
ESTIMATION DES PARAMÈTRES D'UNE COPULE POUR UN MODÈLE DE RISQUE COLLECTIF

SOUS LA SUPERVISION DE
ÉTIENNE MARCEAU

RAPPORT DES TRAVAUX RÉALISÉS

PRÉPARÉ PAR
ALEXANDRE LEPAGE,
DIAMILATOU N'DIAYE,
AMEDEO ZITO

LE 06 JUIN 2019



UNIVERSITÉ
LAVAL

FACULTÉ DES SCIENCES ET DE GÉNIE
ÉCOLE D'ACTUARIAT
UNIVERSITÉ LAVAL
AUTOMNE 2018

Table des matières

1	Introduction	1
2	Notions préliminaires	1
2.1	Modèle collectif du risque	1
2.2	Fonction du maximum de vraisemblance	1
2.3	Copule archimédienne hiérarchique	2
3	Résultats	3
3.1	Copule de Clayton	3
3.2	Copule archimédienne hiérarchique	8
4	Conclusion	20

Liste des illustrations

1	Arbre hiérarchique à un niveau.	3
2	Comparaisons de la distribution marginale des données simulées avec les distributions théoriques pour le scénario 1.	4
3	Nuages de points du scénario 1	5
4	Comparaisons de la distribution marginale des données simulées avec les distributions théoriques pour le scénario 2.	6
5	Nuages de points du scénario 2	7
6	Temps de dérivation d'une copule de Clayton en fonction du nombre de dérivées partielles à effectuer.	8
7	Comparaisons de la distribution marginale des données simulées avec les distributions théoriques pour le scénario 3.	9
8	Nuages de points du scénario 3	10
9	Comparaisons de la distribution marginale des données simulées avec les distributions théoriques pour le scénario 4.	11
10	Nuages de points du scénario 4	12
11	Comparaisons de la distribution marginale des données simulées avec les distributions théoriques pour le scénario 5.	13
12	Nuages de points du scénario 5	14
13	Comparaisons de la distribution marginale des données simulées avec les distributions théoriques pour le scénario 6.	15
14	Nuages de points du scénario 6	16
15	Comparaisons de la distribution marginale des données simulées avec les distributions théoriques pour le scénario 7.	17
16	Nuages de points du scénario 7	18
17	Comparaisons de la distribution marginale des données simulées avec les distributions théoriques pour le scénario 8.	19
18	Nuages de points du scénario 8	20

Liste des tableaux

1	Paramètres initiaux pour la copule de Clayton avec $N \sim \text{Bin}(n, q)$, $X_i \sim X \sim \text{Exp}(\beta)$.	4
2	Sommaire des données simulées pour le scénario 1	4
3	Résultats du scénario 1	5
4	Sommaire des données simulées pour le scénario 2	6
5	Résultats du scénario 2	7
6	Sommaire des données simulées pour le scénario 3	9
7	Résultats du scénario 3	10
8	Sommaire des données simulées pour le scénario 4	11
9	Résultats du scénario 4	12
10	Sommaire des données simulées pour le scénario 5	13
11	Résultats du scénario 5	14
12	Sommaire des données simulées pour le scénario 6	15
13	Résultats du scénario 6	16
14	Sommaire des données simulées pour le scénario 7	17
15	Résultats du scénario 7	18
16	Sommaire des données simulées pour le scénario 8	19
17	Résultats du scénario 8	20

1 Introduction

Le présent projet consiste à estimer les paramètres d'un modèle collectif du risque incorporant une structure de dépendance entre la variable aléatoire de dénombrement (fréquence) et les variables aléatoires représentant les montants individuels de sinistre.

Tout d'abord, la section sur les notions préliminaires expose le modèle de risque avec dépendance, puis introduit la notion d'estimation de paramètres par la méthode du maximum de vraisemblance. Par la suite, on présente les copules archimédiennes hiérarchiques. Finalement, la dernière section expose les résultats d'estimations sur des données simulées.

2 Notions préliminaires

2.1 Modèle collectif du risque

Dans [Cossette et al., 2019b], on propose un modèle collectif du risque dont les composantes sont dépendantes entre elles. Ce modèle s'exprime comme suit.

Soient N , la v.a. du nombre de sinistres, tel que $N \in \mathbb{N}$, et la suite de v.a. $\{X_i, i \in \mathbb{N}_+\}$ représentant les montants de sinistres tel que $X_i \sim X \in \mathbb{R}_+$. On a

$$S = \sum_{i=1}^{\infty} X_i \times \mathbb{1}_{\{N \geq i\}}, \quad (1)$$

où $\{N, X_1, X_2, \dots, X_k\}$ est une suite de v.a. dépendantes. Pour chaque $k \in \mathbb{N}_+$, la loi multivariée de $(N, X_1, X_2, \dots, X_k)$ est définie par une copule archimédienne hiérarchique.

Maintenant, on cherche à estimer les paramètres afférents à un tel modèle. Pour y arriver, [Cossette et al., 2019a] propose une approche de calcul de la vraisemblance par décomposition hiérarchique et une autre méthode, dite plus classique, qui prend l'ensemble du modèle. Dans le cas présent, compte tenu que la variable dénombrante est discrète et que les variables de sinistres sont continues, il est plus simple d'utiliser la méthode de vraisemblance complète. Le présent rapport présente donc les résultats d'une telle approche.

2.2 Fonction du maximum de vraisemblance

Soit une copule de Clayton multivariée dénotée C et représentée par

$$C(u_0, \dots, u_k; \alpha) = (u_0^{-\alpha} + \dots + u_k^{-\alpha} - k)^{-\frac{1}{\alpha}}, \quad (2)$$

$k \in \mathbb{N}_+$. Avec le modèle posé dans la section 2.1, on peut trouver la fonction de densité de la copule C de la manière décrite comme suit.

Soient λ et β les paramètres des lois de fréquence et de sévérité respectivement, ainsi que α , le paramètre de la copule de Clayton. La densité conjointe des v.a. N et (X_1, \dots, X_N) est donnée par

$$\begin{aligned} f_{N, X_1, \dots, X_n}(n, x_1, \dots, x_n; \lambda, \beta, \alpha) &= \frac{\partial^n}{\partial x_1 \dots \partial x_n} C(F_N(n; \lambda), F_{X_1}(x_1; \beta), \dots, F_{X_n}(x_n; \beta); \alpha) \\ &\quad - \frac{\partial^n}{\partial x_1 \dots \partial x_n} C(F_N(n-1; \lambda), F_{X_1}(x_1; \beta), \dots, F_{X_n}(x_n; \beta); \alpha), \end{aligned} \quad (3)$$

$n \in \mathbb{N}_+$. On observe que f_{N, X_1, \dots, X_n} n'est pas une fonction de densité au sens strict du terme, car N est une v.a. discrète et X_1, \dots, X_n sont des v.a. continues.

Puisque la dérivation en chaîne nécessaire pour trouver (3) est excessivement longue à calculer, on peut utiliser R à l'aide de la fonction `Deriv`¹ du package du même nom pour y arriver. Cependant, il est

1. <https://cran.r-project.org/web/packages/Deriv/Deriv.pdf>

important de noter que, même avec cet outil, si la loi de fréquence admet des valeurs supérieures à 5, le temps de calcul peut grimper très rapidement.

Soient n_j , le nombre d'observations où $N = j$, n_{total} , le nombre d'observations total tel que $\sum_{j=1}^k n_j = n_{total}$, et $x_{i,j}$, le j -ième sinistre de la i -ième observation, la fonction de vraisemblance du modèle énoncé dans la section 2.1 est donné par

$$\begin{aligned} \mathcal{L}(\lambda, \beta, \alpha) &= (\Pr(N = 0; \lambda))^{n_0} \\ &\times \prod_{i=1}^{n_1} f_{N, X_1}(1, x_{i,1}; \lambda, \beta, \alpha) \\ &\times \prod_{i=1}^{n_2} f_{N, X_1, X_2}(2, x_{i,1}, x_{i,2}; \lambda, \beta, \alpha) \\ &\dots \\ &\times \prod_{i=1}^{n_k} f_{N, X_1, \dots, X_k}(k, x_{i,1}, \dots, x_{i,k}; \lambda, \beta, \alpha). \end{aligned}$$

Par la suite, afin d'estimer les paramètres, il faut avoir recours à une méthode d'optimisation numérique afin de maximiser le log de la fonction de vraisemblance.

On obtient

$$\begin{aligned} \ell(\lambda, \beta, \alpha) &= n_0 \times \ln(\Pr(N = 0; \lambda)) \\ &+ \sum_{i=1}^{n_1} \ln(f_{N, X_1}(1, x_{i,1}; \lambda, \beta, \alpha)) \\ &+ \sum_{i=1}^{n_2} \ln(f_{N, X_1, X_2}(2, x_{i,1}, x_{i,2}; \lambda, \beta, \alpha)) \\ &\dots \\ &+ \sum_{i=1}^{n_k} \ln(f_{N, X_1, \dots, X_k}(k, x_{i,1}, \dots, x_{i,k}; \lambda, \beta, \alpha)). \end{aligned} \tag{4}$$

Finalement, on calcule numériquement les paramètres en utilisant la fonction `R constrOptim`². Cette dernière cherche à minimiser une fonction passée en argument. La fonction qui sera donc insérée en argument sera le négatif de (4). Afin de s'assurer que le processus d'optimisation numérique soit le plus précis possible, il faut entrer en argument de cette fonction des valeurs initiales adéquates pour chacun des paramètres à estimer. Pour ce faire, la méthode du maximum de vraisemblance de la distribution marginale de chacune des variables est un choix judicieux. Dans le cas où cette méthode s'avère trop laborieuse, on peut compenser avec la méthode des moments ou des quantiles (voir [Klugman et al., 2012]).

2.3 Copule archimédienne hiérarchique

Comme les copules archimédiennes hiérarchiques offrent une bonne flexibilité dans la modélisation de la dépendance, la présente section se penchera sur ce sujet d'intérêt.

Les copules archimédiennes hiérarchiques avec des distributions multivariées composées sont décrites dans [Cossette et al., 2017]. La représentation graphique du modèle collectif du risque tel que décrit dans la section 2.1 et expliquée dans [Cossette et al., 2019b] est présentée dans l'illustration 1.

