Sujets spéciaux III

ACT-7008

Projet 3
Distributions multivariées et dépendance

Par
Olivier Côté

Numéro d'identification
111 250 315

Rapport présenté à

Monsieur

Etienne Marceau

1^{er} mai 2022

Table des matières

1 Question 1

On commence avec un vecteur aléatoire de variables aléatoires discrètes $M = (M_1, \dots, M_n)$ dont la fonction génératrice des probabilité est définie comme suit :

$$\mathcal{P}_{\underline{M}}(s_1, \dots, s_n) = \exp\left\{\frac{\lambda}{1+\alpha}\left(\alpha(s_1-1) + \alpha(s_n-1) + (1-\alpha)\sum_{i=1}^{n}(s_i-1) + \alpha\sum_{i=1}^{n-1}(s_i s_{i+1}-1)\right)\right\} \tag{1}$$

Pour tout $|s_i| \leq 1$ et $i \in A_n = \{1, \dots, n\}$.

On a ensuite un vecteur aléatoire $\underline{X} = (X_1, \dots, X_n)$ dont la transformée Laplace est

$$\mathcal{L}_{\underline{X}}(t_1, \dots, t_n) = \mathcal{P}_{\underline{M}}(\mathcal{L}_B(t_1), \dots, \mathcal{L}_B(t_1)) \tag{2}$$

Pour tout $|t_i| \geq 0$ et $i \in A_n = \{1, \dots, n\}$.

Avec l'équation 2, les hypothèses sont les suivantes :

- Indépendance entre le vecteur aléatoire $\underline{M}$ et le vecteur aléatoire $\underline{B}$.
- Indépendance entre les variables au sein du vecteur aléatoire $\underline{B}$.
- La dépendance entre les variables aléatoires du vecteur $\underline{X}$ est exclusivement introduite par la dépendance au sein du vecteur aléatoire $\underline{M}$.

1.1 Distributions marginales

On peut ensuite identifier la distribution Marginale de M_i. Pour l'exemple, je vais sélectionner $i = 2$, mais on peut facilement démontrer que la distribution est la même pour tout $i \in A_n$.

$$\mathcal{P}_{M_2}(s) = P_{\underline{M}}(1, s, \dots, 1) = \exp\left\{\frac{\lambda}{1+\alpha}\left[\alpha(s-1) + (1-\alpha)(s-1) + \alpha(s-1)\right]\right\}$$

On a donc

$$\mathcal{P}_{M_i}(s) = \exp\{\lambda(s-1)\} \tag{3}$$

Ensuite, pour la distribution marginale de X_i, on peut réarranger l'équation 2 de la sorte :

$$\begin{aligned}
\mathcal{L}_{X_i}(s) &= \mathcal{L}_{\underline{X}}(0, \dots, t, \dots, 0) \\
&= \mathcal{P}_{\underline{M}}(\mathcal{L}_B(0), \dots, \mathcal{L}_B(t), \dots, \mathcal{L}_B(0)) \\
&= \mathcal{P}_{\underline{M}}(1, \dots, \mathcal{L}_B(t), \dots, 1) \\
&= \mathcal{P}_{M_i}(\mathcal{L}_B(t))
\end{aligned}$$

On a donc

$$\mathcal{L}_{X_i}(s) = \mathcal{P}_{M_i}(\mathcal{L}_B(t)) \tag{4}$$

On peut ainsi dire que $X \sim \text{PoisComp}(\lambda, F_B)$. On ne spécifie rien sur la loi de la variable aléatoire B autre qu'elle est strictement positive et que son espérance existe.

1.2 Analyses des moments

On peut calculer l'expérance de la variable aléatoire X_i en se servant de l'indépendance entre M_i et B

$$E(X_i) = E(M_i)E(B)$$

Ensuite, pour la covariance entre X_i et $X_{i'}$, on doit premièrement avoir la covariance entre M_i et $M_{i'}$. Pour se faire, on identifie la distribution conjointe de M_i et $M_{i'}$ à l'aide de la fonction génératrice des probabilités conjointe.

