutos + RCBateaux
2	Motor, other classes	AutoRisquesdivers + BateauxRisquesdivers
3	Marine, aviation and transport	Transports
4	Fire and other damage to property	Incendie + Autresdommagesauxbiens
5	Third-party liability	ResponsabilitéCivile
6	Credit and suretyship	Crédit + Cautions
7	Legal expenses	Protectionjuridique
8	Assistance	Assistance
9	Miscellaneous non-life insurance	Pertesfinancières
21	Accident&Health	Accident
22	Health-Worker Compensation	Maladie

V.2.3. Statistiques exploratoires

Regardons dans un premier temps la charge sinistre ultime pour l'année de survenance 2006 par LoB (Line Of Business) :

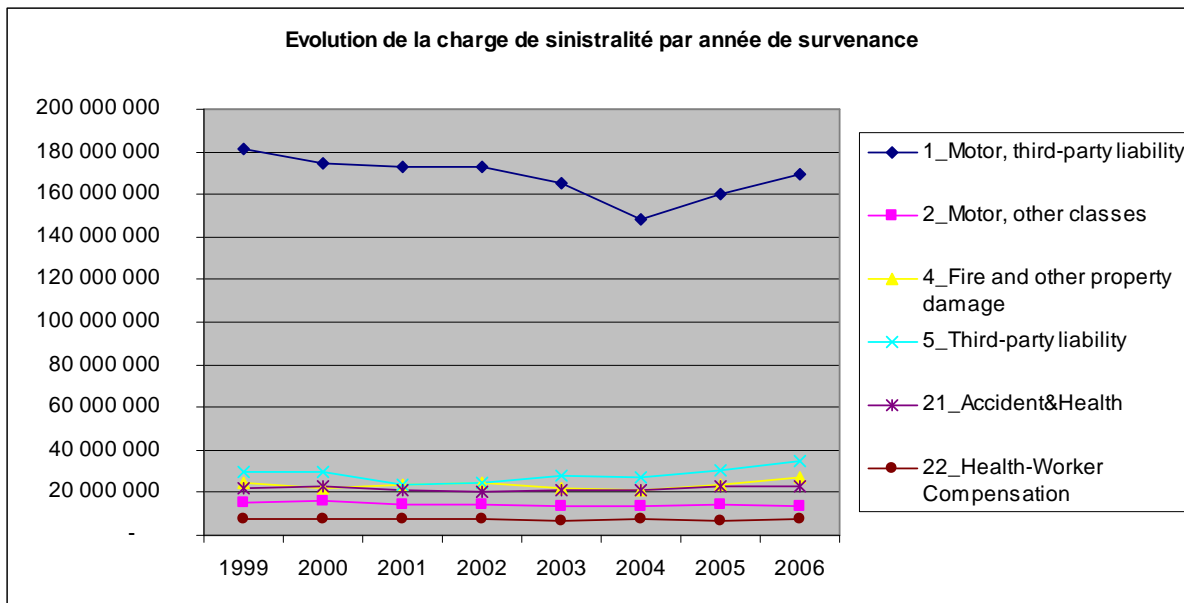


Nous constatons de suite que certaines LoB jouent un rôle négligeable dans la sinistralité globale de la filiale. Aussi, nous nous concentrerons sur les LoB indiquées en gras dans le tableau suivant :

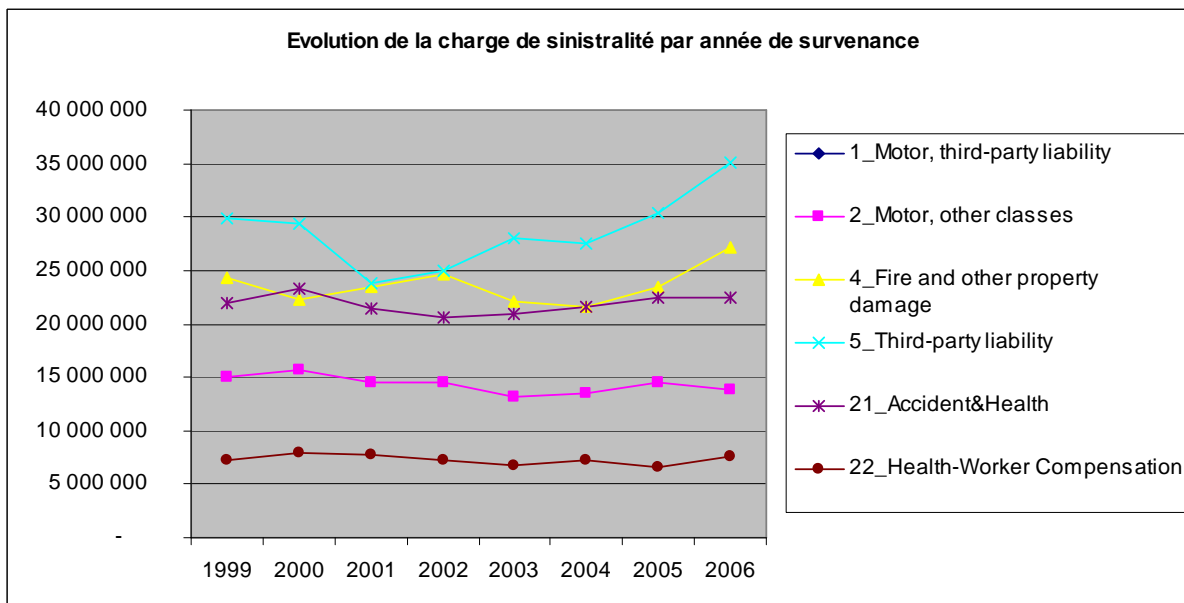
LoB	Charge Sinistre 2006
1_Motor, third-party liability	169 663 392
2_Motor, other classes	13 854 301
3_Marine, aviation, transport (MAT)	1 402 566
4_Fire and other property damage	27 182 169
5_Third-party liability	35 125 120
6_Credit and suretyship	1 521 192
7_Legal expenses	818 457
8_Assistance	183 087
9_Miscellaneous	164 531
21_Accident&Health	22 457 643
22_Health-Worker Compensation	7 634 860

Dans la suite de notre étude, nous considérerons donc seulement 6 LoBs. En effet, en accord avec la règle de proportionnalité, il n'est pas utile de mettre en œuvre une méthode complexe pour des branches présentant des charges de sinistralité relativement faibles.

Par ailleurs, regardons l'évolution des charges de sinistralité « as if » par année de survenance pour voir s'il y a des tendances (« trends ») significatives qu'il pourrait être utile d'identifier.



En zoomant sur la tranche 0€ - 40 M€ en ordonnée, nous obtenons :

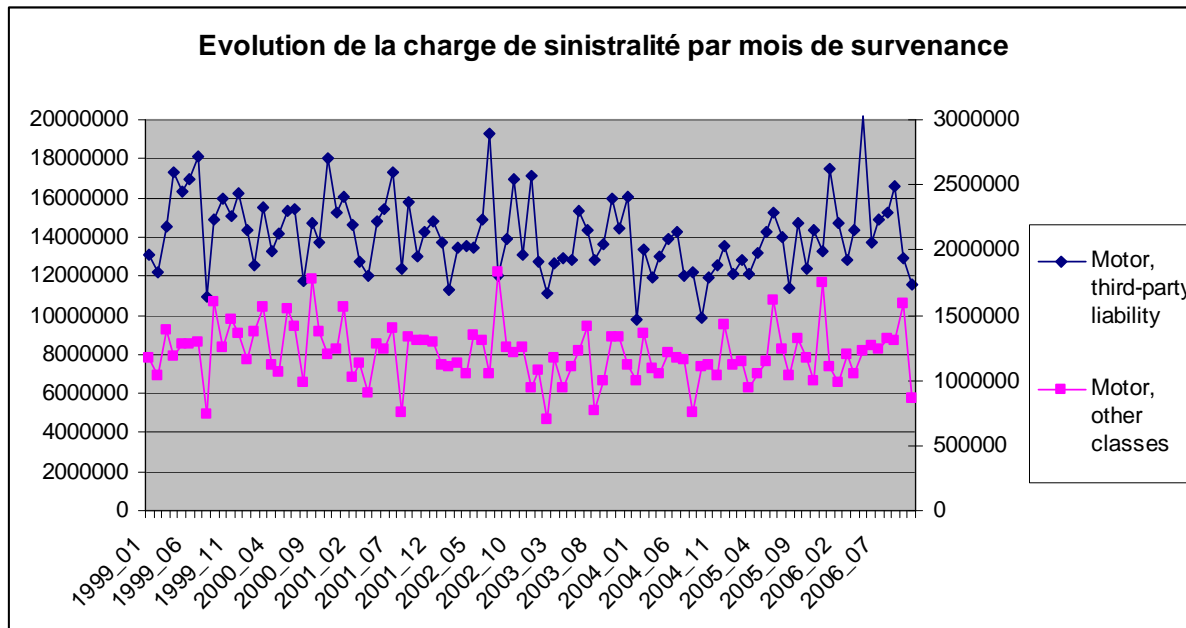


Ces deux graphiques ne révèlent pas de tendance significative.

Calculons la moyenne, l'écart type et le coefficient de variation de la charge annuelle pour les 6 LoB étudiées :

LoB	Moyenne	Ecart-type	CV
1_Motor, third-party liability	168 066 031	10 278 060	0,06
2_Motor, other classes	14 369 986	841 036	0,06
4_Fire and other property damage	23 634 177	1 799 668	0,08
5_Third-party liability	28 644 746	3 505 736	0,12
21_Accident&Health	21 835 728	870 012	0,04
22_Health-Worker Compensation	7 305 511	465 965	0,06

Enfin, à titre d'information voici les évolutions de la charge de sinistralité par mois de survenance, pour 2 des 6 LoB considérées : « Motor, Third-Party liability » et « Motor, other classes »



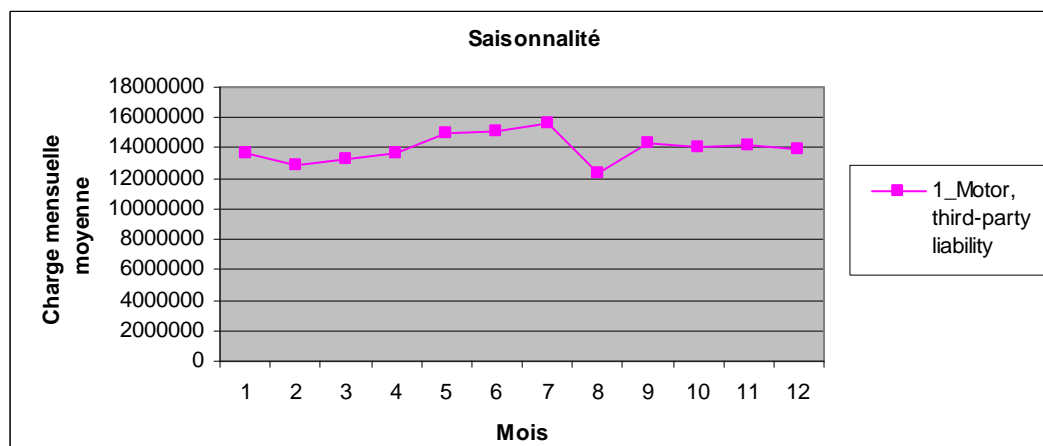
V.3. Exploration de la dépendance observée

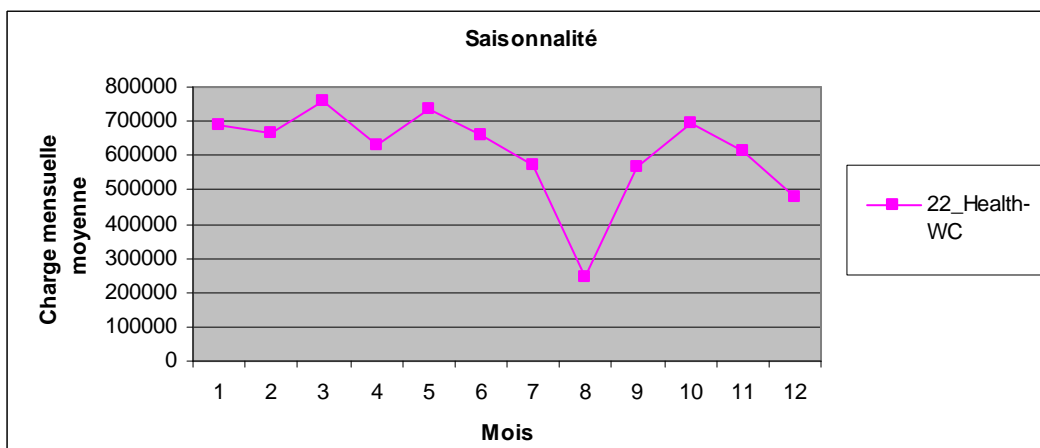
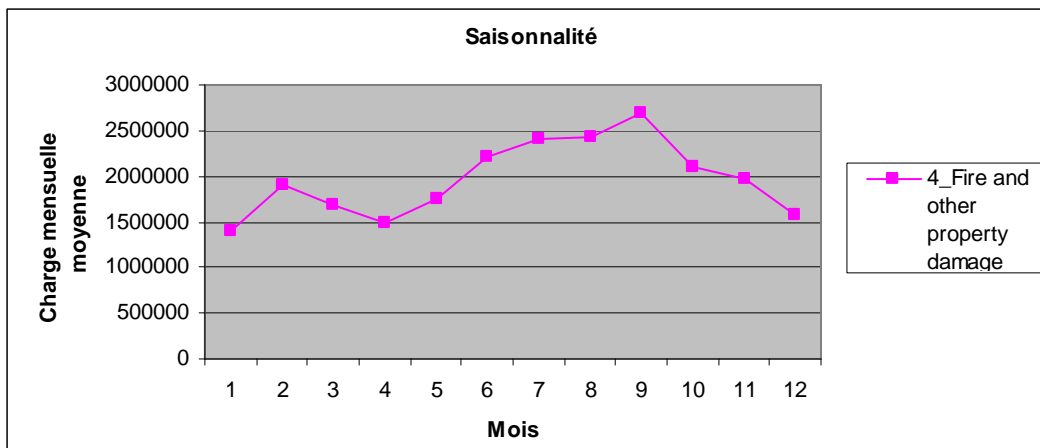
V.3.1. Prise en compte de la saisonnalité

Notre étude va se concentrer dans un premier temps sur la charge de sinistralité mensuelle à l'ultime, pour obtenir en un deuxième temps la charge annuelle.

La vision annuelle est la vision qui nous intéresse in fine.

Aussi, il convient de d'abord de traiter la saisonnalité mensuelle qui n'est pas négligeable. Afin de visualiser l'impact de l'effet de saisonnalité, voici les graphiques représentant les charges annuelles mensuelles moyennes (moyennes calculées sur les 8 années), par mois de survenance:





Pour la RC Auto, nous voyons que la charge mensuelle fluctue de manière non négligeable en fonction du mois dans l'année. Pour le Worker Compensation, il y a une baisse notable pour le mois d'août qui peut être expliquée par les congés estivaux et par le beau temps peu propice aux maladies. Au contraire pour la LoB « Fire and other property damage », l'hivers est moins sinistré que l'été.

Aussi, **nous analyserons les moyennes mobiles sur 12 périodes des charges mensuelles de sinistralité.** L'effet de saisonnalité mensuelle sera ainsi lissé.

V.3.2. Mesures de la dépendance

V.3.2.1. Corrélation linéaire

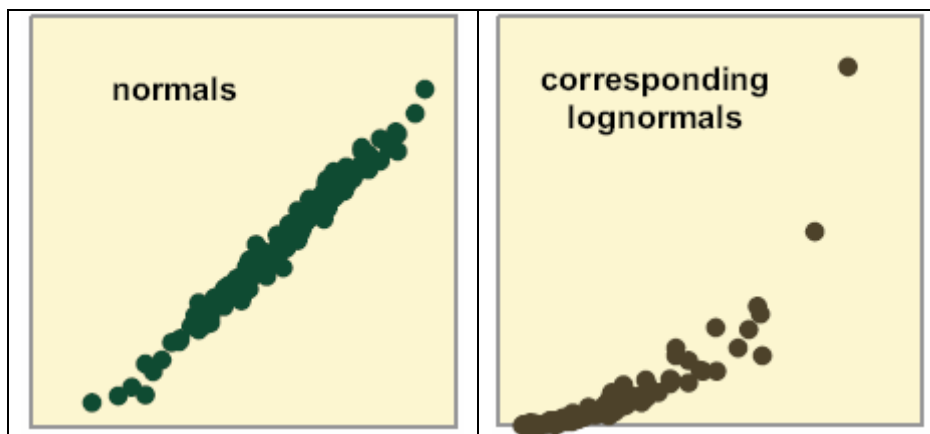
Intéressons-nous d'abord à la matrice de corrélation (au sens traditionnel de Pearson) sur les valeurs des moyennes mobiles sur 12 périodes des charges mensuelles :

	1	2	4	5	21	22
1 - Motor, third-party liability	1,00	0,72	0,38	0,22	0,34	0,63
2 - Motor, other classes	0,72	1,00	0,22	0,17	0,54	0,59
4 - Fire and other property damage	0,38	0,22	1,00	0,27	0,27	0,23
5 - Third-party liability	0,22	0,17	0,27	1,00	0,76	0,26
21 - Accident&Health	0,34	0,54	0,27	0,76	1,00	0,60
22 - Worker Compensation	0,63	0,59	0,23	0,26	0,60	1,00

Il est bon d’avoir en tête que la corrélation linéaire de Pearson n’a de sens que si les lois sous-jacentes sont des gaussiennes. Or, comme dans notre cas, nous nous intéressons à des log-normales, il est peut-être préférable de regarder les résultats obtenus en s’intéressant aux logarithmes des charges de sinistralité.

En effet, il peut y avoir des écarts entre les corrélations linéaires mesurées sur une échelle logarithmique et celles mesurées sur l’échelle d’origine. Par exemple, une corrélation linéaire forte sur l’échelle logarithmique couplée à des coefficients de variation élevés, peut ne donner lieu qu’à une corrélation linéaire faible sur l’échelle d’origine.

Voici un cas où il y a une relation linéaire sur les normales, mais à cause de coefficients de variation importantes, le passage à l’exponentiel déforme la relation linéaire et tend à diminuer le coefficient de relation linéaire :



Regardons alors dans notre cas la matrice de corrélation portant sur les logarithmes des charges de sinistralité :

	1	2	4	5	21	22
1 - Motor, third-party liability	1,00	0,72	0,39	0,19	0,33	0,63
2 - Motor, other classes	0,72	1,00	0,24	0,16	0,55	0,59
4 - Fire and other property damage	0,39	0,24	1,00	0,19	0,24	0,22
5 - Third-party liability	0,19	0,16	0,19	1,00	0,74	0,23
21 - Accident&Health	0,33	0,55	0,24	0,74	1,00	0,59
22 - Worker Compensation	0,63	0,59	0,22	0,23	0,59	1,00

Nous constatons qu’il y a très peu d’écart. Ceci est principalement dû au fait que les coefficients de variation des lois de charge de sinistralité sont faibles :

	Moyenne	Ecart-type	CV
1_Motor, third-party liability	14 024 016	751 476	0,05
2_Motor, other classes	1 200 494	69 954	0,06
4_Fire and other property damage	1 946 658	180 915	0,09
5_Third-party liability	2 387 996	321 930	0,13
21_Accident&Health	1 827 834	91 867	0,05
22_Health-Worker Compensation	614 004	35 393	0,06

N.B. : les CV calculés dans ce tableau correspondent à la variation constatée sur les charges mensuelles (calculée par moyenne mobile)

Les corrélations linéaires étant très proches que l’échelle soit logarithmique ou non, dans la suite de notre étude, nous étudierons les dépendances directement sur les valeurs d’origine, sans passer par la fonction logarithmique.

V.3.2.2. Indicateurs de dépendance

Regardons maintenant les matrices des principaux indicateurs de dépendance (cf [II.4.2. Les principaux indicateurs de dépendance](#))

Voici les coefficients de Kendall :

	1	2	4	5	21	22
1 - Motor, third-party liability	1,00	0,57	0,22	0,13	0,21	0,50
2 - Motor, other classes	0,57	1,00	0,23	0,10	0,38	0,42
4 - Fire and other property damage	0,22	0,23	1,00	-0,06	0,07	0,12
5 - Third-party liability	0,13	0,10	-0,06	1,00	0,43	0,08
21 - Accident&Health	0,21	0,38	0,07	0,43	1,00	0,36
22 - Worker Compensation	0,50	0,42	0,12	0,08	0,36	1,00

Et les coefficients de Spearman :

	1	2	4	5	21	22
1 - Motor, third-party liability	1,00	0,79	0,31	0,20	0,39	0,72
2 - Motor, other classes	0,79	1,00	0,29	0,20	0,56	0,63
4 - Fire and other property damage	0,31	0,29	1,00	-0,06	0,09	0,18
5 - Third-party liability	0,20	0,20	-0,06	1,00	0,63	0,12
21 - Accident&Health	0,39	0,56	0,09	0,63	1,00	0,53
22 - Worker Compensation	0,72	0,63	0,18	0,12	0,53	1,00

V.3.2.3. Conclusion

Les coefficients de corrélation linéaire et les indicateurs de dépendance de Kendall et de Spearman montrent clairement que les différentes LoBs ne sont pas indépendantes. Nous pouvons distinguer des axes de dépendance forts, observés quel que soit le critère de mesure :

- la RC Automobile et le Dommage Auto sont fortement liés (corrélation linéaire de .72, Tau de Kendall de .57 et Rho de Spearman de .79) ; ceci s'explique naturellement par le fait qu'un même accident touche souvent les deux garanties
- les deux LoBs relatives aux SCR santé, c'est-à-dire Accident&Health et Worker compensation sont également fortement liées (corrélation linéaire de .63, Tau de Kendall de .50 et Rho de Spearman de .72) ; en cas d'accident ou de souci de santé, une personne ayant une activité professionnelle est souvent amenée à être en arrêt-maladie.
- les branches liées à l'automobile sont significativement liées à celles de la santé
- enfin, il y a une dépendance notable entre Accident&Health et la RC Générale

V.4. Méthodologie pour la modélisation multivariée de la sinistralité

V.4.1. Calibrage des modèles de chocs communs

Pour chacun des modèles de choc communs, il faut réaliser le calibrage du modèle : il s'agit de trouver les paramètres du modèle qui permettent de modéliser au mieux les données observées.

Pour ce faire, plusieurs méthodes existent et feront l'objet des chapitres suivants.

V.4.1.1. Méthode du maximum de vraisemblance

La méthode la plus commune pour réaliser un calibrage est la méthode du maximum de vraisemblance.

En considérant que nous ayons un échantillon de T données $\{x_{1t}, x_{2t}, \dots, x_{nt}\}_{t=1}^T$ du vecteur aléatoire (X_1, X_2, \dots, X_n) , et que le modèle à calibrer ait k paramètres $\theta = (\theta_1, \theta_2, \dots, \theta_k)$ l'expression de la log-vraisemblance est :

$$l(\theta) = \sum_{t=1}^T \ln(f(x_{1t}, x_{2t}, \dots, x_{nt}))$$

Autrement dit, la log-vraisemblance est assimilable à la probabilité que le modèle ait θ comme jeu de paramètres, et ce, connaissant les T données de l'échantillon.

Le modèle sera calibré avec le jeu de paramètre $\hat{\theta}_{LM}$ qui optimise la log-vraisemblance ; autrement dit :

$$\hat{\theta}_{LM} = \underset{\theta}{\text{Argmax}}(l(\theta))$$

Sous les conditions usuelles de régularité, l'estimateur du maximum de vraisemblance existe, est consistant, et asymptotiquement efficace. Aussi, il vérifie la propriété de normalité asymptotique :

$$\sqrt{n}(\hat{\theta}_{ML} - \theta_0) \xrightarrow{loi} N(0; Fisher^{-1}(\theta_0))$$

avec θ_0 qui représente la vraie valeur des paramètres du modèle et $Fisher(\theta_0)$ représente la matrice d'information de Fisher.

Regardons maintenant la faisabilité pratique de cette méthode :

- pour le modèle multiplicatif log-normal, il y a 5 paramètres à estimer (un pour le choc commun et 4 pour les X_{10} et X_{20})
- pour le modèle multiplicatif log-normal shifté, il y a 6 paramètres à estimer (deux pour le choc commun et 4 pour les X_{10} et X_{20})

Aussi, il faudra réaliser un calcul de maximisation sur 5 ou 6 paramètres, ce qui engendre des temps de calculs inadéquats.

De plus, en appliquant ainsi la méthode du maximum de vraisemblance, nous ne maîtrisons nullement l'adéquation avec les marginales : en effet, nous calibrons ensemble le choc commun et les paramètres liées aux marginales.

V.4.1.2. Adaptation de la méthode du maximum de vraisemblance

L'idée reprend un peu le principe de la méthode IFM utilisée pour les copules : il s'agit de séparer l'estimation des paramètres des marginales de ceux de la copule.

Dans notre cas, nous allons procéder de la manière suivante :

- nous fixons les 4 paramètres des lois X_{10} et X_{20} , de sorte à assurer l'adéquation de l'espérance et de l'écart-type des marginales
- nous appliquons la méthode du maximum de vraisemblance sur le (ou les deux) paramètre(s) du choc commun.

Intéressons-nous à expliciter la fonction de densité $f(x_1, x_2)$ pour le cas d'un choc multiplicatif. Pour rappel ce modèle suppose que :

- la branche 1 a pour sinistralité : $X_1 = \beta \times X_{10}$
- la branche 2 a pour sinistralité : $X_2 = \beta \times X_{20}$

Pour ce faire, partons de la fonction de répartition que nous dériverons par la suite :

$$F_{X_1, X_2}(x_1, x_2) = \Pr(\beta \cdot X_{10} < x_1, \beta \cdot X_{20} < x_2)$$

$$= \int_{\beta=0}^{\infty} F_{X_{10}}\left(\frac{x_1}{\beta}\right) \cdot F_{X_{20}}\left(\frac{x_2}{\beta}\right) \cdot f_{\beta}(\beta) \cdot d\beta$$

Dérivons maintenant cette expression :

$$f_{X_1, X_2}(x_1, x_2) = \frac{\partial F_{X_1, X_2}(x_1, x_2)}{\partial x_1 \partial x_2}$$

$$= \int_{\beta=0}^{\infty} f_{X_{10}}\left(\frac{x_1}{\beta}\right) \cdot f_{X_{20}}\left(\frac{x_2}{\beta}\right) \cdot f_{\beta}(\beta) \cdot \frac{1}{\beta^2} d\beta$$

A partir des expressions obtenues ci-dessus, il s'agit ensuite d'un problème de maximisation à une ou deux dimensions. Ce qui, en pratique, est réalisable avec des temps de calculs raisonnables.

V.4.1.3. Méthode basée sur les indicateurs de dépendance

Cette méthode consiste à utiliser les mesures de dépendance. Elle est particulièrement adaptée dans le cas où le modèle présente une mesure de dépendance qui est directement exprimable en fonction des paramètres du modèle.