2. <https://stat.ethz.ch/R-manual/R-devel/library/stats/html/constrOptim.html>

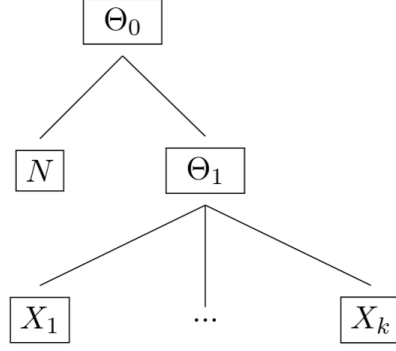


Illustration 1 – Arbre hiérarchique à un niveau.

Ici, Θ_0 représente une variable aléatoire définie sur \mathbb{N}_+ qui sert à créer un lien de dépendance entre la variable N et les $\{X_i, i = 1, \dots, k\}$. Ensuite, on a $\theta_1 = \sum_{i=1}^{\theta_0} B_i$ qui sert à créer un lien de dépendance entre les X_i . Les B_i sont i.i.d. et indépendants de θ_0 . B_i peut appartenir à \mathbb{R}_+ comme à \mathbb{N}_+ .

Sous cette représentation, on obtient une copule archimédienne hiérarchique s'exprimant comme

$$C(u_0, u_1, \dots, u_n) = \mathcal{L}_{\theta_0} \left(\mathcal{L}_{\theta_0}^{-1}(u_0) - \ln \left(\mathcal{L}_B \left(\sum_{i=1}^n \mathcal{L}_B^{-1}(\exp(-\mathcal{L}_{\theta_0}^{-1}(u_i))) \right) \right) \right), \quad (5)$$

où \mathcal{L} correspond à la transformée de Laplace-Stieltjes d'une variable aléatoire.

À ce point-ci, afin de trouver les paramètres d'une telle copule, il suffit d'appliquer (3) et (4), puis d'optimiser numériquement.

3 Résultats

La présente section explique les scénarios testés ainsi que les résultats obtenus avec la méthodologie expliquée dans la section 2.2.

Pour les fins du présent travail, les estimations sont faites sur des données simulées. Pour faire ces simulations, le module **R** nommé `copula`³ offre une fonction `rCopula` avec laquelle, il est possible de simuler différentes copules connues, dont la copule de Clayton.

Pour les copules archimédiennes hiérarchique, le module `nCopula`⁴ offre une fonction nommée `rCompCop` qui permet de simuler de la même façon que `rCopula`, mais en entrant comme argument la structure hiérarchique.

3.1 Copule de Clayton

Scénario 1 *Binomial* :

Pour débiter gentiment, prenons un modèle binomial-exponentiel avec une copule de Clayton dont la représentation est exprimée en (2). Soient $N \sim \text{Binom}(n, q)$ et $X_i \sim X \sim \text{Exp}(\beta)$ et le paramètre de dépendance de la copule est dénotée α .

Pour le premier scénario, on utilise des données simulées avec les paramètres initiaux présentés dans le tableau 1.

3. <https://cran.r-project.org/web/packages/copula/copula.pdf>

4. <https://cran.r-project.org/web/packages/nCopula/nCopula.pdf>

n	q	α	β
5.00	0.40	1.50	0.01

Tableau 1 – Paramètres initiaux pour la copule de Clayton avec $N \sim \text{Bin}(n, q)$, $X_i \sim X \sim \text{Exp}(\beta)$.

Le sommaire des données simulées aux fins de l'estimation sont présentées dans le tableau 2. On y voit que la moyenne du nombre de sinistres est de 2 et que la moyenne des x_i est très proche de 100; ce qui correspond aux attentes. De plus, les quantiles des X_i sont très proches l'un de l'autre; ce qui signifie que les variables X_i sont identiquement distribués. Visuellement parlant, l'illustration 2 présente l'adéquation des données empiriques avec les lois théoriques. On voit donc que les variables simulées ont un comportement très similaire à la distribution marginale théorique des variables.

N	X_1	X_2
Min. :0.000	Min. : 0.0049	Min. : 0.0066
1st Qu. :1.000	1st Qu. : 28.8101	1st Qu. : 29.1251
Median :2.000	Median : 70.8847	Median : 71.0354
Mean :1.991	Mean :101.0603	Mean : 101.3232
3rd Qu. :3.000	3rd Qu. :140.7053	3rd Qu. : 139.3937
Max. :5.000	Max. :864.5799	Max. :1071.7262
X_3	X_4	X_5
Min. : 0.0068	Min. : 0.0057	Min. : 0.007
1st Qu. : 28.6798	1st Qu. : 29.1230	1st Qu. : 29.076
Median : 69.3332	Median : 70.2568	Median : 69.112
Mean : 99.4368	Mean : 101.6533	Mean : 99.038
3rd Qu. :137.4674	3rd Qu. : 141.3571	3rd Qu. :135.610
Max. :988.0162	Max. :1001.1827	Max. :840.130

Tableau 2 – Sommaire de 10 000 simulations de $\{N, X_1, \dots, X_5\}$ en supposant que F_{N, X_1, \dots, X_5} est définie avec une copule de Clayton($\alpha = 1.5$), $N \sim \text{Bin}(5, 2/5)$ et $X_i \sim X \sim \text{Exp}(1/100)$, pour $i = 1, \dots, 5$.

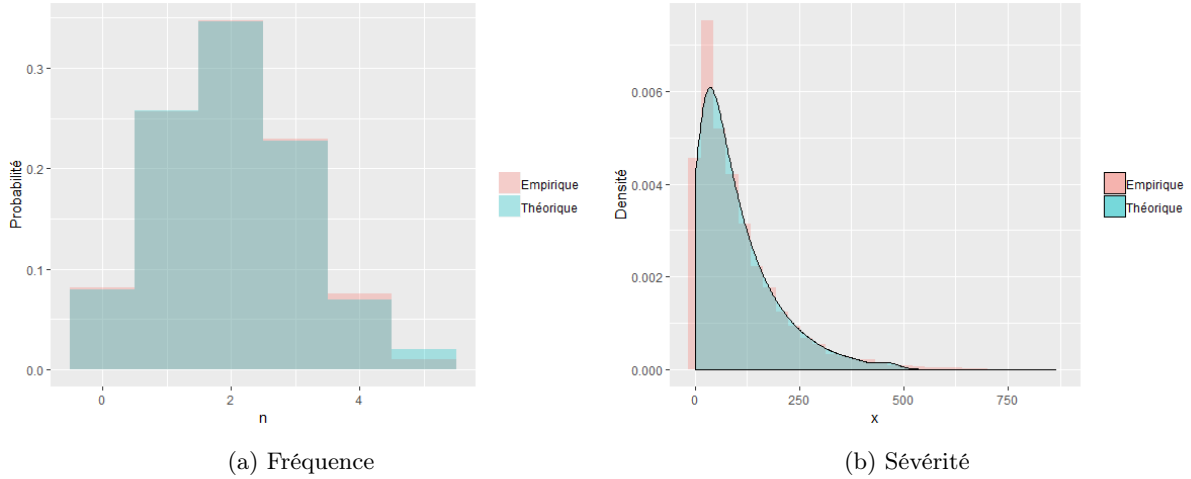


Illustration 2 – Comparaisons de la distribution marginale des données simulées avec les distributions théoriques pour le scénario 1.

Pour ce qui est du comportement entre les variables, l'illustration 3 présente les nuages de points représentant les corrélations. On y voit que le ρ de Spearman est le même pour tous les X_i . De plus, la forme des nuages de points pour les variables continues est caractéristique de la copule de Clayton. Cela indique que la simulation sort des résultats adéquats.

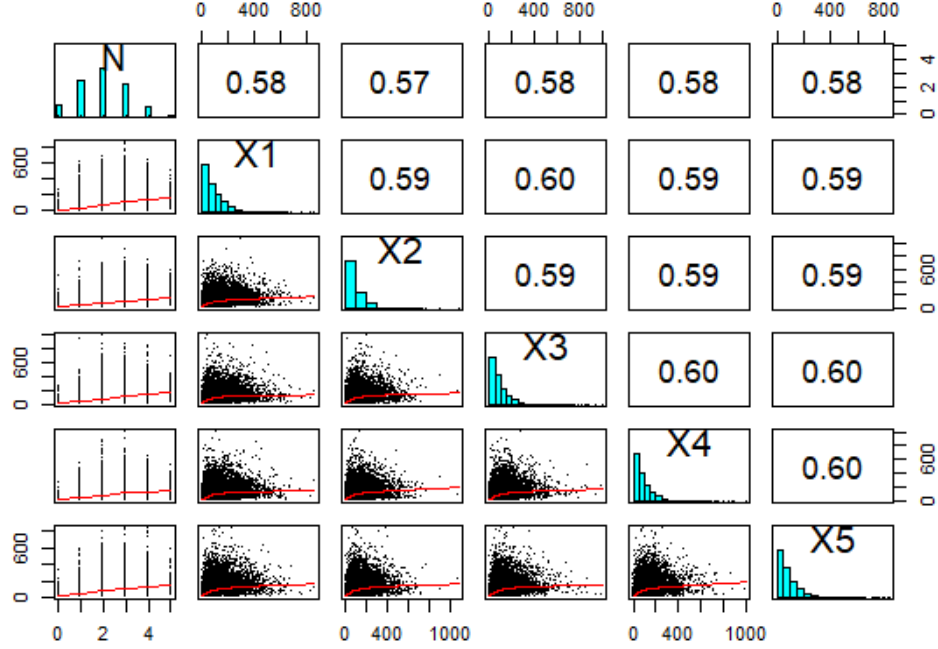


Illustration 3 – Nuages de points avec copule de Clayton ($\alpha = 1.5$), $N \sim \text{Binom}(5, 2/5)$ et $X_i \sim X \sim \text{Exp}(1/100)$: En bas de la diagonale se trouve les nuages de points illustrant la corrélation des différentes variables. En haut, on indique les coefficients de corrélation de Spearman. À noter que ce graphique est produit avec la fonction `R pairs.panels` du module `psych`. Dans cette dernière, le calcul des coefficients de corrélation sont calculés selon l'hypothèse que les variables sont continues. Cela implique que la première ligne n'est pas valide.