Premièrement, on peut démontrer facilement à l'aide de l'équation 1 que si $|i-i'| > 1$, on a indépendance entre M_i et $M_{i'}$:

$$\mathcal{P}_{M_i,M_{i'}}(s_i, s_{i'}) = \mathcal{P}_{M_i}(s_i)\mathcal{P}_{M_{i'}}(s_{i'})$$

Par contre, pour $|i - i'| = 1$, on aura (on suppose $i, i' \in \{2, \dots, n-1\}$ sans perte de généralité) :

$$\begin{aligned}
\mathcal{P}_{M_i,M_{i'}}(s_i, s_{i'}) &= \mathcal{P}_{\underline{M}}(1, \dots, s_i, s_{i'}, \dots, 1) \\
&= \exp\left\{\frac{\lambda}{1+\alpha}\left[(1-\alpha)(s_i - 1) + (1-\alpha)(s_{i'} - 1) + \alpha(s_i - 1) + \alpha(s_{i'} - 1) + \alpha(s_i s_{i'} - 1)\right]\right\} \\
&= e^{\frac{\lambda}{1+\alpha}(s_i - 1)} e^{\frac{\lambda}{1+\alpha}(s_{i'} - 1)} e^{\frac{\alpha\lambda}{1+\alpha}(s_i s_{i'} - 1)}
\end{aligned}$$

À la vue de la dernière équation, on reconnait un choc commun. On peut donc exprimer le couple aléatoire $(M_i, M_{i'})$ de la sorte :

$$(M_i, M_{i'}) = J_{ii'} + (J_i, J_{i'}) \tag{5}$$

Où

$$\begin{aligned}
J_i \sim J_{i'} &\sim Pois\left(\frac{\lambda}{1+\alpha}\right) \\
J_{ii'} &\sim Pois\left(\frac{\alpha\lambda}{1+\alpha}\right) \\
J_i, J_{i'}, J_{ii'} &\text{ mutuellement indépendants}
\end{aligned}$$

On peut maintenant calculer la covariance désirée pour $|i - i'| > 1$:

$$Cov(M_i, M_{i'}) = Cov(J_i + J_{ii'}, J_{i'} + J_{ii'})$$
$$Cov(J_{ii'}, J_{ii'}) = Var(J_{ii'}) = \frac{\alpha\lambda}{1+\alpha}$$

On peut donc écrire, de manière plus générale :

$$Cov(M_i, M_{i'}) = \begin{cases} 0, & |i - i'| > 1 \\ \frac{\alpha\lambda}{1+\alpha}, & |i - i'| = 1 \end{cases} \tag{6}$$

On peut maintenant calculer la covariance entre X_i et $X_{i'}$

$$\begin{aligned}
Cov(X_i, X_{i'}) &= E[Cov(X_i, X_{i'}|\underline{M})] + Cov(E[X_i|\underline{M}], E[X_{i'}|\underline{M}]) \\
&= E[\sum_{j=1}^{M_i} \sum_{k=1}^{M_{i'}} Cov(B_{ji}, B_{ki})] + Cov(M_i E(B_i), M_{i'} E(B_{i'})) \\
&= 0 + E(B)^2 Cov(M_i, M_{i'})
\end{aligned}$$

En combinant cette dernière expression avec l'équation 6, on obtient

$$Cov(X_i, X_{i'}) = \begin{cases} 0, & |i - i'| > 1 \\ E(B)^2 \frac{\alpha\lambda}{1+\alpha}, & |i - i'| = 1 \end{cases} \tag{7}$$

1.3 Convergence d'une variable aléatoire

Soit W_n s'écrivant de plusieurs manières différentes :

$$\begin{aligned}
W_n = \frac{1}{n} S_n &= \frac{1}{n} \sum_{i=1}^{n} X_i \\
&= \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{M_i} B_{ji}
\end{aligned}$$

Avec la convention $\sum_{i=1}^{0} = 0$.

Pour l'espérance de W_n, on a donc

$$E[W_n] = \frac{1}{n} \sum_{i=1}^{n} E(X_i) = \lambda a$$

où $E(B) = a$. Ensuite, pour la variance, on a (en supposant $E(B^2) < \infty$

$$\begin{aligned}
Var(W_n) &= \frac{1}{n^2} Var(\sum_{i=1}^{n} X_i) \\
&= \frac{1}{n^2} \left\{ \sum_{i=1}^{n} Var(X_i) + 2 \sum_{i=2}^{n} \sum_{j=1}^{i-1} Cov(X_i, X_j) \right\} \\
&= \frac{1}{n^2} \left\{ \sum_{i=1}^{n} Var(X_i) + 2 \sum_{i=2}^{n} Cov(X_i, X_{i-1}) \right\} \\
&= \frac{1}{n^2} \left\{ n\lambda(B^2) + 2(n-1)E(B)^2 \frac{\alpha\lambda}{1+\alpha} \right\} \\
&= \frac{1}{n} \left\{ \lambda(B^2) + 2(1 - \frac{1}{n})E(B)^2 \frac{\alpha\lambda}{1+\alpha} \right\}
\end{aligned}$$

La variance a donc une forme similaire à la fonction $g(.)$ suivante

$$g(n) = \frac{1}{n} \left(b + c \cdot (1 - \frac{1}{c}) \right)$$

où b et c sont des constantes.

On remarque donc rapidement que

$$\lim_{n\to\infty} g(n) = 0$$

et donc que

$$\lim_{n\to\infty} Var(W_n) = 0 \tag{8}$$

On aura besoin de cette relation plus tard.