Par exemple, dans le cas d'un choc commun multiplicatif log-normal, nous obtenons :

$$\tau_{Kendall} = \frac{2}{\pi} \cdot \text{Arc sin}(\rho)$$

$$\rho_{Spearman} = \frac{6}{\pi} \cdot \text{Arc sin}\left(\frac{\rho}{2}\right)$$

De ce fait, nous pouvons calibrer le modèle en obtenons ρ à l'aide d'une des égalités suivantes :

$$\rho = \sin\left(\frac{\pi}{2} \cdot \tau_{Kendall}\right)$$

$$\rho = 2 \cdot \sin\left(\frac{\pi}{6} \cdot \rho_{Spearman}\right)$$

V.4.1.4. Choix de la méthode et mise en application

Parmi les trois méthodes étudiées, nous retiendrons l'adaptation de la méthode de la vraisemblance aux modèles de chocs communs.

En effet, elle garantit l'adéquation de l'espérance et de l'écart-type des marginales et offre une mise en pratique possible, avec des temps de calculs réalistes.

La première méthode (méthode générale du maximum de vraisemblance) ne remplit aucune de ces deux conditions et la troisième méthode n'est guère réalisable dès lors que nous ne sommes plus dans le cas du choc multiplicatif log-normal (pour lequel il existe des expressions simples liant les indicateurs de dépendance aux paramètres du modèle).

La mise en pratique de la méthode retenue ne pose aucun problème dans le cas de modèle avec un seul paramètre pour le choc commun (cas du choc multiplicatif log-normal simple). Dans le cas où il y a deux paramètres (cas du choc multiplicatif log-normal avec shift), nous maximiserons en balayant les deux dimensions suivantes du choc commun :

- la corrélation linéaire induite par le choc (nommée cor)
- le second paramètre : C pour le choc multiplicatif log-normal shifté

V.4.2. Adéquation et comparaison des différents modèles de chocs communs

Une fois le calibrage réalisé, nous obtenons un modèle calibré pour chaque famille de modèle étudiée. Il convient alors de vérifier l'adéquation de ces modèles avec les données observées, et notamment :

- de déterminer lequel de ces modèles correspond le mieux aux données observées
- de vérifier que ce modèle représente les données observées bien plus fidèlement qu'un modèle supposant l'indépendance des branches.

Afin de répondre à ces deux objectifs, nous allons présenter les principaux moyens de procéder.

V.4.2.1. Examen graphique

Une première possibilité est de comparer visuellement la densité de la copule empirique correspondant aux données observées à celle associée au modèle de choc commun calibré. La copule de Deheuvels (cf chapitre [III.4.2.2. Construction de la copule empirique de Deheuvels](#)) est utilisée pour les données observées.

Une simple lecture visuelle permet de vérifier l'adéquation quant aux aspects suivants :

- La symétrie ou asymétrie de la structure de dépendance est-elle bien respectée par le modèle : la fonction de densité présente-t-elle une symétrie suivant la diagonale (0;1)-(1;0) ?
- Le décalage entre le modèle et les données de l'historique portent-elles essentiellement sur les sinistres extrêmement forts, faibles ou sur l'ensemble du spectre de la sinistralité ?

C'est une approche simple et qui permet de rapidement identifier des modèles totalement inadaptés.

Egalement, si la copule empirique associée aux données de l'historique semble être plate (ou avec des variations minimales) sur le carré $[0;1]^2$, cela signifie que nous sommes très proches d'un cas d'indépendance et qu'il n'y a guère d'intérêt à recourir à des modèles de chocs communs.

V.4.2.2. Valeurs des log-vraisemblances maximales et critère d'Akaike

Une autre approche consiste à comparer les valeurs des log-vraisemblances, i.e. les valeurs de $l\left(\hat{\theta}_{LM}\right)$.

Toutefois, certains des modèles étudiés contiennent d'avantage de paramètres que d'autres. Ces modèles permettront mécaniquement de mieux coller aux données observées, mais sans pour autant garantir un gain en prédictibilité du même acabit.

Un cas extrême illustrant ce principe serait un modèle qui a autant de paramètres qu'il y a d'échantillons observés ; il serait alors possible de calibrer un modèle très complexe qui collerait parfaitement au passé, mais qui risque d'aboutir à des résultats de pauvre qualité quand il s'agira de prédire le futur. Il y aurait alors « sur-apprentissage » : la forte adéquation avec les données passées masque en fait une capacité de prédiction des données futures (ce qui nous intéresse in fine) affaiblie.

Aussi, nous utiliserons le critère d'Akaike qui permet de prendre en compte ce phénomène : en effet, il est défini ainsi :

$$AIC = 2K - 2 \cdot \log(V)$$

avec K le nombre de paramètres du modèle, et V la vraisemblance. Le modèle présentant le AIC le plus faible sera retenu. Avec ce critère nous voyons bien que plus le nombre de paramètres est élevé, plus la log-vraisemblance devra être forte pour obtenir un même niveau de AIC.

V.4.2.3. Distances entre les classes

Il s'agit de la méthode étudiée dans le chapitre [III.6.2.2. Distance entre les classes](#). Etant donné le volume des données, nous découperons le carré $[0;1]^2$ en les 25 ($5*5$) classes suivantes :

$$\left(\left[\frac{i-1}{k}; \frac{i}{k} \right] \times \left[\frac{j-1}{k}; \frac{j}{k} \right] \right)_{1 \leq i, j \leq 5}$$

V.4.2.4. Probabilité simultanée d'excès de quantiles

En nous rappelant que le but final de la modélisation de la dépendance est l'estimation de la VaR à 99.5%, il est intéressant de chercher à comparer les probabilités simultanées d'excès de quantiles entre l'historique et le modèle.

Pour ce faire nous utiliserons les indicateurs de dépendance de queue définis dans le chapitre [II.4.3. Les indicateurs de dépendance de queue](#) : La upper tail concentration function R et la La lower tail concentration function L définies par :

$$R(z) = P(X > F_X^{-1}(z) / Y > F_Y^{-1}(z)) \quad L(z) = P(X < F_X^{-1}(z) / Y < F_Y^{-1}(z))$$

Nous disposons d'un historique de 84 points, ce qui ne donne même pas 1 observation correspondant à un quantile de 99.5%. Aussi, cette analyse ne peut être réalisée que sur la upper tail concentration à 80%.

V.4.2.5. Choix de la méthode

Chacune des quatre méthodes exposées précédemment possède son avantage. Généralement, nous utiliserons l'examen graphique pour rejeter rapidement les modèles inadaptés et comprendre visuellement la qualité de l'adéquation.

Puis, pour réaliser une comparaison basée sur des éléments quantitatifs, le critère d’Akaike sera utilisé. En effet, il s’appuie sur une approche standard en incluant le maximum de vraisemblance et permet de choisir le modèle ayant le meilleur compromis complexité/pouvoir de prédiction.

V.5. Application dans le cas de 2 branches

Trois applications seront étudiées :

- la dépendance entre la RC Auto et le Dommage Auto
- la dépendance entre Accident&Health et Worker Compensation
- la dépendance entre la somme des deux branches automobiles et « Fire and other Property Damage »

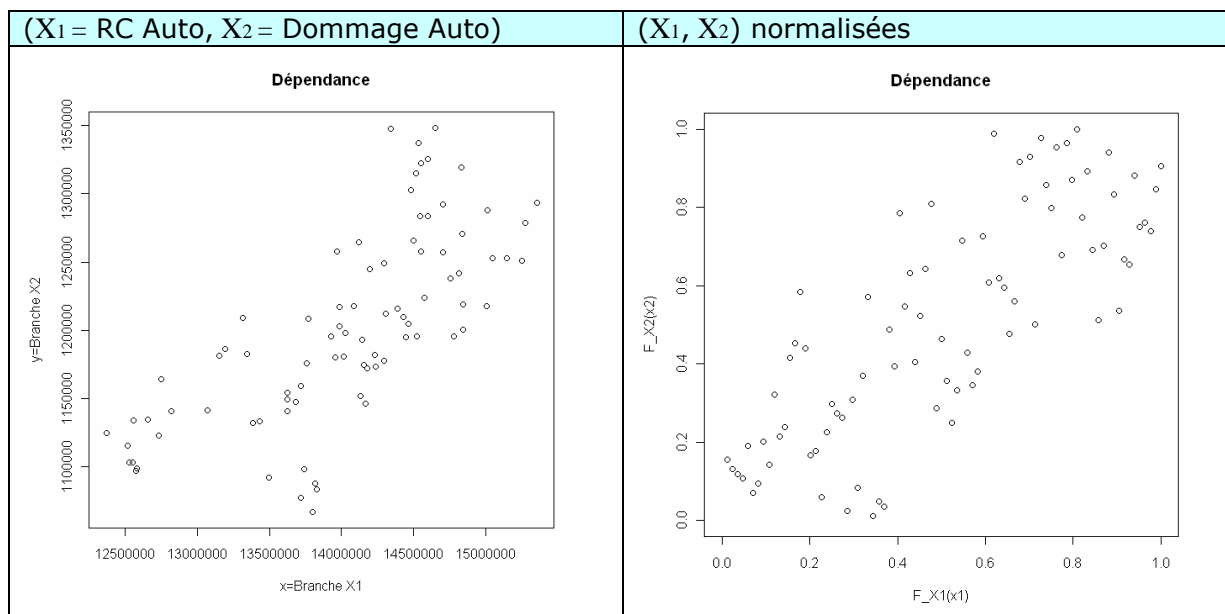
Pour chacune des applications, nous testerons les 3 modèles suivants :

Nom du modèle	Modèle	Lois utilisées dans le modèle
modèle à choc commun multiplicatif log-normal	$\begin{cases} X_1 = X_{10} \times \beta \\ X_2 = X_{20} \times \beta \end{cases}$	X_1, X_2 sont des lognormales β est une lognormale
modèle à choc commun multiplicatif log-normal shifté	$\begin{cases} X_1 = X_{10} \times \beta \\ X_2 = X_{20} \times \beta \end{cases}$	X_1, X_2 sont des lognormales β est une lognormale shiftée
modèle à choc commun diviseur log-normal shifté	$\begin{cases} X_1 = X_{10} / \beta \\ X_2 = X_{20} / \beta \end{cases}$	X_1, X_2 sont des lognormales β est une lognormale shiftée

Egalement, nous incluons dans notre analyse le modèle d’indépendance.

V.5.1. Modélisation de la dépendance entre la RC Auto et le Dommage Auto

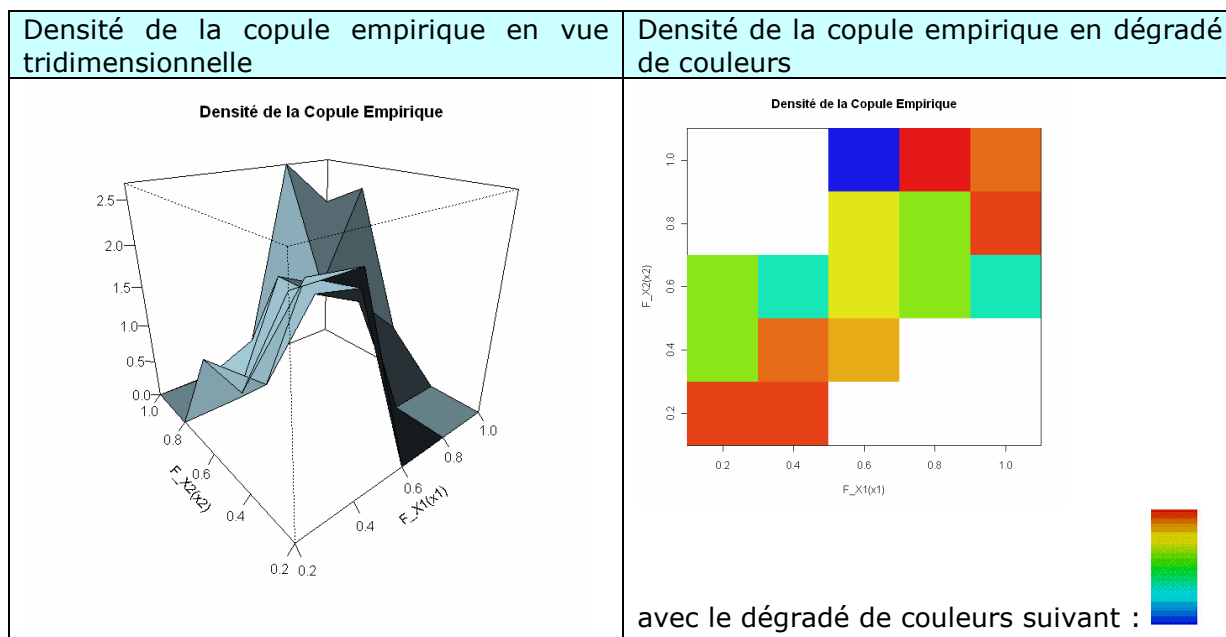
Regardons dans un premier temps le nuage des points correspondant aux 84 observations pour les deux branches :



P.S. : Le lecteur est invité à se reporter à la partie [III.4.1. Visualisation suite à la « normalisation »](#) des données pour la normalisation de (X_1, X_2)

Nous voyons sur ces deux graphes une forte dépendance entre les branches RC Auto et Dommage Auto : les points ont tendance à être alignés sur la première bissectrice du carré $[0;1]^2$. De plus, les concentrations de points semblent être du même niveau dans le coin supérieur droite et dans le coin inférieur gauche.

Puis, voici la forme des densités des copules empiriques obtenues pour les deux branches (se reporter à : [III.4.2. Visualisation de la structure de dépendance à l'aide de copule empirique](#) pour de plus amples informations) :



La densité de la copule empirique tend à confirmer nos observations précédentes. D'après cette analyse graphique, nous pourrions avoir l'intuition que le modèle le plus adapté serait un modèle à choc commun multiplicatif log-normal simple (c'est-à-dire non shifté).

Nous pouvons par ailleurs remarquer que nous sommes loin des densités de copule empiriques obtenues dans le cas de l'analyse théorique des modèles de choc commun ; en effet, la quantité de points est bien moindre.

Pour rappel, les propriétés des deux branches sont les suivantes :

	Moyenne	Ecart-type	CV
1_Motor, third-party liability	14 024 016	751 476	0,05
2_Motor, other classes	1 200 494	69 954	0,06

Avec un coefficient de corrélation linéaire observé de 0.72, nous sommes bien en-dessous de la limite relative au modèle à choc commun multiplicatif log-normal (cf [IV.1.9. Conclusion sur le choc multiplicatif log-normal](#)) :

$$0.72 = \rho = Cor[X_1, X_2] \leq \text{Min} \left(\frac{CV_1}{CV_2}; \frac{CV_2}{CV_1} \right) = 0.92$$

En calibrant les 3 modèles suivant la méthode adaptée de la maximisation de log-vraisemblance (cf [V.3.1. Calibrage des modèle de chocs communs](#)), nous obtenons les résultats suivants pour les paramètres des trois modèles :

Modèle	X10		X20		Choc		
	Moyenne	Ecart-type	Moyenne	Ecart-type	Moyenne	Ecart-type	Shift
Multiplicatif log-normal	14 024 016	303 779	1 200 494	37 814	1	0,0490	aucun
Multiplicatif log-normal shifté	14 024 016	303 779	1 200 494	37 814	1	0,0490	0,00
Diviseur log-normal shifté	14 024 016	336 463	1 200 494	39 791	1	0,0479	0,86

Regardons maintenant les log-vraisemblances et le critère d'Akaike pour choisir le modèle qui présente la meilleure adéquation avec nos données ([cf V.3.2. Adéquation et comparaison des différents modèles de chocs communs](#)) :

Modèle	Choix du modèle	
	-LogV	AIC/2
Indépendance, sans choc commun	2 311,97	2 311,97
Choc commun multiplicatif log-normal	2 278,34	2 279,34
Choc commun multiplicatif log-normal shifté	2 278,34	2 280,34
Choc commun diviseur log-normal shifté	2 278,92	2 280,92

Déjà, nous remarquons que la log-vraisemblance est significativement meilleure avec un modèle à choc commun que sans (c'est-à-dire en supposant les deux branches indépendantes).

Par ailleurs, la log-vraisemblance présente des niveaux similaires pour les trois modèles de choc commun. Aussi, le critère d'Akaike va favoriser le modèle le plus simple, le modèle avec le plus petit nombre de paramètres, i.e. le choc commun multiplicatif log-normal simple, sans shift.

De ce fait, les critères qualitatifs utilisés amènent à la même conclusion que l'analyse préliminaire visuelle.

Enfin, intéressons-nous à la sinistralité globale aux deux branches :

Modèle	Sinistralité totale = somme des branches		
	Moyenne	Ecart-type	SCR
Indépendance, sans choc commun	15 224 511	754 755	2 063 578
Choc commun multiplicatif log-normal	15 224 511	806 983	2 200 852
Choc commun multiplicatif log-normal shifté	15 224 511	806 983	2 200 852
Choc commun diviseur log-normal shifté	15 224 511	779 430	1 755 619
<i>Données historiques constatées sur la somme des deux branches</i>	15 224 511	803 515	2 191 114

Le SCR correspond à la Var à 99.5% à laquelle on a retranché la moyenne. (les valeurs affichées ici étant des valeurs mensuelles obtenues par moyenne mobile sur 12 mois, il conviendrait de les multiplier par 12 afin d'obtenir leurs équivalents annuels)

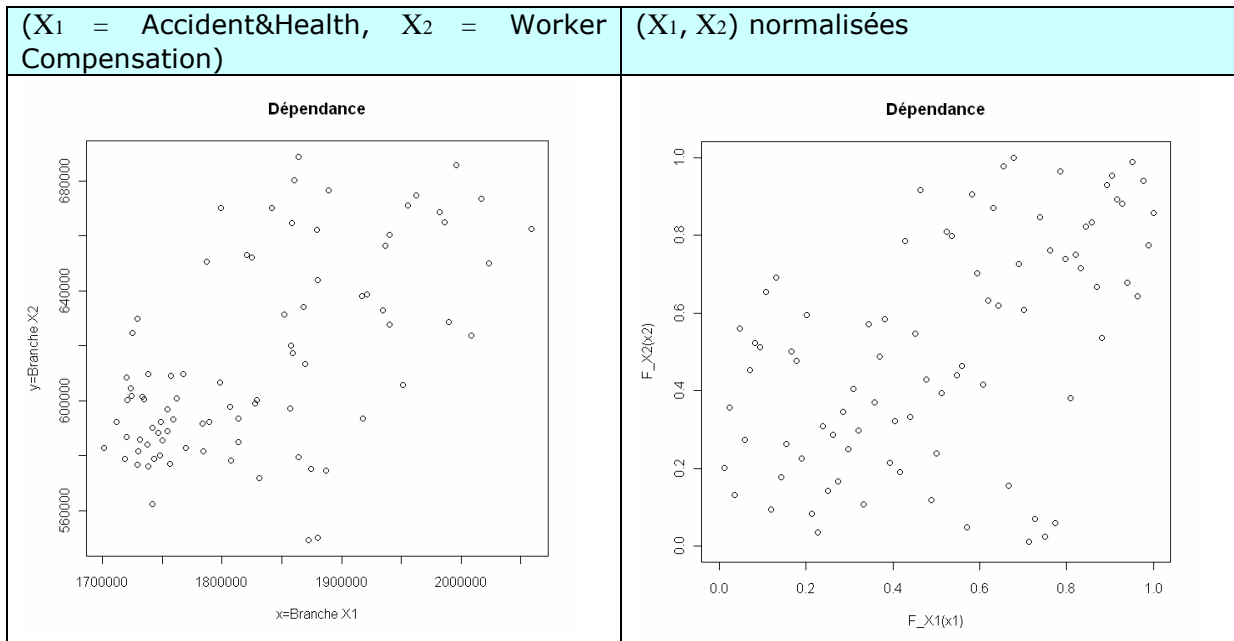
Fort logiquement, avec le modèle à choc commun multiplicatif lognormal, nous obtenons un écart-type et un SCR supérieurs à ceux qui auraient été obtenus dans le cas d'indépendance.

La dernière ligne du tableau indique la moyenne et l'écart-type observés sur l'historique de la somme des deux branches. La VaR à 99.5% est calculée avec l'hypothèse que la somme suit une loi log-normale.

Il est intéressant de remarquer que le modèle retenu présente un écart-type et une VaR de montants similaires à ceux de la dernière ligne.

V.5.2. Modélisation de la dépendance entre Accident&Health et Worker Compensation

Regardons dans un premier temps le nuage des points correspondant aux 84 observations pour les deux branches :

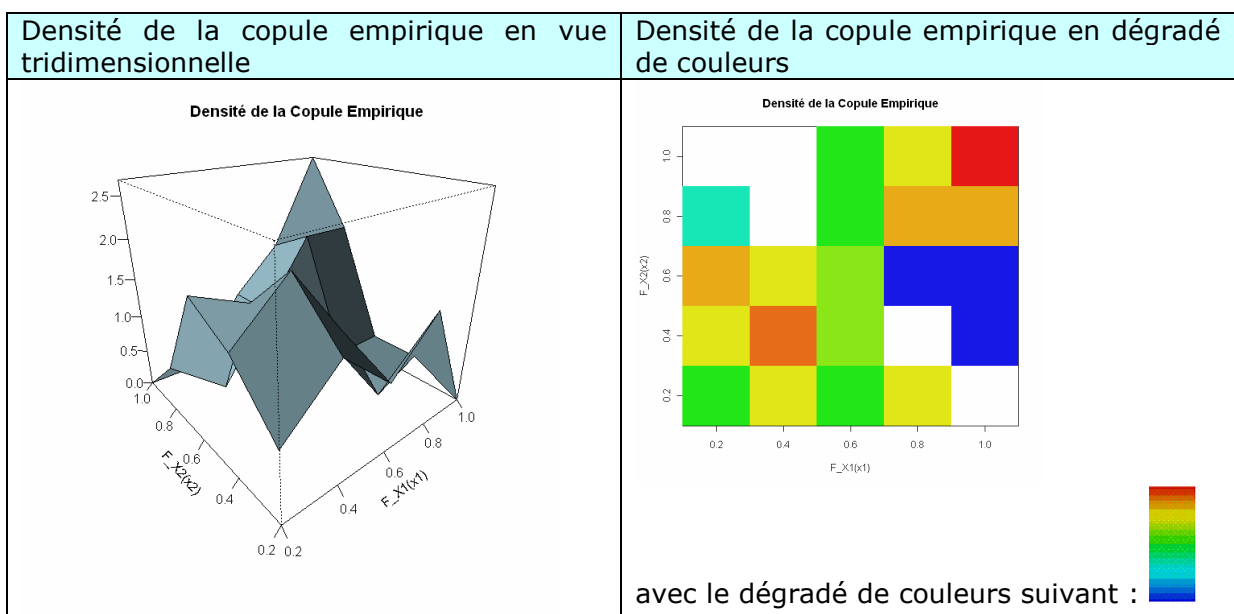


P.S. : Le lecteur est invité à se reporter à la partie [III.4.1. Visualisation suite à la « normalisation »](#) des données pour la normalisation de (X₁, X₂)

Nous voyons sur ces deux graphes une dépendance significative entre les branches « Accident&Health » et « Worker Compensation » : les points ont tendance à être alignés sur la première bissectrice du carré [0;1]². De plus, si nous regardons les sinistralités extrêmes, la concentration dans le cadran [.8;1]² semble plus forte que dans le cadran [0;.2]² :

- dans le cadran [.8;1]², il y a 9 points
- dans le cadran [0;.2]², il n’y a que 3 points

Puis, voici la forme des densités des copules empiriques obtenues pour les deux branches (se reporter à : [III.4.2. Visualisation de la structure de dépendance à l’aide de copule empirique](#) pour de plus amples informations) :



La densité de la copule empirique tend à confirmer nos observations précédentes. D'après cette analyse graphique, nous pourrions avoir l'intuition que le modèle le plus adapté serait un modèle à choc commun multiplicatif log-normal shifté. Pour rappel, les propriétés des deux branches sont les suivantes :

	Moyenne	Ecart-type	CV
21 Accident&Health	1 827 834	91 867	0,05
22_Health-Worker Compensation	614 004	35 393	0,06

Avec un coefficient de corrélation linéaire observé de 0.60, nous sommes bien en-dessous de la limite relative au modèle à choc commun multiplicatif log-normal (cf [IV.1.9. Conclusion sur le choc multiplicatif log-normal](#)) :

$$0.60 = \rho = Cor[X_1, X_2] \leq \text{Min}\left(\frac{CV_1}{CV_2}; \frac{CV_2}{CV_1}\right) = 0.87$$

En calibrant les 3 modèles suivant la méthode adaptée de la maximisation de log-vraisemblance (cf [V.3.1. Calibrage des modèle de chocs communs](#)), nous obtenons les résultats suivants pour les paramètres des trois modèles :

Modèle	X10		X20		Choc		
	Moyenne	Ecart-type	Moyenne	Ecart-type	Moyenne	Ecart-type	Shift
Multiplicatif log-normal	1 827 834	53 091	614 004	24 858	1	0,0410	aucun
Multiplicatif log-normal shifté	1 827 834	47 518	614 004	23 549	1	0,0430	0,92
Diviseur log-normal shifté	1 827 834	50 415	614 004	24 220	1	0,0420	0,02

Regardons maintenant les log-vraisemblances et le critère d'Akaike pour choisir le modèle qui présente la meilleure adéquation avec nos données (cf [V.3.2. Adéquation et comparaison des différents modèles de chocs communs](#)) :

Modèle	Choix du modèle	
	-LogV	AIC/2
Indépendance, sans Choc commun	2 074,93	2 074,93
Choc commun multiplicatif log-normal	2 056,28	2 057,28
Choc commun multiplicatif log-normal shifté	2 053,05	2 055,05
Choc commun diviseur log-normal shifté	2 056,73	2 058,73

Déjà, nous remarquons que la log-vraisemblance est significativement meilleure avec un modèle à choc commun que sans (c'est-à-dire en supposant les deux branches indépendantes).

Par ailleurs, la log-vraisemblance présente un avantage notable pour le choc commun multiplicatif log-normal shifté. Selon le critère d'Akaike, ce gain en qualité d'adéquation est suffisant pour compenser la perte liée à une complexité accrue par l'ajout d'un paramètre supplémentaire : la valeur du shift. Aussi, selon le critère d'Akaike, le modèle à retenir est celui à choc commun multiplicatif log-normal shifté.

De ce fait, les critères qualitatifs utilisés rejoignent la conclusion issue de l'analyse préliminaire visuelle.

Enfin, intéressons-nous à la sinistralité globale aux deux branches :

Modèle	Sinistralité totale = somme des branches		
	Moyenne	Ecart-type	SCR
Indépendance, sans Choc commun	2 441 837	98 651	266 461
Choc commun multiplicatif log-normal	2 441 837	116 069	314 443
Choc commun multiplicatif log-normal shifté	2 441 837	117 595	452 391
Choc commun diviseur log-normal shifté	2 441 837	117 077	322 117
<i>Données historiques constatées sur la somme des deux branches</i>	2 441 837	116 485	315 929

Le SCR correspond à la Var à 99.5% à laquelle on a retranché la moyenne. (les valeurs affichées ici étant des valeurs mensuelles obtenues par moyenne mobile sur 12 mois, il conviendrait de les multiplier par 12 afin d'obtenir leurs équivalents annuels)

Fort logiquement, avec le modèle à choc commun multiplicatif lognormal shifté, nous obtenons un écart-type et une VaR à 99.5% supérieurs à ceux qui auraient été obtenus dans le cas d'indépendance.

