

MATHS APPRO - MATHS II 2016

La simulation de vecteurs aléatoires dont les composantes ne sont pas indépendantes intervient dans l'évaluation de risques cumulés dans des domaines tels que l'assurance, la finance, la médecine ou l'écologie. On résume les liaisons entre les composantes à l'aide de fonctions de plusieurs variables appelées copules.

L'objet du problème consiste à présenter cette notion de copule dans le cadre de la simulation d'un vecteur aléatoire à deux composantes.

On suppose que toutes les variables aléatoires et tous les vecteurs aléatoires qui interviennent dans ce problème sont définis sur un même espace probabilisé $(\Omega, \mathcal{A}, P)$.

On rappelle que la loi d'un vecteur aléatoire (X, Y) à valeurs dans $\mathbb{R}^2$ est caractérisée par la fonction $F_{(X,Y)}$ définie sur $\mathbb{R}^2$ par : $\forall (x, y) \in \mathbb{R}^2, F_{(X,Y)}(x, y) = P([X \leq x] \cap [Y \leq y])$. On dit que $F_{(X,Y)}$ est la *fonction de répartition conjointe* de X et Y .

Partie I - Simulation d'une variable aléatoire à densité

1. (a) Justifier la convergence de l'intégrale $\int_0^{+\infty} \frac{dx}{\sqrt{x(1+x)}}$.
- (b) Soit V une variable aléatoire telle que $V(\Omega) = [0, \pi/2[$ suivant la loi uniforme sur $[0, \pi/2[$.
On pose : $X = \tan^2(V)$. Montrer que X est une variable aléatoire à densité.
- (c) En déduire que la fonction $f : x \mapsto \begin{cases} \frac{1}{\pi\sqrt{x(1+x)}} & \text{si } x > 0 \\ 0 & \text{sinon} \end{cases}$ est une densité de probabilité.
2. (a) Compléter le code Python de la fonction **simulX** suivante de sorte que son application à l'entier N ($N \geq 2$) fournisse un tableau **numpy** contenant N simulations indépendantes de la variable aléatoire X .

```
1 import numpy as np
import numpy.random as rd

def simulX(N):
5     u = rd.random(...)
    x = np.ones(len(u)) # Tableau de meme taille que u
    for i in range(0, N):
        x[i] = ...
    return x
```

- (b) Après avoir affecté une valeur entière supérieure ou égale à 2 à la variable N , on exécute les commandes suivantes :

```
1 x = simulX(N)
y = 0
for i in range(0, N):
    if x[i] > 1:
5         y = y+1
q = y/N
```

Trouver la loi d'une variable aléatoire dont la valeur de y est, en fin de boucle, une simulation. De quel nombre peut-on s'attendre que q soit proche lorsque la valeur affectée à N est grande et pourquoi ?

3. Soit X une variable aléatoire à densité dont la fonction de répartition est notée F_X .
Soit p un réel de $]0, 1[$. On pose : $K_p = \{x \in \mathbb{R}, F_X(x) = p\}$ et $J_p = \{x \in \mathbb{R}, F_X(x) < p\}$.
- (a) Justifier que les deux ensembles J_p et K_p ne sont pas vides et montrer que si $a \in J_p$ et $b \in K_p$, on a nécessairement : $a < b$.
- (b) On pose : $G_X(p) = \inf(K_p)$. Justifier l'existence de $G_X(p)$ et établir l'égalité : $F_X(G_X(p)) = p$.
- (c) En déduire pour tout x réel, l'équivalence : $x < G_X(p) \Leftrightarrow F_X(x) < p$.
- (d) Soit U une variable aléatoire telle que $U(\Omega) =]0, 1[$ et qui suit la loi uniforme sur $]0, 1[$.
Montrer que la variable aléatoire $G_X(U)$ suit la même loi que X .

- (e) On suppose que X admet pour densité la fonction f définie dans la question 1c et que p n'est plus fixé. Déterminer la fonction $G_X : p \mapsto G_X(p)$ définie sur $]0, 1[$.