On rappelle l'inégalité de Markov, démontrée par moi-même pour le projet 1 :

$$P(X \geq a) \leq \frac{E(g(X))}{g(a)} \tag{9}$$

Pour l'appliquer à notre situation, on remplace X par $(W_n - E(W_n))^2$, g(.) par la fonction identité et a par ε^2 (où $\varepsilon > 0$). On obtient alors

$$\begin{aligned}
0 \leq P((W_n - E(W_n))^2 \geq \ \varepsilon^2) &\leq \frac{E((W_n - E(W_n))^2)}{\varepsilon^2} \\
0 \leq P(|W_n - E(W_n)| \geq \ \varepsilon) &\leq \frac{Var(W_n)}{\varepsilon^2} \\
\lim_{n\to\infty} 0 \leq \lim_{n\to\infty} P(|W_n - E(W_n)| \geq \ \varepsilon) &\leq \lim_{n\to\infty} \frac{Var(W_n)}{\varepsilon^2} \\
0 \leq \lim_{n\to\infty} P(|W_n - E(W_n)| \geq \ \varepsilon) &\leq 0
\end{aligned}$$

La dernière relation provient de l'équation 8. Par le théorème du sandwich, on déduit

$$\lim_{n\to\infty} P(|W_n - E(W_n)| \geq \ \varepsilon) = 0$$

Ce qui est la condition nécessaire et suffisante pour dire que

$$W_n \xrightarrow{P} E(W_n) = a\lambda$$

1.4 Distribution globale

On peut s'intéresser à la distribution globale du nombre de sinistre, que l'on dénotera N_n et à la distribution globale du montant de sinistre S_n.

$$N_n = \sum_{i=1}^{n} M_i$$

On peut donc réécrire S_n

$$S_n = \sum_{i=1}^{N_n} B_i$$

Avec la convention $\sum_{i=1}^{0} = 0$.

On peut s'intéresser à la distribution de N_n

$$\begin{aligned}
\mathcal{P}_{N_n}(s) &= \mathcal{P}_{\underline{M}}(s, \dots, s) \\
&= \exp\left\{\frac{\lambda}{1+\alpha}\left[2(s-1) + (1-\alpha)n(s-1) + \alpha(n-1)(s^2-1)\right]\right\} \\
&= \exp\left\{\frac{n\lambda}{1+\alpha}\left[(\alpha - \frac{\alpha}{n})s^2 + (\frac{2}{n} + (1-\alpha))s + \frac{\alpha-2}{n} - 1\right]\right\} \\
&= \mathcal{P}_F\left(\mathcal{P}_K(s)\right)
\end{aligned}$$

où

$$\begin{aligned}
\mathcal{P}_F(s) &= e^{\frac{n\lambda}{1+\alpha}(s-1)} \\
\mathcal{P}_K(s) &= \left(\alpha - \frac{\alpha}{n}\right)s^2 + \left(\frac{2}{n} + (1-\alpha)\right)s + \frac{\alpha-2}{n}
\end{aligned}$$

On peut vérifier que les coefficients de la fonction $\mathcal{P}_K(s)$ somment bien à 1. On déduit donc que

$$N_n \sim \text{Poiscomp}\left(\frac{n\lambda}{1+\alpha}, F_K\right)$$

Pour interpréter ce résultat, on pourrait dire que F est le nombre de sinistres, alors que K est le nombre de réclamant pour un sinistre. N_n serait donc le nombre total de réclamant. Il s'agit évidemment seulement d'un exemple d'interprétation.

Pour S_n, c'est un peut plus facile puisqu'on sait que la fréquence $\underline{M}$ et la sévérité $\underline{B}$ sont indépendantes. On a

$$\begin{aligned}
\mathcal{L}_{S_n}(s) &= \mathcal{P}_{N_n}\left(\mathcal{L}_B(s)\right) \\
&= \mathcal{P}_F\left(\mathcal{P}_K\left(\mathcal{L}_B(s)\right)\right)
\end{aligned}$$

On a donc

$$S_n \sim \text{Poiscomp}\left(\frac{n\lambda}{1+\alpha}, F_C\right)$$

où

$$\mathcal{P}_C(s) = \mathcal{P}_K\left(\mathcal{P}_B(s)\right)$$

1.5 Borne basé sur la mesure entropique

On commence par définir la mesure entropique

$$\psi_{W_n}(\rho) = \frac{1}{\rho}\ln\left(M_{W_n}(\rho)\right), \quad \rho \in (0, t_0)$$