Le SCR obtenu pour le modèle retenu est de l'ordre de 50% supérieur à celui obtenu par les valeurs historiques constatées sur la somme des deux branches. En revenant au monde de l'assurance, cela correspondrait à un besoin en capital extrêmement élevé.

Avec un shift de 0.92, ce modèle tend à avoir un upper tail concentration très élevé (cf [IV.2.5. Analyse quantitative de la structure de dépendance obtenue](#) pour les chocs multiplicatifs log-normaux shiftés) : quand il y a une sinistralité extrêmement forte sur la première branche, il y a de très fortes chances pour qu'il en soit de même pour la deuxième branche. La VaR qui correspond à la somme de ces deux branches est donc très fortement impactée.

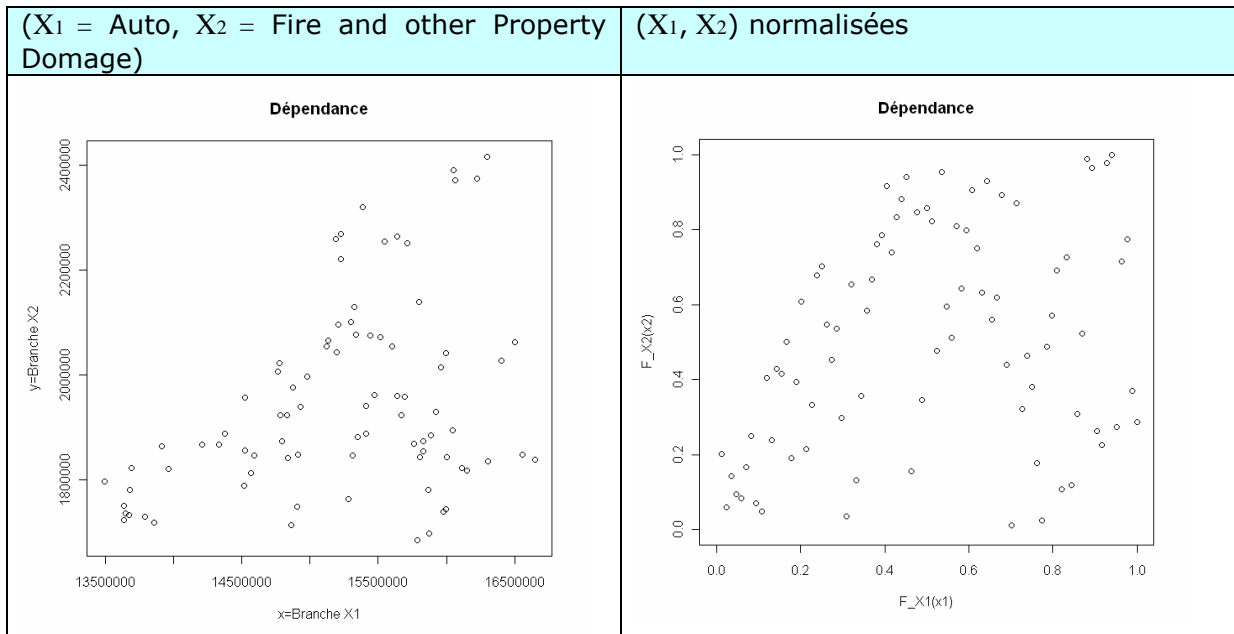
Ce qui nous amène à la limite du modèle face au nombre très restreint de données. En effet, nous n'avons pas assez de données pour observer un nombre de fois raisonnable l'apparition simultanée dans les deux branches de quantiles supérieurs à 90% (et encore moins à 99,5% !). Même pour des quantiles à 80%/20% :

- il n'y a que 9 observations correspondant à des quantiles supérieurs à 80% pour les 2 branches
- il n'y a que 3 observations correspondant à des quantiles inférieurs à 20% pour les 2 branches

De plus, un calibrage basé sur un maximum de vraisemblance cherche à coller au mieux pour l'ensemble des points sans se soucier particulièrement des quantiles extrêmes simultanés.

V.5.3. Modélisation de la dépendance entre la somme des deux branches automobiles et « Fire and other Property Damage »

Regardons dans un premier temps le nuage des points correspondant aux 84 observations pour les deux branches :

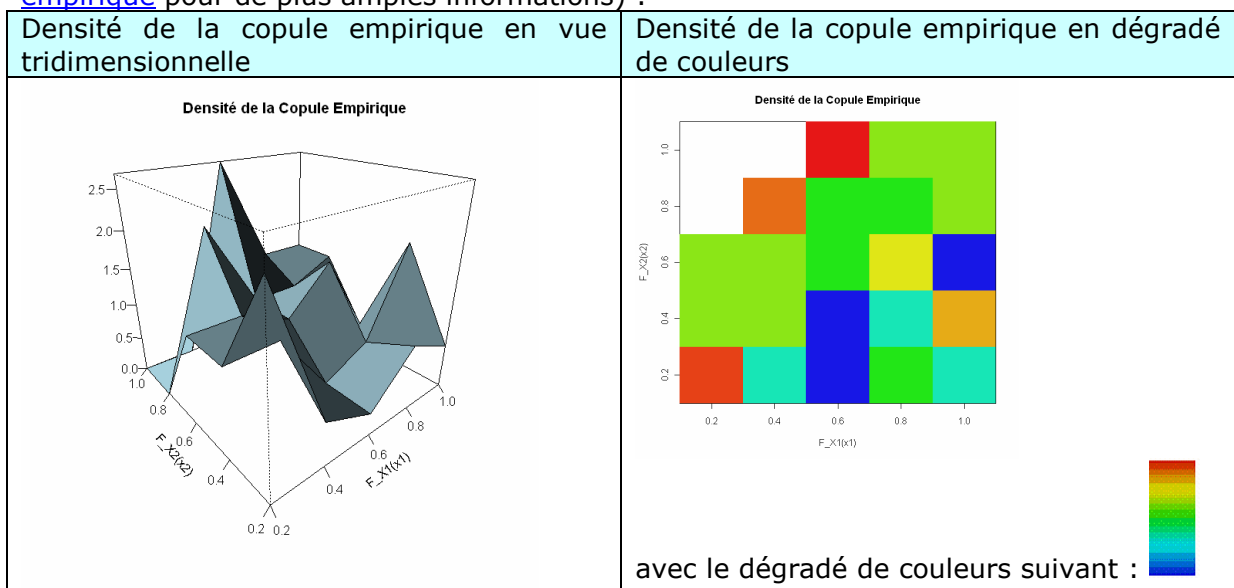


P.S. : Le lecteur est invité à se reporter à la partie [III.4.1. Visualisation suite à la « normalisation »](#) des données pour la normalisation de (X₁, X₂)

Nous voyons sur ces deux graphes une dépendance significative entre les branches Accident&Health et Worker Compensation : les points ont tendance à être alignés sur la première bissectrice du carré [0;1]². De plus, si nous regardons les sinistralités extrêmes, la concentration dans le cadran [.8;1]² semble moins forte que dans le cadran [0;.2]² :

- dans le cadran [.8;1]², il y a 4 points
- dans le cadran [0;.2]², il y a 8 points

Puis, voici la forme des densités des copules empiriques obtenues pour les deux branches (se reporter à : [III.4.2. Visualisation de la structure de dépendance à l'aide de copule empirique](#) pour de plus amples informations) :



La densité de la copule empirique tend à confirmer nos observations précédentes. D'après cette analyse graphique, nous pourrions avoir l'intuition que le modèle le plus adapté serait un modèle à choc commun diviseur log-normal shifté. Pour rappel, les propriétés des deux branches sont les suivantes :

	Moyenne	Ecart-type	CV
4. Fire and other property damage	1 946 658	180 915	0,09
Automobile (1+2)	15 224 511	803 515	0,05

Avec un coefficient de corrélation linéaire observé de 0.37, nous sommes bien en-dessous de la limite relative au modèle à choc commun multiplicatif log-normal (cf [IV.1.9. Conclusion sur le choc multiplicatif log-normal](#)) :

$$0.37 = \rho = \text{Cor}[X_1, X_2] \leq \text{Min}\left(\frac{CV_1}{CV_2}; \frac{CV_2}{CV_1}\right) = 0.57$$

En calibrant les 3 modèles suivant la méthode adaptée de la maximisation de log-vraisemblance (cf [V.3.1. Calibrage des modèle de chocs communs](#)), nous obtenons les résultats suivants pour les paramètres des trois modèles :

Modèle	X10		X20		Choc		
	Moyenne	Ecart-type	Moyenne	Ecart-type	Moyenne	Ecart-type	Shift
Multiplicatif log-normal	15 224 511	478 049	1 946 658	160 845	1	0,0424	aucun
Multiplicatif log-normal shifté	15 224 511	330 337	1 946 658	154 620	1	0,0481	0,03
Diviseur log-normal shifté	15 224 511	175 503	1 946 658	150 398	1	0,0515	0,88

Nous pouvons remarquer que le modèle diviseur log-normal shifté donne un shift de .88 qui est inférieur au shift de .92 obtenu dans l'application précédente. Ce qui est en phase avec le fait que, dans la première application, il y avait 3 fois plus de points dans le cadran $[\cdot 8; 1]^2$ que dans le cadran $[0; \cdot 2]^2$, alors que dans la deuxième application, il n'y a que 2 fois plus de points dans le cadran $[0; \cdot 2]^2$ que dans le cadran $[\cdot 8; 1]^2$.

Regardons maintenant les log-vraisemblances et le critère d'Akaike pour choisir le modèle qui présente la meilleure adéquation avec nos données (cf [V.3.2. Adéquation et comparaison des différents modèles de chocs communs](#)) :

Modèle	Choix du modèle	
	-LogV	AIC/2
Indépendance, sans Choc commun	2 394,71	2 394,71
Choc commun multiplicatif log-normal	2 387,76	2 388,76
Choc commun multiplicatif log-normal shifté	2 387,53	2 389,53
Choc commun diviseur log-normal shifté	2 383,21	2 385,21

Déjà, nous remarquons que la log-vraisemblance est significativement meilleure avec un modèle à choc commun que sans (c'est-à-dire en supposant les deux branches indépendantes).

Par ailleurs, la log-vraisemblance présente un avantage notable pour le choc commun diviseur log-normal shifté. Selon le critère d'Akaike, ce gain en qualité d'adéquation est suffisant pour compenser la perte liée à une complexité accrue par l'ajout d'un paramètre supplémentaire : la valeur du shift. Aussi, selon le critère d'Akaike, le modèle à retenir est celui à choc commun diviseur log-normal shifté.

De ce fait, les critères qualitatifs utilisés rejoignent la conclusion issue de l'analyse préliminaire visuelle.

Enfin, intéressons-nous à la sinistralité globale aux deux branches :

Modèle	Sinistralité totale = somme des branches		
	Moyenne	Ecart-type	SCR
Indépendance, sans Choc commun	17 171 169	824 195	2 238 710
Choc commun multiplicatif log-normal	17 171 169	886 389	2 416 164
Choc commun multiplicatif log-normal shifté	17 171 169	903 036	2 465 376
Choc commun diviseur log-normal shifté	17 171 169	868 587	1 718 538
<i>Données historiques constatées sur la somme des deux branches</i>	17 171 169	887 624	2 417 675

Le SCR correspond à la Var à 99.5% à laquelle on a retranché la moyenne. (les valeurs affichées ici étant des valeurs mensuelles obtenues par moyenne mobile sur 12 mois, il conviendrait de les multiplier par 12 afin d'obtenir leurs équivalents annuels)

Le modèle à choc commun diviseur lognormal shifté, nous obtenons logiquement un écart-type supérieur à celui du cas d'indépendance. En revanche, nous obtenons une VaR inférieure.

La valeur du choc $1/\beta$ est capée à $1/0.88=1.14$. De ce fait, la volatilité du choc aura tendance à se concentrer sur les valeurs faibles au détriment des valeurs extrêmement fortes sur lesquelles se basent la VaR à 99.5%.

Les variables X_{10} et X_{20} peuvent toujours donner lieu à des valeurs extrêmement élevées mais avec une volatilité diminuée : en effet, une partie importante de leur variance est transférée au choc commun.

D'où une VaR et un SCR bien moindres pour le choc commun log-normal shifté diviseur.

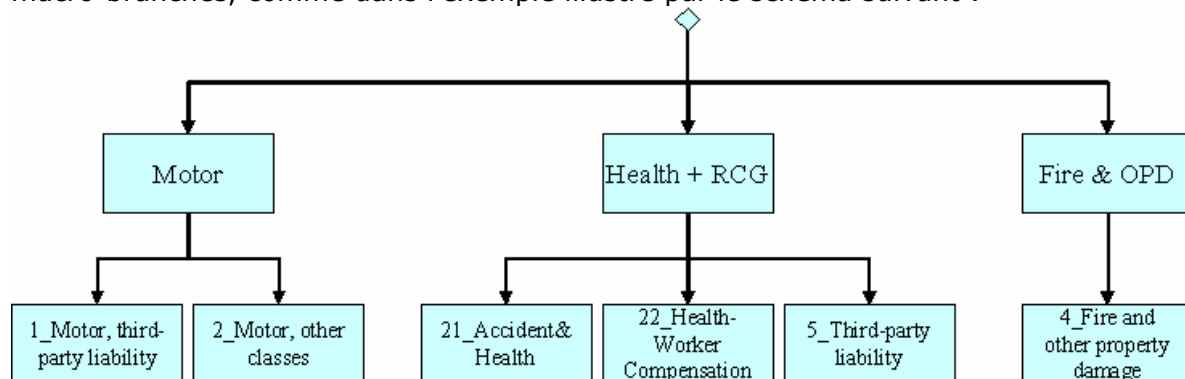
V.6. Application à l'ensemble du portefeuille pour calculer le SCR

V.6.1. Approche utilisée

Au vu des volumes des différentes branches (cf [V.1.3. Statistiques exploratoires](#)), nous nous concentrerons sur les 6 branches les plus significatives, à savoir :

- 1_Motor, third-party liability
- 2_Motor, other classes
- 4_Fire and other property damage
- 5_Third-party liability
- 21_Accident&Health
- 22_Health-Worker Compensation

Nous utiliserons alors une structure hiérarchique : nous regrouperons les 6 branches en 3 macro-branches, comme dans l'exemple illustré par le schéma suivant :



Ainsi, la macro-branche « Motor » sera obtenue par addition des LoB 1 et 2.

Puis, nous réaliserons un calibrage de modèle de choc commun sur les macro-branches.

Si l'objectif avait été de modéliser intégralement toutes les branches, il aurait également fallu réaliser des modèles de choc commun sur les LoB à l'intérieur de chaque macro-branche ; mais ce n'est pas le cas.

Dans la suite, nous étudierons trois différents modèles hiérarchiques, en commençant par celui illustré dans le schéma ci-dessus.

V.6.2. Résultats

V.6.2.1. 1^{er} modèle hiérarchique

Le premier modèle hiérarchique est composé des macro-branches suivantes :

Macro-Branche	Branches composant la macro-branche
Motor	1_Motor, third-party liability 2_Motor, other classes
Health + RCG	21_Accident&Health 22_Health-Worker Compensation 5_Third-party liability
Fire&OPD	4_Fire and other property damage

Les caractéristiques des macro-branches sont les suivantes :

	Moyenne	Ecart-type	CV
Motor	15 224 511	803 515	0,053
Health + RCG	4 829 833	409 922	0,085
Fire&OPD	1 946 658	180 915	0,093

Les corrélations de Pearson des macro-branches sont les suivantes :

	All Motors	Health + RCG	Fire&OPD
All Motors	1,00	0,31	0,38
Health + RCG	0,31	1,00	0,29
Fire&OPD	0,38	0,29	1,00

Et voici pour le Tau de Kendall :

	All Motors	Health + RCG	Fire&OPD
All Motors	1,00	0,22	0,23
Health + RCG	0,22	1,00	-0,06
Fire&OPD	0,23	-0,06	1,00

Et voici pour le Rho de Spearman :

	All Motors	Health + RCG	Fire&OPD
All Motors	1,00	0,36	0,31
Health + RCG	0,36	1,00	-0,06
Fire&OPD	0,31	-0,06	1,00

Voici les résultats obtenus :

Modèle	Choix du modèle	Sinistralité totale = somme des branches		
	AIC/2	Moyenne	Ecart-type	SCR
Indépendance, sans Choc commun	3 595,77	22 001 002	919 592	2 481 760
Choc commun multiplicatif log-normal	3 585,83	22 001 002	1 113 606	3 031 872
Choc commun multiplicatif log-normal shifté	3 586,81	22 001 002	1 108 307	3 015 754
Choc commun diviseur log-normal shifté	3 586,55	22 001 002	1 100 702	2 726 683
<i>Données historiques constatées sur la somme des branches</i>		22 001 002	1 096 704	2 981 288

Le SCR correspond à la Var à 99.5% à laquelle on a retranché la moyenne. (les valeurs affichées ici étant des valeurs mensuelles obtenues par moyenne mobile sur 12 mois, il conviendrait de les multiplier par 12 afin d'obtenir leurs équivalents annuels)

Ce premier modèle hiérarchique retient un modèle à choc commun multiplicatif log-normal.

Le SCR (en tant que VaR à 99,5% moins la moyenne) est de l'ordre de 3 millions d'euros, ce qui représente à peu près 13.8% de la charge ultime de sinistres.

V.6.2.2. 2^{ème} modèle hiérarchique

Le deuxième modèle hiérarchique est composé des macro-branches suivantes :

Macro-Branche	Branches composant la macro-branche
Motor	1_Motor, third-party liability 2_Motor, other classes
Health	21_Accident&Health 22_Health-Worker Compensation
RCG+Fire&OPD	4_Fire and other property damage 5_Third-party liability

Les caractéristiques des macro-branches sont les suivantes :

	Moyenne	Ecart-type	CV
Motor	15 224 511	803 515	0,053
Health	2 441 837	116 485	0,048
RCG+Fire&OPD	4 334 654	409 071	0,094

Les corrélations de Pearson des macro-branches sont les suivantes :

	All Motors	Health	RCG+Fire&OPD
All Motors	1,00	0,49	0,34
Health	0,49	1,00	0,66
RCG+Fire&OPD	0,34	0,66	1,00

Et voici pour le Tau de Kendall :

	All Motors	Health	RCG+Fire&OPD
All Motors	1,00	0,29	0,24
Health	0,29	1,00	0,26
RCG+Fire&OPD	0,24	0,26	1,00

Et voici pour le Rho de Spearman :

	All Motors	Health	RCG+Fire&OPD
All Motors	1,00	0,52	0,34
Health	0,52	1,00	0,40
RCG+Fire&OPD	0,34	0,40	1,00

Voici les résultats obtenus :

Modèle	Choix du modèle	Sinistralité totale = somme des branches		
	AIC/2	Moyenne	Ecart-type	SCR
Indépendance, sans Choc commun	3 557,73	22 001 002	909 238	2 457 139
Choc commun multiplicatif log-normal	3 536,33	22 001 002	1 045 748	2 843 251
Choc commun multiplicatif log-normal shifté	3 534,40	22 001 002	1 074 901	3 600 315
Choc commun diviseur log-normal shifté	3 536,26	22 001 002	1 076 181	2 955 262
<i>Données historiques constatées sur la somme des branches</i>		22 001 002	1 096 704	2 981 288

Le SCR correspond à la Var à 99.5% à laquelle on a retranché la moyenne. (les valeurs affichées ici étant des valeurs mensuelles obtenues par moyenne mobile sur 12 mois, il conviendrait de les multiplier par 12 afin d'obtenir leurs équivalents annuels)

Cette fois-ci, c'est le modèle à choc commun multiplicatif log-normal shifté qui est retenu. Il s'en suit une VaR à 99.5% nettement supérieure à celle obtenue dans le cas précédent : d'où un SCR à 3,6M€. Ce qui ramène le SCR à 16.4% de la moyenne de la sinistralité totale (charge ultime).

La valeur du shift du choc commun β vaut 0.91.

Or, le shift de β joue un rôle très sensible sur la VaR puisque celui-ci conditionne la probabilité d'apparition simultanée de sinistralité extrêmement forte. Toutefois, pour que son calibrage ait du sens, par définition, il faudrait assez de données pour observer plusieurs cas de sinistralité extrêmement forte. Or avec une petite centaine de points, nous sommes loin de ce cas idéal.

Par ailleurs, ce deuxième modèle souffre du fait que, concernant la corrélation entre la deuxième et la troisième macro-branche, nous ne respectons pas la limite relative au modèle à choc commun multiplicatif log-normal (cf [IV.1.9. Conclusion sur le choc multiplicatif log-normal](#)) pour le couple des deuxième et troisième macro-branches :

$$.66 = \rho = Cor[X_2, X_3] \text{ n'est pas inférieur à : } \min\left(\frac{CV_3}{CV_2}; \frac{CV_2}{CV_3}\right) = .51$$

De ce fait, la corrélation linéaire sur les deux dernières macro-branches induite par le modèle est nettement inférieure à celle mesurée sur les données historiques :

Corrélation	(X1,X2)	(X1,X3)	(X2,X3)
Constaté sur les données historiques	0,34	0,49	0,66
Induit par le modèle	0,29	0,57	0,32

Enfin, nous pouvons voir que le SCR qui correspond au modèle à choc commun multiplicatif non-shifté est de 2.84M€, ce qui est en retrait par rapport aux 3M€ du premier modèle. De même, les écart-types des modèles à choc sont légèrement plus faibles que ceux du premier modèle hiérarchique. Ces écarts sont dus à la limitation évoquée ci-dessus sur les corrélations : en effet, cette dernière bride la corrélation entre X2 et X3, et donc, le SCR.

V.6.2.3. 3^{ème} modèle hiérarchique

Le troisième modèle hiérarchique est composé des macro-branches suivantes :

Macro-Branche	Branches composant la macro-branche
Motor + WC	1_Motor, third-party liability 2_Motor, other classes 22_Health-Worker Compensation
Accident + RCG	21_Accident&Health 5_Third-party liability
Fire&OPD	4_Fire and other property damage

Les caractéristiques des macro-branches sont les suivantes :

	Moyenne	Ecart-type	CV
Motor + WC	15 838 514	826 705	0,052
Accident + RCG	4 215 830	396 092	0,094
Fire&OPD	1 946 658	180 915	0,093

Les corrélations de Pearson des macro-branches sont les suivantes :

	Motor + WC	Accident+ RCG	Fire&OPD
Motor + WC	1,00	0,27	0,38
Accident+ RCG	0,27	1,00	0,28
Fire&OPD	0,38	0,28	1,00

Et voici pour le Tau de Kendall :

	Motor + WC	Accident+ RCG	Fire&OPD
Motor + WC	1,00	0,17	0,23
Accident+ RCG	0,17	1,00	-0,07
Fire&OPD	0,23	-0,07	1,00

Et voici pour le Rho de Spearman :

	Motor + WC	Accident+ RCG	Fire&OPD
Motor + WC	1,00	0,27	0,31
Accident+ RCG	0,27	1,00	-0,07
Fire&OPD	0,31	-0,07	1,00

Voici les résultats obtenus :

Modèle	Choix du modèle	Sinistralité totale = somme des branches		
	AIC/2	Moyenne	Ecart-type	SCR
Indépendance, sans Choc commun	3 595,09	22 001 002	934 228	2 521 272
Choc commun multiplicatif log-normal	3 586,29	22 001 002	1 115 174	3 035 901
Choc commun multiplicatif log-normal shifté	3 587,29	22 001 002	1 108 872	3 017 261
Choc commun diviseur log-normal shifté	3 586,80	22 001 002	1 102 254	2 667 143
<i>Données historiques constatées sur la somme des branches</i>		22 001 002	1 096 704	2 981 288

Le SCR correspond à la Var à 99.5% à laquelle on a retranché la moyenne. (les valeurs affichées ici étant des valeurs mensuelles obtenues par moyenne mobile sur 12 mois, il conviendrait de les multiplier par 12 afin d'obtenir leurs équivalents annuels)

Tout comme pour le premier modèle, nous tendons vers un modèle à choc commun multiplicatif log-normal. Il donne un SCR de l'ordre de 3 millions d'euros, ce qui représente à peu près 14% de l'espérance moyenne de la charge ultime des sinistres.

V.6.3. Conclusion

Parmi les 3 modèles hiérarchiques étudiés, 2 ont retenu un modèle à choc commun multiplicatif non shifté et 1 a retenu un modèle à choc commun multiplicatif shifté.

Modèle	Type du choc commun log-normal	Sinistralité totale		
		Moyenne	Ecart-type	VaR 99,5%
1er modèle hiérarchique	Multiplicatif non shifté	22 001 002	1 113 606	3 031 872
2e modèle hiérarchique	Multiplicatif shifté	22 001 002	1 074 901	3 600 315
3e modèle hiérarchique	Multiplicatif non shifté	22 001 002	1 115 174	3 035 901

Les 2 modèles à choc multiplicatif non shifté donnent des résultats similaires tant au niveau de l'écart-type de la sinistralité que de la VaR à 99.5% qui se situe aux alentours de 13.8% de l'espérance moyenne de la charge ultime de sinistres.

En revanche, le modèle à choc multiplicatif shifté présente une VaR à 99.5% supérieure de 20%. Cet écart est clairement dû à la nature du choc commun qui en favorisant l'apparition simultanée de scénarios extrêmes sur toutes les branches gonfle fortement la VaR ; et ce d'autant plus, que le paramètre β retenu lors du calibrage présente une valeur élevée à .91.

Le risque de modèle lié au choix de la structure hiérarchique semble donc réel, et ce d'autant plus, qu'un choix peut se heurter à la limite relative au modèle à choc commun multiplicatif log-normal (cf [IV.1.9. Conclusion sur le choc multiplicatif log-normal](#)), comme c'est le cas pour le deuxième modèle hiérarchique.