Avec les données simulées, on désire maintenant estimer les paramètres. On pose donc les valeurs de départ pour l'optimisation à l'estimateur du maximum de vraisemblance des distributions marginales des variables aléatoires. On a donc que $\tilde{q} = \bar{N}/n$, $\tilde{\beta} = 1/\bar{X}$, tel que \bar{N} et \bar{X} sont les moyennes empiriques de N et X respectivement. Pour ce qui est du α de la Clayton, il faut regarder les nuages de points de l'illustration 3. Plus le nuage est dense, plus le paramètre est grand. En revanche, si le nuage est large, plus α est près de 0.

Pour ce scénario, on a comme valeur de départ $\tilde{q} = 0.398720$, $\tilde{\beta} = 0.010031$. Pour α , posons la valeur initiale à 1. On obtient donc les estimateurs présentés dans le tableau 3. On y voit que les résultats de l'estimation sont relativement très précises. De plus, pour une fréquence maximale de 5, les temps de dérivation et d'estimation sont très rapides.

	q	β	α		
Valeurs de départ	0.3987	0.0100	1.0000	Temps de dérivation	4.22
Estimateurs	0.3997	0.0100	1.5051	Temps d'estimation	63.17
Vrais paramètres	0.4000	0.0100	1.5000		

Tableau 3 – Résultats de l'estimation des paramètres avec une copule de Clayton(α), $N \sim \text{Binom}(n, q)$ et $X \sim \text{Exp}(\beta)$ suite à 10 000 simulations.

Scénario 2 Poisson :

À la suite des résultats concluants du premier exemple, un modèle avec une loi de fréquence ayant un support infini peut ajouter un défi en terme de temps de calcul. Pour ce scénario, posons que $N \sim \text{Pois}(\lambda = 1)$ et $X_i \sim X \sim \text{Exp}(\beta = 1/100)$ et que le paramètre de dépendance de la Copule

de Clayton est $\alpha = 1.5$.

Comme pour le scénario 1, un sommaire des données simulées est présenté dans le tableau 4, l'illustration 5 compare la distribution marginale des données empiriques avec celle des lois théorique et l'illustration 5 présente les nuages de points afin de voir les corrélations.

N	X_1	X_2	X_3	X_4
Min. :0.0000	Min. : 0.006	Min. : 0.0076	Min. : 0.004	Min. : 0.0047
1st Qu. :0.0000	1st Qu. : 29.296	1st Qu. : 28.6055	1st Qu. : 29.217	1st Qu. : 28.8844
Median :1.0000	Median : 68.978	Median : 70.1421	Median : 70.539	Median : 69.1708
Mean :0.9987	Mean : 99.922	Mean : 100.1551	Mean : 99.760	Mean : 98.9900
3rd Qu. :2.0000	3rd Qu. :138.052	3rd Qu. : 136.9883	3rd Qu. : 137.060	3rd Qu. :136.3721
Max. :6.0000	Max. :914.272	Max. :1041.9811	Max. :1050.119	Max. :953.2574
	X_5	X_6	X_7	X_8
Min. : 0.0068	Min. : 0.008	Min. : 0.0072	Min. : 0.0093	
1st Qu. : 29.5422	1st Qu. : 28.664	1st Qu. : 29.3327	1st Qu. : 28.8089	
Median : 70.7936	Median : 71.331	Median : 69.5485	Median : 69.3149	
Mean : 100.4916	Mean :102.201	Mean :101.1266	Mean : 99.5730	
3rd Qu. : 139.2436	3rd Qu. :141.000	3rd Qu. :139.8154	3rd Qu. :137.2429	
Max. :1078.6322	Max. :894.636	Max. :961.4753	Max. :824.4899	

Tableau 4 – Sommaire de 10 000 simulations de $\{N, X_1, \dots, X_8\}$ en supposant que F_{N, X_1, \dots, X_8} est définie avec une copule de Clayton($\alpha = 1.5$), $N \sim Pois(1)$ et $X_i \sim X \sim Exp(1/100)$, pour $i = 1, \dots, 8$.

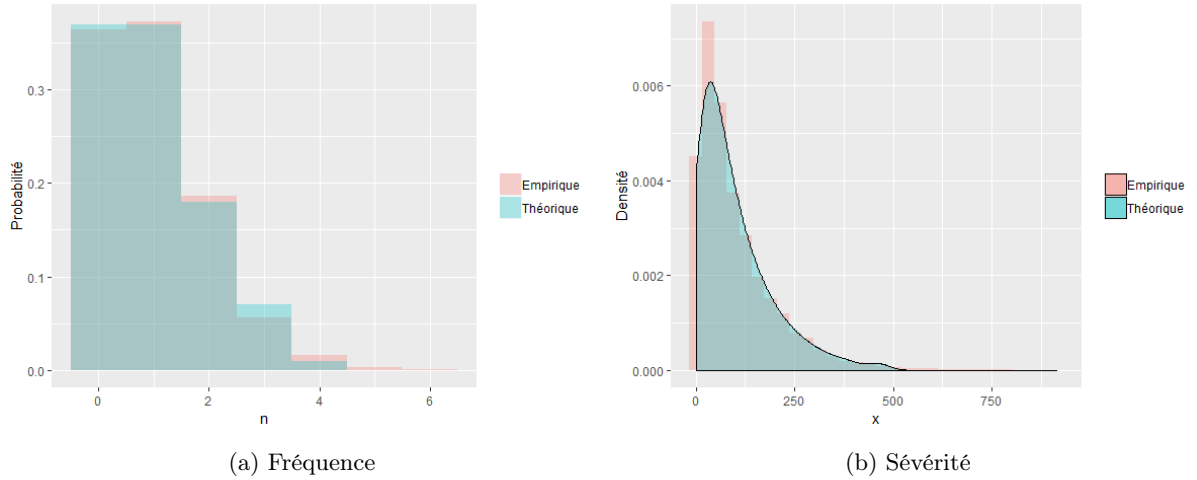


Illustration 4 – Comparaisons de la distribution marginale des données simulées avec les distributions théoriques pour le scénario 2.

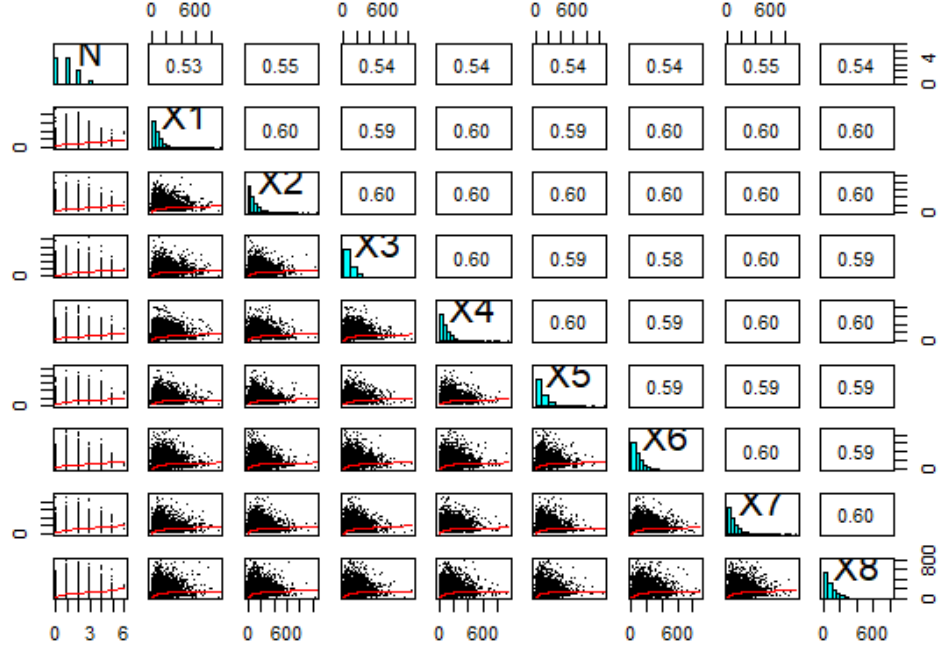


Illustration 5 – Nuages de points avec copule de Clayton ($\alpha = 1.5$), $N \sim Pois(1)$ et $X_i \sim X \sim Exp(1/100)$: En bas de la diagonale se trouve les nuages de points illustrant la corrélation des différentes variables. En haut, on indique les coefficients de corrélation de Spearman. À noter que ce graphique est produit avec la fonction `R pairs.panels` du module `psych`. Dans cette dernière, le calcul des coefficients de corrélation sont calculés selon l'hypothèse que les variables sont continues. Cela implique que la première ligne n'est pas valide.

Comme pour le scénario 1, le tableau 4 et les illustrations 5 démontrent que les résultats de la simulation sont adéquats.

Pour ce qui est de l'estimation des paramètres, commençons par définir les paramètres initiaux. On pose donc $\tilde{\lambda} = \bar{N} = 0.998700$, $\tilde{\beta} = 1/\bar{X} = 0.009972$ et comme pour l'exemple précédent, prenons 1 comme valeur de départ pour α . Les estimations obtenus sont présentés dans le tableau 5.

	λ	β	α		
Valeurs de départ	0.9987	0.0100	1.0000	Temps de dérivation	12.60
Estimateurs	1.0040	0.0101	1.5415	Temps d'estimation	75.27
Vrais paramètres	1.0000	0.0100	1.5000		

Tableau 5 – Résultats de l'estimation des paramètres avec une copule de Clayton(α), $N \sim Pois(\lambda)$ et $X \sim Exp(\beta)$ suite à 10 000 simulations.

Dans cet exemple, l'estimation n'est pas plus longue puisque le paramètre de la poisson n'est pas très grand. Cependant, si on l'augmente un tant soit peu, on obtient des quantiles élevés qui font en sorte que le temps de dérivation augmente de façon exponentielle.

Par exemple, si $N \sim Pois(2)$, le temps de calcul dépasse 12 heures puisque le 99,9999 percentile de cette loi est de 12. Ainsi la méthode du maximum de vraisemblance nécessite de générer une liste de douze fonctions qui sont dérivées jusqu'à douze fois. Même avec la fonction `Deriv` de R, ce processus est extrêmement long. À cet effet, l'illustration 6 présente le temps de dérivation selon le nombre de dérivées partielles à effectuer sur une copule de Clayton.