Si on récupère l'équation 9 avec $g(x) = e^{\rho x}$, $X = W_n$ et $a = \psi_{W_n}(\rho) + u$, alors on a

$$\begin{aligned}
P\left(W_n > \psi_{W_n}(\rho) + u\right) &\leq M_{W_n}(\rho)e^{-\rho\left(\frac{1}{\rho}\ln M_{W_n}(\rho) + u\right)} \\
&= M_{W_n}(\rho)M_{W_n}(\rho)^{-1}e^{-\rho u} \\
&= e^{-\rho u}
\end{aligned}$$

Maintenant, si on veut développer la mesure entropique pour W_n, il faut spécifier la distribution de B. Si $B \sim Gamma(\eta, \beta)$, on peut débuter.

$$\begin{aligned}
M_{W_n}(t) &= M_S\left(\frac{t}{n}\right) \\
&= \mathcal{P}_F\left(\mathcal{P}_K\left(M_B\left(\frac{t}{n}\right)\right)\right) \\
&= \exp\left\{\frac{n\lambda}{1+\alpha}\left[\mathcal{P}_K\left(\left(\frac{n\beta}{n\beta-t}\right)^{\eta}\right)-1\right]\right\}
\end{aligned}$$

On a donc la condition $t \in [-\infty, \beta n]$. Cette condition se transposera directemetn dans notre mesure entropique pour ρ.

Maintenant, on peut écrire la forme explicite pour $\eta = 1$:

$$\mathcal{P}_K\left(\frac{n\beta}{n\beta-t}\right) = \left(\alpha - \frac{\alpha}{n}\right)\left(\frac{n\beta}{n\beta-t}\right)^2 + \left(\frac{2}{n} + (1-\alpha)\right)\frac{n\beta}{n\beta-t} + \frac{\alpha-2}{n}$$

On peut insérer cette expression dans les expression précédentes pour obtenir $M_{W_n}(t)$.

En sautant les développements [1], on obtient :

$$\psi_{W_n}(\rho) = \frac{n\lambda}{\rho(1+\alpha)}\left[\left(\alpha - \frac{\alpha}{n}\right)\left(\frac{n\beta}{n\beta-t}\right)^2 + \left(\frac{2}{n} + (1-\alpha)\right)\frac{n\beta}{n\beta-t} + \frac{\alpha-2}{n} - 1\right]$$

1.6 Exemple : Mélange d'erlang

On fait maintenant un exemple plus complet avec $B \sim MixErl(\underline{\gamma}, \beta)$ et $\gamma_j = P(J = j) \quad \forall j \in \{1, 2, ...\}$. Voici une représentation de D

$$\mathcal{L}_B(t) = \mathcal{P}_J\left(\mathcal{L}_D(t)\right) \quad t \leq 0$$

où $D \sim Exp(\beta)$.

Pour S_n, on a donc

$$\begin{aligned}
\mathcal{L}_{S_n} &= \mathcal{P}_{N_n}\left(\mathcal{L}_B(t)\right) \\
&= \mathcal{P}_F\left(\mathcal{P}_K\left(\mathcal{P}_J\left(\mathcal{L}_D(t)\right)\right)\right) \\
&= \mathcal{P}_{N^*}\left(\mathcal{L}_D(t)\right)
\end{aligned}$$

On a donc une « fréquence modifiée »

$$\mathcal{P}_{N^*}(t) = \mathcal{P}_F\left(\mathcal{P}_K\left(\mathcal{P}_J\left(t\right)\right)\right)$$

Ainsi, soit le vecteur de poids $\underline{\nu}$, où :

$$\nu_j = P\left(N^* = j\right) = P\left(\sum_{i=1}^{F}\sum_{k=1}^{K_i} J_{ik} = j\right)$$

On peut interpréter N^* de la sorte :

1. malheureusement

— F est le nombre de sinistre

— K est le nombre de réclamant sur un dossier

— J est le nombre de blessure pour un réclamant sur un dossier

Ainsi, N^* serait la quantité totale de blessure, et D serait la sévérité d'une blessure.

On arrive naturellement à un mélange d'erlang :

$$F_{S_n}(x) = \nu_0 + \sum_{k=1}^{n} \nu_k H(x, k, \beta)$$

Ensuite, puisque

$$M_{Wn}(t) = M_{S_n}(\frac{t}{n})$$

nous allons avoir simplement $D^* \sim Exp(n\beta)$. On touve ainsi la fonction de répartition suivante pour W_n :

$$F_{W_n}(x) = \nu_0 + \sum_{k=1}^{n} \nu_k H(x, k, n\beta) \tag{10}$$

Finalement, pour notre mélange d'Erlang, on peut développer la mesure entropique [2]

$$\psi_{W_n}(\rho) = \frac{n\lambda}{\rho(1+\alpha)} \left[\left(\alpha - \frac{\alpha}{n}\right) \mathcal{P}_J \left(\frac{n\beta}{n\beta - t}\right)^2 + \left(\frac{2}{n} + (1-\alpha)\right) \mathcal{P}_J \left(\frac{n\beta}{n\beta - t}\right) + \frac{\alpha - 2}{n} - 1 \right]$$

1.7 Résultats numériques

On débute avec certains paramètres pour l'évaluation des métriques et pour le mélange d'erlang.