En conclusion, en l'absence d'un grand nombre de données (par exemple, un millier d'observations pour avoir 5 points correspondants à un quantile 99.5%) il semble préférable de se limiter au choc commun non shifté. En effet :

- un choc commun diviseur shifté tend à baisser significativement le SCR, ce qui, à moins de disposer d'arguments concrets spécifiques aux LoBs étudiées, semble se heurter au principe de prudence de l'assurance, et également être difficile à justifier auprès du régulateur
- un choc commun lognormal shifté (multiplicatif ou diviseur) requerrait un grand nombre de données pour obtenir un calibrage fiable et robuste du paramètre β qui a un impact fort dans le SCR ; or en pratique, il est difficile de disposer d'autant de données

Enfin, dans le cas où il y a une forte présomption de dépendance pour les sinistralités extrêmes (et donc d'un fort upper tail concentration), il y a une alternative à l'utilisation de modèles à choc commun shifté. L'alternative consisterait à procéder en deux niveaux, en décomposant la sinistralité en deux parties : les sinistres graves et les sinistres attritionnels ; puis, il faudrait :

- étudier les sinistres graves à une maille individuelle ; il est alors possible d'appliquer des modèles de choc commun traditionnels basés sur un choc portant sur les lois de Poisson modélisant le nombre de sinistres graves
- modéliser la dépendance pour la sinistralité attritionnelle avec un modèle de choc commun lognormal multiplicatif non-shifté

VI. CONCLUSION

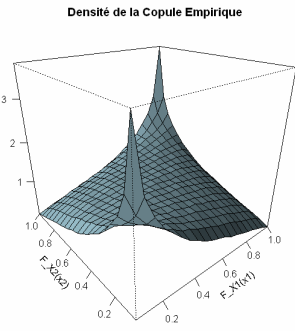
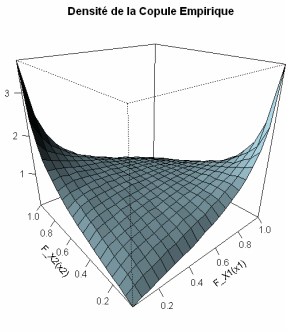
VI.1. Résumé

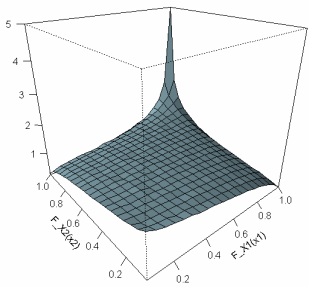
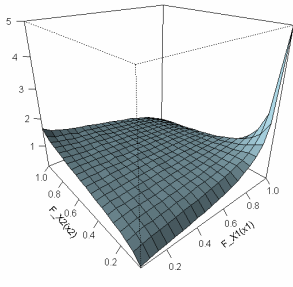
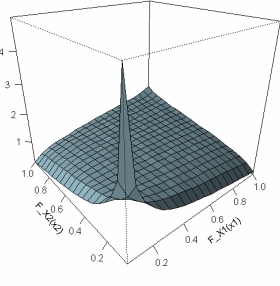
Les modèles à chocs communs log-normaux multiplicatifs permettent de produire différents types de structure de dépendance. Le tableau ci-dessous résume les possibilités offertes par ces modèles, dans le cas de deux branches :

#	Modèle	Type de structures engendrées	Comparaison avec les copules
1	$\begin{cases} X_1 = X_{10} \times \beta \\ X_2 = X_{20} \times \beta \end{cases}$	Structure symétrique Dépendance positive	Mathématiquement identique à une copule gaussienne
2	$\begin{cases} X_1 = X_{10} \times \beta \\ X_2 = X_{20} / \beta \end{cases}$	Structure symétrique Dépendance négative	Mathématiquement identique à une copule gaussienne
3	$\begin{cases} X_1 = X_{10} \times (\beta + c) \\ X_2 = X_{20} \times (\beta + c) \end{cases}$	Structure asymétrique avec le poids sur les sinistralités fortes Dépendance positive	Assimilable à une copule de Gumbel ou HRT pour des valeurs de choc élevées (#.9)
4	$\begin{cases} X_1 = X_{10} \times (\beta + c) \\ X_2 = X_{20} / (\beta + c) \end{cases}$	Structure asymétrique Dépendance négative	Pour valeurs de choc élevées (#.9), la structure de dépendance est assimilable à celle d'une copule HRT, Gumbel ou Clayton ayant subi une rotation de +/-90° par rapport au centre du carré [0;1] ²
5	$\begin{cases} X_1 = X_{10} / (\beta + c) \\ X_2 = X_{20} / (\beta + c) \end{cases}$	Structure asymétrique avec le poids sur les sinistralités faibles Dépendance positive	Assimilable à une copule de Clayton pour des valeurs de choc élevées (#.9)

Dans ce tableau , X_1, X_2 sont des lognormales et β est une lognormale ; c est une constante (prise supérieure à 0.5 ; sinon on retombe dans un cas très similaire à celui d'un choc sans shift), et les modèles qui incluent cette constante sont dits « shiftés »

A titre d'illustration voici les représentation tri-dimensionnelles des structures de dépendance (se reporter à : [III.4.2. Visualisation de la structure de dépendance à l'aide de copule empirique](#) pour de plus amples informations) :

#	Structure de dépendance	#	Structure de dépendance
(1)		(2)	

(3)	<p style="text-align: center;">Densité de la Copule Empirique</p> 	(4)	<p style="text-align: center;">Densité de la Copule Empirique</p> 
(5)	<p style="text-align: center;">Densité de la Copule Empirique</p> 		

L'introduction d'un shift dans la loi lognormale du choc commun permet d'élargir le spectre de structure de dépendance des modèles. En effet, notre étude montre qu'il est alors possible de jouer sur deux paramètres : le shift et la variance de la loi log-normale. Le paramètre du shift a deux effets :

- au niveau de la dépendance de queue : augmentation de l'asymétrie de la structure de dépendance, ce qui est également visible via les indicateurs de dépendance de queue
- au niveau de la dépendance générale : légère diminution de la dépendance de rang.

Quant au paramètre lié à la variance de la loi log-normale, en l'augmentant, il y a deux impacts :

- augmentation de la corrélation linéaire induite par le modèle
- accentuation de la forme de la dépendance produite, couplée avec l'augmentation des indicateurs de rang

Toutefois, le recours au shift entraîne également une perte de robustesse qui rend le modèle un peu plus difficile à justifier. En effet, le shift conditionne l'apparition simultanée de sinistralité extrêmes ; il faudrait donc un très grand nombre de données pour faire un calibrage robuste, ce qui arrive peu souvent en pratique.

Par ailleurs, les impacts en SCR peuvent être soit une augmentation très forte (+20% par rapport à un modèle sans shift dans le cas [V.5.2.2. 2ème modèle hiérarchique](#) ; +50% dans le cas [V.4.2. Modélisation de la dépendance entre Accident&Health et Worker Compensation](#)), soit un SCR inférieur à celui du cas d'indépendance ; le deuxième cas requiert une solide justification.

Egalement, il est bon de noter que la portée du modèle peut être limitée dans le cas où les marginales ont des coefficients de variations très différents, et ce par la relation suivante :

$$\rho = Cor[X_1, X_2] \leq \text{Min}\left(\frac{CV_1}{CV_2}; \frac{CV_2}{CV_1}\right)$$

Dans le cas où il y a plusieurs branches, l'utilisation de modèle à chocs communs devient plus délicate. La solution retenue consiste à faire des modèles hiérarchiques, i.e. à regrouper plusieurs branches dans des macro-branches. Mais apparaît alors le souci du risque de modèle. Ce risque tend à diminuer significativement si on se restreint aux modèles à choc log-normal non shifté.

Malgré ces quelques inconvénients, les modèles à choc communs présentent des avantages certains qui seront présentés dans la partie suivante.

VI.2. Synthèse sur la comparaison avec les copules

L'avantage principal des modèles à chocs communs multiplicatifs par rapport aux copules réside dans sa facilité d'explication. En effet, il est plus simple d'expliquer un modèle basé sur une multiplication des sinistralités des branches par un facteur commun que des copules qui influent sur des distributions jointes multidimensionnelles.

Aussi, dans ce sens, ils peuvent être également plus faciles à justifier auprès d'un régulateur.

De plus, la simulation des modèles à chocs communs multiplicatifs est plus simple et ne nécessite pas d'algorithme et/ou d'outils dédiés.

Toutefois, dans le cas où il y a plus que trois branches, ils requièrent l'utilisation de modèle hiérarchique ce qui se heurte au risque de modèle et à la limite théorique :

$$\rho = Cor[X_1, X_2] \leq Min\left(\frac{CV_1}{CV_2}; \frac{CV_2}{CV_1}\right)$$

Dans le cas où le ou les facteurs de dépendance ne sont pas clairement explicables, les copules peuvent être une bonne option dans le sens où elles constituent une démarche bien connue et fréquemment utilisée.

En revanche, s'il est possible d'explicitier un facteur de dépendance (par exemple un indice de prix, comme le coût de la Tierce Personne), alors les modèles à choc commun pourront être préférés. Et le présent mémoire apporte des éléments de réponse pour mieux appréhender la dépendance créée par l'utilisation de tels modèles.

VII. BIBLIOGRAPHIE

- Mémoire d'actuariat de 2007 « Provisionnement et Corrélation entre Branches » (KRAUTH G.)
- Livre « Assurance Non-Vie : Modélisation, Simulation » (Christian Partrat et Jean-Luc Besson)
- Article « The Common Shock Model for Correlated Insurance Losses » (Glenn G. Meyers)
- Cours 2009 de l'ISFA « MESURES et CONCEPTS de DEPENDANCE - ρ de Spearman, τ de Kendall concepts PQD, PUOD et PLOD » (Magda BOLEWSKA, Anisa CAJA, Marie DOITTEAU, Claire RUGA, Jean-Christophe VANNEAUX, Javier VARGAS)
- Cours 2009 de l'ISFA « Tail Dependence » (Loïc ALANDOU d'ARBOUSSIER, Jonathan AQUEREBURU, Andrei BEJENARI, Selma JAZIRI, Thierry MOUDIKI et Mohamed TALFI)
- Support de la conférence actuarielle du 28 Septembre 2008 « 33rd ANNUAL GIRO CONVENTION - Top down / Bottom up Correlation » (Colin Kerley / Simon Margetts)
- Article « Aggregation of Correlated Risk Portfolios: Models & Algorithms » (Shaun S. Wang, Ph.D.)
- Article « The Report of the Research Working Party on Correlations and Dependencies Among All Risk Sources » (Casualty Actuarial Society _Forum, Winter 2006)
- Mémoire d'actuariat de 2008 « Tempêtes : prise en compte de la dépendance spatiale avec la théorie des copules » (ANDRIANTAVY C.)
- Mémoire d'actuariat de 2007 « Un modèle stochastique non-vie utilisant des copulas elliptique de dimension n pour évaluer des besoins de fonds propres » (BAGARRY M.)
- Mémoire d'actuariat de 2002 "Copules, besoin en fonds propres et allocation de capital pour une compagnie d'assurance non vie multibranches"
- QIS4 : Technical specifications - CEIOPS

VIII. ANNEXES

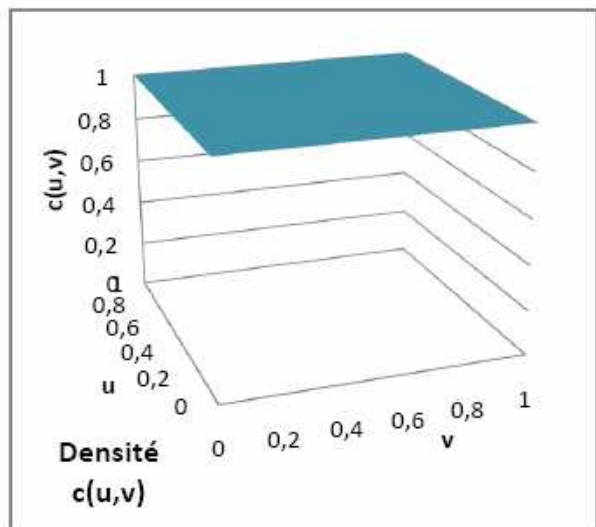
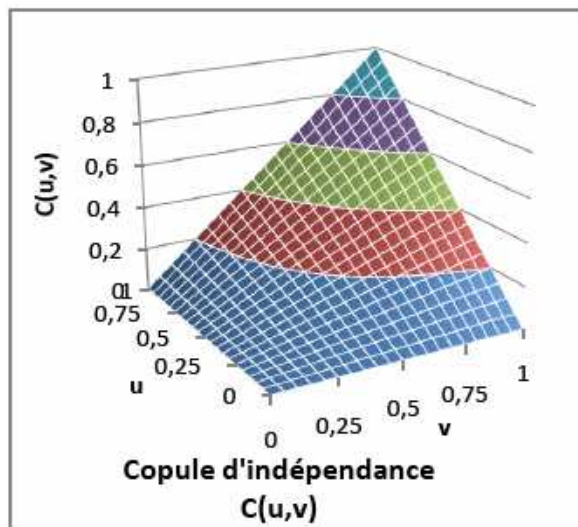
VIII.1. Les copules de référence

VIII.1.1. Copule d'indépendance

Il s'agit de la copule correspondant au cas où les branches sont indépendantes. En dimension 2, nous obtenons donc :

$$C(u, v) = u \cdot v \text{ et } c(u, v) = 1$$

Voici la représentation graphique de la fonction de densité et de la fonction de répartition :



VIII.1.2. Copule normale

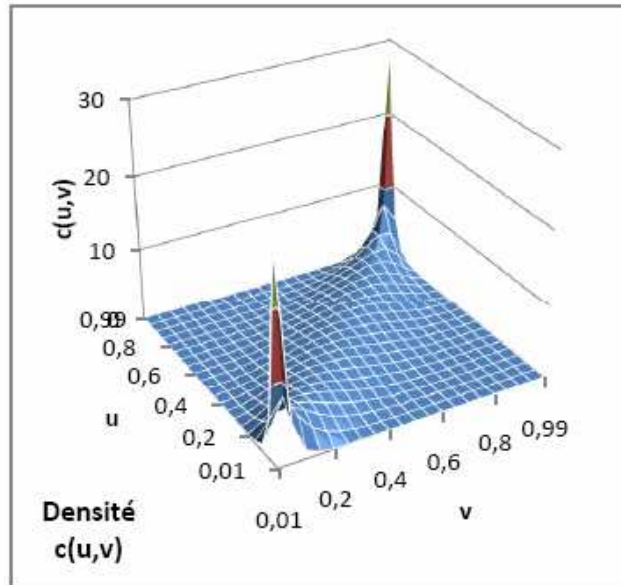
Avec R une matrice symétrique définie positive avec $\text{diag}(R) = (1, 1, \dots, 1)$ et ϕ_R la fonction de répartition de la loi normale multivariée centrée ayant R comme matrice de corrélation, la copule Normale est définie par :

$$C_R^{\text{Normale}}(u, v) = \phi_R(\phi^{-1}(u), \phi^{-1}(v))$$

où ϕ^{-1} désigne l'inverse de la fonction de répartition de la loi normale centrée réduite ϕ

Il n'existe pas d'expression analytique simple pour la fonction de densité.

Voici la représentation graphique de la fonction de densité:



$$\tau = .7 \Leftrightarrow \rho = .89$$

Par ailleurs, nous avons les relations suivantes concernant les indicateurs de dépendance :

$$\tau_{Kendall} = \frac{2}{\pi} \cdot \text{Arc sin}(\rho)$$

$$\rho_{Spearman} = \frac{6}{\pi} \cdot \text{Arc sin}\left(\frac{\rho}{2}\right)$$

VIII.1.3. Copule de Gumbel

La couple de Gumbel est une copule dite « archimédienne ». Elle est construite à partir d’une fonction continue, décroissante et convexe φ :

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

Pour la copule de Gumbel, nous avons :

$$\varphi_\theta(t) = (-\ln(t))^\theta \quad \text{avec } \theta > 1$$

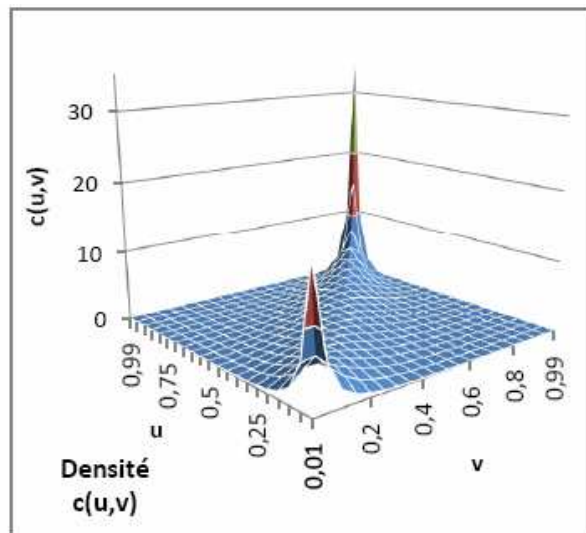
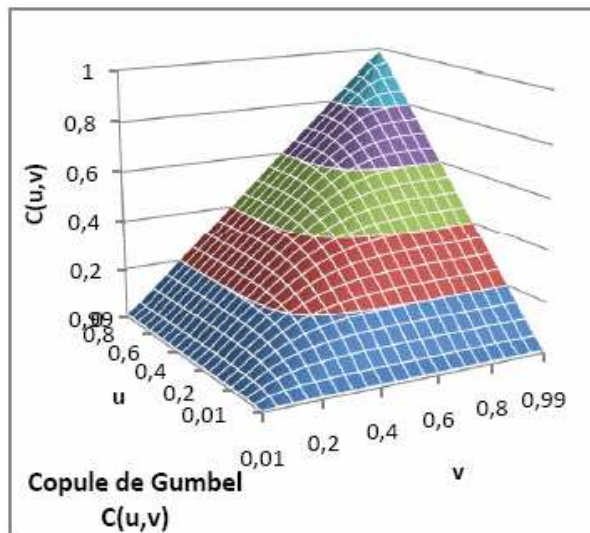
Et donc :

$$C(u_1, u_2, \dots, u_d) = \exp\left[-\left(\sum_{i=1}^d (-\ln(u_i))^\theta\right)^{1/\theta}\right]$$

ce qui donne en dimension 2 :

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

Voici la représentation graphique de la fonction de densité et de la fonction de répartition :



$$\tau = .7 \Leftrightarrow \theta = 3.33$$

Le paramètre θ de la copule de Gumbel étant nécessairement supérieur à 1, celle-ci n'appréhende que les dépendances positives. Elle représente bien les risques ayant une structure de dépendance plus accentuée que la queue. C'est pourquoi elle est particulièrement adaptée en assurance et en finance, notamment lorsque l'on étudie la dépendance entre les événements de forte intensité.

Enfin, nous avons la relation suivante pour le Tau de Kendall : $\tau(\theta) = 1 - \frac{1}{\theta}$

VIII.1.4. Copule de Clayton

Tout comme la copule de Gumbel, la copule de Clayton est une copule archimédienne. Son générateur est le suivant :

$$\varphi_{\theta}(t) = \frac{(t^{-\theta} - 1)}{\theta} \quad \text{avec } \theta > 0$$

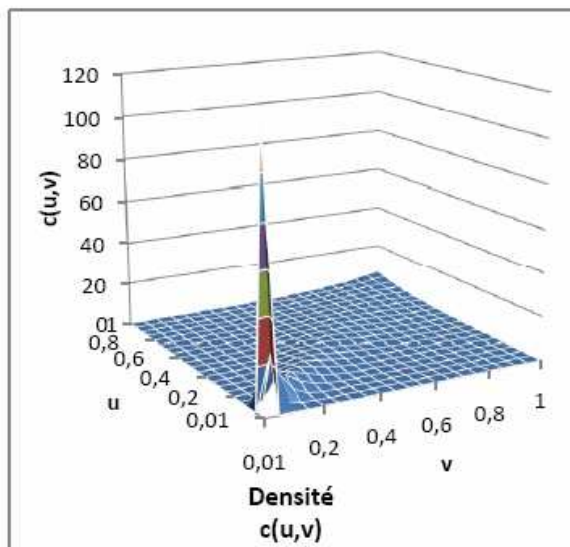
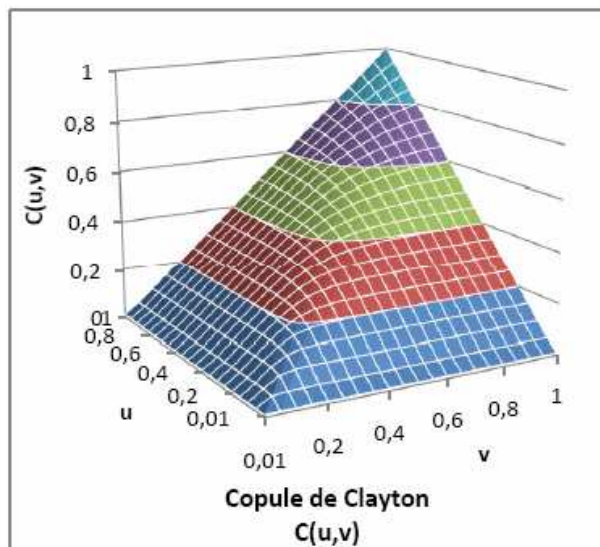
Et la copule de Clayton est donc :

$$C(u_1, u_2, \dots, u_d) = \left(\sum_{i=1}^d u_i^{-\theta} - d + 1 \right)^{\frac{1}{\theta}}$$

Et en dimension 2, nous avons :

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

Voici la représentation graphique de la fonction de densité et de la fonction de répartition :



$$\tau = .7 \Leftrightarrow \theta = 4.67$$

Comme la copule de Gumbel, la copule de Clayton ne permet également de ne modéliser les dépendances positives. En revanche, elle permet d'étudier la dépendance entre les événements de faible intensité.

Enfin, nous avons la relation suivante pour le Tau de Kendall : $\tau(\theta) = \frac{\theta}{\theta + 2}$

VIII.1.5. Copule HRT

Nous avons vu que la copule de Clayton a tendance à corrélérer entre eux les petits sinistres. L'idée de la copule HRT est de construire une copule dont les propriétés sont opposées à celles de la copule de Clayton. Pour ce faire, la copule HRT est définie comme la copule de survie de la copule de Clayton.

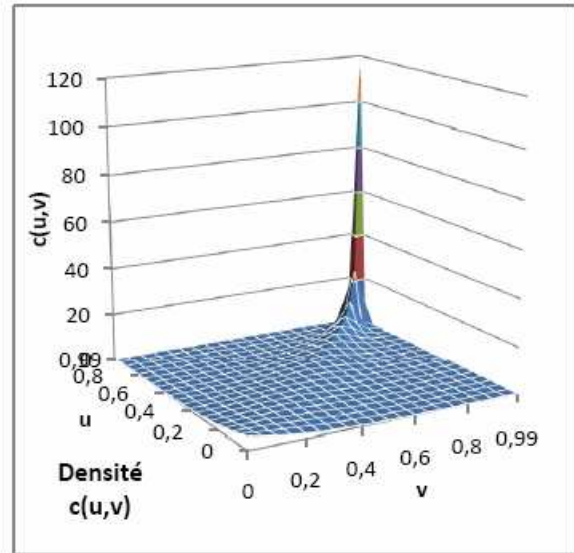
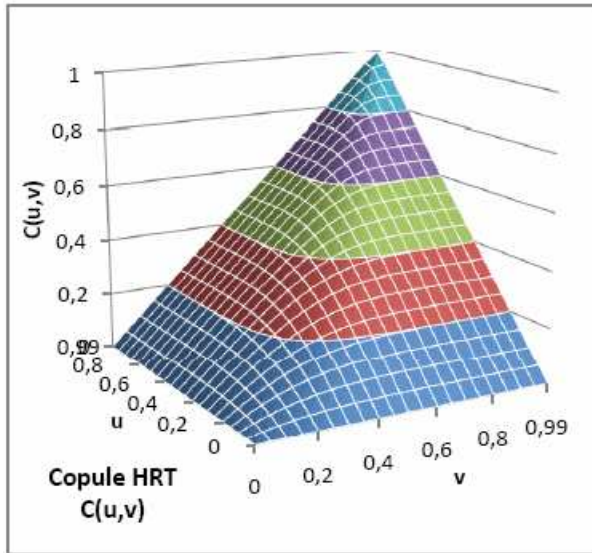
L'expression suivante donne l'expression de la copule de survie en fonction de la copule :

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

D'où, l'expression de la copule HRT en dimension 2 :

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

Voici la représentation graphique de la fonction de densité et de la fonction de répartition :



$$\tau = .7 \Leftrightarrow \theta = 4.67$$

Enfin, nous avons la relation suivante pour le Tau de Kendall : $\tau(\theta) = \frac{\theta}{\theta + 2}$

VIII.2. Résultats obtenus pour le choc commun multiplicatif log-normal

no	Intensité du choc	Corrélation Pearson théorique	Corrélation Pearson mesurée	écart en %	Tau de Kendall mesuré	Tau de Kendall théorique	écart en %	Rho de Spearman mesuré	Rho de Spearman théorique	écart en %	Tail dependance 90%	Tail dependance 99.5%
1	0,025	0,03985	0,04101	2,9%	0,02752	0,02611	-5,1%	0,03913	0,03916	0,1%	0,11301	0,00780
2	0,05	0,14112	0,14036	-0,5%	0,08863	0,08965	1,2%	0,13477	0,13415	-0,5%	0,14925	0,01320
3	0,075	0,26725	0,26698	-0,1%	0,17812	0,17205	-3,4%	0,25782	0,25571	-0,8%	0,20160	0,03020
4	0,1	0,38991	0,38993	0,0%	0,24799	0,25500	2,8%	0,37761	0,37476	-0,8%	0,26375	0,05940
5	0,125	0,49610	0,49627	0,0%	0,33407	0,33059	-1,0%	0,48279	0,47890	-0,8%	0,32362	0,10199
6	0,15	0,58305	0,58267	-0,1%	0,40078	0,39598	-1,2%	0,56968	0,56459	-0,9%	0,38042	0,13979
7	0,175	0,65253	0,65212	-0,1%	0,45721	0,45224	-1,1%	0,63990	0,63432	-0,9%	0,43293	0,19025
8	0,2	0,70762	0,70807	0,1%	0,51062	0,50087	-1,9%	0,69680	0,69115	-0,8%	0,48021	0,23932
9	0,225	0,75135	0,75127	0,0%	0,55359	0,54111	-2,3%	0,74198	0,73545	-0,9%	0,51883	0,28605
10	0,25	0,78628	0,78649	0,0%	0,57677	0,57621	-0,1%	0,77865	0,77189	-0,9%	0,55493	0,32691
11	0,275	0,81439	0,81451	0,0%	0,61397	0,60598	-1,3%	0,80799	0,80108	-0,9%	0,58644	0,36598
12	0,3	0,83723	0,83723	0,0%	0,63558	0,63165	-0,6%	0,83222	0,82490	-0,9%	0,61291	0,40264
13	0,325	0,85597	0,85614	0,0%	0,66103	0,65429	-1,0%	0,85236	0,84484	-0,9%	0,63736	0,43704
14	0,35	0,87148	0,87144	0,0%	0,68024	0,67362	-1,0%	0,86862	0,86103	-0,9%	0,65700	0,46824
15	0,375	0,88443	0,88449	0,0%	0,69761	0,69099	-0,9%	0,88280	0,87491	-0,9%	0,67562	0,49417
16	0,4	0,89534	0,89549	0,0%	0,71807	0,70635	-1,6%	0,89430	0,88664	-0,9%	0,69341	0,51317
17	0,425	0,90460	0,90453	0,0%	0,72752	0,71956	-1,1%	0,90440	0,89630	-0,9%	0,70841	0,53863
18	0,45	0,91252	0,91236	0,0%	0,73932	0,73148	-1,1%	0,91292	0,90469	-0,9%	0,72178	0,55090
19	0,475	0,91934	0,91931	0,0%	0,74753	0,74250	-0,7%	0,92045	0,91216	-0,9%	0,73403	0,56963
20	0,5	0,92524	0,92539	0,0%	0,76521	0,75253	-1,7%	0,92706	0,91871	-0,9%	0,74392	0,59109

Nous pouvons constater que les écarts entre valeur théoriques et valeurs mesurées sont très faibles : les écarts relatifs sont inférieurs à 1% sauf pour la première simulation pour laquelle ils sont plus importants tout en restant à des valeurs très correctes. Cette tendance peut s'expliquer par le fait que pour un choc d'intensité très faible, les indicateurs de dépendance seront faibles, d'où une différence relative plus forte.

