Feuille de TP 2 – Copules archimédiennes

Indication générale : lorsqu'on demande N simulations d'un vecteur X de dimension d, on fera attention à donner les simulations en ligne : les simulations seront stockées dans une matrice de dimension $N \times d$ où chaque ligne est une simulation du vecteur X.

1 Simulation de copules avec R

Pour R, il existe différents packages qui permettent de simuler des copules. En voici quelques uns :
- `copula` : assez simple d'utilisation.
- `fCopulae` : beaucoup plus complet (mais un peu plus difficile à utiliser).
- `gumbel` : spécifique pour la copule de Gumbel (très utilisée en assurance).

Le fichier pdf décrivant le package `fCopulae` est disponible sur la page UMTICE du module. Il regroupe les copules par famille
- `rarchmCopula` : simule une copule archimédienne. Il y en a 22 possibles, mais les principales sont celles de Clayton (1), de Gumbel (4), et de Franck (5). Le descriptif complet se trouve aussi sur UMTICE.
- `rellipticalCopula` : simule une copule elliptique avec trois types possibles dont "norm" (gaussienne), "t" (Student).

Exercice 1.1.
1. Utiliser la commande `rarchmCopula` pour simuler toujours (X, Y) mais où la corrélation est donnée par une copule de Clayton, de Gumbel, puis une copule de Franck. Tracer le nuage de points.
2. Dans les trois cas, que se passe-t-il lorsque le paramètre θ augmente ? Les variables sont-elles corrélées positivement ou négativement ?
3. (Auto-Apprentissage) Utiliser la commande rellipticalCopula pour simuler un vecteur (X, Y) avec la copule adéquate et comparer avec le code obtenu dans les exercices 2.1, 2.2.

2 Une application en assurance

La copule de Gumbel est utilisée pour valoriser les couvertures indicielles catastrophe. Ces contrats sont des dérivés climatiques adaptés à la réassurance d'événements catastrophe (tempête, vague de froid, ...) basé sur un indice climatique (force du vent, température, ...). L'indice climatique doit refléter au mieux les caractérisques des montants des sinistres associés au risque météo pour diminuer le risque de base. En général, on choisit un panier de n stations (peu éloignées des régions assurées) dans lesquelles on mesure la variable climatique $X_i(t)$ au cours de la période $[t-1, t]$. Ensuite, l'indice journalier d'une station i est construit par $I_i(t) = \min(L_i - K_i, X_i(t) - K_i)$ où K_i et L_i sont le seuil et la limite par station. Sur une période T, l'indice d'une station est donc défini par $S_i(T) = \sum_{t=1}^{T} I_i(t)$ et l'indice cumulé par $S_T = \sum_{i=1}^{n} p_i S_i(T)$ pour une pondération $p_1, \ldots, p_n$. Enfin le flux engendré par la couverture indicielle est celui d'un call spread :

$$C_T = \mathbb{E}\left[N \times \min(L - K, (S_T - K)^+)\right],$$

où K et L sont la franchise et la limite du contrat, et N le montant nominal.

Pour notre exemple, on traite les risques tempête en Rhône-Alpes. $X_i(t)$ désigne donc la force maximale du vent (en m/s) par jour. Nous avons choisi deux stations Saint Martin en Haut (variable X) et Echirolles (variable Y) avec les seuils respectifs 10 et 9, et les limites 16 et 15. On prend T = 600 jours, N = 1, K = 50 et L = 200.

Exercice 2.1. On suppose que X et Y suivent une loi gamma de paramètres respectifs $(\alpha_X, \lambda_X) = (7,135; 1,178)$ et $(\alpha_Y, \lambda_Y) = (5,038; 0,763)$ et que la corrélation est donnée par une copule de Gumbel de paramètre 1,472.

1. Simuler le vecteur aléatoire (X, Y).
2. Estimer la valeur de C_T par la méthode de Monte Carlo. Comparer avec le résultat historique : $C_T = 80,64$.
3. En faisant varier le paramètre de la copule de -25%, -10%, $+10\%$ et $+25\%$, que constate-t-on sur le payoff ? Calculer dans chaque cas, les quantiles à 75% et à 90% (appelés VaR pour Value at Risk).

3 Copules archimédiennes : méthodes de simulation

3.1 Méthode avec fonction générateur d'une copule archimédienne.

Il s'agit de la méthode décrite rapidement en cours pour simuler une copule archimédienne C de générateur la fonction ϕ. On utilise le fait que si le vecteur aléatoire (U, V) est donné par la copule C, alors les variables aléatoires $C(U, V)$ et $\dfrac{\phi(U)}{\phi(U) + \phi(V)}$ sont deux variables indépendantes, la première de fonction de répartition F, où $F(t) = t - \dfrac{\phi(t)}{\phi'(t)}$ et la seconde est uniforme sur $[0, 1]$. Donc on a l'algorithme :

1. Simulation U_1 et U_2 de loi uniforme et indépendantes.
2. On génère T de loi F, par inversion : $T = F^{-1}(U_1)$ où F^{-1} est la fonction quantile associée à F.
3. $U = \phi^{-1}(U_2\phi(T))$ et $V = \phi^{-1}((1 - U_2)\phi(T))$.