— $\lambda = \frac{1}{\tan(\pi/8)}$

— $\alpha \in \{0, 0.5, 0.99\}$

— $\beta = 0.1$

— $\rho \in \{0.01, 0.05\}$

— $n \in \{1, 10, 100, 1000\}$

— $(\gamma_1, \dots, \gamma_5) = (0.3, 0.2, 0.1, 0.15, 0.25)$

— $\kappa \in \{0.01, 0.5, 0.99\}$

On commence par l'espérance de W_n. On rappelle que l'espérance de W_n est égale à $\lambda E(B)$. Cette quantité est donc invariante à n et α. On obtient $E(W_n) = 68.81$.

Ensuite, on rappelle que l'expression de la variance $Var(W_n)$ variait selon n et α. On peut observer la valeur de $Var(W_n)$ selon les différents paramètres au tableau 1. On constate que pour un n fixé (et $n > 1$), la variance augmente avec α. On constate aussi que pour un α fixé, la variance diminue lorsque n augmente.

On peut aussi évaluer la mesure entropique $\psi_{W_n}(\rho)$ aux valeurs $\rho \in \{0.01, 0.05\}$. La mesure $\psi_{W_n}(\rho)$ varie selon α et n. Le tableau 2 présente les résultats de $\psi_{W_n}(0.01)$ selon α et n, alors que le tableau 3 présente les résultats de $\psi_{W_n}(0.05)$ selon α et n

Tableau 1 – $Var(W_n)$ selon α et n

		α		
		0	0.5	0.99
n	1	3259.19	3259.19	3259.19
	10	325.92	443.58	501.52
	100	32.59	45.53	51.91
	1000	3.26	4.57	5.21

Tableau 2 – $\psi_{W_n}(0.01)$ selon α et n

		α		
		0	0.5	0.99
n	1	267.56	148.64	90.08
	10	84.56	75.78	71.46
	100	70.35	69.49	69.07
	1000	68.96	68.87	68.83

Ensuite, en se servant de l'équation 10, on peut évaluer la $VaR_\kappa(W_n)$ à l'aide d'outils numériques. La $VaR_\kappa(W_n)$ varie selon κ, α et n. Le tableau 4 présente les résultats de $VaR_{0.01}(W_n)$ selon α et , le tableau 5 présente les résultats de $VaR_{0.5}(W_n)$ selon α et et le tableau 6 présente les résultats de $VaR_{0.99}(W_n)$ selon α et .

Finalement, on peut calculer la $TVaR_\kappa(W_n)$ selon les différentes valeurs de κ, α et n. On rappelle que, pour W_n, l'expression de $TVaR_\kappa(W_n)$ est la suivante

$$TVaR_\kappa(W_n) = \frac{1}{1-\kappa}\sum_{k=1}^{\infty}\nu_k \frac{k}{\beta n}\overline{H}(VaR_\kappa(W_n); k+1, \beta n)$$

On peut voir la $TVaR_{0.01}(W_n)$, $TVaR_{0.5}(W_n)$ et la $TVaR_{0.99}(W_n)$ en fonction de α et n aux tableaux 7, 8 et 9, respectivement. Les valeurs qui ne fonctionnent pas et qui aurait mérité une plus grande investigation ont été identifiés en rouge.

2. Partiellement, j'ai laissé la fonction génératrice des probabilités de J pour alléger la notation.

TABLEAU 3 – $\psi_{W_n}(0.05)$ selon α et n

		α		
		0	0.5	0.99
n	1	1680.29	933.50	565.73
	10	93.45	86.83	83.57
	100	71.02	70.42	70.13
	1000	69.02	68.97	68.94

TABLEAU 4 – $VaR_{0.01}(W_n)$ selon α et n

		α		
		0	0.5	0.99
n	1	26.63	0.00	0.00
	10	41.30	29.59	23.81
	100	57.23	54.16	52.74
	1000	64.78	10.92	63.56

TABLEAU 5 – $VaR_{0.5}(W_n)$ selon α et n

		α		
		0	0.5	0.99
n	1	196.01	104.13	58.80
	10	81.54	72.02	67.37
	100	70.08	69.12	68.66
	1000	179.51	180.06	179.44

TABLEAU 6 – $VaR_{0.99}(W_n)$ selon α et n

		α		
		0	0.5	0.99
n	1	478.72	327.60	243.63
	10	603.85	720.11	603.85
	100	1179.21	1080.97	1174.53
	1000	319.95	320.73	321.02