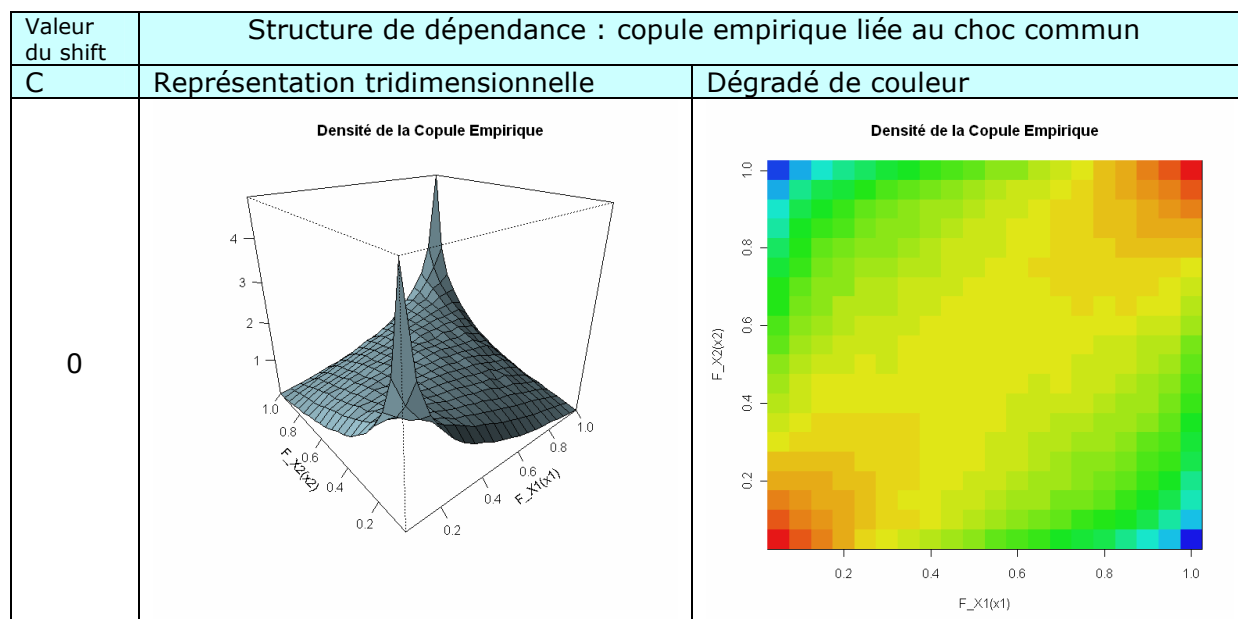
VIII.3. Résultats obtenus pour le choc commun multiplicatif log-normal shifté

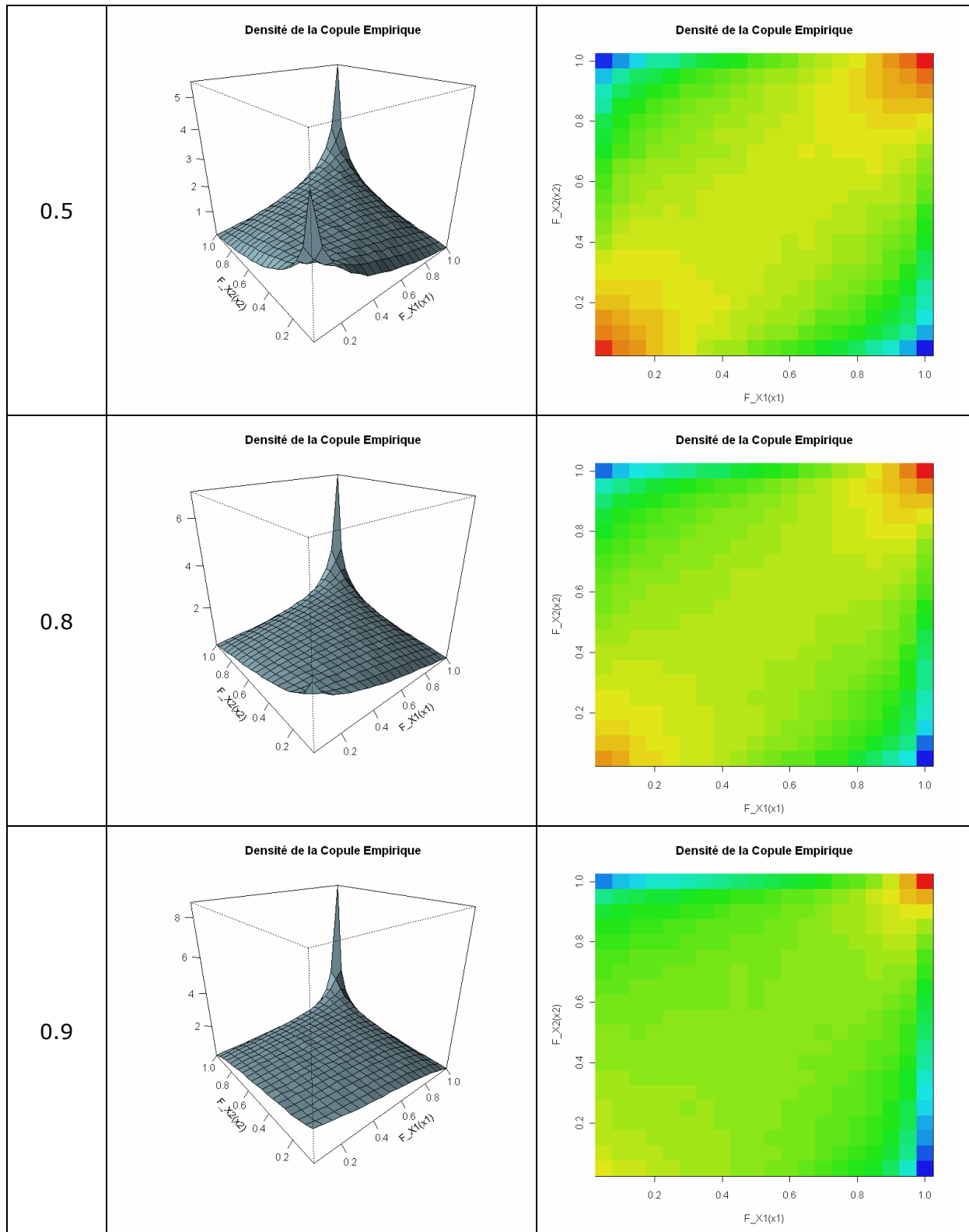
VIII.3.1. Cas de cor = .3

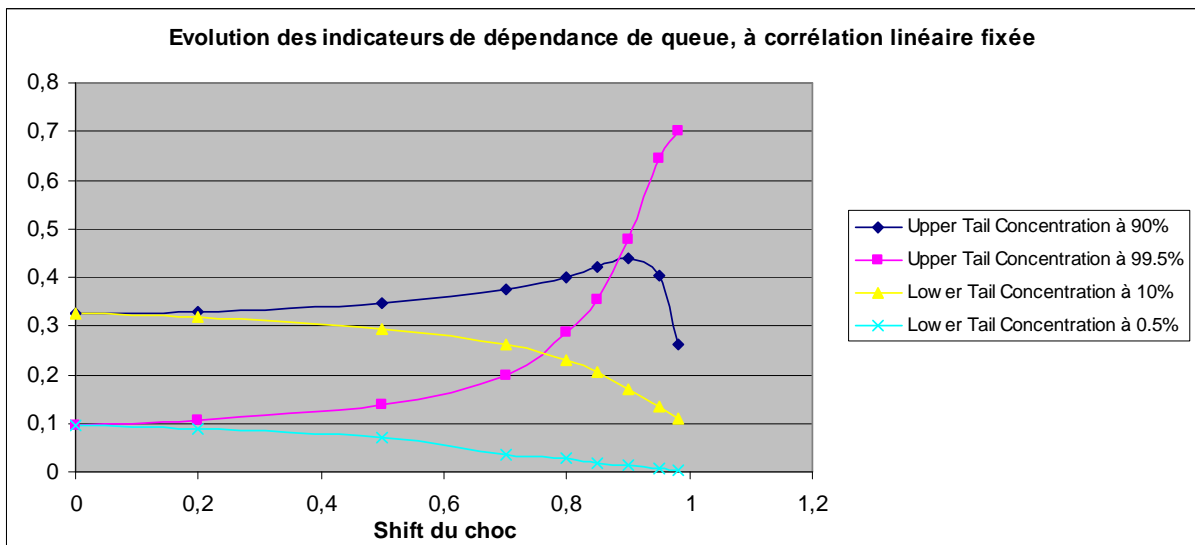
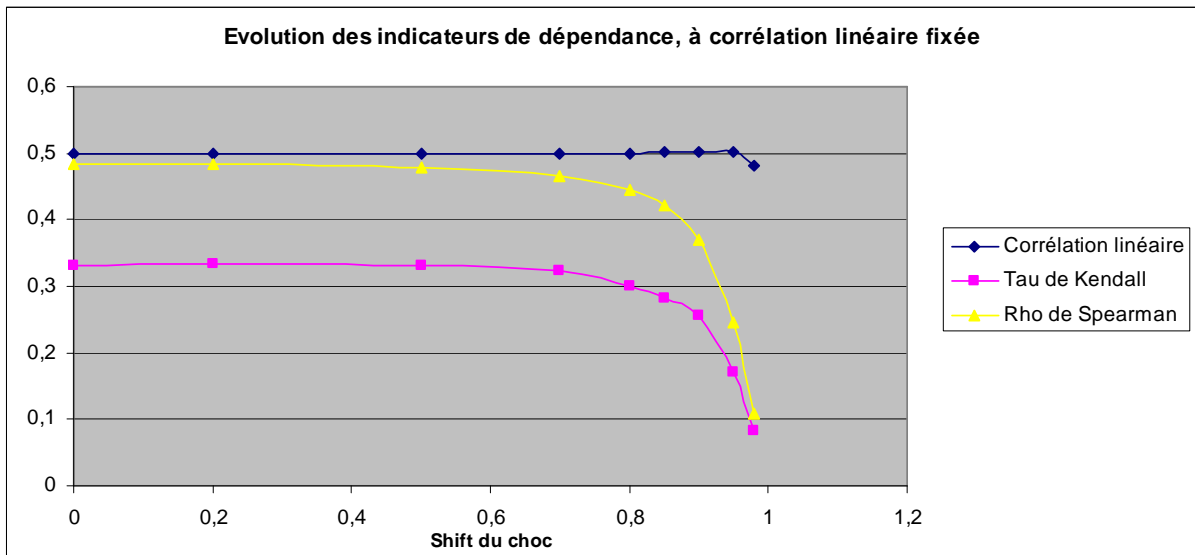
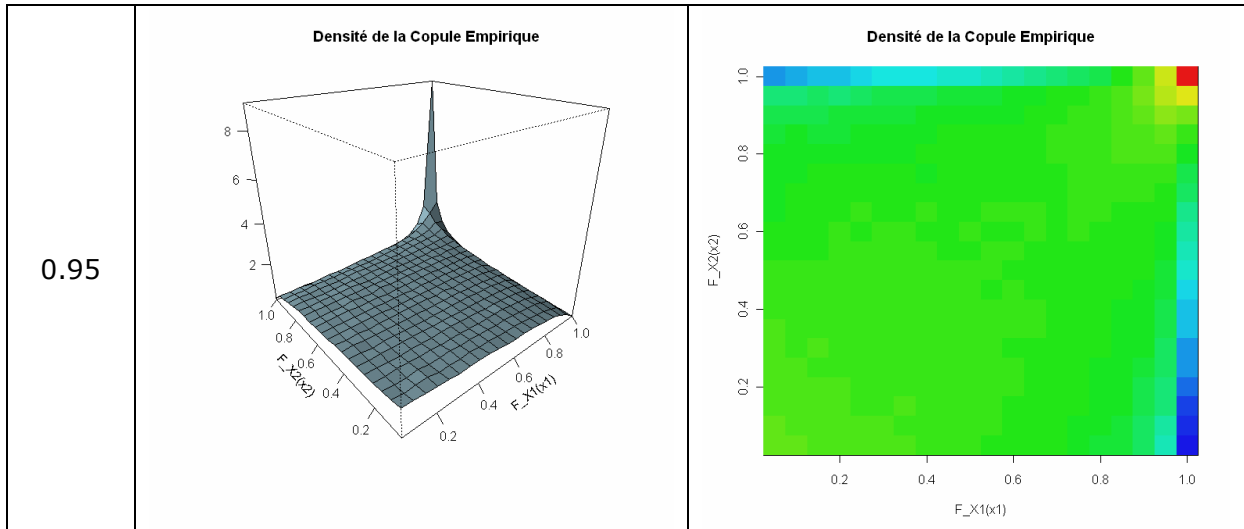
Shift	Corrélation Pearson	Tau de Kendall	Rho de Spearman	Upper Tail Concentration à 90%	Upper Tail Concentration à 99.5%	Lower Tail Concentration à 10%	Lower Tail Concentration à .5%
0,000	0,299	0,196	0,289	0,216	0,038	0,214	0,041
0,200	0,298	0,188	0,286	0,218	0,037	0,211	0,031
0,500	0,302	0,193	0,288	0,227	0,043	0,204	0,028
0,700	0,300	0,189	0,281	0,239	0,063	0,191	0,021
0,800	0,301	0,182	0,274	0,254	0,097	0,179	0,018
0,850	0,300	0,177	0,263	0,269	0,131	0,168	0,015
0,900	0,301	0,163	0,242	0,288	0,225	0,151	0,011
0,950	0,301	0,118	0,176	0,290	0,442	0,125	0,005
0,980	0,307	0,059	0,084	0,219	0,550	0,110	0,005

Shift	D avec copule d'indépendance	D avec copule gaussienne	D avec copule Gumbel	D avec copule Clayton	D avec copule HRT
0,000	8,769	0,009	1,844	3,527	3,459
0,200	8,637	0,013	1,540	3,406	2,934
0,500	8,794	0,045	1,354	4,063	2,696
0,700	8,573	0,178	0,921	4,739	1,940
0,800	8,523	0,462	0,498	5,471	1,135
0,850	8,418	0,853	0,286	6,248	0,646
0,900	8,111	1,655	0,158	7,043	0,156
0,950	6,088	3,298	0,531	6,135	0,335
0,980	2,020	5,265	0,394	2,162	0,395

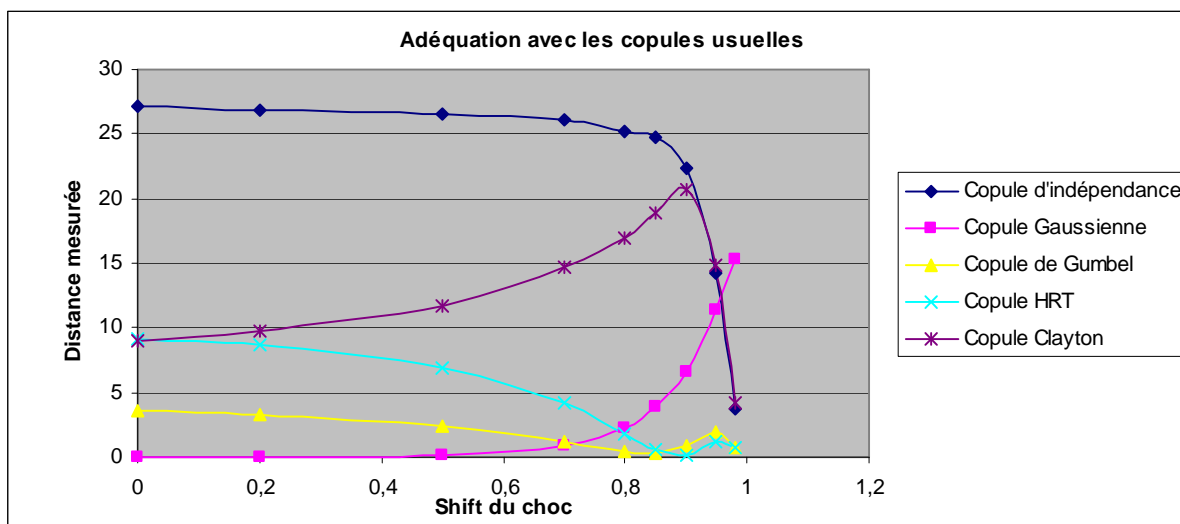
VIII.3.2. Cas de cor = .5







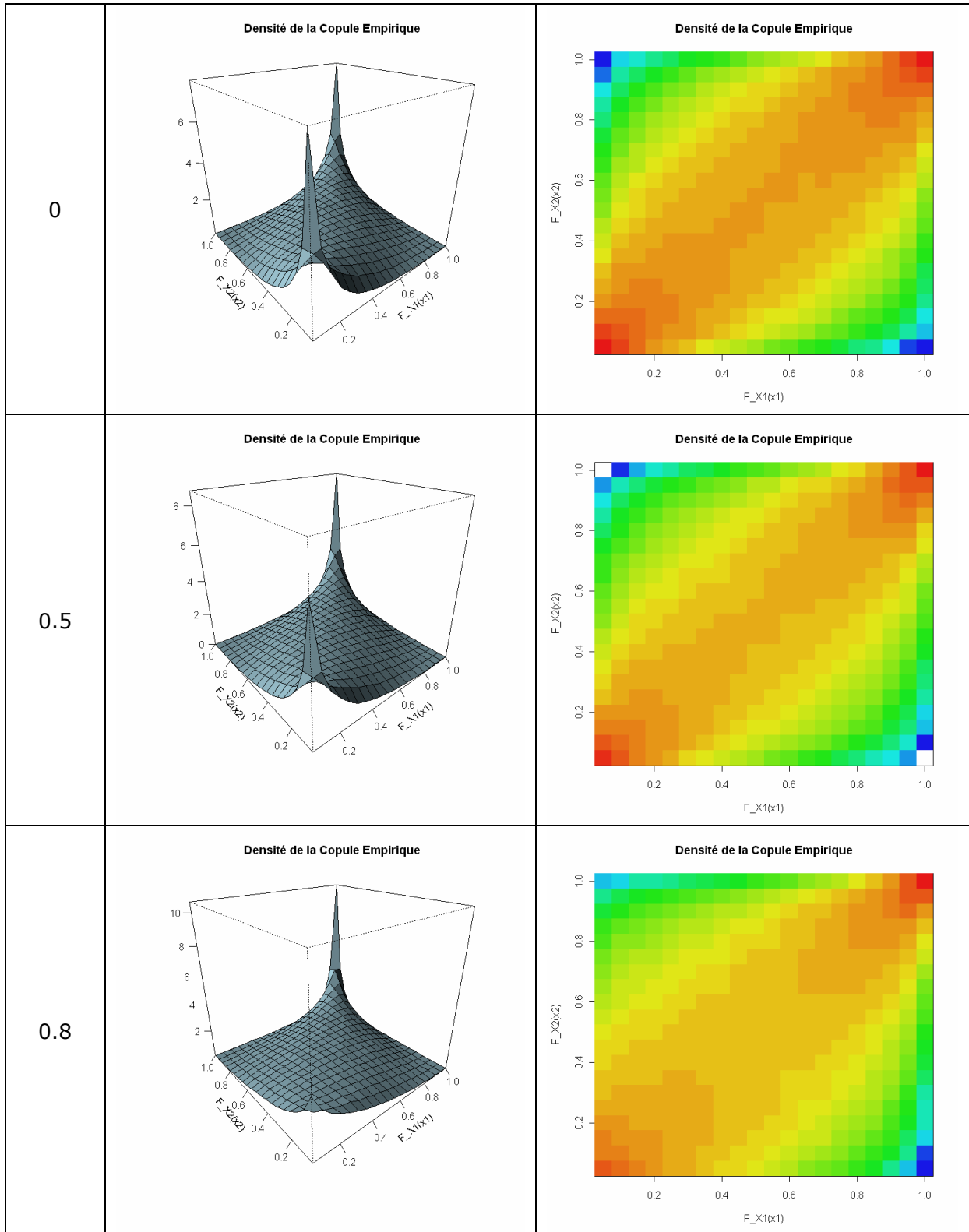
Shift	Corrélation Pearson	Tau de Kendall	Rho de Spearman	Upper Tail Concentration à 90%	Upper Tail Concentration à 99.5%	Lower Tail Concentration à 10%	Lower Tail Concentration à .5%
0,000	0,499	0,332	0,483	0,325	0,094	0,327	0,094
0,200	0,499	0,334	0,482	0,330	0,106	0,318	0,087
0,500	0,499	0,332	0,478	0,347	0,137	0,296	0,071
0,700	0,500	0,324	0,467	0,374	0,197	0,264	0,035
0,800	0,500	0,301	0,446	0,398	0,286	0,229	0,028
0,850	0,500	0,281	0,422	0,423	0,353	0,206	0,018
0,900	0,500	0,255	0,370	0,438	0,478	0,170	0,013
0,950	0,502	0,171	0,247	0,404	0,644	0,134	0,008
0,980	0,481	0,082	0,109	0,263	0,700	0,111	0,005

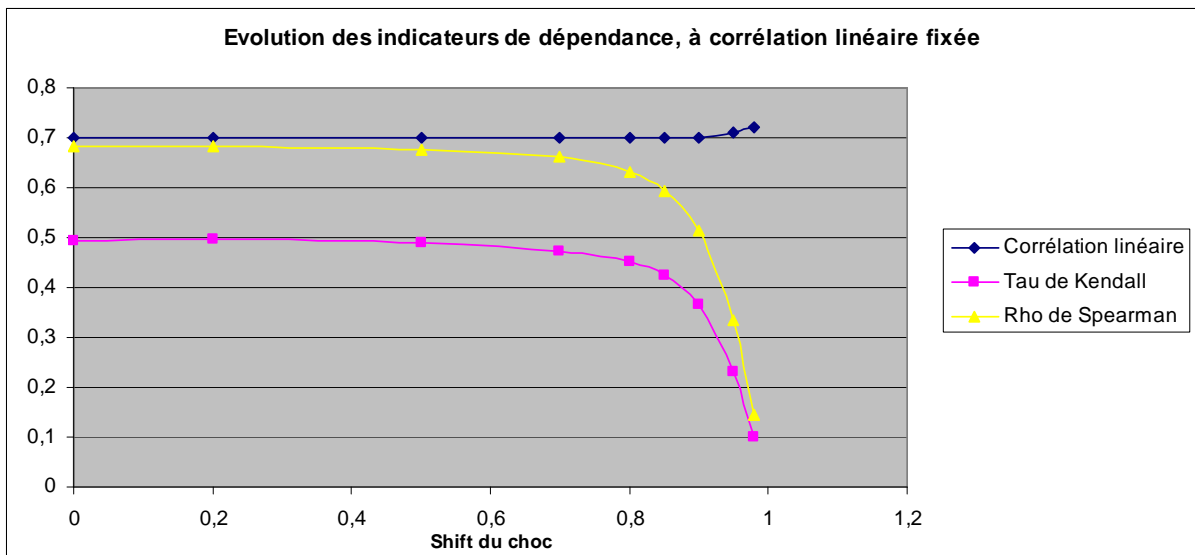
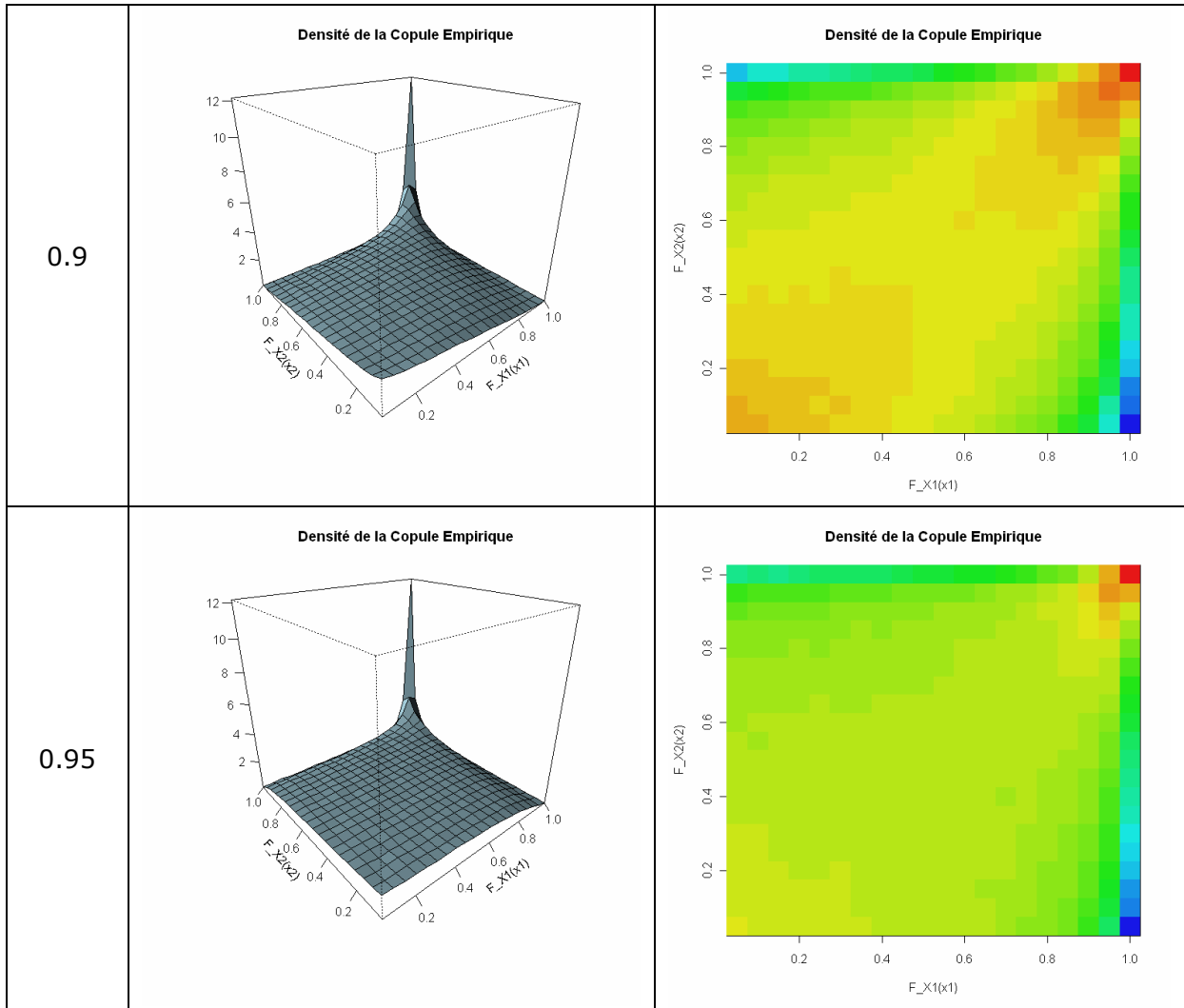