Partie II - Fonction de répartition conjointe de deux variables aléatoires de lois uniformes

4. Soit (X, Y) un couple de variables aléatoires et F la fonction de répartition conjointe de X et Y . Soit $x \in \mathbb{R}$.
- Montrer que la fonction $y \mapsto F(x, y)$ est croissante sur $\mathbb{R}$.
 - Établir l'égalité : $[X \leq x] = \bigcup_{n \in \mathbb{N}^*} ([X \leq x] \cap [Y \leq n])$.
 - Montrer que $F(x, y)$ tend vers $P([X \leq x])$ lorsque y tend vers $+\infty$.
 - Quelle est la limite de $F(x, y)$ lorsque y tend vers $-\infty$?
 - Soit a, a', b, b' des réels vérifiant : $a \leq a'$ et $b \leq b'$. On pose $A = [a < X \leq a']$ et $B = [b \leq Y \leq b']$.
 - Exprimer la probabilité $P(A \cap B)$ en fonction de $P([X \leq a] \cap B)$ et $P([X \leq a'] \cap B)$.
 - Établir l'égalité : $P(A \cap B) = F(a', b') - F(a', b) - F(a, b') + F(a, b)$.

Dans les questions 5 et 6, on note U et V deux variables aléatoires suivant chacune la loi uniforme sur $[0, 1]$ et $F_{(U,V)}$ leur fonction de répartition conjointe.

On note C la restriction de $F_{(U,V)}$ à $[0, 1]^2$: $\forall (u, v) \in [0, 1]^2$, $C(u, v) = P([U \leq u] \cap [V \leq v])$.

Pour tout couple $(u, v) \in [0, 1]^2$, on pose $C_+(u, v) = \min\{u, v\}$ et $C_-(u, v) = \max\{u + v - 1, 0\}$.

Pour $(u, v) \in [0, 1]^2$, on note $\overline{[U > u] \cup [V > v]}$ l'événement contraire de l'événement $[U > u] \cup [V > v]$.

On rappelle que si deux vecteurs aléatoires (X_1, Y_1) et (X_2, Y_2) ont même loi et si g est une fonction continue sur $\mathbb{R}^2$ à valeurs dans $\mathbb{R}$, alors les variables aléatoires $g(X_1, Y_1)$ et $g(X_2, Y_2)$ ont même loi.

5. (a) Comparer pour tout $(u, v) \in [0, 1]^2$, les trois événements : $\overline{[U > u] \cup [V > v]}$, $[U \leq u] \cap [V \leq v]$ et $[U \leq u]$.
- (b) Justifier pour tout $(u, v) \in [0, 1]^2$, la double inégalité : $u + v - 1 \leq C(u, v) \leq u$.
- (c) En déduire l'encadrement suivante : $\forall (u, v) \in [0, 1]^2$, $C_-(u, v) \leq C(u, v) \leq C_+(u, v)$.
6. (a) Calculer $F_{(U,U)}(x, y)$ selon les valeurs du couples (x, y) de $\mathbb{R}^2$.
- (b) Représenter dans le plan rapporté à un repère orthonormé une ligne de niveau pour la fonction de deux variables $F_{(U,U)}$ correspondant à une valeur de la fonction strictement comprise entre 0 et 1. Hachurer sur la même figure, l'ensemble des couples $(x, y) \in \mathbb{R}^2$ pour lesquels $F_{(U,U)}(x, y) = x$.
- (c) Montrer que C est égale à C_+ si et seulement si les variables aléatoires U et V sont égales presque sûrement.
- (d) Calculer la fonction de répartition conjointe $F_{(U, 1-U)}$ et donner une condition nécessaire et suffisante portant sur U et V pour que C soit égale à C_- .