Tableau 7 – $TVaR_{0.01}(W_n)$ selon α et n

		α		
		0	0.5	0.99
n	1	205.73	115.83	70.20
	10	83.03	73.88	41.79
	100	70.33	69.44	69.00
	1000	68.99	69.55	68.87

Tableau 8 – $TVaR_{0.5}(W_n)$ selon α et n

		α		
		0	0.5	0.99
n	1	284.62	172.55	114.06
	10	98.31	90.55	86.69
	100	74.78	74.66	74.56
	1000	0.00	0.00	0.00

Tableau 9 – $TVaR_{0.99}(W_n)$ selon α et n

		α		
		0	0.5	0.99
n	1	529.36	369.42	279.66
	10	0.00	0.00	0.00
	100	0.00	0.00	0.00
	1000	0.00	0.00	0.00

2 Question 2

Pour ce problèmes on commence avec un vecteur aléatoire

$$\underline{X} = (X_1, X_2, X_3, X_4, X_5)$$

où $X_i \sim Bin(10, 0.05i)$ On a aussi la fonction de répartition conjointe suivante

$$F_{\underline{X}}(x_1, x_2, x_3, x_4, x_5) = C_0\left(C_1\left(F_{X_1}(x_1), F_{X_2}(x_2)\right), C_2\left(F_{X_3}(x_3), F_{X_4}(x_4), F_{X_5}(x_5)\right)\right) \tag{11}$$

où C_j correspond à la copule AMH avec $\alpha_j = 0.1j + 0.2$. On rappelle que, pour une copule archimédienne, on a indépendance conditionnelle par rapport à une variable aléatoire Θ. Pour une copule AMH, on peut avoir indépendance entre les composantes si on conditionne par rapport à Θ, où $\Theta \sim Geo(1-\alpha)$ (nombre d'essais).

On a donc 5 variables aléatoires dont la structure de dépendance est définie par des copules archimédiennes imbriquées.

On définit $S = X_1 + ... + X_5$. On s'intéresse maintenant à la fonction de masse de probabilité de S, soit $f_S(s)$.

2.1 Fonction de masse de probabilité de S

Premièrement, grâce à la loi des probabilités totales, on peut facilement écrire ceci

$$f_S(k) = \sum_{\theta_0=1}^{\infty} f_{S|\Theta_0=\theta_0}(k) f_{\Theta_0}(\theta_0)$$

Par la suite, en se servant du fait que la variable aléatoire Θ_0 régit la relation de dépendance de C_0 dans l'équation 11, on peut écrire

$$f_S(k) = \sum_{\theta_0=1}^{\infty} f_{X_1+X_2|\Theta_0=\theta_0} * f_{X_3+X_4+X_5|\Theta_0=\theta_0}(k) f_{\Theta_0}(\theta_0) \tag{12}$$

Cette équation servira à calculer la fonction de masse de probabilité de S. Par contre, il manque à identifier des expressions manipulables pour $f_{X_1+X_2|\Theta_0=\theta_0}$ et $f_{X_3+X_4+X_5|\Theta_0=\theta_0}$ afin que l'équation 12 nous permettent réellement de calculer $f_S(s)$.

Pour $f_{X_1+X_2|\Theta_0=\theta_0}$, la loi des probabilités totales nous permet d'écrire

$$f_{X_1+X_2|\Theta_0=\theta_0}(k) = \sum_{\theta_{0,1}=1}^{\infty} f_{X_1+X_2|\Theta_0=\theta_0,\Theta_{0,1}=\theta_{0,1}}(k) f_{\Theta_{0,1}}(\theta_{0,1})$$

où $Theta_{0,1}$ est la variable aléatoire qui régi la structure de dépendance de la copule AMH C_1 de l'équation 11 conditionnellement à la réalisation de la variable aléatoire Θ_0 qui régit la structure de dépendance de C_0. Les variables X_1 et X_2 sont donc indépendantes conditionnellement à $\Theta_{0,1}$ et Θ. Dans notre cas, $(\Theta_{0,1}|\Theta_0 = \theta_0) \sim \text{BinNeg}(\theta_0, q^* = \frac{1-\alpha_i}{1-\alpha_0})$. On peut donc réécrire :

$$f_{X_1+X_2|\Theta_0=\theta_0}(k) = \sum_{\theta_{0,1}=1}^{\infty} f_{X_1|\Theta_0=\theta_0,\Theta_{0,1}=\theta_{0,1}} * f_{X_2|\Theta_0=\theta_0,\Theta_{0,1}=\theta_{0,1}}(k) f_{\Theta_{0,1}}(\theta_{0,1}) \tag{13}$$