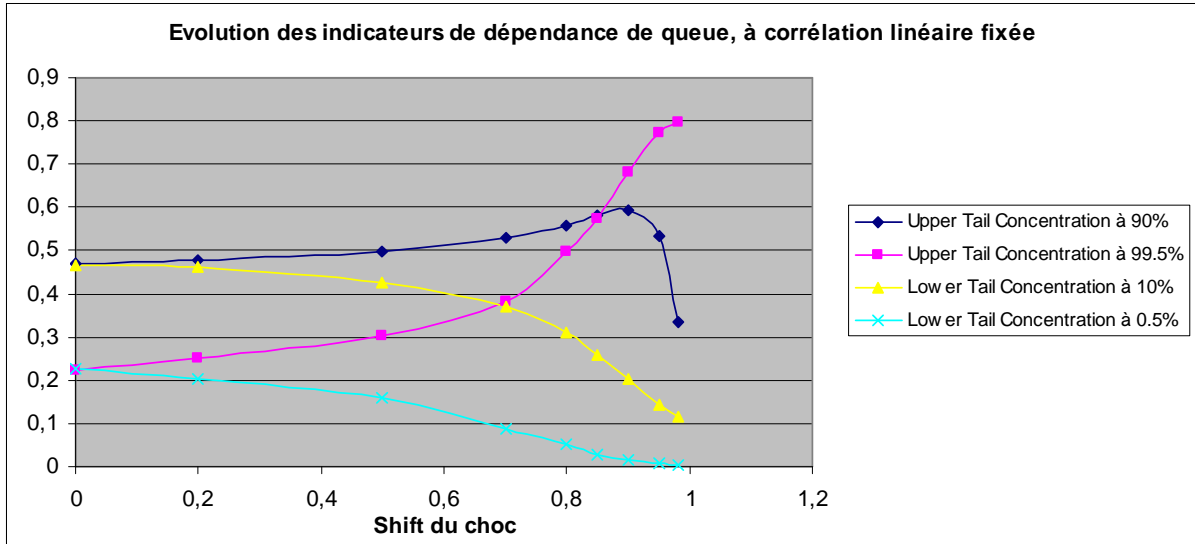
Shift	D avec copule d'indépendance	D avec copule gaussienne	D avec copule Gumbel	D avec copule Clayton	D avec copule HRT
0,000	27,098	0,008	3,557	8,948	9,198
0,200	26,898	0,016	3,257	9,736	8,655
0,500	26,497	0,190	2,327	11,738	6,830
0,700	26,068	0,895	1,172	14,753	4,238
0,800	25,163	2,213	0,416	17,012	1,832
0,850	24,688	3,838	0,332	18,899	0,611
0,900	22,294	6,636	0,913	20,731	0,170
0,950	14,266	11,412	2,024	14,897	1,246
0,980	3,724	15,291	0,796	4,236	0,737

VIII.3.3. Cas de cor = .7

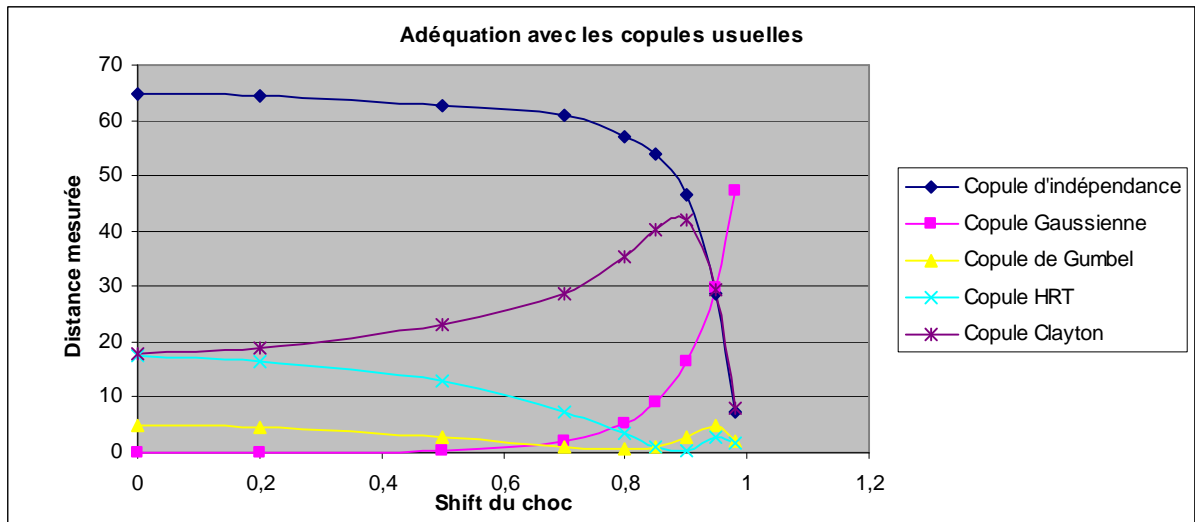
Valeur du shift	Structure de dépendance : copule empirique liée au choc commun	
C	Représentation tridimensionnelle	Dégradé de couleur







Shift	Corrélation Pearson	Tau de Kendall	Rho de Spearman	Upper Tail Concentration à 90%	Upper Tail Concentration à 99.5%	Lower Tail Concentration à 10%	Lower Tail Concentration à .5%
0,000	0,700	0,492	0,684	0,471	0,223	0,468	0,228
0,200	0,700	0,495	0,682	0,477	0,251	0,460	0,203
0,500	0,700	0,490	0,677	0,498	0,303	0,424	0,160
0,700	0,701	0,471	0,661	0,531	0,381	0,371	0,089
0,800	0,699	0,451	0,629	0,559	0,497	0,310	0,051
0,850	0,700	0,424	0,593	0,580	0,572	0,260	0,028
0,900	0,701	0,364	0,513	0,593	0,682	0,203	0,015
0,950	0,709	0,232	0,335	0,533	0,772	0,143	0,008
0,980	0,721	0,100	0,146	0,333	0,798	0,116	0,006



Shift	D avec copule d'indépendance	D avec copule gaussienne	D avec copule Gumbel	D avec copule Clayton	D avec copule HRT
0,000	64,724	0,010	4,795	17,748	17,366
0,200	64,260	0,030	4,411	19,017	16,609
0,500	62,815	0,424	2,866	22,989	12,840
0,700	60,830	2,050	1,049	28,693	7,363
0,800	57,155	5,262	0,601	35,478	3,393
0,850	53,793	9,275	1,105	40,237	1,202
0,900	46,615	16,323	2,748	42,061	0,415
0,950	28,556	29,676	4,953	29,553	2,813
0,980	7,482	47,139	2,091	7,931	1,906

VIII.4. Programme « R » utilisé

```
#####
### -pour les graphes de dépendance en fonction de var(beta)
###   -prendre simu=0, corr=.2/.3/.4/.5
###   -changer cas_etude=1/2 (2 c'est pour le second cas étudié)
###   -ne pas oublier de charger les packages
#####

rm(list=ls(all=TRUE))

#####
### Graphical Functions
#####

jauge_rainbow <- function (n){
  z0 <- array(0,c(4,n))
  for (i in 1:4){
    for (j in 1:n){
      z0[i,j]=-j
    }
  }
  x0=seq(1,4,1)
  y0=seq(1,n,1)
  image(x0, y0, z0, col = rainbow(n,start=0/6,end=4/6,s=0.9,v=0.9), axes = FALSE, main = "Légende")
}

troisd <- function (x,y,nbre_pas_x,nbre_pas_y){
  #fonction de densité empirique
  z <- array(0,c(nbre_pas_x,nbre_pas_y))
  for(k in 1:length(x)){
    i = as.integer(x[k]*nbre_pas_x)+1
    j = as.integer(y[k]*nbre_pas_y)+1
    if (x[k]==1) i=nbre_pas_x
    if (y[k]==1) j=nbre_pas_y
    z[i,j]=z[i,j]+1
  }
  for(i in 1:nbre_pas_x){
    for (j in 1:nbre_pas_y){
      z[i,j]=z[i,j]/length(x)
    }
  }
  #fonction de répartition empirique
  z2 <- array(0,c(nbre_pas_x,nbre_pas_y))
  for(i in 1:nbre_pas_x){
    if (i==1) z2[i,1]=z[i,1] else z2[i,1]=z[i,1]+z2[i-1,1]
  }
  for(j in 1:nbre_pas_y){
    if (j==1) z2[1,j]=z[1,j] else z2[1,j]=z[1,j]+z2[1,j-1]
  }
  for(i in 2:nbre_pas_x){
    for(j in 2:nbre_pas_y){
      z2[i,j]=z[i,j]+z2[i-1,j]+z2[i,j-1]-z2[i-1,j-1]
    }
  }
  list(x=seq(1,nbre_pas_x,1)/nbre_pas_x,y=seq(1,nbre_pas_y)/nbre_pas_y,z=z,z2=z2)
}

#####
### calcule les matrices de corrélation
### en input : matrice, contient pour chaque lob (1 lob = 1 colonne) des observations
#####

calcul_cor <- function (matrice,type_cor){
```

```

#fonction de densité empirique
long = length(matrice[1:1,])
correl <- matrix(nrow=long,ncol=long)

for(i in 1:long){
  for (j in 1:long){
    correl[i,j]=cor(matrice[,i:i], matrice[,j:j] , method = type_cor)
  }
}
list(correl=correl)
}

#####
### Statistical Functions
#####

##### fonction de génération d'échantillons suivant une pareto
rpareto <- fonction (n,xmin,alpha){
  x=runif(n,0,1)
  y=seq(1,n,1)
  for(i in 1:n)
  {
    y[i]=xmin/((1-x[i])^(1/alpha))
  }
  list(x=y)
}

##### fonction d'erreur et son inverse
erf <- fonction(x) 2 * pnorm(x * sqrt(2)) - 1
erf_inv <- fonction(x) 1/sqrt(2)* qnorm(x/2+1/2)

##### pour calibre mu et sigma d'une lognormale pour avoir une moyenne et sigma donné
mu_lognormal <- fonction(moy,sig) log(moy)-0.5*log(1+sig^2/moy^2)
sig_lognormal <- fonction(moy,sig) sqrt(log(1+sig^2/moy^2))

##### pour calibre k et teta d'une gamma pour avoir une moyenne et sigma donné
k_gamma <- fonction(moy,sig) (moy/sig)*(moy/sig)
teta_gamma <- fonction(moy,sig) sig*sig/moy

##### pour calculer un tau de kendall #####
tauKendall <- fonction (x,y) {

  res=0
  n=length(x)
  for(i2 in 2:n)
  {
    for(j2 in 1:(i2-1))
    {
      if ( ((x[i2]>x[j2])&(y[i2]>y[j2])) || ((x[i2]<x[j2])&(y[i2]<y[j2])) ) {
        res = res +1
      }
      else
      {
        res = res-1
      }
    }
  }

  list(res=2*res/(n*(n-1)))
}

#pour tester
#nb_kendall=10000
#cor(SynthSimu$x[0:nb_kendall], SynthSimu$y[0:nb_kendall] , method = "kendall")
#nb_kendall=10000
#Kendall(SynthSimu$x[0:nb_kendall], SynthSimu$y[0:nb_kendall])$tau * 2

```

```
#####
```

```
#####
#####
#### Fonctions à base de MC
#####
#####
#####

#####
#### Simulation of a Multiplicative shock
#####

#### en paramètre les n valeurs du x0, y0 et du choc
#### retourne : x0, y0, x, Y, Fx_x, Fy_y
syntheseSimu <- fonction (x0,y0,x,y,choc) {

    print (mean(choc))

    Fx <- ecdf(x)
    Fy <- ecdf(y)
    Fx_x = Fx(x)
    Fy_y = Fy(y)

    list(x0=x0,y0=y0,choc=choc,x=x,y=y,Fx_x=Fx_x,Fy_y=Fy_y)
}

#### retourne la valeur de la loi jointe avec en paramètre synthSimu
fonctionJointe <- fonction(SynthSimu,x0,y0){
    x=SynthSimu$x
    y=SynthSimu$y
    nb1=0
    nb2=0
    for(i in 1:length(x)){
        if(x[i]<=x0) {
            nb1=nb1+1
            if(y[i]<=y0) {nb2=nb2+1}
        }
    }
    if(nb1==0) v=0 else v=nb2/length(x)
    list(valeur=v)
}

#### pour visualiser le choc
traceChoc<- fonction(SynthSimu, window) {
    if (window ==1) {windows()}
    if (type_loi==1)
    {

        plot(density(SynthSimu$choc,from=0.5,to=2,bw=0.005),col="red",main=paste("Shock"),xlab="beta",
ylab="Density")
    }
    if (type_loi>=2)
    {
        plot(density(SynthSimu$choc),col="red",main=paste("Shock"),xlab="beta",ylab="Density")
    }
}

#### pour visualiser les lois des marginales
traceLoiMarginales<- fonction(SynthSimu, window) {
    x =SynthSimu$x
    y =SynthSimu$y
    if (window ==1) {windows()}
    plot(density(x),col="blue",xlim=c(50,170),main=paste("Marginales"),xlab="x1(blue) x2(green)")
    lines(density(y),col="green")
}

#### pour visualiser les point des marginales
tracePoints<- fonction(SynthSimu, window, nb_pts){
    if (window ==1) {windows()}
    x =SynthSimu$x
```

```

y =SynthSimu$y
plot(x[1:nb_pts],y[1:nb_pts],main=paste("Dépendance"),xlab="x=Branche X1", ylab="y=Branche
X2")
}

### pour visualiser les Fpoint des marginales
traceFPoints<- fonction(SynthSimu, window, nb_pts){
  if (window ==1) {windows()}
  Fx_x =SynthSimu$Fx_x
  Fy_y =SynthSimu$Fy_y
  if (window ==1) {windows()}
  plot(Fx_x[1:nb_pts],Fy_y[1:nb_pts],main=paste("Dépendance"),xlim=c(0,1),ylim=c(0,1),xlab="F_X1(
x1)", ylab="F_X2(x2)")
}

#####
### Determine la copule à partir d'une simulation MC du choc
#####

### en paramètre : le SynthSimu calculé précédemment et le nombre de pas de x et de y
### retourne : la fonction de densité et de répartition (dans les z) pour les (x,y)
copuleEmpirique <- fonction (SynthSimu, nb_pts_x, nb_pts_y){
  cop <- troisd(SynthSimu$Fx_x,SynthSimu$Fy_y,nb_pts_x,nb_pts_y)
  list(cDensite_x=cop$x,cDensite_y=cop$y,cDensite_z=cop$z,cRepartition_x=cop$x,cRepartition_y=cop
$y,cRepartition_z=cop$z,valeurs=cop$z2, SynthSimu=SynthSimu)
}

### pour visualiser la fonction de densité d'une copule empirique
traceCopuleDensite<- fonction(cEmpirique, window){
  if (window ==1) {windows()}
  persp(cEmpirique$cDensite_x, cEmpirique$cDensite_y, cEmpirique$cDensite_z*nb_pts_x*nb_pts_y,
theta = -40, phi = 25, expand=.95, scale = TRUE, axes = TRUE, xlab="F_X1(x1)", ylab="F_X2(x2)", zlab="",
col = "lightblue", shade = 0.75, ticktype = "detailed", main="Densité de la Copule Empirique")
}

### pour visualiser la fonction de densité d'une copule empirique
traceCopuleDensiteCouleur<- fonction(cEmpirique, window, jauge){
  if (window ==1) {windows()}
  image(cEmpirique$cDensite_x, cEmpirique$cDensite_y, -log(cEmpirique$cDensite_z), col =
rainbow(40,start=0/6,end=4/6,s=0.9,v=0.9),axes = TRUE, xlab="F_X1(x1)", ylab="F_X2(x2)", main="Densité
de la Copule Empirique")
  if (window ==1) {windows()}
  if (jauge==1) {jauge_arcenciel <- jauge_rainbow(40)}
}

### pour visualiser la fonction de répartition d'une copule empirique
traceCopuleRepartition<- fonction(cEmpirique, window){
  if (window ==1) {windows()}
  persp(cEmpirique$cRepartition_x, cEmpirique$cRepartition_y, cEmpirique$cRepartition_z, theta = -40,
phi = 25, expand=.95, scale = TRUE, axes = TRUE, xlab="F_X1(x1)", ylab="F_X2(x2)", zlab="", col =
"lightblue", shade = 0.75, ticktype = "detailed", main="Densité de la Copule Empirique")
}

#####
### Estime la dépendance de queue MC
#####

### en paramètre : le SynthSimu calculé précédemment et le niveau z
### retourne : valeur
depQueueEmp <- fonction(SynthSimu,z){
  Fx_x=SynthSimu$Fx_x
  Fy_y=SynthSimu$Fy_y
  nb1=0
  nb2=0
  for(i in 1:length(Fx_x)){
    if(Fx_x[i]>=z) {
      nb1=nb1+1
      if(Fy_y[i]>=z) {nb2=nb2+1}
    }
  }
}

```

```

        if(nb1==0) v=0 else v=nb2/nb1
        list(valeur=v)
    }

### en paramètre : le SynthSimu calculé précédemment et le niveau z
### retourne : valeur
depQueueBasseEmp <- fonction(SynthSimu,z){
    Fx_x=SynthSimu$Fx_x
    Fy_y=SynthSimu$Fy_y
    nb1=0
    nb2=0
    for(i in 1:length(Fx_x)){
        if(Fx_x[i]<=z) {
            nb1=nb1+1
            if(Fy_y[i]<=z) {nb2=nb2+1}
        }
    }
    if(nb1==0) v=0 else v=nb2/nb1
    list(valeur=v)
}

```

```

#####
#####
#####
##### Fonctions à base de calcul théoriques ou numériques
#####
#####
#####
#####

```

```

#####
##### Determine la copule à partir d'une simulation MC du choc
#####
#####

```

```

#####
### Copule théorique pour choc multiplicatif avec des lognormales partout
### en paramètre implicites : mu_ et sigma_ pour (1,2,c)
### retourne : la fonction de répartition en (u,v)
CopuleTheorique_1 <- fonction (u,v){
    F_x0 <- fonction(x) plnorm(x,mu_1,sigma_1)
    F_y0 <- fonction(x) plnorm(x,mu_2,sigma_2)
    F_x_inv <- fonction(x) qlnorm(x,mu_1+mu_c,sqrt(sigma_1^2+sigma_c^2))
    F_y_inv <- fonction(x) qlnorm(x,mu_2+mu_c,sqrt(sigma_2^2+sigma_c^2))
    f_beta <- fonction(x) dlnorm(x,mu_c,sigma_c)
    f <- fonction(beta) F_x0(F_x_inv(u)/beta)*F_y0(F_y_inv(v)/beta)*f_beta(beta)
    #x=integrate(f,lower=0,upper=moyc+4*sigma_c)
    x=integrate(f,lower=0,upper=qlnorm(.99999,mu_c,sigma_c)+sigma_c)
    list(valeur=x$value)
}

```

```

#####
##### Determine la copule à partir d'une simulation MC du choc
#####
#####

```

```

### en paramètre : le CopuleThéorique calculée précédemment et le nombre de pas de x et de y
### retourne : la fonction de répartition (sous forme de tableau) pour les (x,y)
copuleTheorique <- fonction (cTheorique, nb_pts_x, nb_pts_y){
    z0 <- array(0,c(10,10))
    for(i in 1:10){
        for(j in 1:10){
            test <- cTheorique(i/10,j/10)
            z0[i,j]= test$value
        }
    }
    list(valeurs=z0)
}

```

```
#####
### Estime la dépendance de queue à partir de la formule avec les copules
#####

depQueueTeh <-function(CopuleTheo,z){
  cTheo <- CopuleTheo (z, z)
  cc=cTheo$valeur
  list(valeur=(1-z-z+cc)/(1-z))
}

#####
#####
#####
### Fonctions pour fitter les copules
#####
#####
#####

#####
### Renvoie une matrice nb_pts_x*nb_pts_y de l'écart entre une copule théorique et la copule générée
#####

ecartCopule <- function(CopuleTheo,SynthSimu0, param, nb_pts_x, nb_pts_y){
  cEmpirique <- copuleEmpirique (SynthSimu0, nb_pts_x, nb_pts_y)

  m0 <- matrix(nrow=nb_pts_x,ncol=nb_pts_y)
  for (i in 1:nb_pts_x){
    for (j in 1:nb_pts_y){
      tt <- CopuleTheo(i/nb_pts_x,j/nb_pts_y,param)
      m0[i,j]=tt$r
    }
  }
  m2 <- matrix(nrow=nb_pts_x,ncol=nb_pts_y)
  m2[1,1]=m0[1,1]
  for (i in 2:nb_pts_x){m2[i,1]=m0[i,1]-m0[i-1,1]}
  for (j in 2:nb_pts_x){m2[1,j]=m0[1,j]-m0[1,j-1]}
  for (i in 2:nb_pts_x){
    for (j in 2:nb_pts_y){
      m2[i,j]=m0[i,j]-m0[i,j-1]-m0[i-1,j]+m0[i-1,j-1]
    }
  }

  list(copuleGeneree = cEmpirique$cDensite_z, copuleTheorique = m2, ecart = (cEmpirique$cDensite_z-
m2)/m2, ecart2 = (cEmpirique$cDensite_z-m2))
}

#ecartCopule(copuleGaussienneRepartition,SynthSimu,9,10,10)$copuleTheorique

#####
### Fais un test de chi2 entre la copule générée et la copule théorique en divisant le  $[0,1]^2$  en nb_x*nb_y
#####
testChi2 <- fonction(CopuleTheo,SynthSimu0, rho, nb_pts_x, nb_pts_y,nb_pts_simules){
  cEmpirique <- copuleEmpirique (SynthSimu0, nb_pts_x, nb_pts_y)

  m0 <- matrix(nrow=nb_pts_x,ncol=nb_pts_y)
  for (i in 1:nb_pts_x){
    for (j in 1:nb_pts_y){
      tt <- CopuleTheo(i/nb_pts_x,j/nb_pts_y,rho)
      m0[i,j]=tt$r
    }
  }
  m2 <- matrix(nrow=nb_pts_x,ncol=nb_pts_y)
  m2[1,1]=m0[1,1]
  for (i in 2:nb_pts_x){m2[i,1]=m0[i,1]-m0[i-1,1]}
  for (j in 2:nb_pts_x){m2[1,j]=m0[1,j]-m0[1,j-1]}
  for (i in 2:nb_pts_x){
    for (j in 2:nb_pts_y){
      m2[i,j]=m0[i,j]-m0[i,j-1]-m0[i-1,j]+m0[i-1,j-1]
    }
  }
}
```

```

}

e1 = as.integer(cEmpirique$cDensite_z*nb_pts_simules)
e2 = as.integer(m2*nb_pts_simules)
e = (e1-e2)*(e1-e2)/e2

df = nb_pts_x*nb_pts_y-1
#print (sum(e))
#print (paste("p-value : ", 1-pchisq(sum(e),df)))
print(chisq.test(e1, p = e2, rescale.p = TRUE))
print ((cEmpirique$cDensite_z-m2)/m2)
print (e1)

list(copuleGeneree = cEmpirique$cDensite_z, copuleTheorique = m2, ecart = (cEmpirique$cDensite_z-
m2)/m2,p_value =1-pchisq(sum(e),df) )
}
#testChi2(copuleGaussienneRepartition,SynthSimu,.7076,10,10)$copuleGeneree
#testChi2(copuleGaussienneRepartition,SynthSimu,.7076,10,10)$copuleTheorique

#####
#### Fonction de répartition d'une copule gaussienne avec en paramètre le rho de la matrice de corr
#####

copuleGaussienneRepartition <- fonction (u,v,rho){
  m <- 2
  sigma <- diag(2)
  sigma[2,1] <- rho
  sigma[1,2] <- rho
  r=pmvnorm(lower=rep(-Inf, m), upper=c(qnorm(u),qnorm(v)), mean=rep(0, m), corr=sigma)
  list(r=r)
}

#####
#### Fonction de répartition d'une copule de student avec en paramètre le rho de la matrice de corr
#####

copuleStudentRepartition <- fonction (u,v,rho){
  m <- 2
  sigma <- diag(2)
  sigma[2,1] <- rho
  sigma[1,2] <- rho

  r=pmvt(lower=rep(-Inf, m), upper=c(qt(u),qt(v)), delta=rep(0, m), corr=sigma)
  list(r=r)
}

m <-copuleStudentRepartition(.8,.5,.57)

#rho=.57
#m <- 2
#sigma <- diag(2)
#sigma[2,1] <- rho
#sigma[1,2] <- rho
#u=.8
#v=.5
#s=pmvt(lower=rep(-Inf, m), upper=c(qt(u),qt(v)), delta=rep(0, m), corr=sigma)
#s

#m0 <- matrix(nrow=10,10)
#for (i in 1:10){
#for (j in 1:10){
#  m0[i,j]=copuleStudentRepartition(i/10,j/10,.5)$r
#}
#}

#####
#### Fonction de répartition d'une copule indépendante
#####

```

```

copuleIndepRepartition <- fonction (u,v,rho){
  list(r=u*v)
}

#####
#### Fonction de répartition d'une copule de Gumber avec en paramètre teta
#####

copuleGumbelRepartition <- fonction (u,v,teta){
  cop = exp(- ( -log(u))^teta + (-log(v))^teta )^(1/teta)
  list(r=cop)
}

#####
#### Fonction de répartition d'une copule de Clayton avec en paramètre le teta
#####

copuleClaytonRepartition <- fonction (u,v,teta){
  cop = ( u^(-teta) + v^(-teta) - 1 )^(-1/teta)
  list(r=cop)
}

#####
#### Fonction de répartition d'une copule de Frank avec en paramètre le teta
#####

copuleFrankRepartition <- fonction (u,v,teta){
  cop = (-1/teta) * log(1 - ( (1-exp(-teta*u))*(1-exp(-teta*v)) )/(1-exp(-teta) ) )
  list(r=cop)
}

#####
#### Fonction de répartition d'une copule HRT avec en paramètre le teta
#####

copuleHRTRepartition <- fonction (u,v,teta){
  cop = u+v-1 + ( (1-u)^(-teta) + (1-v)^(-teta) - 1 )^(-1/teta)
  list(r=cop)
}

#####
#### Fonction de répartition d'une copule d'indépendance (le paramètre teta n'est pas utilisé)
#####

copuleIndependanceRepartition <- fonction (u,v,teta){
  cop = u*v
  list(r=cop)
}

#####
#### Dessiner la fonction de densité de la copule, en divisant le [0,1]² en nb_x*nb_y
#####
dessineCopule <- fonction(CopuleTheo,nb_pts_x, nb_pts_y,parametre,titre, repartition)
{
  rho=parametre

  m0 <- matrix(nrow=nb_pts_x,ncol=nb_pts_y)
  for (i in 1:nb_pts_x){
    for (j in 1:nb_pts_y){
      tt <- CopuleTheo(i/nb_pts_x,j/nb_pts_y,rho)
      m0[i,j]=tt*r
    }
  }

  m2 <- matrix(nrow=nb_pts_x,ncol=nb_pts_y)
  if (repartition==1)
  {
    m2=m0
  }
}