Partie III - Copules

On appelle *copule* toute fonction Φ définie sur $[0, 1]^2$, à valeurs réelles, vérifiant les trois propriétés suivantes :

- $\forall u \in [0, 1]$, $\Phi(u, 0) = \Phi(0, u) = 0$;
- $\forall u \in [0, 1]$, $\Phi(u, 1) = \Phi(1, u) = u$;
- $\forall (u, u', v, v') \in [0, 1]^4$, $u \leq u'$ et $v \leq v' \Rightarrow \Phi(u', v') - \Phi(u', v) - \Phi(u, v') + \Phi(u, v) \geq 0$.

On appelle *copule à densité* toute copule Φ dont la restriction à l'ouvert $]0, 1[^2$ est de classe $\mathcal{C}^2$ sur $]0, 1[^2$.

7. *Exemples.* On reprend le contexte et les notations du préambule des questions 5 et 6.
- Vérifier que C est une copule. Dans la suite (partie IV), on l'appelle la *copule associée au couple (U, V)* .
 - En déduire que C_+ , C_- ainsi que la fonction Π définie sur $[0, 1]^2$ par $\Pi(u, v) = uv$ sont des copules.
8. Soit Φ une copule à densité et $(a, b) \in]0, 1[^2$.
- Pour tout couple (h, k) de réels non nuls tels que $(a + h, b + k) \in]0, 1[^2$, on pose :

$$G(h, k) = \frac{1}{hk} (\Phi(a + h, b + k) - \Phi(a + h, b) - \Phi(a, b + k) + \Phi(a, b)).$$

- Soit h un réel non nul tel que $a + h \in]0, 1[$. Justifier que $G(h, k)$ admet une limite $H(h)$ lorsque k tend vers 0 et exprimer $H(h)$ à l'aide de la dérivée partielle $\partial_2(\Phi)$ de Φ par rapport à sa seconde variable.
- On note $\partial_{1,2}^2(\Phi)$ la dérivée partielle seconde croisée de Φ sur $]0, 1[^2$ et on rappelle que $\partial_{1,2}^2(\Phi) = \partial_{2,1}^2(\Phi)$. Trouver la limite de $H(h)$ lorsque h tend vers 0 et en déduire que $\partial_{1,2}^2(\Phi)(a, b) \geq 0$.

9. Soit φ une fonction définie sur $[0, 1]^2$, à valeurs réelles, continue sur $[0, 1]^2$ et de classe $\mathcal{C}^2$ sur $]0, 1[^2$.
 Pour tout $(u, u', v, v') \in [0, 1]^4$, on pose : $\Psi(u, u', v, v') = \varphi(u', v') - \varphi(u', v) - \varphi(u, v') + \varphi(u, v)$.
- (a) Pour tout $(u, u', v, v') \in]0, 1[^4$, justifier l'égalité : $\Psi(u, u', v, v') = \int_v^{v'} \left(\int_u^{u'} \partial_{1,2}^2(\varphi)(x, y) dx \right) dy$.
- (b) Soit u et u' des réels tels que : $0 \leq u \leq u' \leq 1$. Pour tout $n \in \mathbb{N}^*$, on pose $\begin{cases} u_n &= \frac{1}{3n} + \left(1 - \frac{1}{n}\right)u \\ u'_n &= \frac{2}{3n} + \left(1 - \frac{1}{n}\right)u' \end{cases}$.
 Vérifier pour tout $n \in \mathbb{N}^*$, les inégalités strictes suivantes : $0 < u_n < u'_n < 1$.
- (c) On pose : $\begin{cases} T_0 &= \{(u, u', v, v') \in \mathbb{R}^4 ; 0 < u < u' < 1, 0 < v < v' < 1\} \\ T &= \{(u, u', v, v') \in \mathbb{R}^4 ; 0 \leq u \leq u' \leq 1, 0 \leq v \leq v' \leq 1\} \end{cases}$.
 On suppose que la fonction Ψ est positive ou nulle sur T_0 . Montrer que Ψ est positive ou nulle sur T .
- (d) En déduire que, pour que la fonction φ soit une copule, il suffit qu'elle vérifie les trois propriétés suivantes :
- $\forall u \in [0, 1], \varphi(u, 0) = \varphi(0, u) = 0$;
 - $\forall u \in [0, 1], \varphi(u, 1) = \varphi(0, 1) = 1$;
 - $\forall (x, y) \in]0, 1[^2, \partial_{1,2}^2(\varphi)(x, y) \geq 0$.