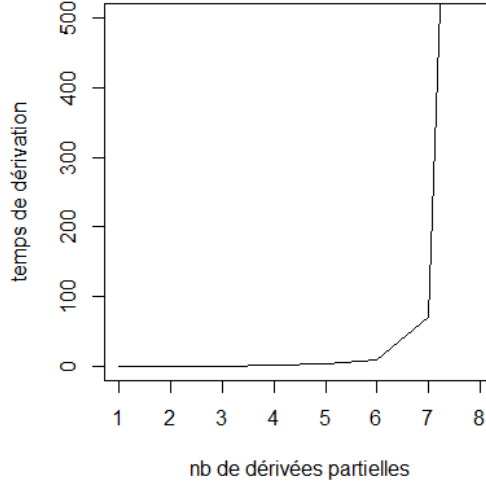


Illustration 6 – Temps de dérivation d’une copule de Clayton en fonction du nombre de dérivées partielles à effectuer.

Ce que l’on peut observer avec ces deux résultats, c’est, d’une part, qu’avec 10 000 simulations, on obtient des résultats très adéquats. D’autre part, on peut observer que le temps de calcul augmente significativement si N peut prendre des valeurs supérieures à 7 et si le nombre de paramètres à estimer est grand. Le temps de calcul est donc un enjeu important.

3.2 Copule archimédienne hiérarchique

Désormais, prenons la copule qui nous intéresse vraiment : la copule archimédienne hiérarchique. Dans la présente section un scénario est fait pour chacune des six familles de copule archimédienne hiérarchique.

Scénario 3 *Logarithmique-Logarithmique*

Soit $\Theta_0 \sim \text{Logarithmique}(\gamma_0)$ avec $\mathcal{L}_{\Theta_0} = \frac{\ln(1-\gamma_0 e^{-t})}{\ln(1-\gamma_0)}$ et $\mathcal{L}_{\Theta_0}^{-1} = -\ln\left(\frac{1-(1-\gamma_0)^u}{\gamma_0}\right)$. Soit $B_i \sim B \sim \text{Logarithmique}(\gamma_1)$ avec $\mathcal{L}_B = \frac{\ln(1-\gamma_1 e^{-t})}{\ln(1-\gamma_1)}$ et $\mathcal{L}_B^{-1} = -\ln\left(\frac{1-(1-\gamma_1)^u}{\gamma_1}\right)$. Soient $\alpha_0 = -\ln(1-\gamma_0)$ et $\alpha_1 = -\ln(1-\gamma_1)$. Avec $\alpha_0 = 0.5$, $\alpha_1 = 0.8$, $N \sim \text{Bin}(n=5, q=2/5)$ et $X_i \sim X \sim \text{Exp}(\beta=1/100)$, on obtient le sommaire des simulations présenté dans le tableau 6 ainsi que les illustrations 7 et 8.

N	X_1	X_2
Min. :0.000	Min. : 0.0005	Min. : 0.0003
1st Qu. :1.000	1st Qu. : 28.8718	1st Qu. : 28.7527
Median :2.000	Median : 70.2264	Median : 69.5752
Mean :1.993	Mean : 99.0743	Mean : 101.3211
3rd Qu. :3.000	3rd Qu. : 137.8503	3rd Qu. : 140.6853
Max. :5.000	Max. :1159.3514	Max. :1276.8609
X_3	X_4	X_5
Min. : 0.0039	Min. : 0.0003	Min. : 0.0055
1st Qu. : 27.8579	1st Qu. : 29.2991	1st Qu. : 28.8589
Median : 69.3043	Median : 69.3083	Median : 69.9943
Mean : 99.3564	Mean : 99.3880	Mean :100.8576
3rd Qu. :138.5018	3rd Qu. :136.2992	3rd Qu. :140.1445
Max. :867.0918	Max. :852.5330	Max. :997.9467

Tableau 6 – Sommaire de 10 000 simulations de $\{N, X_1, \dots, X_5\}$ en supposant que F_{N, X_1, \dots, X_5} est définie avec une copule de archimédienne hiérarchique logarithmique-logarithmique($\alpha_0 = 0.5, \alpha_1 = 0.8$), $N \sim \text{Bin}(5, 2/5)$ et $X_i \sim X \sim \text{Exp}(1/100)$, pour $i = 1, \dots, 5$.

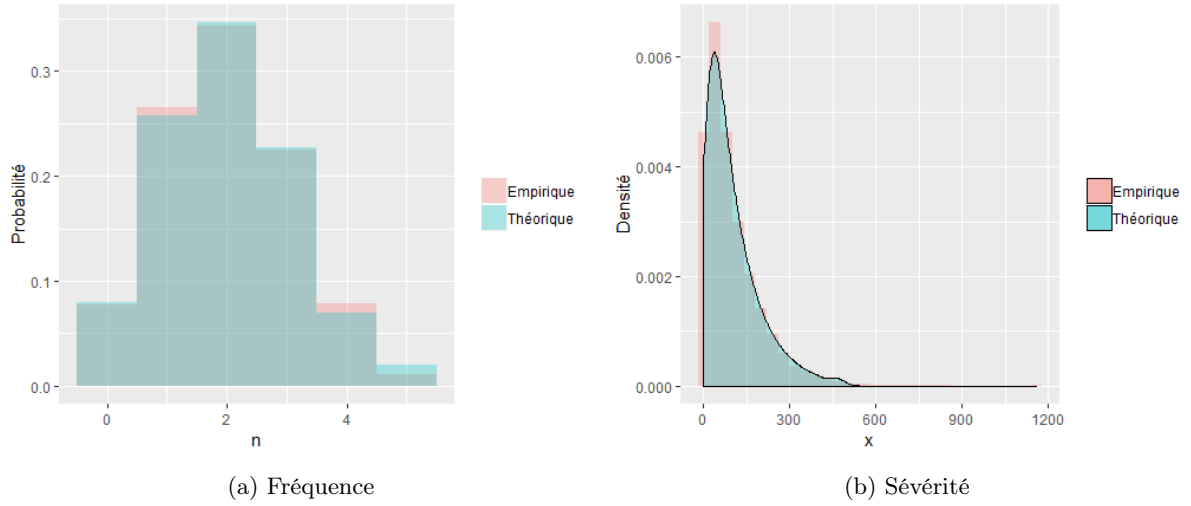


Illustration 7 – Comparaisons de la distribution marginale des données simulées avec les distributions théoriques pour le scénario 3.

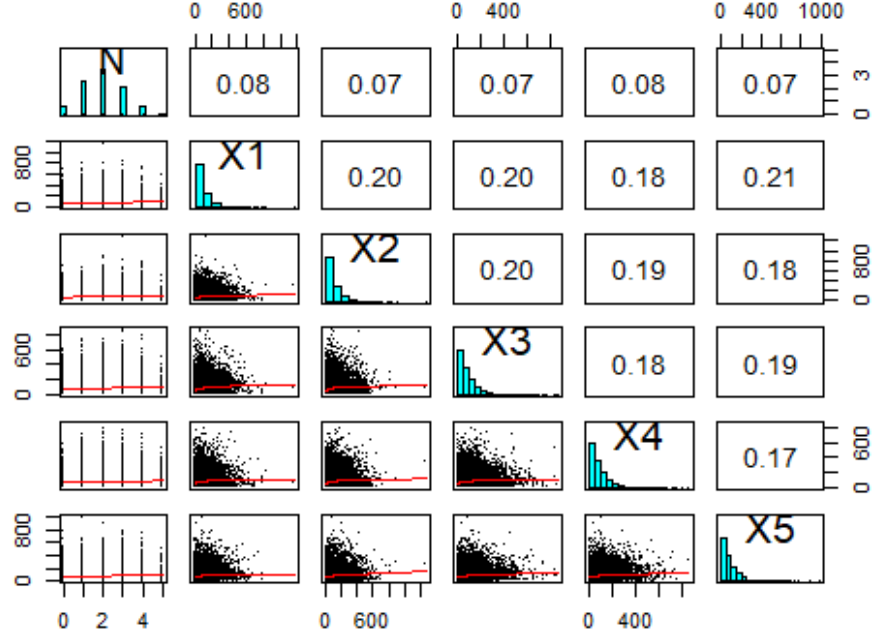


Illustration 8 – Nuages de points avec copule archimédienne hiérarchique logarithmique-logarithmique($\alpha_0 = 0.5, \alpha_1 = 0.8$), $N \sim \text{Bin}(5, 2/5)$ et $X_i \sim X \sim \text{Exp}(1/100)$, pour $i = 1, \dots, 5$. : En bas de la diagonale se trouve les nuages de points illustrant la corrélation des différentes variables. En haut, on indique les coefficients de corrélation de Spearman. À noter que ce graphique est produit avec la fonction `R pairs.panels` du module `psych`. Dans cette dernière, le calcul des coefficients de corrélation sont calculés selon l’hypothèse que les variables sont continues. Cela implique que la première ligne n’est pas valide.

Pour l’estimation des paramètres initiaux, avec la méthode du maximum de vraisemblance, on a $\tilde{q} = \bar{N}/n = 0.398600$ et $\tilde{\beta} = 1/\bar{X} = 0.010070$. Cependant, pour les paramètres de dépendance, il n’est pas aisé de trouver des valeurs de départ adéquates. Pour le moment, de façon arbitraire, posons $\alpha_0 = \alpha_1 = 0.4$. On obtient donc les estimateurs présentés dans le tableau 7

	q	β	α_0	α_1		
Valeurs de départ	0.3986	0.0100	0.4000	0.4000	Temps de dérivation	2199.69
Estimateurs	0.3986	0.0100	0.4499	0.7700	Temps d’estimation	254.75
Vrais paramètres	0.4000	0.0100	0.5000	0.8000		

Tableau 7 – Résultats de l’estimation des paramètres avec une copule archimédienne hiérarchique logarithmique-logarithmique($\alpha_0 = 0.5, \alpha_1 = 0.8$), $N \sim \text{Bin}(5, 2/5)$ et $X_i \sim X \sim \text{Exp}(1/100)$, pour $i = 1, \dots, 5$, à la suite de 10000 simulations.

À la lecture du tableau 7, on voit que les estimateurs de α_0 et α_1 sont moins près des vrais valeurs que pour les autres. Cela est explicable du fait que, pour ces paramètres, les valeurs de départ ne sont pas optimales ce qui augmente les probabilités que la fonction `R constrOptim` ne trouve pas le minimum absolu de la fonction de score, mais plutôt qu’elle trouve un minimum local.