Voici un exemple. Il s'agit d'une copule archimédienne C_θ dont la fonction ϕ_θ est donnée par

$$\forall u \in]0, 1], \quad \phi_\theta(u) = \left(\frac{1}{u} - 1\right)^\theta,$$

pour $\theta \geqslant 1$.

Exercice 3.1.

1. Vous vérifierez que cette fonction est bien strictement décroissante et convexe sur $]0, 1]$. Vous calculerez également la solution y de l'équation : $x = y - \frac{\phi(y)}{\phi'(y)}$ et expliciterez la fonction ϕ^{-1}.
2. En utilisant l'algorithme précédent, simuler deux variables aléatoires uniformes sur $[0, 1]$ dont la dépendance est donnée par la copule C_θ.
3. Enfin simuler un vecteur (X, Y) de marginales respectives une loi exponentielle de paramètre 2 et une loi de Pareto de paramètre 3, et dont la dépendance est donnée par cette copule archimédienne de paramètre $\theta = 3$.

Dans ce cas, on peut expliciter toutes les fonctions nécessaires à l'algorithme. Mais ce n'est pas possible dans des cas plus généraux, pour lesquels on pourra résoudre des équations numériquement (fonction `uniroot` de R) lorsqu'il n'y a pas de formule explicite pour les fonctions à calculer.

3.2 Méthode des distributions conditionnelles

On peut simuler un vecteur $U = (U_1, U_2)$ dont la distribution est donnée par une copule C par la méthode dite des distributions conditionnelles. Soit $C_{2|1}$ la distribution de U_2 conditionnellement à U_1, $C_{2|1}(u, u_1) = \mathbb{P}(U_2 \leqslant u | U_1 = u_1)$.

Exercice 3.2. Montrer que

$$C_{2|1}(u_2, u_1) = \left.\frac{\partial C(x_1, u_2}{\partial x_1}\right|_{x_1=u_1} = \partial_1 C(u_1, u_2).$$

L'algorithme est le suivant :

1. Simuler deux variables uniformes sur $[0, 1]$ indépendantes V_1 et V_2.
2. Poser $U_1 = V_1$.
3. Poser $U_2 = C_{2|1}^{-1}(V_2, U_1)$.

On remarque que

1. $C_{2|1}^{-1}(V_2, U_1)$ est, lorsqu'elle existe, la fonction inverse (ordinaire) de $x \mapsto C_{2|1}(x, U1)$, c'est-à-dire
$$U_2 = C_{2|1}^{-1}(V_2, U_1) \Longleftrightarrow C_{2|1}(U2, U1) = V2.$$
2. Dans cet algorithme on utilise la fonction inverse (ordinaire) de la fonction $x \mapsto C_{2|1}(x, U1)$. La fonction inverse n'est pas toujours définie. Dans le cas des copules de Clayton ou de Frank la fonction inverse est bien définie. Dans le cas de la copule de Gumbel ce n'est pas le cas ! On peut dans ce cas soit passer par une inversion numérique en utilisant `uniroot` de R, soit utiliser la définition de fonction "quasi-inverse" qui généralise la fonction inverse.

Exercice 3.3. Simuler deux variables aléatoires uniformes sur $[0, 1]$ dont la dépendance est donnée par la copule de Clayton par la méthode des distributions conditionnelles.

Exercice 3.4. Simuler deux variables aléatoires uniformes sur $[0, 1]$ dont la dépendance est donnée par la copule de Frank par la méthode des distributions conditionnelles.

3.3 Méthode analytique spécifique pour la copule de Clayton

On se propose de simuler un vecteur aléatoire (U, V) selon la procédure suivante :

1. Simuler deux variables aléatoires indépendantes X_1 et X_2, suivant une loi exponentielle de paramètre 1.
2. Simuler une variable aléatoire Z, indépendante de X_1 et X_2, suivant une loi Gamma de densité
$$f(x) = \frac{1}{\Gamma(1/\theta)} x^{1/\theta - 1} e^{-x} \mathbf{1}_{]0,+\infty[}(x)$$
3. Poser $U = \left(1 + \frac{X_1}{Z}\right)^{-\frac{1}{\theta}}$ et $V = \left(1 + \frac{X_2}{Z}\right)^{-\frac{1}{\theta}}$.

Exercice 3.5.

1. Simuler $N = 1000$ réalisations du vecteur (U, V) en prenant $\theta = 3$, et représenter le nuage de point obtenu. Calculer numériquement la corrélation linéaire entre U et V.
   Indication : on pourra simuler les réalisations de Z à l'aide de la commande `rgamma` en prenant comme paramètres shape $= 1/\theta$ et scale $= 1$.
2. Montrer numériquement que la loi du couple (U, V) a bien une copule pour fonction de répartition.
   Indication : on pourra regarder les marginales...
3. Reprendre la question 2 en faisant varier $\theta \in]0, 20]$. Analyser (en détaillant) les résultats obtenus.
4. Simuler N réalisations d'un vecteur aléatoire dont la dépendance est modélisée par une copule de Clayton de paramètre $\theta = 3$. Comparer vos résultats avec ceux de la question 2. Que peut-on en conclure ?
   Indication : on pourra utiliser la commande `rarchmCopula` du package `fCopulae` avec type "1".