```

```

else
{
  m2[1,1]=m0[1,1]*(nb_pts_x*nb_pts_y)
  for (i in 2:nb_pts_x){m2[i,1]=(m0[i,1]-m0[i-1,1])*(nb_pts_x*nb_pts_y)}
  for (j in 2:nb_pts_x){m2[1,j]=(m0[1,j]-m0[1,j-1])*(nb_pts_x*nb_pts_y)}
  for (i in 2:nb_pts_x){
    for (j in 2:nb_pts_y){
      m2[i,j]=(m0[i,j]-m0[i,j-1]-m0[i-1,j]+m0[i-1,j-1])*(nb_pts_x*nb_pts_y)
    }
  }
}

x=seq(1,nb_pts_x,1)/nb_pts_x
y=seq(1,nb_pts_y,1)/nb_pts_y
persp(x, y, m2, theta = -40, phi = 25, expand=.95, scale = TRUE, axes = TRUE, xlab="F_X1(x1)",
ylab="F_X2(x2)", zlab="", col = "lightblue", shade = 0.75, ticktype = "detailed", main=titre)
}

#a<-dessineCopule(copuleIndependanceRepartition,25,25,.2,"Copule Gaussienne",0)

#####
#####
#####
#### Fonctions pour fitter les modèles de choc commun
#####
#####
#####

densiteJointe_Mult_Lognormale <- fonction (xx,yy){
  f_x0 <- fonction(x) dlnorm(x,mu_1,sigma_1)
  f_y0 <- fonction(x) dlnorm(x,mu_2,sigma_2)
  f_beta <- fonction(x) dlnorm(x,mu_c,sigma_c)
  f <- fonction(beta) f_x0(xx/(beta+min_c))*f_y0(yy/(beta+min_c))*f_beta(beta)/((beta+min_c)^2)
  x=integrate(f,lower=0,upper=qlnorm(.99999,mu_c,sigma_c)+sigma_c)
  list(v=x$value)
}

densiteJointe_Add_Gamma <- fonction (xx,yy){
  f_x0 <- fonction(x) dgamma(x,shape=k_1, scale = teta_1)
  f_y0 <- fonction(x) dgamma(x,shape=k_2, scale = teta_2)
  f_beta <- fonction(x) dgamma(x,shape=k_c, scale = teta_c)
  f <- fonction(beta) f_x0(xx-beta)*f_y0(yy-beta*etype_m2/etype_m1)*f_beta(beta)
  x=integrate(f,lower=0,upper=qgamma(.9999,shape=k_c, scale = teta_c))
  list(v=x$value)
}

#####
#####
#####
#### Zone de lancement
#####
#####
#####

#####
#### Zone de lancement // cor =.2
#####
#####

#####
#### choix du type de simulation
# -1: on lit des résultats d'un fichier
# 0 : une seule simulation
# 1 : on fait varier les paramètres du choc min
simu=0

```

onfaitlefit=0 #0: rien ; 1:fit sur les copues; 2 :fit sur les modèles de choc commun
 type_loi = 10 #1 :log-normal, 2:gamma mult, 3: choc sur paramètre, 4:gamma simple, 5:gamma large a cor
 fixée

```

if (simu<1){param1 <-array(0,c(1))}
if (simu==1){param1 =seq(0.02,0.2,.02)}
if (simu==1){param1 =c(seq(.1,.9,.1),seq(1,10,1),seq(15,35,5),seq(40,100,10))}

if (simu==1 & type_loi==5){param1 =c(seq(.25,2,.25),seq(3,10,1),seq(11,19,1))}
#if (simu==1 & type_loi==5){param1 =c(seq(.5,2,.5),seq(3,10,1),seq(11,19,2))}

##### pour le choc gamma additif
###cor=.1
if (simu==1 & type_loi==5){param1 =c(seq(.125,2,.125),seq(2,5,.5),seq(6,18,2),seq(20,155,5))}
###cor=.2
if (simu==1 & type_loi==5){param1 =c(seq(.25,2,.25),seq(3,10,1),seq(11,19,1),seq(20,77,3))}
###cor=.3
if (simu==1 & type_loi==5){param1 =c(seq(.25,2.5,.25),seq(3,10,1),seq(11,19,2),seq(20,50,3))}
###cor=.4
if (simu==1 & type_loi==5){param1 =c(seq(.25,2.5,.25),seq(3,10,1),seq(11,19,2),seq(20,38,3))}
###cor=.5
if (simu==1 & type_loi==5){param1 =c(seq(.25,2.5,.25),seq(3,10,1),seq(11,19,2),seq(20,32,3))}
###cas d'étude = 2
if (simu==1 & type_loi==5){param1 =c(seq(.25,1.75,.25),seq(2,7.5,.5))}
###cas d'étude = 3
if (simu==1 & type_loi==5){param1 =c(seq(.125,1.75,.125),seq(2,13,.5))}

##### pour le choc lognormal shifté
###cas d'étude = 3
if (simu==1 & type_loi==10){param1 =c(0,.2,.5,.7,.8,.85,.9,.95,.98)}

###cor=.2
if (simu==1 & type_loi==5){param1 =c(seq(.25,2,.25),seq(3,10,1),seq(11,19,1),seq(20,77,3))}

nb_kendall = 1000 #nbre de points sur lesquels sont calculés le tau de kendall

corr_theo <- array(0,c(length(param1)))
corr_simu <- array(0,c(length(param1)))
corr_kendall <- array(0,c(length(param1)))
corr_spearman <- array(0,c(length(param1)))
dep_queue_0_9_theo <- array(0,c(length(param1)))
dep_queue_0_995_theo <- array(0,c(length(param1)))
dep_queue_0_9_simu <- array(0,c(length(param1)))
dep_queue_0_995_simu <- array(0,c(length(param1)))
dep_queue_0_1_simu <- array(0,c(length(param1)))
dep_queue_0_05_simu <- array(0,c(length(param1)))
ecart_copule_ind_max <- array(0,c(length(param1)))
ecart_copule_ind_moyen <- array(0,c(length(param1)))
distance_copule_ind <- array(0,c(length(param1)))
ecart_copule_gaus_max <- array(0,c(length(param1)))
ecart_copule_gaus_moyen <- array(0,c(length(param1)))
distance_copule_gaus <- array(0,c(length(param1)))
ecart_copule_gumbel_max <- array(0,c(length(param1)))
ecart_copule_gumbel_moyen <- array(0,c(length(param1)))
distance_copule_gumbel <- array(0,c(length(param1)))
ecart_copule_clayton_max <- array(0,c(length(param1)))
ecart_copule_clayton_moyen <- array(0,c(length(param1)))
distance_copule_clayton <- array(0,c(length(param1)))
ecart_copule_hrt_max <- array(0,c(length(param1)))
ecart_copule_hrt_moyen <- array(0,c(length(param1)))
distance_copule_hrt <- array(0,c(length(param1)))
k1 <- array(0,c(length(param1)))
k2 <- array(0,c(length(param1)))
kc <- array(0,c(length(param1)))
teta1 <- array(0,c(length(param1)))
teta2 <- array(0,c(length(param1)))
tetac <- array(0,c(length(param1)))

nomFichierInput = "c:/mes documents/01-Mémoire/données (application numérique)/v3 - charges/03 -
Data_par_mois_avec_mm.csv"

```

```

nomFichierOutputPearson = "c:/mes documents/01-Mémoire/données (application numérique)/v3 -
charges/cor_pearson.csv"
nomFichierOutputKendall = "c:/mes documents/01-Mémoire/données (application numérique)/v3 -
charges/cor_kendall.csv"
nomFichierOutputSpearman = "c:/mes documents/01-Mémoire/données (application numérique)/v3 -
charges/cor_spearman.csv"
nomFichierOutputSimus = "c:/mes documents/01-Mémoire/données (etude des chos)/res_simus_30_bis.csv"

```

```

no_lob1=12
no_lob2=4
no_lob3=20
no_lob4=15

```

```
for(i in 1:length(param1)){
```

```

#####
### paramètres de MC
nb_pts = 4000000

nb_pts_x = 20 ## nbre_pas pour la copule empirique calculée à partir de MC
nb_pts_y = 20

fenetre = 1 ## si =1, chaque graphe a sa fenetre, si=0, tous les graphes sont dans la même fenetre

existeCopuleTheorique = 0
nomCopuleTheorique = CopuleTheorique_1

if (simu>-1)
{
##### on fait de la simulation de monte carlo #####
### détermination de paramètres intermédiaires
if (type_loi == 1) # Pour un choc log-normal
{
moy_10 = 100
etype_10 = 10
moy_20 = 100
etype_20 = 15

etype_c = .15
min_c = 0
moy_c= 1
#if(simu==1) {min_c = param1[i]}
#if(simu==1) {etype_c = param1[i]}

mu_1=mu_lognormal(moy_10,etype_10)
sigma_1=sig_lognormal(moy_10,etype_10)
mu_2=mu_lognormal(moy_20,etype_20)
sigma_2=sig_lognormal(moy_20,etype_20)
mu_c=mu_lognormal(moy_c-min_c,etype_c)
sigma_c=sig_lognormal(moy_c-min_c,etype_c)

x0 = rlnorm(nb_pts,mu_1,sigma_1)
y0 = rlnorm(nb_pts,mu_2,sigma_2)
choc = rlnorm(nb_pts,mu_c,sigma_c)+min_c

x=x0*(choc)
y=y0*(choc)
}
if (type_loi == 10) # Pour un choc log-normal, vision calibrage
{
moy_1 = 100
etype_1 = 10
moy_2 = 100
etype_2 = 12

moy_c= 1
cor =.3
min_c = 0.9

```

```

if(simu==1) {min_c = param1[i]}

etype_c = sqrt(cor*(etype_1/moy_1)*(etype_2/moy_2))*moy_c

moy_10=moy_1/moy_c
moy_20=moy_2/moy_c
etype_10=sqrt((etype_1^2-etype_c^2*moy_10^2)/(etype_c^2+moy_c^2))
etype_20=sqrt((etype_2^2-etype_c^2*moy_20^2)/(etype_c^2+moy_c^2))

mu_1=mu_lognormal(moy_10,etype_10)
sigma_1=sig_lognormal(moy_10,etype_10)
mu_2=mu_lognormal(moy_20,etype_20)
sigma_2=sig_lognormal(moy_20,etype_20)
mu_c=mu_lognormal(moy_c-min_c,etype_c)
sigma_c=sig_lognormal(moy_c-min_c,etype_c)

x0 = rlnorm(nb_pts,mu_1,sigma_1)
y0 = rlnorm(nb_pts,mu_2,sigma_2)
choc = rlnorm(nb_pts,mu_c,sigma_c)+min_c

x=x0/(choc)
y=y0/(choc)
}
if (type_loi == 11) # Pour un choc log-normal, vision calibrage avec 4 branches
{
  moy_1 = 100
  etype_1 = 10
  moy_2 = 100
  etype_2 = 24
  moy_3 = 100
  etype_3 = 20
  moy_4 = 100
  etype_4 = 8

  moy_c= 1

  etype_c = .09
  min_c = 0

  #if(simu==1) {min_c = param1[i]}
  #if(simu==1) {etype_c = param1[i]}

  moy_10=moy_1/moy_c
  moy_20=moy_2/moy_c
  moy_30=moy_3/moy_c
  moy_40=moy_4/moy_c
  etype_10=sqrt((etype_1^2-etype_c^2*moy_10^2)/(etype_c^2+moy_c^2))
  etype_20=sqrt((etype_2^2-etype_c^2*moy_20^2)/(etype_c^2+moy_c^2))
  etype_30=sqrt((etype_3^2-etype_c^2*moy_30^2)/(etype_c^2+moy_c^2))
  etype_40=sqrt((etype_4^2-etype_c^2*moy_40^2)/(etype_c^2+moy_c^2))

  mu_1=mu_lognormal(moy_10,etype_10)
  sigma_1=sig_lognormal(moy_10,etype_10)
  mu_2=mu_lognormal(moy_20,etype_20)
  sigma_2=sig_lognormal(moy_20,etype_20)
  mu_3=mu_lognormal(moy_30,etype_30)
  sigma_3=sig_lognormal(moy_30,etype_30)
  mu_4=mu_lognormal(moy_40,etype_40)
  sigma_4=sig_lognormal(moy_40,etype_40)
  mu_c=mu_lognormal(moy_c-min_c,etype_c)
  sigma_c=sig_lognormal(moy_c-min_c,etype_c)

  x0 = rlnorm(nb_pts,mu_1,sigma_1)
  y0 = rlnorm(nb_pts,mu_2,sigma_2)
  z0 = rlnorm(nb_pts,mu_3,sigma_3)
  w0 = rlnorm(nb_pts,mu_4,sigma_4)
  choc = rlnorm(nb_pts,mu_c,sigma_c)+min_c

  x=x0*(choc)
  y=y0*(choc)
  z=z0*(choc)
}

```

```

w=w0*(choc)
x_calib=x
y_calib=y
z_calib=z
w_calib=w
}
if (type_loi == 2) # Pour un choc gamma multiplicatif
{
    moy_m1 = 100
    etype_m1 = 200
    moy_m2 = 100
    etype_m2 = 205

    etype_c = 1.51
    moy_c=1
    min_c = 0.8

    moy_1=moy_m1/moy_c
    moy_2=moy_m2/moy_c
    etype_1=sqrt(etype_m1^2-moy_1^2*etype_c^2)/moy_c
    etype_2=sqrt(etype_m2^2-moy_2^2*etype_c^2)/moy_c

    k_1=k_gamma(moy_1,etype_1)
    teta_1=teta_gamma(moy_1,etype_1)
    k_2=k_gamma(moy_2,etype_2)
    teta_2=teta_gamma(moy_2,etype_2)
    k_c=k_gamma(moy_c,etype_c)
    teta_c=teta_gamma(moy_c,etype_c)

    x0 = rgamma(nb_pts,shape=k_1, scale = teta_1)
    y0 = rgamma(nb_pts,shape=k_2, scale = teta_2)
    choc = rgamma(nb_pts,shape=k_c, scale = teta_c)+min_c

    x=x0*choc
    y=y0/choc
}

if (type_loi == 4) # Pour un choc gamma aditif simple
{
    moy_m1 = 100
    etype_m1 = 10*5
    moy_m2 = 100
    etype_m2 = 15*5

    k_a=k_gamma(moy_m1,etype_m1)
    teta_a=teta_gamma(moy_m1,etype_m1)
    k_b=k_gamma(moy_m2,etype_m2)
    teta_b=teta_gamma(moy_m2,etype_m2)

    k_max = min(k_a,k_b)
    k_c=k_max/7
    etype_c=k_c

    k_1=k_a-k_c
    k_2=k_b-k_c

    x0 = rgamma(nb_pts,shape=k_1, scale = 1)
    y0 = rgamma(nb_pts,shape=k_2, scale = 1)
    choc = rgamma(nb_pts,shape=k_c, scale = 1)

    x=teta_a*(x0+choc)
    y=teta_b*(y0+choc)
}

if (type_loi == 5) # Pour un choc gamma aditif plus complexe à cor fixée
{
    cas_etude=1
    if(cas_etude==1)
    {
        moy_m1 = 100
        etype_m1 = 20
        moy_m2 = 100
    }
}

```

```

        etype_m2 = 25
    }
    if(cas_etude==2)
    {
        moy_m1 = 100
        etype_m1 = 60
        moy_m2 = 100
        etype_m2 = 75
    }
    if(cas_etude==3)
    {
        moy_m1 = 100
        etype_m1 = 15
        moy_m2 = 100
        etype_m2 = 60
    }
    #on fixe la corrélation
    cor=.15
    k_c_max = (1/cor)*min((moy_m1/etype_m1)^2,(moy_m2/etype_m2)^2)
    #on fait varier le k
    k_c = 20
    if(simu==1) {k_c=param1[i]}
    teta_c=etype_m1*sqrt(cor/k_c)

    moy_c=k_c*teta_c
    etype_c=sqrt(k_c*teta_c^2)

    moy_1=moy_m1-moy_c
    moy_2=moy_m2-moy_c*(etype_m2/etype_m1)
    etype_1=sqrt(etype_m1^2-etype_c^2)
    etype_2=sqrt(etype_m2^2-(etype_c*etype_m2/etype_m1)^2)

    k_1=k_gamma(moy_1,etype_1)
    teta_1=teta_gamma(moy_1,etype_1)
    k_2=k_gamma(moy_2,etype_2)
    teta_2=teta_gamma(moy_2,etype_2)

    x0 = rgamma(nb_pts,shape=k_1, scale = teta_1)
    y0 = rgamma(nb_pts,shape=k_2, scale = teta_2)
    choc = rgamma(nb_pts,shape=k_c, scale = teta_c)

    x=x0+choc
    y=y0+etype_m2/etype_m1*choc
}
aditif plus complexe
if ((type_loi==2)||((type_loi==4)||((type_loi==5) ||(type_loi==6))) # Pour un choc gamma
{
    k1[i]=k_1
    k2[i]=k_2
    kc[i]=k_c
    teta1[i]=teta_1
    teta2[i]=teta_2
    tetac[i]=teta_c
}

#####
### simulation de choc sur paramètre
if (type_loi==3){
    x0 = rlnorm(nb_pts,mu_1,sigma_1)
    y0 = rlnorm(nb_pts,mu_2,sigma_2)
    #choc = rnorm(nb_pts,moy_c-min_c,etype_c)+min_c
    choc = rlnorm(nb_pts,mu_c,sigma_c)+min_c
    x<- array(0,c(nb_pts))
    y<- array(0,c(nb_pts))
    for(i2 in 1:nb_pts)
    {
        mu_1=mu_lognormal(moy_1*choc[i2],etype_1)
        sigma_1=sig_lognormal(moy_1*choc[i2],etype_1)
        mu_2=mu_lognormal(moy_2*choc[i2],etype_2)
        sigma_2=sig_lognormal(moy_2*choc[i2],etype_2)
        x[i2] = rlnorm(1,mu_1,sigma_1)

```

```

        y[i2] = rlnorm(1,mu_2,sigma_2)
    }
}

### synthétise la simulation
SynthSimu <- syntheseSimu(x0,y0,x,y,choc)
}
else
{
##### on lit les résultats depuis un fichier #####
x2lu <- read.table(nomFichierInput, header = TRUE, sep = ";", row.names = 1)
#x2lu=log(x2lu)
dep <- calcul_cor(x2lu,"pearson")
write.table(dep$correl, file = nomFichierOutputPearson, sep = ";", dec = ",", col.names = NA,
qmethod = "double")
dep <- calcul_cor(x2lu,"kendall")
write.table(dep$correl, file = nomFichierOutputKendall, sep = ";", dec = ",", col.names = NA,
qmethod = "double")
dep <- calcul_cor(x2lu,"spearman")
write.table(dep$correl, file = nomFichierOutputSpearman, sep = ";", dec = ",", col.names =
NA, qmethod = "double")

SynthSimu
syntheseSimu(x2lu[,no_lob1:no_lob1],x2lu[,no_lob2:no_lob2],x2lu[,no_lob1:no_lob1],x2lu[,no_lob2:no_lob2],
x2lu[,no_lob2:no_lob2])
}

### trace les graphes
if(simu>-1)
{
par(mfrow=c(2,4))
graphe <- traceChoc (SynthSimu,fenetre)
graphe <- traceLoiMarginales(SynthSimu,fenetre)
graphe <- tracePoints(SynthSimu,fenetre,200)
graphe <- traceFPoints(SynthSimu,fenetre,200)
}
else
{
par(mfrow=c(2,4))
graphe <- traceChoc (SynthSimu,fenetre)
graphe <- traceLoiMarginales(SynthSimu,fenetre)
graphe <- tracePoints(SynthSimu,fenetre,200)
graphe <- traceFPoints(SynthSimu,fenetre,200)
}

#####
### calcule la copule empirique à partir de MC
cEmpirique <- copuleEmpirique (SynthSimu, nb_pts_x, nb_pts_y)
print (cEmpirique$valeurs)

graphe <- traceCopuleDensite(cEmpirique,fenetre)
graphe <- traceCopuleDensiteCouleur(cEmpirique,fenetre,1)
graphe <- traceCopuleRepartition(cEmpirique,fenetre)

#####
### calcule la valeur théorique
if(existeCopuleTheorique==1){
cTheorique <- copuleTheorique (nomCopuleTheorique, nb_pts_x, nb_pts_y)
print (cTheorique$valeurs)

### sort le tableau des écarts entre théorique et numérique + écart_max + moy_écarts
(cEmpirique$valeurs-cTheorique$valeurs)/cTheorique$valeurs
paste("Ecart max entre copule théo et num (en %): ", 100*max(abs((cEmpirique$valeurs-
cTheorique$valeurs)/cTheorique$valeurs)), "%")
paste("Ecart moyen entre copule théo et num (en %): ", 100*mean(abs((cEmpirique$valeurs-
cTheorique$valeurs)/cTheorique$valeurs)), "%")
}

#####

```

```

### calcule corrélation théorique

b=etype_c^2
cor_t=(b*moy_1*moy_2)/ sqrt( (etype_1^2*(1+b) + b*moy_1^2) * ( etype_2^2*(1+b) +
b*moy_2^2))
print(paste("Corrélation de Pearson théorique : ",cor_t ))
corr_theo[i]=cor_t

### calcule corrélation à partir de MC
cor_mc = cor(SynthSimu$x, SynthSimu$y , method = "pearson")
print(paste("Corrélation de Pearson calculée avec MC : ",cor_mc))
corr_simu[i]=cor_mc
corr_kendall[i]=cor(SynthSimu$x[0:nb_kendall], SynthSimu$y[0:nb_kendall] , method = "kendall")
#corr_kendall[i]=Kendall(SynthSimu$x[0:nb_kendall], SynthSimu$y[0:nb_kendall])$tau
corr_spearman[i]=cor(SynthSimu$x, SynthSimu$y , method = "spearman")

#####
### calcule dépendance de queue MC
z=0.9
dep <- depQueueEmp(SynthSimu,z)
dep_queue_0_9_simu[i] = dep$valeur
z=0.995
dep <- depQueueEmp(SynthSimu,z)
dep_queue_0_995_simu[i] = dep$valeur
print(paste("Dep de queue 90% = ",dep_queue_0_9_simu[i]))
print(paste("Dep de queue 99.5% = ",dep_queue_0_995_simu[i]))
### idem mais pour lower tail
z=0.1
dep <- depQueueBasseEmp(SynthSimu,z)
dep_queue_0_1_simu[i] = dep$valeur
z=0.005
dep <- depQueueBasseEmp(SynthSimu,z)
dep_queue_0_05_simu[i] = dep$valeur
print(paste("Dep de queue 10% = ",dep_queue_0_1_simu[i]))
print(paste("Dep de queue 0.5% = ",dep_queue_0_05_simu[i]))

### calcule dependance de queue théo
z=0.9
if(existeCopuleTheorique==1){
  dep2 <- depQueueTeh(nomCopuleTheorique,z)
  dep2$valeur
  dep_queue_0_9_theo[i] = dep2$valeur
}
z=0.995
if(existeCopuleTheorique==1){
  dep2 <- depQueueTeh(nomCopuleTheorique,z)
  dep2$valeur
  dep_queue_0_995_theo[i] = dep2$valeur
}

if(onfaitlefit==1)
{
#####
#####
### regarde le fit avec une copule d'indépendance
ecar = ecartCopule(copuleIndépendanceRepartition,SynthSimu, cor_mc,10,10)
print("Ecart avec la copule d'indépendance")
print(paste("Ecart moyen :",mean(abs(ecar$ecart))))
print(paste("Ecart max :",max(abs(ecar$ecart))))
print(paste("Distance :",sum((ecar$ecart2*100)^2)))

ecart_copule_ind_max[i]=max(abs(ecar$ecart))
ecart_copule_ind_moyen[i]=mean(abs(ecar$ecart))
distance_copule_ind[i]=sum((ecar$ecart2*100)^2)

#####
#####

```

```

### regarde le fit avec une gaussienne
#print("Copule gaussienne")
ecar = ecartCopule(copuleGaussienneRepartition,SynthSimu, cor_mc,10,10)
#print(ecar$copuleTheorique)
print("Ecart avec la copule gaussienne")
#print(ecar$ecart)
print(paste("Ecart moyen :",mean(abs(ecar$ecart))))
print(paste("Ecart max :",max(abs(ecar$ecart))))
print(paste("Distance :",sum((ecar$ecart2*100)^2)))
print("")
#print(paste("p-value          du          test          du          Chi^2          :          ",
testChi2(copuleGaussienneRepartition,SynthSimu,cor_mc,4,4,nb_pts)$p_value))

ecart_copule_gaus_max[i]=max(abs(ecar$ecart))
ecart_copule_gaus_moyen[i]=mean(abs(ecar$ecart))
distance_copule_gaus[i]=sum((ecar$ecart2*100)^2)

#####
#####
### regarde le fit avec une Gumbel
tau = corr_kendall[i]
ecar = ecartCopule(copuleGumbelRepartition,SynthSimu, 1/(1-tau),10,10)
print("Ecart avec la copule Gumbel")
print(paste("Ecart moyen :",mean(abs(ecar$ecart))))
print(paste("Ecart max :",max(abs(ecar$ecart))))
print(paste("Distance :",sum((ecar$ecart2*100)^2)))
print("")

ecart_copule_gumbel_max[i]=max(abs(ecar$ecart))
ecart_copule_gumbel_moyen[i]=mean(abs(ecar$ecart))
distance_copule_gumbel[i]=sum((ecar$ecart2*100)^2)

#####
#####
### regarde le fit avec une Clayton
tau = corr_kendall[i]
ecar = ecartCopule(copuleClaytonRepartition,SynthSimu, (2*tau)/(1-tau),10,10)
print("Ecart avec la copule Clayton")
print(paste("Ecart moyen :",mean(abs(ecar$ecart))))
print(paste("Ecart max :",max(abs(ecar$ecart))))
print(paste("Distance :",sum((ecar$ecart2*100)^2)))
print("")

ecart_copule_clayton_max[i]=max(abs(ecar$ecart))
ecart_copule_clayton_moyen[i]=mean(abs(ecar$ecart))
distance_copule_clayton[i]=sum((ecar$ecart2*100)^2)

#####
#####
### regarde le fit avec une HRT
tau = corr_kendall[i]
ecar = ecartCopule(copuleHRTRepartition,SynthSimu, (2*tau)/(1-tau),10,10)
print("Ecart avec la copule HRT")
print(paste("Ecart moyen :",mean(abs(ecar$ecart))))
print(paste("Ecart max :",max(abs(ecar$ecart))))
print(paste("Distance :",sum((ecar$ecart2*100)^2)))
print("")

ecart_copule_hrt_max[i]=max(abs(ecar$ecart))
ecart_copule_hrt_moyen[i]=mean(abs(ecar$ecart))
distance_copule_hrt[i]=sum((ecar$ecart2*100)^2)

#### affiche les copule fittées (suivant le tau de kendall ou le coefficient de pearson
if(simu<1)
{
windows()
par(mfrow=c(1,1))
tau = corr_kendall[i]
a<-dessineCopule(copuleGaussienneRepartition,25,25,cor_mc,"Copule Gaussienne",0)