Scénario 4 Logarithmique-Gamma

Soit $\Theta_0 \sim \text{Logarithmique}(\gamma)$ avec $\mathcal{L}_{\Theta_0} = \frac{\ln(1-\gamma e^{-t})}{\ln(1-\gamma)}$ et $\mathcal{L}_{\Theta_0}^{-1} = -\ln\left(\frac{1-(1-\gamma)^u}{\gamma}\right)$. Soit $\alpha_0 = -\ln(1-\gamma)$. Soit $B_i \sim B \sim \text{gamma}(1/\alpha_1, 1)$ avec $\mathcal{L}_B = (1+t)^{-1/\alpha_1}$ et $\mathcal{L}_B^{-1} = u^{-\alpha_1} - 1$. Avec $\alpha_0 = 0.5$, $\alpha_1 = 5$, $N \sim \text{Bin}(n = 5, q = 2/5)$ et $X_i \sim X \sim \text{Exp}(\beta = 1/100)$, on obtient le sommaire des simulations présenté dans le tableau 8 ainsi que les illustrations 9 et 10.

N	X_1	X_2
Min. :0	Min. : 0.0114	Min. : 0.0121
1st Qu. :1	1st Qu. : 28.3559	1st Qu. : 27.984
Median :2	Median : 68.5183	Median : 67.6682
Mean :2	Mean : 98.2412	Mean : 99.0285
3rd Qu. :3	3rd Qu. :136.8315	3rd Qu. : 138.0584
Max. :5	Max. :888.6908	Max. :1008.9367
X_3	X_4	X_5
Min. : 0.0113	Min. : 0.0273	Min. : 0.0118
1st Qu. : 28.1662	1st Qu. : 28.0790	1st Qu. : 28.1698
Median : 69.8290	Median : 68.1602	Median : 68.6309
Mean : 99.3519	Mean : 98.5785	Mean : 98.8991
3rd Qu. :139.2965	3rd Qu. :138.2672	3rd Qu. : 136.6441
Max. :781.9900	Max. :950.1376	Max. :1095.3264

Tableau 8 – Sommaire de 10 000 simulations de $\{N, X_1, \dots, X_5\}$ en supposant que F_{N, X_1, \dots, X_5} est définie avec une copule de archimédienne hiérarchique logarithmique-gamma($\alpha_0 = 0.5, \alpha_1 = 5$), $N \sim \text{Bin}(5, 2/5)$ et $X_i \sim X \sim \text{Exp}(1/100)$, pour $i = 1, \dots, 5$.

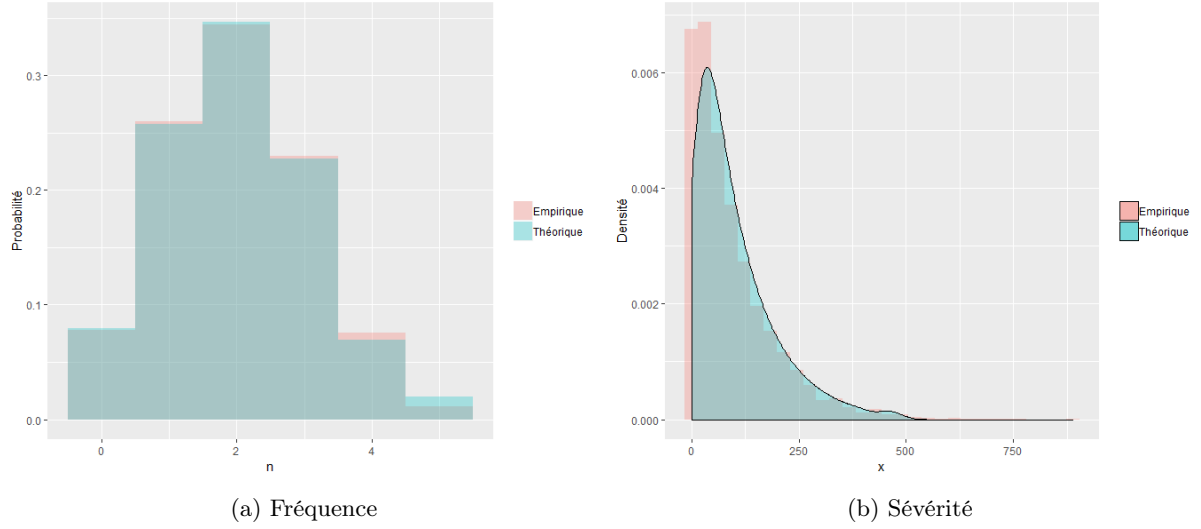


Illustration 9 – Comparaisons de la distribution marginale des données simulées avec les distributions théoriques pour le scénario 4.

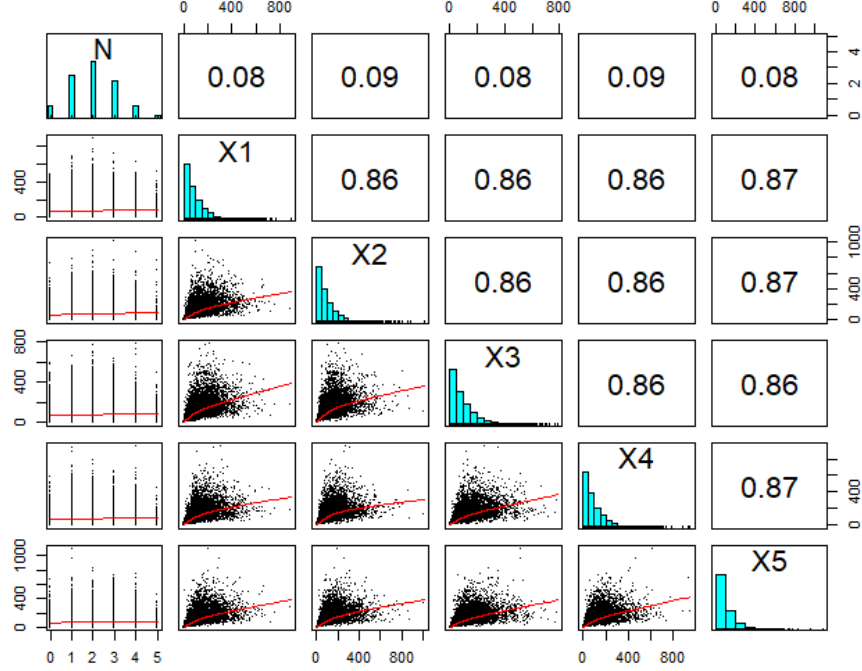


Illustration 10 – Nuages de points avec copule archimédienne hiérarchique logarithmique-gamma($\alpha_0 = 0.5, \alpha_1 = 5$), $N \sim \text{Bin}(5, 2/5)$ et $X_i \sim X \sim \text{Exp}(1/100)$, pour $i = 1, \dots, 5$. : En bas de la diagonale se trouve les nuages de points illustrant la corrélation des différentes variables. En haut, on indique les coefficients de corrélation de Spearman. À noter que ce graphique est produit avec la fonction `R pairs.panels` du module `psych`. Dans cette dernière, le calcul des coefficients de corrélation sont calculés selon l’hypothèse que les variables sont continues. Cela implique que la première ligne n’est pas valide.

Les valeurs de départ de l’optimisation numérique ainsi que les résultats de l’estimation sont présentés dans le tableau 9.

	q	β	α_0	α_1		
Valeurs de départ	0.4001	0.0097	0.5000	4.0000	Temps de dérivation	50.38
Estimateurs	0.3970	0.00983	0.5133	5.0751	Temps d’estimation	2714.00
Vrais paramètres	0.4000	0.0100	0.5000	5.0000		

Tableau 9 – Résultats de l’estimation des paramètres avec une copule archimédienne hiérarchique logarithmique-gamma($\alpha_0 = 0.5, \alpha_1 = 5$), $N \sim \text{Bin}(5, 2/5)$ et $X_i \sim X \sim \text{Exp}(1/100)$, pour $i = 1, \dots, 5$, à la suite de 10 000 simulations.

Scénario 5 *Logarithmique-Géométrique*

Soit $\Theta_0 \sim \text{Logarithmique}(\gamma)$ avec $\mathcal{L}_{\Theta_0} = \frac{\ln(1-\gamma e^{-t})}{\ln(1-\gamma)}$ et $\mathcal{L}_{\Theta_0}^{-1} = -\ln\left(\frac{1-(1-\gamma)^u}{\gamma}\right)$. Soit $\alpha_0 = -\ln(1-\gamma)$. Soit $B_i \sim B \sim \text{Geo}(1 - \alpha_1)$ avec $E[B] = \frac{1}{1-\alpha_1}$, $\mathcal{L}_B = \frac{1-\alpha_1}{e^t - \alpha_1}$ et $\mathcal{L}_B^{-1} = \ln\left(\frac{1-\alpha_1}{u} + \alpha_1\right)$. Avec $\alpha_0 = 0.5$, $\alpha_1 = 0.3$, $N \sim \text{Bin}(n = 5, q = 2/5)$ et $X_i \sim X \sim \text{Exp}(\beta = 1/100)$, on obtient le sommaire des simulations présenté dans le tableau 10 ainsi que les illustrations 11 et 12.

N	X_1	X_2
Min. :0.000	Min. : 0.0152	Min. : 0.0023
1st Qu. :1.000	1st Qu. : 28.3205	1st Qu. : 27.6623
Median :2.000	Median : 68.6348	Median : 68.6183
Mean :1.994	Mean : 98.9724	Mean : 98.2568
3rd Qu. :3.000	3rd Qu. :138.8460	3rd Qu. :136.5162
Max. :5.000	Max. :959.2391	Max. :857.1796
X_3	X_4	X_5
Min. : 0.0015	Min. : 0.0052	Min. : 0.0033
1st Qu. : 28.5541	1st Qu. : 29.3939	1st Qu. : 28.1749
Median : 68.5017	Median : 71.0794	Median : 68.6236
Mean : 100.1266	Mean :100.1761	Mean : 97.9566
3rd Qu. : 138.9653	3rd Qu. :138.0791	3rd Qu. :136.7516
Max. :1195.4707	Max. :754.0023	Max. :924.5513

Tableau 10 – Sommaire de 10 000 simulations de $\{N, X_1, \dots, X_5\}$ en supposant que F_{N, X_1, \dots, X_5} est définie avec une copule de archimédienne hiérarchique logarithmique-géométrique($\alpha_0 = 0.5, \alpha_1 = 0.3$), $N \sim \text{Bin}(5, 2/5)$ et $X_i \sim X \sim \text{Exp}(1/100)$, pour $i = 1, \dots, 5$.