Finalement, pour identifier $f_{X_1|\Theta_0=\theta_0,\Theta_{0,1}=\theta_{0,1}}$ et $f_{X_2|\Theta_0=\theta_0,\Theta_{0,1}=\theta_{0,1}}$ dans cette expression, on a besoin de l'expression suivante de Cossette et collab. (2018)

$$F_{X_i|\Theta_0=\theta_0,\Theta_{0,1}=\theta_{0,1}}(x) = e^{-\theta_{0,1}\mathcal{L}^{-1}_{\Theta_1}\left(F_{X_i}(x)\right)} \quad \forall i \in \{1,2\} \tag{14}$$

On peut obtenir des expressions similaires pour calculer récursivement les valeurs de la fonction de masse de probabilité de $f_{X_3+X_4+X_5|\Theta_0=\theta_0}$. Ainsi, en en commeçant seulement avec les marges $F_{X_i}(x)$ et la structure de dépendance de l'équation 11, on a pu identifier une méthodologie qui permet de calculer les valeurs de la fonction de masse de probabilité de S tout en faisant usange des algorithmes d'agrégations connus.

On commence avec les marges $F_{X_i}(x_i)$, on calcule ensuite les marges conditionnelles $f_{X_i|\Theta_0=\theta_0,\Theta_{0,1}}$ à l'aide de l'équation 14. Par la suite, on peut monter d'un étage dans notre agrégation des risques à l'aide de l'équation 13. Finalement, avec les fonctions de masses de probabilités agrégées, on peut identifier $f_S(s)$ à l'aide de l'équation 12. Cette méthodologie s'implante très bien en R et ne cause aucun enjeu computationnel puisque nous sommes en faibles dimensions (nous n'avons que 5 variables aléatoires dont le support ne contient que 11 valeurs).

À chaque étape où une convolution devait être faire, j'ai utilisé l'algorithme fft. La fonction de masse de probabilité obtenue est celle de la figure 1.

2.2 Covariance entre deux variables X_i et X_j

On s'intéresse maintenant à la covariance entre deux variables aléatoires X_i et X_j pour lesquelles $i \neq j$ (sinon, c'est sans intérêt).

La covariance s'écrit comme suit

$$Cov(X_i, X_j) = E(X_iX_j) - E(X_i)E(X_j)$$

Le seul élément qui posera problème est $E(X_iX_j)$. Nous devons identifier la fonction de répartition conjointe $F_{X_i,X_j}(x_i, x_j)$ pour arriver à calculer les quantités désirées.

On peut identifier cette fonction de répartition conjointe de la sorte :

$$\begin{aligned}
F_{X_1,X_2}(x_1,x_2) &= \lim_{(x_3,x_4,x_5)\to(\infty,\infty,\infty)} F_{\underline{X}}(x_1,x_2,x_3,x_4,x_5) \\
&= \lim_{(x_3,x_4,x_5)\to(\infty,\infty,\infty)} C_0\left(C_1\left(F_{X_1}(x_1),F_{X_2}(x_2)\right),C_2\left(F_{X_3}(x_3),F_{X_4}(x_4),F_{X_5}(x_5)\right)\right) \\
&= C_0\left(C_1\left(F_{X_1}(x_1),F_{X_2}(x_2)\right),C_2\left(1,1,1\right)\right) \\
&= C_0\left(C_1\left(F_{X_1}(x_1),F_{X_2}(x_2)\right),1\right) \\
&= C_1\left(F_{X_1}(x_1),F_{X_2}(x_2)\right)
\end{aligned}$$