```

```

windows()
a<-dessineCopule(copuleGumbelRepartition,25,25,1/(1-tau),"Copule Gumbel",0)
windows()
a<-dessineCopule(copuleClaytonRepartition,25,25,(2*tau)/(1-tau),"Copule Clayton",0)
windows()
#a<-dessineCopule(copuleFrankRepartition,25,25,1)
a<-dessineCopule(copuleHRTRepartition,25,25,(2*tau)/(1-tau),"Copule HRT",0)
}

##### on fait le fit sur les modèles de choc commun
if(onfaitlefit==2)
{
}
print(i)
}

#####
### trace les courbes synthétisant les jeux de simulation

if(simu>0){
  plot(param1,corr_simu,main=paste("Corrélation",tau,"mesurée"),xlab="x=Min_choc",
  ylab="y=Corr",col="green",type="l",ylim=c(0,1))
  lines(corr_theo,col="blue",type="l")
  plot(param1,dep_queue_0_9_simu,main=paste("Dep_queue_0.9"),xlab="x=Min_choc",
  ylab="y=Dep_0.9",col="green",type="l",ylim=c(0,1))
  lines(param1,dep_queue_0_9_theo,col="blue")
  plot(param1,dep_queue_0_995_simu,main=paste("Dep_queue_0.995"),xlab="x=Min_choc",
  ylab="y=Dep_0.995",col="green",type="l",ylim=c(0,1))
  lines(param1,dep_queue_0_995_theo,col="blue")
  plot(param1,corr_kendall,main=paste("Corrélation de Kendall",tau,"mesurée"),xlab="x=Min_choc",
  ylab="y=Corr_Kendall",col="green",type="l",ylim=c(0,1))
  plot(param1,corr_spearman,main=paste("Corrélation de Spearman",tau,"mesurée"),xlab="x=Min_choc",
  ylab="y=Corr_Spearman",col="green",type="l",ylim=c(0,1))

  if ((type_loi == 1) || (type_loi == 10))
  {
    mdat <-matrix
(c(param1,corr_theo,corr_simu,corr_kendall,corr_spearman,ecart_copule_ind_max,ecart_copule_ind_moyen,di
stance_copule_ind,ecart_copule_gaus_max,ecart_copule_gaus_moyen,distance_copule_gaus,ecart_copule_gu
mbel_max,ecart_copule_gumbel_moyen,distance_copule_gumbel,ecart_copule_clayton_max,ecart_copule_clay
ton_moyen,distance_copule_clayton,ecart_copule_hrt_max,ecart_copule_hrt_moyen,distance_copule_hrt,dep_
queue_0_9_simu,dep_queue_0_995_simu,dep_queue_0_1_simu,dep_queue_0_05_simu),ncol=24)
    mdat
    colo=c("i","param1","corr_theo","corr_simu","tau_kendall","rho_spearman","cop_ind_max","cop_ind_
moy","d_cop_ind","cop_gaus_max","cop_gaus_moy","d_cop_gaus","cop_gumb_max","cop_gumb_moy","d_cop
_gumb","cop_clay_max","cop_clay_moy","d_cop_clay","cop_hrt_max","cop_hrt_moy","d_cop_hrt","dep_queue
_90%","dep_queue_99.5%","dep_queue_10%")
    colo
  }
  if ((type_loi == 2)|| (type_loi == 4)|| (type_loi == 5) || (type_loi==6)) #gamma
  {
    mdat <-matrix
(c(param1,corr_theo,corr_simu,corr_kendall,corr_spearman,ecart_copule_ind_max,ecart_copule_ind_moyen,di
stance_copule_ind,ecart_copule_gaus_max,ecart_copule_gaus_moyen,distance_copule_gaus,ecart_copule_gu
mbel_max,ecart_copule_gumbel_moyen,distance_copule_gumbel,ecart_copule_clayton_max,ecart_copule_clay
ton_moyen,distance_copule_clayton,ecart_copule_hrt_max,ecart_copule_hrt_moyen,distance_copule_hrt,dep_
queue_0_9_simu,dep_queue_0_995_simu,dep_queue_0_1_simu,dep_queue_0_05_simu,k1,teta1,k2,teta2,kc,t
etac),ncol=30)
    mdat
    colo=c("i","param1","corr_theo","corr_simu","tau_kendall","rho_spearman","cop_ind_max","cop_ind_
moy","d_cop_ind","cop_gaus_max","cop_gaus_moy","d_cop_gaus","cop_gumb_max","cop_gumb_moy","d_cop
_gumb","cop_clay_max","cop_clay_moy","d_cop_clay","cop_hrt_max","cop_hrt_moy","d_cop_hrt","dep_queue
_90%","dep_queue_99.5%","dep_queue_10%","dep_queue_.5%","k1","teta1","k2","teta2","kc")
    colo
  }
}

```

```
write.table(mdat, file = nomFichierOutputSimus, sep = ";", dec = ",", col.names = colo, qmethod = "double")
```

```
}
```

```
#a<-dessineCopule(copuleGaussienneRepartition,25,25,.2,"Copule Gaussienne",0)
#a<-dessineCopule(copuleGumbelRepartition,25,25,1/(1-tau),"Copule Gumbel",0)
#a<-dessineCopule(copuleClaytonRepartition,25,25,(2*tau)/(1-tau),"Copule Clayton",0)
#a<-dessineCopule(copuleHRTRepartition,25,25,(2*tau)/(1-tau),"Copule HRT",0)
```

```
#####
#####
##### pour chequer les fonctions de densité jointes :
##### d'abord on vérifie qu'on retrouve bien la fonction de répartition jointe
##### en comparant le résultant avec fonctionJointe(SynthSimu,110,120)$v (résulta issu de simulation
MC)
##### puis, on intègre la fonction de densité jointe (avec une méthode numérique approchée
##### pour vérifier qu'on retombe bien sur la fonction de répartition jointe
#####
#####
```

```
xx=120
yy=80
```

```
fonctionJointe(SynthSimu,xx,yy)$v
```

```
if (1==2)
{
```

```
  fonctionJointeTheorique_1 <- function (xx,yy){
    F_x0 <- function(x) plnorm(x,mu_1,sigma_1)
    F_y0 <- function(x) plnorm(x,mu_2,sigma_2)
    f_beta <- function(x) dlnorm(x,mu_c,sigma_c)
    f <- function(beta) F_x0(xx/(beta+min_c))*F_y0(yy/(beta+min_c))*f_beta(beta)
    x=integrate(f,lower=0,upper=qlnorm(.99999,mu_c,sigma_c)+sigma_c)
    list(v=x$value)
  }
  print(fonctionJointeTheorique_1(xx,yy)$v)
```

```
  densiteJointeTheorique_1 <- function (xx,yy){
    f_x0 <- function(x) dlnorm(x,mu_1,sigma_1)
    f_y0 <- function(x) dlnorm(x,mu_2,sigma_2)
    f_beta <- function(x) dlnorm(x,mu_c,sigma_c)
    f <- function(beta) <- function(beta)
    f_x0(xx/(beta+min_c))*f_y0(yy/(beta+min_c))*f_beta(beta)/((beta+min_c)^2)
    x=integrate(f,lower=0,upper=qlnorm(.99999,mu_c,sigma_c)+sigma_c)
    list(v=x$value)
  }
  densiteJointeTheorique_1(xx,yy)$v
```

```
}else{
```

```
  fonctionJointeTheorique_2 <- function (xx,yy){
    F_x0 <- function(x) pgamma(x,shape=k_1, scale = teta_1)
    F_y0 <- function(x) pgamma(x,shape=k_2, scale = teta_2)
    f_beta <- function(x) dgamma(x,shape=k_c, scale = teta_c)
    f <- function(beta) F_x0(xx-beta)*F_y0(yy-beta*etype_m2/etype_m1)*f_beta(beta)
    x=integrate(f,lower=0,upper=qgamma(.9999,shape=k_c, scale = teta_c))
    list(v=x$value)
  }
  print(fonctionJointeTheorique_2(xx,yy)$v)
```

```
  densiteJointeTheorique_2 <- function (xx,yy){
    f_x0 <- function(x) dgamma(x,shape=k_1, scale = teta_1)
    f_y0 <- function(x) dgamma(x,shape=k_2, scale = teta_2)
    f_beta <- function(x) dgamma(x,shape=k_c, scale = teta_c)
    f <- function(beta) f_x0(xx-beta)*f_y0(yy-beta*etype_m2/etype_m1)*f_beta(beta)
    x=integrate(f,lower=0,upper=qgamma(.9999,shape=k_c, scale = teta_c))
    list(v=x$value)
  }
```



```

logVmin=9e999

if(nbre_branches<=3){w=x} #on remplit artificiellement w et z pour éviter que ça plante après
if(nbre_branches<=2){z=x}

moy_1=mean(x)
moy_2=mean(y)
moy_3=mean(z)
moy_4=mean(w)
etype_1=sd(x)
etype_2=sd(y)
etype_3=sd(z)
etype_4=sd(w)

#on boucle sur chaque valeur possible de la corrélation
for(i in 1:length(valeursCor)){
  cor=valeursCor[i]
  ok=1

  #vérifie si la cor ne dépasse pas la limite des modèles lognormaux
  if(nbre_branches==2){
    cv1=etype_1/moy_1
    cv2=etype_2/moy_2
    if( (cor>min(cv1/cv2,cv2/cv1)) && typeModele<=2) {ok=0}
    if(typeModele==3) {k_c_max =
(1/cor)*min((moy_m1/etype_m1)^2,(moy_m2/etype_m2)^2)}
  }

  if (ok==1){
    #on boucle sur chaque valeur possible du second paramètre
    for(j in 1:spectre_2parametre){

      #1/ d'abord, on détermine les paramètres du modèle en fonction de cor et du 2nd
paramètre
      if (typeModele<=2){#modèle log-normale
        moy_c= 1
        if (typeModele==1) {pom=+1} else {pom=-1}
        if(nbre_branches==200){
          etype_c = sqrt(cor*(etype_1/moy_1)*(etype_2/moy_2))*moy_c
        }else{
          etype_c = valeursCor[i]
        }
        min_c = (j-1)/spectre_2parametre

        moy_10=moy_1/moy_c
        moy_20=moy_2/moy_c
        if(nbre_branches>=3){moy_30=moy_3/moy_c}
        if(nbre_branches>=4){moy_40=moy_4/moy_c}
        etype_10=sqrt((etype_1^2-
etype_c^2*moy_10^2)/(etype_c^2+moy_c^2))
        etype_20=sqrt((etype_2^2-
etype_c^2*moy_20^2)/(etype_c^2+moy_c^2))
        if(nbre_branches>=3){etype_30=sqrt((etype_3^2-
etype_c^2*moy_30^2)/(etype_c^2+moy_c^2))}
        if(nbre_branches>=4){etype_40=sqrt((etype_4^2-
etype_c^2*moy_40^2)/(etype_c^2+moy_c^2))}

        mu_1=mu_lognormal(moy_10,etype_10)
        sigma_1=sig_lognormal(moy_10,etype_10)
        mu_2=mu_lognormal(moy_20,etype_20)
        sigma_2=sig_lognormal(moy_20,etype_20)
        if(nbre_branches>=3){
          mu_3=mu_lognormal(moy_30,etype_30)
          sigma_3=sig_lognormal(moy_30,etype_30)
        }
        if(nbre_branches>=4){
          mu_4=mu_lognormal(moy_40,etype_40)
          sigma_4=sig_lognormal(moy_40,etype_40)
        }
        mu_c=mu_lognormal(moy_c-min_c,etype_c)
        sigma_c=sig_lognormal(moy_c-min_c,etype_c)
      }
    }
  }
}

```

```

print(paste(etype_c, " ", min_c, " ", mu_c, " ", sigma_c))

densiteJointe_Mult_Lognormale <- function (xx,yy,zz,ww){
  f_x0 <- function(x) dlnorm(x,mu_1,sigma_1)
  f_y0 <- function(x) dlnorm(x,mu_2,sigma_2)
  if(nbre_branches>=3){f_z0 <- function(x)
dlnorm(x,mu_3,sigma_3)}
  if(nbre_branches>=4){f_w0 <- function(x)
dlnorm(x,mu_4,sigma_4)}
  f_beta <- function(x) dlnorm(x,mu_c,sigma_c)
  if(nbre_branches==2){
    if(pom==+1){
      f <- function(beta)
f_x0(xx/(beta+min_c))*f_y0(yy/(beta+min_c))*f_beta(beta)/((beta+min_c)^2)
    }else{
      f <- function(beta)
f_x0(xx*(beta+min_c))*f_y0(yy*(beta+min_c))*f_beta(beta)*((beta+min_c)^2)
    }
  }
  if(nbre_branches==3){
    if(pom==+1){
      f <- function(beta)
f_x0(xx/(beta+min_c))*f_y0(yy/(beta+min_c))*f_z0(zz/(beta+min_c))*f_beta(beta)/((beta+min_c)^3)
    }else{
      f <- function(beta)
f_x0(xx*(beta+min_c))*f_y0(yy*(beta+min_c))*f_z0(zz*(beta+min_c))*f_beta(beta)*((beta+min_c)^3)
    }
  }
  if(nbre_branches==4){
    if(pom==+1){
      f <- function(beta)
f_x0(xx/(beta+min_c))*f_y0(yy/(beta+min_c))*f_z0(zz/(beta+min_c))*f_w0(ww/(beta+min_c))*f_beta(beta)
/((beta+min_c)^4)
    }else{
      f <- function(beta)
f_x0(xx*(beta+min_c))*f_y0(yy*(beta+min_c))*f_z0(zz*(beta+min_c))*f_w0(ww*(beta+min_c))*f_beta(beta)
*((beta+min_c)^4)
    }
  }
}

x=integrate(f,lower=0,upper=qlnorm(.99999,mu_c,sigma_c)+sigma_c)
list(v=x$value)
}

if (typeModele==3){#modèle gamma

  k_c = (k_c_max^.9)^(j/spectre_2parametre)
  teta_c=etype_m1*sqrt(cor/k_c)

  moy_c=k_c*teta_c
  etype_c=sqrt(k_c*teta_c^2)

  moy_1=moy_m1-moy_c
  moy_2=moy_m2-moy_c*(etype_m2/etype_m1)
  etype_1=sqrt(etype_m1^2-etype_c^2)
  etype_2=sqrt(etype_m2^2-(etype_c*etype_m2/etype_m1)^2)

  k_1=k_gamma(moy_1,etype_1)
  teta_1=teta_gamma(moy_1,etype_1)
  k_2=k_gamma(moy_2,etype_2)
  teta_2=teta_gamma(moy_2,etype_2)

print(paste(k_1," ",k_2," ",k_c, "/", k_c_max, " ", teta_1, " ", teta_2, " ", teta_c))

densiteJointe_Add_Gamma <- function (xx,yy){
  f_x0 <- function(x) dgamma(x,shape=k_1, scale = teta_1)
  f_y0 <- function(x) dgamma(x,shape=k_2, scale = teta_2)
  f_beta <- function(x) dgamma(x,shape=k_c, scale = teta_c)
  f <- function(beta) f_x0(xx-beta)*f_y0(yy-
beta*etype_m2/etype_m1)*f_beta(beta)
}

```



```

        x3=rlnorm(nb_pts,mu3,sig3)
    }
    if(typeModele==1){
        if(nbre_branches==2){s=(x1+x2)*(x_c+meilleur2param)}           else
{s=(x1+x2+x3)*(x_c+meilleur2param)}
        corM=cor((x1)*(x_c+meilleur2param),(x2)*(x_c+meilleur2param))
    }else{
        if(nbre_branches==2){s=(x1+x2)/(x_c+meilleur2param)}           else
{s=(x1+x2+x3)/(x_c+meilleur2param)}
        corM=cor((x1)/(x_c+meilleur2param),(x2)/(x_c+meilleur2param))
    }
    var = quantile(s,probs=.995)- param1 - param3
    if(nbre_branches>=3){var=var-param5}
}
else
{
    var=0
    etype=0
}

list(var=var,etype=sd(s),corM=corM,resultats2=resultats2,x_res=valeursCor,y_res=(seq(1,spectre_2p
arametre,1)-
1)/spectre_2param,meilleurCor=meilleurCor,meilleur2param=meilleur2param,logVmin=logVmin,param11=
meilleur2param,param10=meilleurCor,param1=param1,param2=param2,param3=param3,param4=param4,pa
ram5=param5,param6=param6,param7=param7,param8=param8)

    #print (resultats)
    #print (meilleurCor)
    #print (meilleur2param)

    #print(valeursCor)
    #print((seq(1,spectre_2parametre,1)-1)/spectre_2parametre)
    #print(resultats2)

}

if(test==1){
    nb=1500
    ss
fitModeleChocCommun(2,x_calib[0:nb],y_calib[0:nb],z_calib[0:nb],w_calib[0:nb],1,c(seq(.035,.065,.01)),1) <-
}else{
    ss <- fitModeleChocCommun(3,x_calib,y_calib,z_calib,w_calib,1,c(seq(.025,(.052),.0001)),1)
}

windows()
par(mfrow=c(1,1))
image(ss$x_res, ss$y_res, (ss$resultats2), col = rainbow(40,start=0/6,end=4/6,s=0.9,v=0.9),axes = TRUE,
xlab="cor/etype_c", ylab="min_c", main="LogVraisemblance")
#print(ss$resultats2)
print(ss$meilleurCor)
print(ss$meilleur2param)
print(paste("Meilleure LogV: ",ss$logVmin))
print(paste("Param1",ss$param1))
print(paste("Param2",ss$param2))
print(paste("Param3",ss$param3))
print(paste("Param4",ss$param4))
print(paste("Param10",ss$param10))
print(paste("Param11",ss$param11))
print(paste("Param5",ss$param5))
print(paste("Param6",ss$param6))
print(paste("Var=",ss$var))
print(paste("Etype=",ss$etype))
print(paste("CorModèle=",ss$corM))

#####
#####

```

```
#####
#####
calculerLogVindendant <-function (nbre_branches,x,y,z,w,typeModele)
{
  moy_1=mean(x)
  moy_2=mean(y)
  moy_3=mean(z)
  moy_4=mean(w)
  etype_1=sd(x)
  etype_2=sd(y)
  etype_3=sd(z)
  etype_4=sd(w)

  if(typeModele==1){
    mu_1=mu_lognormal(moy_1,etype_1)
    sigma_1=sig_lognormal(moy_1,etype_1)
    mu_2=mu_lognormal(moy_2,etype_2)
    sigma_2=sig_lognormal(moy_2,etype_2)
    if(nbre_branches>=3){
      mu_3=mu_lognormal(moy_3,etype_3)
      sigma_3=sig_lognormal(moy_3,etype_3)
    }
    if(nbre_branches>=4){
      mu_4=mu_lognormal(moy_4,etype_4)
      sigma_4=sig_lognormal(moy_4,etype_4)
    }

    densiteJointe_Mult_Lognormale <- function (xx,yy,zz,ww){
      f_x0 <- function(x) dlnorm(x,mu_1,sigma_1)
      f_y0 <- function(x) dlnorm(x,mu_2,sigma_2)
      if(nbre_branches>=3){f_z0 <- function(x) dlnorm(x,mu_3,sigma_3)}
      if(nbre_branches>=4){f_w0 <- function(x) dlnorm(x,mu_4,sigma_4)}
      if(nbre_branches==2){ valeur= f_x0(xx)*f_y0(yy) }
      if(nbre_branches==3){ valeur= f_x0(xx)*f_y0(yy)*f_z0(zz) }
      if(nbre_branches==4){ valeur= f_x0(xx)*f_y0(yy)*f_z0(zz)*f_w0(ww) }
      list(v=valeur)
    }

    logV=0
    for(k in 1:length(x)){
      if (typeModele==1){logV=logV-log(densiteJointe_Mult_Lognormale(x[k],y[k],z[k],w[k])$v) }
      if (typeModele==2){logV=logV-log(densiteJointe_Add_Gamma(x[k],y[k])$v) }
    }
    list(logV=logV)
  }
}
#calculerLogVindendant(4,x_calib[0:nb],y_calib[0:nb],z_calib[0:nb],w_calib[0:nb],1)$logV
calculerLogVindendant(3,x_calib,y_calib,z_calib,w_calib,1)$logV

#####
#####
#####
#####

print(paste("x: moy=", mean(x_calib) , " e_type=" , sd(x_calib)))
print(paste("y: moy=", mean(y_calib) , " e_type=" , sd(y_calib)))
print(paste("z: moy=", mean(z_calib) , " e_type=" , sd(z_calib)))
print(paste("x+y: moy=", mean(y_calib+x_calib) , " e_type=" ,sd(y_calib+x_calib)))
print(paste("x+y+z: moy=", mean(y_calib+x_calib+z_calib) , " e_type=" ,sd(y_calib+x_calib+z_calib)))

print(paste("cor(x,y)=",cor(x_calib,y_calib)))

#####
#####
#####
##### calcul de VaR
VaR1branche <- function(moy,etype)
{
  mu=mu_lognormal(moy,etype)
```

```

sigma=sig_lognormal(moy,etype)
var = qlnorm(.995,mu,sigma)- moy
list(var=var)
}
print(paste("Var(x+y histo)=",VaR1branche(mean(y_calib+x_calib),sd(y_calib+x_calib))$var))
print(paste("Var(x+y+z
histo)=",VaR1branche(mean(y_calib+x_calib+z_calib),sd(y_calib+x_calib+z_calib))$var))

VaR2branchesIndep <- function(nbre_branches,nb_pts,moy1,etype1,moy2,etype2,moy3,etype3,moy4,etype4)
{
  mu1=mu_lognormal(moy1,etype1)
  sig1=sig_lognormal(moy1,etype1)
  mu2=mu_lognormal(moy2,etype2)
  sig2=sig_lognormal(moy2,etype2)
  x1=rlnorm(nb_pts,mu1,sig1)
  x2=rlnorm(nb_pts,mu2,sig2)
  s=x1+x2
  if(nbre_branches>=3){
    mu3=mu_lognormal(moy3,etype3)
    sig3=sig_lognormal(moy3,etype3)
    x3=rlnorm(nb_pts,mu3,sig3)
    s=s+x3
  }
  if(nbre_branches>=4){
    mu4=mu_lognormal(moy4,etype4)
    sig4=sig_lognormal(moy4,etype4)
    x4=rlnorm(nb_pts,mu4,sig4)
    s=s+x4
  }
  if(nbre_branches==2){var = quantile(x1+x2,probs=.995)- moy1 - moy2}
  if(nbre_branches==3){var = quantile(x1+x2+x3,probs=.995)- moy1 - moy2 - moy3}
  if(nbre_branches==4){var = quantile(x1+x2+x3+x4,probs=.995)- moy1 - moy2 - moy3 - moy4}
  list(var=var,etype=sd(s))
}
ss <-
VaR2branchesIndep(3,1000000,mean(x_calib),sd(x_calib),mean(y_calib),sd(y_calib),mean(z_calib),sd(z_calib)
,0,0)
print(paste("Var(x+y(+z) ind)=",ss$var," ecart-type=",ss$etype))

#####
#####
#####
##### cas de 3 branches avec 2 chocs

if(superTest==1){

TroisBranches2chocs <-
function(x,y,z,nb_pts,moy1,etype1,moy2,etype2,moy3,etype3,moyc1,etypec1,minc1,moyc2,etypec2,minc2)
{
  mu1=mu_lognormal(moy1,etype1)
  sig1=sig_lognormal(moy1,etype1)
  mu2=mu_lognormal(moy2,etype2)
  sig2=sig_lognormal(moy2,etype2)
  mu3=mu_lognormal(moy3,etype3)
  sig3=sig_lognormal(moy3,etype3)
  muc1=mu_lognormal(moyc1,etypec1)
  sigc1=sig_lognormal(moyc1,etypec1)
  muc2=mu_lognormal(moyc2,etypec2)
  sigc2=sig_lognormal(moyc2,etypec2)

  f_x0 <- function(x) dlnorm(x,mu1,sigma1)
  f_y0 <- function(x) dlnorm(x,mu2,sigma2)
  f_z0 <- function(x) dlnorm(x,mu3,sigma3)
  f_alpha <- function(x) dlnorm(x,muc1,sigmac1)
  f_beta <- function(x) dlnorm(x,muc2,sigmac2)

  ##fonction de la densité jointe pour des valeurs alpha et beta fixées
  pom=+1

```

```

    if(pom==+1){
      f <- function(xx,yy,zz,a,b)
      f_x0(xx/(a+minc1)/(b+minc2))*f_y0(yy/(a+minc1))*f_z0(zz/(b+minc2))*f_alpha(a)*f_beta(b)/((a+minc1)^2
    )/((b+minc2)^2)
    }else{
      f <- function(xx,yy,zz,a,b)
      f_x0(xx/(a+minc1)/(b+minc2))*f_y0(yy/(a+minc1))*f_z0(zz/(b+minc2))*f_alpha(a)*f_beta(b)/((a+minc1)^2
    )/((b+minc2)^2)
    }

    ##on calcule la logVraisemblance sur l'ensemble des points
    logV=0
    for(k in 1:length(x)){

      ### on calcule à la mano la double intégrale
      n=500
      max=5
      valeur = 0
      for(i in 1:n)
      {
        for(j in 1:n)
        {
          valeur=
          densiteJointeTheorique(x[k],y[k],z[k],max*i/n,yy*max*j/n)$v
          valeur
          +
        }
        valeur = valeur *xx*yy/ ((max/n)^2)
        logVpoint=valeur
        logV=logV+logVpoint
      }
      x1=rlnorm(nb_pts,mu1,sig1)
      x2=rlnorm(nb_pts,mu2,sig2)
      x3=rlnorm(nb_pts,mu3,sig3)
      list(logV=logV)
    }

    ss <-
    TroisBranches2chocs(x_calib,y_calib,z_calib,1000000,mean(x_calib),sd(x_calib),mean(y_calib),sd(y_calib),mea
n(z_calib),sd(z_calib),0,0,0,0)
    print(paste("LogV=",ss$logV))

  }

  moy1=100
  moy2=100
  etype1=10
  etype2=10
  min_c=0.95
  etypec=.1
  nb_pts=1000000

  mu1=mu_lognormal(moy1,etype1)
  sig1=sig_lognormal(moy1,etype1)
  mu2=mu_lognormal(moy2,etype2)
  sig2=sig_lognormal(moy2,etype2)
  muc=mu_lognormal(1-min_c,etypec)
  sigc=sig_lognormal(1-min_c,etypec)
  x1=rlnorm(nb_pts,mu1,sig1)
  x2=rlnorm(nb_pts,mu2,sig2)
  xc=rlnorm(nb_pts,muc,sigc)+min_c

  a=(x1*xc)
  b=(x2*xc)

  cor(a,b)
  sd(a+b)

  #cor(a, b , method = "spearman")

```

```
moy1=1
etype1=.5
    mu1=mu_lognormal(moy1,etype1)
    sig1=sig_lognormal(moy1,etype1)
x1=rlnorm(nb_pts,mu1,sig1)
mean(x1)
```