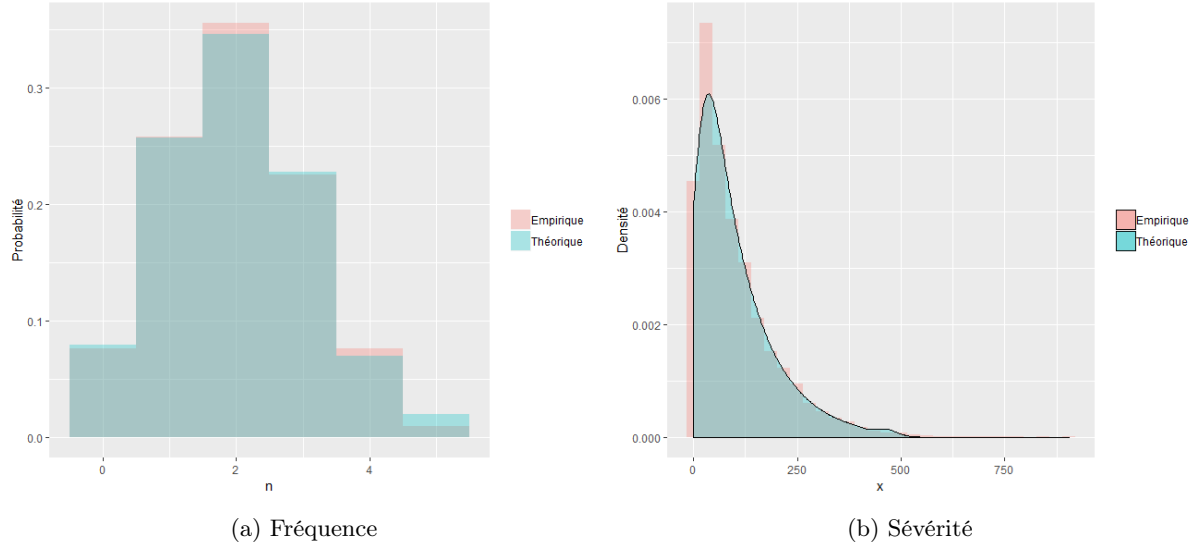


Illustration 11 – Comparaisons de la distribution marginale des données simulées avec les distributions théoriques pour le scénario 5.

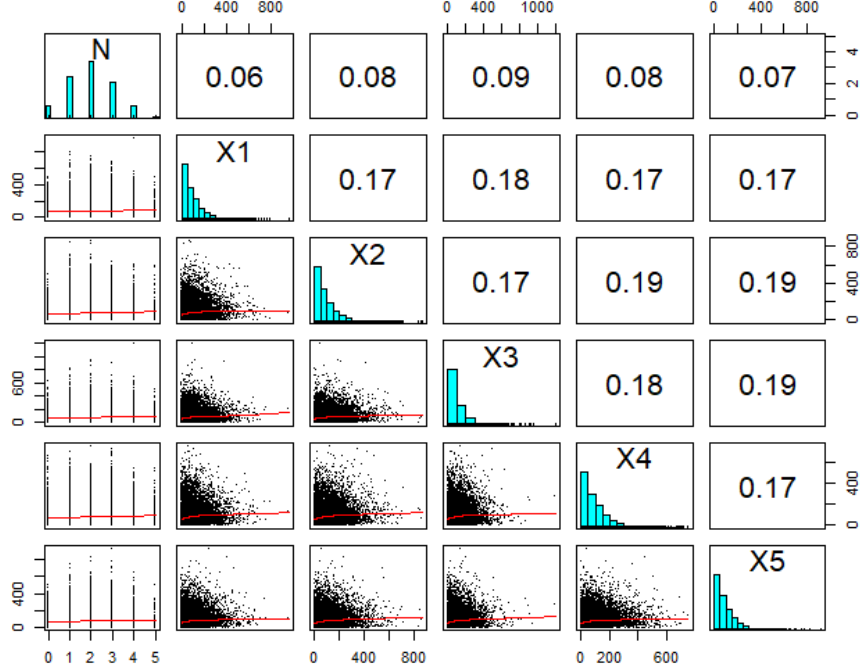


Illustration 12 – Nuages de points avec copule archimédienne hiérarchique logarithmique-géométrique($\alpha_0 = 0.5, \alpha_1 = 0.3$), $N \sim \text{Bin}(5, 2/5)$ et $X_i \sim X \sim \text{Exp}(1/100)$, pour $i = 1, \dots, 5$. : En bas de la diagonale se trouve les nuages de points illustrant la corrélation des différentes variables. En haut, on indique les coefficients de corrélation de Spearman. À noter que ce graphique est produit avec la fonction `R pairs.panels` du module `psych`. Dans cette dernière, le calcul des coefficients de corrélation sont calculés selon l’hypothèse que les variables sont continues. Cela implique que la première ligne n’est pas valide.

Les valeurs de départ de l’optimisation numérique ainsi que les résultats de l’estimation sont présentés dans le tableau 11.

	q	β	α_0	α_1		
Valeurs de départ	0.3987	0.0097	0.5000	0.5000	Temps de dérivation	109.45
Estimateurs	0.3999	0.0100	0.5028	0.5059	Temps d’estimation	1001.81
Vrais paramètres	0.4000	0.0100	0.5000	0.3000		

Tableau 11 – Résultats de l’estimation des paramètres avec une copule archimédienne hiérarchique logarithmique-géométrique($\alpha_0 = 0.5, \alpha_1 = 0.3$), $N \sim \text{Bin}(5, 2/5)$ et $X_i \sim X \sim \text{Exp}(1/100)$, pour $i = 1, \dots, 5$, à la suite de 10 000 simulations.

Scénario 6 Géométrique-Logarithmique

Soit $\Theta_0 \sim \text{Geo}(1 - \alpha_0)$ avec $E[M] = \frac{1}{1 - \alpha_0}$, $\mathcal{L}_{\Theta_0} = \frac{1 - \alpha_0}{e^t - \alpha_0}$ et $\mathcal{L}_{\Theta_0}^{-1} = \ln\left(\frac{1 - \alpha_0}{u} + \alpha_0\right)$. Soit $B_i \sim B \sim \text{Logarithmique}(\gamma)$ avec $\mathcal{L}_B = \frac{\ln(1 - \gamma e^{-t})}{\ln(1 - \gamma)}$ et $\mathcal{L}_B^{-1} = -\ln\left(\frac{1 - (1 - \gamma)^u}{\gamma}\right)$. Soit $\alpha_1 = -\ln(1 - \gamma)$. Avec $\alpha_0 = 0.5, \alpha_1 = 0.8, N \sim \text{Bin}(n = 4, q = 2/5)$ et $X_i \sim X \sim \text{Exp}(\beta = 1/100)$, on obtient le sommaire des simulations présenté dans le tableau 12 ainsi que les illustrations 13 et 14.

N	X_1	X_2	X_3	X_4
Min. :0.000	Min. : 0.0023	Min. : 0.0073	Min. : 0.0121	Min. : 0.0065
1st Qu. :1.000	1st Qu. : 28.0900	1st Qu. : 27.9637	1st Qu. : 28.2663	1st Qu. : 29.1747
Median :2.000	Median : 67.8528	Median : 69.5458	Median : 67.9348	Median : 69.4055
Mean :1.595	Mean : 98.5520	Mean : 99.2869	Mean : 99.0303	Mean :100.4566
3rd Qu. :2.000	3rd Qu. :137.5109	3rd Qu. :137.5743	3rd Qu. :136.4480	3rd Qu. :139.4227
Max. :4.000	Max. :837.1782	Max. :964.5771	Max. :977.3574	Max. :854.9710

Tableau 12 – Sommaire de 10 000 simulations de $\{N, X_1, \dots, X_4\}$ en supposant que F_{N, X_1, \dots, X_4} est définie avec une copule de archimédienne hiérarchique géométrique-logarithmique($\alpha_0 = 0.5, \alpha_1 = 0.8$), $N \sim \text{Bin}(4, 2/5)$ et $X_i \sim X \sim \text{Exp}(1/100)$, pour $i = 1, \dots, 4$.

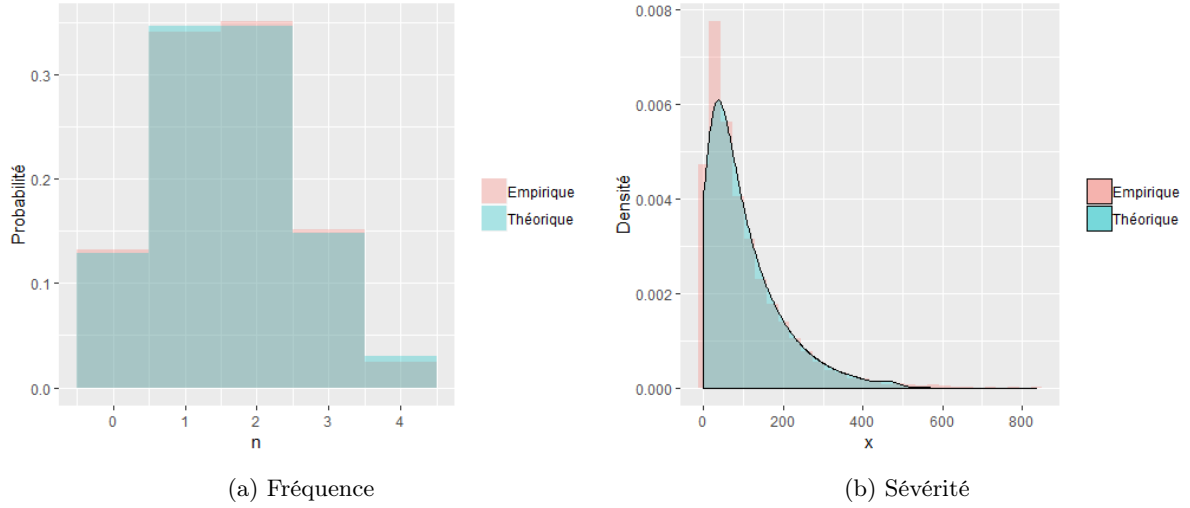


Illustration 13 – Comparaisons de la distribution marginale des données simulées avec les distributions théoriques pour le scénario 6.