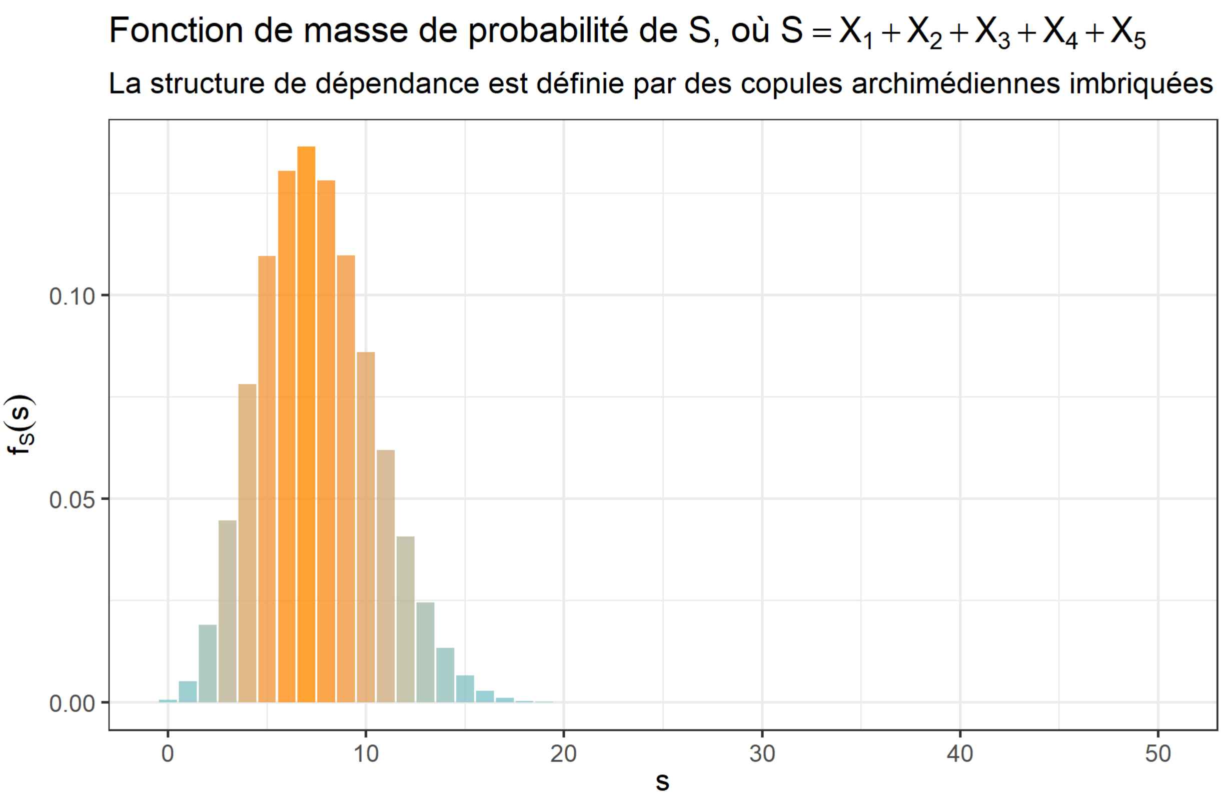

FIGURE 1 – fonction de masse de probabilité de S

On obtient donc des expressions sympathiques à travailler. On peut obtenir des expressions similaires pour les autres couples de variables aléatoires (X_i, X_j). Par exemple, on a

$$\begin{aligned} F_{X_1,X_3}(x_1,x_3) &= \lim_{(x_2,x_4,x_5)\to(\infty,\infty,\infty)} F_{\underline{X}}(x_1,x_2,x_3,x_4,x_5) \\ &= C_0\left(C_1\left(F_{X_1}(x_1),1\right), C_2\left(F_{X_3}(x_3),1,1\right)\right) \\ &= C_0\left(F_{X_1}(x_1), F_{X_3}(x_3), 1\right) \end{aligned}$$

Avec ces fonctions de répartitions conjointes en main, on peut facilement calculer la matrice de covariance entre X_i et X_j. Les valeurs calculées sont présentées au tableau 10

TABLEAU 10 – Matrice $Cov(X_i, X_j)$ pour le problème 2.

		i				
		1	2	3	4	5
	1	0.4750	0.0518	0.0418	0.0474	0.0517
	2	0.0518	0.9000	0.0637	0.0723	0.0789
j	3	0.0418	0.0637	1.2750	0.1860	0.2032
	4	0.0474	0.0723	0.1860	1.6000	0.2311
	5	0.0517	0.0789	0.2032	0.2311	1.8750

2.3 Mesures de risque pour S

Maintenant qu'on connait la fonction de masse de probabilité de S, on peut s'intéresser à calculer diverses quantités d'intérêt pour S.

On débute avec l'espérance de S, on obtient une valeur de 7.5, ce qui semble cohérent avec le graphique de la figure 1.

On calcule ensuite $\sqrt{Var(S)}$, ce qui nous donne environ 8.1807.

Ensuite, on peut calculer la $VaR_\kappa(S)$ et la $TVaR_\kappa(S)$ pour diverses valeurs de κ. Ces quantités sont calculés au tableau 11.

TABLEAU 11 – Quantités d'intérêt pour la variable aléatoire S

	κ	$VaR_\kappa(S)$	$TVaR_\kappa(S)$
0.5	0.50	7	9.79
0.9	0.90	11	12.82
0.95	0.95	12	13.85
0.99	0.99	15	15.74
0.995	0.99	15	16.47
0.9995	1.00	18	18.52

Malheureusement, manque de temps.

3 Question 3

Malheureusement, manque de temps.

Références

Cossette, H., E. Marceau, I. Mtalai et D. Veilleux. 2018, « Dependent risk models with archimedean copulas : A computational strategy based on common mixtures and applications », *Insurance : Mathematics and Economics*, vol. 78, p. 53–71. 13