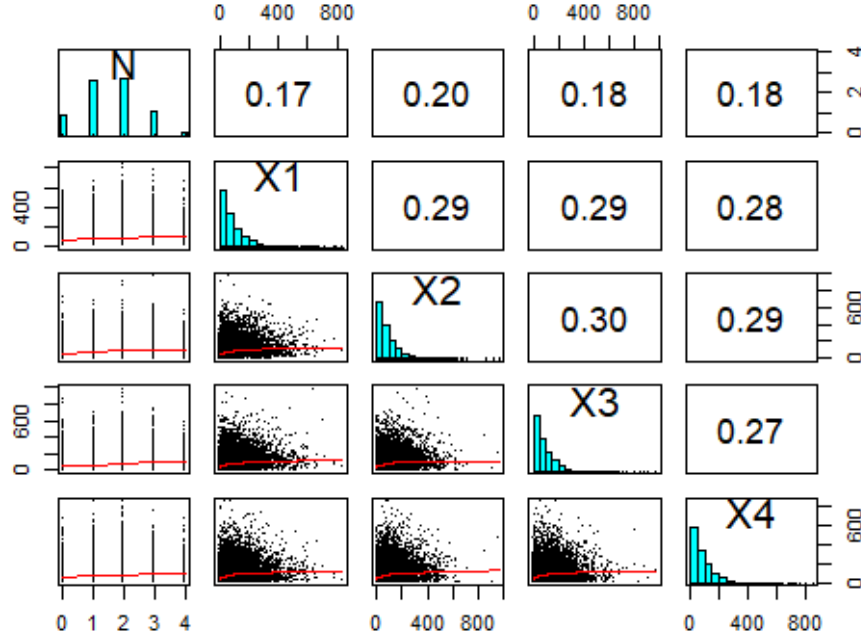


Illustration 14 – Nuages de points avec copule archimédienne hiérarchique géométrique-logarithmique($\alpha_0 = 0.5, \alpha_1 = 0.8$), $N \sim \text{Bin}(4, 2/5)$ et $X_i \sim X \sim \text{Exp}(1/100)$, pour $i = 1, \dots, 4$. : En bas de la diagonale se trouve les nuages de points illustrant la corrélation des différentes variables. En haut, on indique les coefficients de corrélation de Spearman. À noter que ce graphique est produit avec la fonction `R pairs.panels` du module `psych`. Dans cette dernière, le calcul des coefficients de corrélation sont calculés selon l’hypothèse que les variables sont continues. Cela implique que la première ligne n’est pas valide.

Les valeurs de départ de l’optimisation numérique ainsi que les résultats de l’estimation sont présentés dans le tableau 13.

	q	β	α_0	α_1		
Valeurs de départ	0.3987	0.0101	0.5000	0.5000	Temps de dérivation	57.38
Estimateurs	0.3990	0.0101	0.4971	0.8712	Temps d’estimation	149.72
Vrais paramètres	0.4000	0.0100	0.5000	0.8000		

Tableau 13 – Résultats de l’estimation des paramètres avec une copule archimédienne hiérarchique géométrique-logarithmique($\alpha_0 = 0.5, \alpha_1 = 0.8$), $N \sim \text{Bin}(4, 2/5)$ et $X_i \sim X \sim \text{Exp}(1/100)$, pour $i = 1, \dots, 4$, à la suite de 10 000 simulations.

Scénario 7 Géométrie-Gamma

Soit $\Theta_0 \sim \text{Geo}(1 - \alpha_0)$ avec $E[M] = \frac{1}{1 - \alpha_0}$, $\mathcal{L}_{\Theta_0} = \frac{1 - \alpha_0}{e^t - \alpha_0}$ et $\mathcal{L}_{\Theta_0}^{-1} = \ln\left(\frac{1 - \alpha_0}{u} + \alpha_0\right)$. Soit $B_i \sim B \sim \text{gamma}(1/\alpha_1, 1)$ avec $\mathcal{L}_B = (1 + t)^{-1/\alpha_1}$ et $\mathcal{L}_B^{-1} = u^{-\alpha_1} - 1$. Avec $\alpha_0 = 0.5, \alpha_1 = 0.8, N \sim \text{Bin}(n = 5, q = 2/5)$ et $X_i \sim X \sim \text{Exp}(\beta = 1/100)$, on obtient le sommaire des simulations présenté dans le tableau 14 ainsi que les illustrations 15 et 16.

N	X_1	X_2
Min. :0.000	Min. : 0.0006	Min. : 0.0005
1st Qu. :1.000	1st Qu. : 29.2505	1st Qu. : 29.1884
Median :2.000	Median : 69.6601	Median : 70.7418
Mean :2.002	Mean :100.0527	Mean :100.3462
3rd Qu. :3.000	3rd Qu. :138.5964	3rd Qu. :138.9114
Max. :5.000	Max. :908.6793	Max. :924.0011
X_3	X_4	X_5
Min. : 0.0005	Min. : 0.0005	Min. : 0.0008
1st Qu. : 29.2432	1st Qu. : 29.1898	1st Qu. : 29.4361
Median : 69.1627	Median : 70.8558	Median : 69.3439
Mean : 100.4765	Mean :101.3061	Mean :101.1562
3rd Qu. : 139.4879	3rd Qu. :139.4048	3rd Qu. :139.4248
Max. :1208.1788	Max. :916.6048	Max. :846.3202

Tableau 14 – Sommaire de 10 000 simulations de $\{N, X_1, \dots, X_5\}$ en supposant que F_{N, X_1, \dots, X_5} est définie avec une copule de archimédienne hiérarchique géométrique-gamma($\alpha_0 = 0.5, \alpha_1 = 0.8$), $N \sim \text{Bin}(5, 2/5)$ et $X_i \sim X \sim \text{Exp}(1/100)$, pour $i = 1, \dots, 5$.

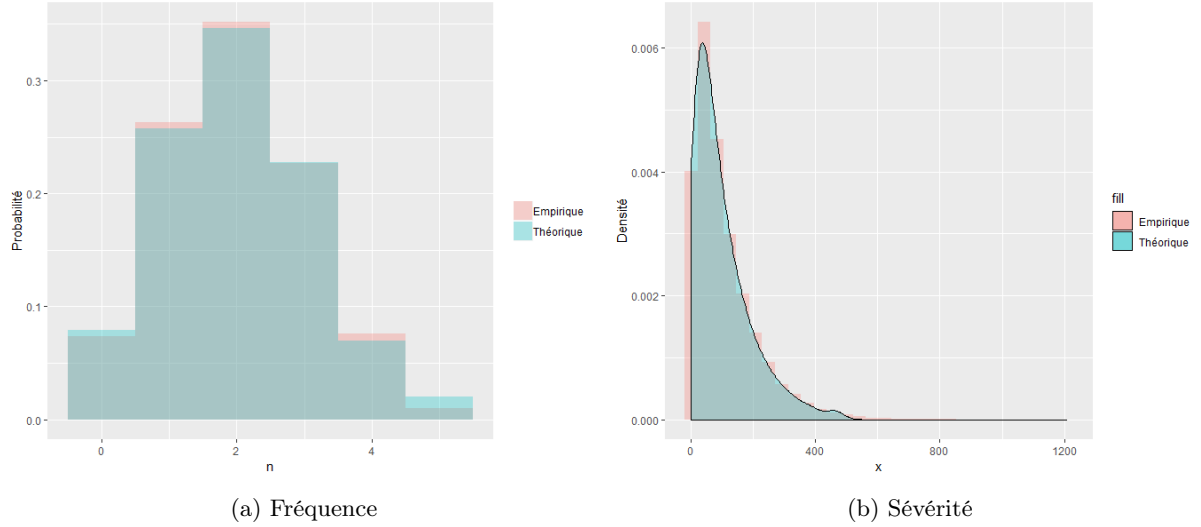


Illustration 15 – Comparaisons de la distribution marginale des données simulées avec les distributions théoriques pour le scénario 7.

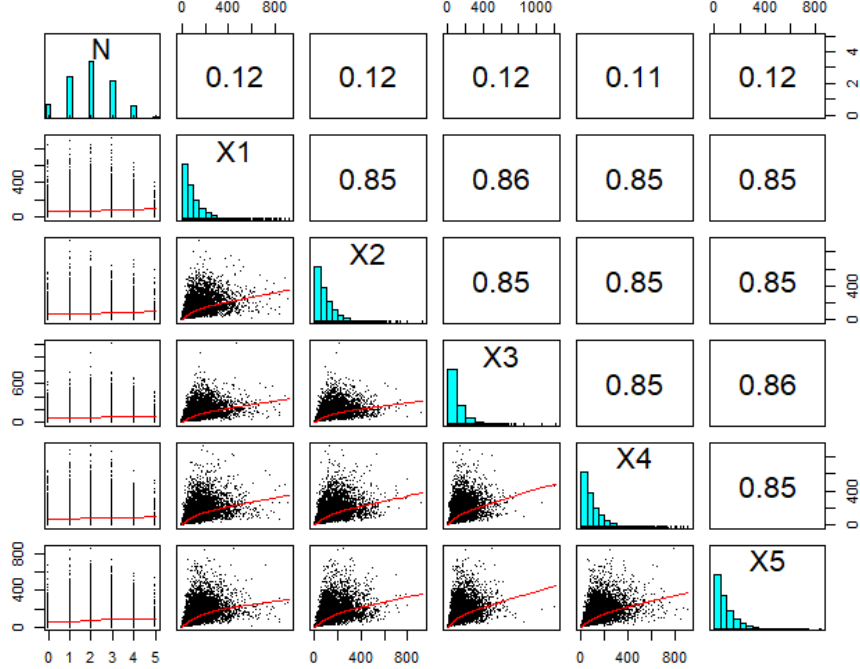


Illustration 16 – Nuages de points avec copule archimédienne hiérarchique géométrique-gamma($\alpha_0 = 0.5, \alpha_1 = 0.8$), $N \sim \text{Bin}(5, 2/5)$ et $X_i \sim X \sim \text{Exp}(1/100)$, pour $i = 1, \dots, 5$. : En bas de la diagonale se trouve les nuages de points illustrant la corrélation des différentes variables. En haut, on indique les coefficients de corrélation de Spearman. À noter que ce graphique est produit avec la fonction `R pairs.panels` du module `psych`. Dans cette dernière, le calcul des coefficients de corrélation sont calculés selon l’hypothèse que les variables sont continues. Cela implique que la première ligne n’est pas valide.

Les valeurs de départ de l’optimisation numérique ainsi que les résultats de l’estimation sont présentés dans le tableau 15.

	q	β	α_0	α_1		
Valeurs de départ	0.4005	0.0095	0.5000	4.0000	Temps de dérivation	57.68
Estimateurs	0.4019	0.0157	0.2503	6.1103	Temps d’estimation	1846.39
Vrais paramètres	0.4000	0.0100	0.3000	5.0000		

Tableau 15 – Résultats de l’estimation des paramètres avec une copule archimédienne hiérarchique géométrique-gamma($\alpha_0 = 0.5, \alpha_1 = 0.8$), $N \sim \text{Bin}(5, 2/5)$ et $X_i \sim X \sim \text{Exp}(1/100)$, pour $i = 1, \dots, 5$, à la suite de 10 000 simulations.

Scénario 8 Géométrie-Géométrie

Soit $\Theta_0 \sim \text{Geo}(1 - \alpha_0)$ avec $E[M] = \frac{1}{1 - \alpha_0}$, $\mathcal{L}_{\Theta_0} = \frac{1 - \alpha_0}{e^t - \alpha_0}$ et $\mathcal{L}_{\Theta_0}^{-1} = \ln\left(\frac{1 - \alpha_0}{u} + \alpha_0\right)$. Soit $B_i \sim B \sim \text{Geo}(1 - \alpha_1)$ avec $E[B] = \frac{1}{1 - \alpha_1}$, $\mathcal{L}_B = \frac{1 - \alpha_1}{e^t - \alpha_1}$ et $\mathcal{L}_B^{-1} = \ln\left(\frac{1 - \alpha_1}{u} + \alpha_1\right)$. Avec $\alpha_0 = 0.5$, $\alpha_1 = 0.3$, $N \sim \text{Bin}(n = 4, q = 2/5)$ et $X_i \sim X \sim \text{Exp}(\beta = 1/100)$, on obtient le sommaire des simulations présenté dans le tableau 16 ainsi que les illustrations 17 et 18.

N	X_1	X_2	X_3	X_4
Min. :0.000	Min. : 0.0087	Min. : 0.0062	Min. : 0.0068	Min. : 0.0003
1st Qu. :1.000	1st Qu. : 29.1523	1st Qu. : 28.7089	1st Qu. : 28.6573	1st Qu. : 27.7252
Median :2.000	Median : 70.0952	Median : 68.7772	Median : 70.0095	Median : 67.3220
Mean :1.602	Mean : 100.5380	Mean :100.13362	Mean : 99.6027	Mean : 99.9894
3rd Qu. :2.000	3rd Qu. : 140.2621	3rd Qu. :140.8914	3rd Qu. : 139.3071	3rd Qu. : 139.6741
Max. :4.000	Max. :1471.4156	Max. :936.3070	Max. :1080.3850	Max. :1012.0754

Tableau 16 – Sommaire de 10 000 simulations de $\{N, X_1, \dots, X_4\}$ en supposant que F_{N, X_1, \dots, X_4} est définie avec une copule de archimédienne hiérarchique géométrique-géométrique($\alpha_0 = 0.5, \alpha_1 = 0.3$), $N \sim Bin(4, 2/5)$ et $X_i \sim X \sim Exp(1/100)$, pour $i = 1, \dots, 4$.

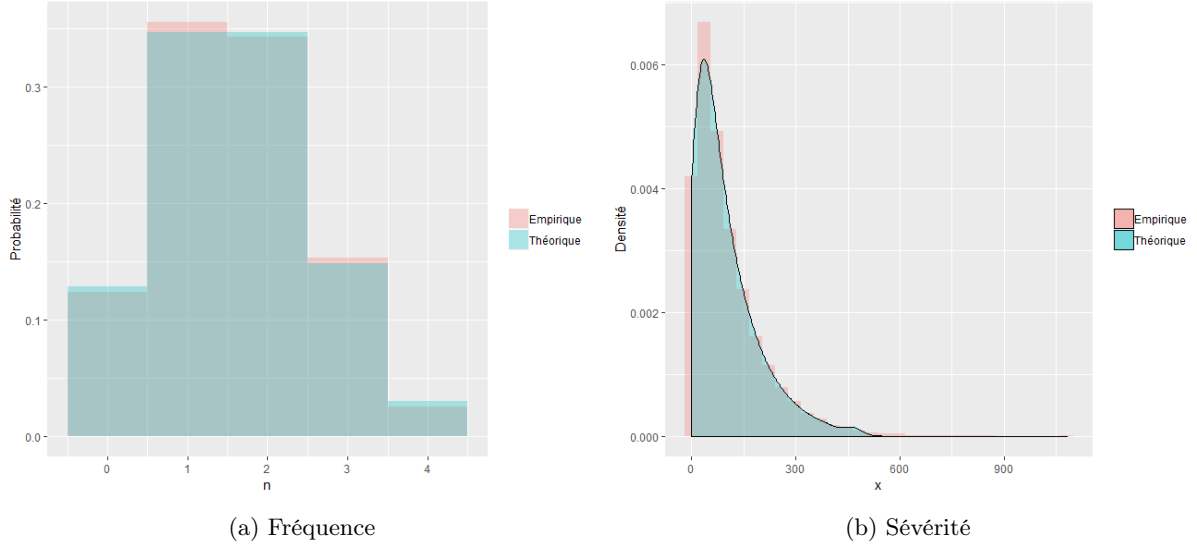


Illustration 17 – Comparaisons de la distribution marginale des données simulées avec les distributions théoriques pour le scénario 8.

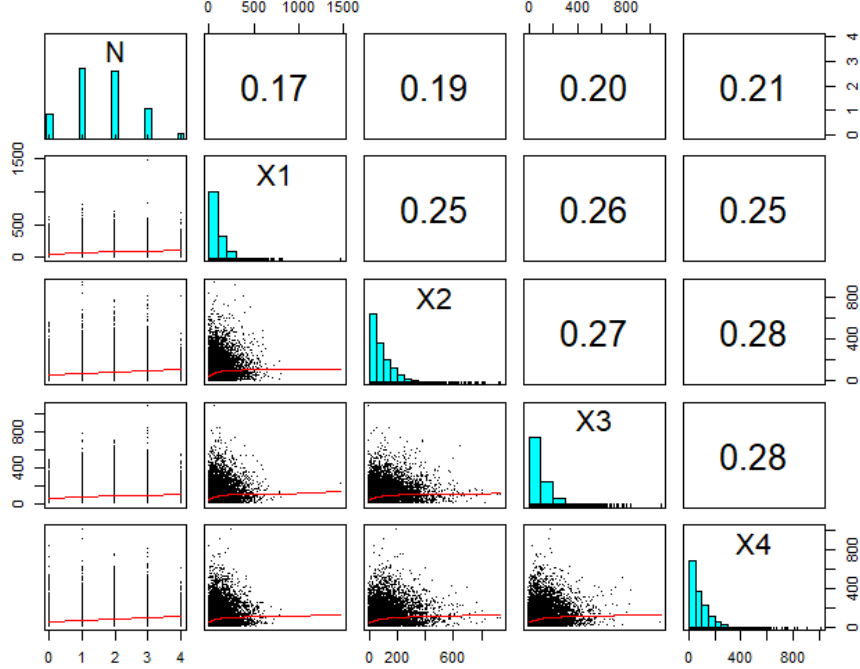


Illustration 18 – Nuages de points avec copule archimédienne hiérarchique géométrique-géométrique($\alpha_0 = 0.5, \alpha_1 = 0.3$), $N \sim \text{Bin}(4, 2/5)$ et $X_i \sim X \sim \text{Exp}(1/100)$, pour $i = 1, \dots, 4$. : En bas de la diagonale se trouve les nuages de points illustrant la corrélation des différentes variables. En haut, on indique les coefficients de corrélation de Spearman. À noter que ce graphique est produit avec la fonction `R pairs.panels` du module `psych`. Dans cette dernière, le calcul des coefficients de corrélation sont calculés selon l’hypothèse que les variables sont continues. Cela implique que la première ligne n’est pas valide.

Les valeurs de départ de l’optimisation numérique ainsi que les résultats de l’estimation sont présentés dans le tableau 17.

	q	β	α_0	α_1		
Valeurs de départ	0.4005	0.0091	0.5000	0.5000	Temps de dérivation	15.87
Estimateurs	0.4008	0.0099	0.5677	0.4564	Temps d’estimation	305.33
Vrais paramètres	0.4000	0.0100	0.5000	0.3000		

Tableau 17 – Résultats de l’estimation des paramètres avec une copule archimédienne hiérarchique géométrique-géométrique($\alpha_0 = 0.5, \alpha_1 = 0.3$), $N \sim \text{Bin}(4, 2/5)$ et $X_i \sim X \sim \text{Exp}(1/100)$, pour $i = 1, \dots, 4$, à la suite de 10 000 simulations.

4 Conclusion

Pour conclure, le modèle collectif du risque présenté dans [Cossette et al., 2019b] présente un défi dans l’estimation des paramètres par la méthode du maximum de vraisemblance tel que présentée dans la section 2.2 puisque la dérivation en chaîne sur un grand nombre de variables pose un problème de temps de calcul.

À cet effet, une solution envisageable pourrait être d’utiliser la vraisemblance par décomposition hiérarchique proposé dans [Cossette et al., 2019a] afin de trouver le paramètre de dépendance entre N et X_1 , puis de se limiter à un nombre restreint de X_i (disons 5) afin d’estimer le paramètre de dépendance entre les X_i . Cependant, avec cette méthode, le fait de travailler avec des variables continues et discrètes cause un problème.

De plus, la difficulté à trouver une valeur de départ adéquate pour α_0 et α_1 fait en sorte que les estimateurs de ces paramètres manquent de précision.

Références

- [Cossette et al., 2017] Cossette, H., Gadoury, S.-P., Marceau, E., and Mtalai, I. (2017). Hierarchical archimedean copulas through multivariate compound distributions. *Insurance : Mathematics and Economics*, 76.
- [Cossette et al., 2019a] Cossette, H., Gadoury, S.-P., Marceau, E., and Robert, C. Y. (2019a). Composite likelihood estimation method for hierarchical archimedean copulas defined with multivariate compound distributions. *Journal of Multivariate Analysis*, 172.
- [Cossette et al., 2019b] Cossette, H., Marceau, E., and Mtalai, I. (2019b). Collective risk models with dependance.
- [Klugman et al., 2012] Klugman, S., Panjer, H., and Willmot, G. (2012). *Loss Models : From Data to Decisions*. Wiley, fourth edition.