Partie IV - Familles de copules et simulation

10. Soit M la fonction définie sur $[0, 1]^2$ par : $\forall (u, v) \in [0, 1]^2, M(u, v) = uv(u + v - uv)$.
- (a) Montrer que la fonction $S : (u, v) \mapsto u + v - 2uv$ admet sur $[0, 1]^2$ un minimum global c et un maximum global d et les calculer.
- (b) Montrer que M est une copule à densité.
- (c) Soit $\theta \in \mathbb{R}$ et M_θ la fonction définie sur $[0, 1]^2$ par : $\forall (u, v) \in [0, 1]^2, M_\theta(u, v) = (1 - \theta)uv + \theta M(u, v)$.
 Pour quelles valeurs de θ la fonction M_θ est-elle une copule ?
11. Soit N une variable aléatoire suivant la loi de Bernoulli de paramètre $p \in]0, 1[$ et U_0, V_0, U_1, V_1 quatre variables aléatoires suivant chacune la loi uniforme sur $[0, 1]$.
 On suppose que (U_0, V_0) et (U_1, V_1) et N sont mutuellement indépendants, autrement dit, on suppose que pour tout $(u_0, v_0, u_1, v_1, x) \in \mathbb{R}^5$, on a :

$$P([U_0 \leq u_0] \cap [V_0 \leq v_0] \cap [U_1 \leq u_1] \cap [V_1 \leq v_1] \cap [N \leq x]) = F_{(U_0, V_0)}(u_0, v_0) F_{(U_1, V_1)}(u_1, v_1) P([N \leq x]).$$

- (a) Pour tout $\omega \in \Omega$, on pose $U_N(\omega) = \begin{cases} U_0(\omega) & \text{si } N(\omega) = 0 \\ U_1(\omega) & \text{si } N(\omega) = 1 \end{cases}$ et $V_N(\omega) = \begin{cases} V_0(\omega) & \text{si } N(\omega) = 0 \\ V_1(\omega) & \text{si } N(\omega) = 1 \end{cases}$.
 Montrer que U_N et V_N sont des variables aléatoires et suivent chacune la loi uniforme sur $[0, 1]$.
- (b) Exprimer la copule associée au couple (U_N, V_N) à l'aide des copules associées aux deux couples (U_0, V_0) et (U_1, V_1) .
12. Soit $p \in]0, 1[$ et C_p la fonction définie sur $[0, 1]^2$ par : $\forall (u, v) \in [0, 1]^2, C_p(u, v) = puv + (1 - p) \min\{u, v\}$.
- (a) Montrer que C_p est une copule.
- (b) Proposer une méthode de simulation d'un couple aléatoire (U, V) auquel est associée la copule C_p et donner le code **Python** d'une fonction **simulc** réalisant cette simulation pour toute valeur donnée de p .
 Cette fonction aura pour seul argument le paramètre p et retournera le couple (u, v) .
13. Soit X et Y deux variables aléatoires à densité, de fonctions de répartition respectives F_X et F_Y .
 Pour tout $p \in]0, 1[$, on pose : $\begin{cases} G_X(p) &= \inf(\{x \in \mathbb{R} ; F_X(x) = p\}) \\ G_Y(p) &= \inf(\{x \in \mathbb{R} ; F_Y(x) = p\}) \end{cases}$.
- Soit C la fonction définie sur $[0, 1]^2$ par $C(u, v) = \begin{cases} F_{(X, Y)}(G_X(u), G_Y(v)) & \text{si } (u, v) \in]0, 1[^2 \\ 0 & \text{si } uv = 0 \\ u & \text{si } v = 1 \\ v & \text{si } u = 1 \end{cases}$.
- Montrer que C est une copule. En déduire un procédé de simulation du couple (X, Y) à partir de la simulation (u, v) d'un couple (U, V) auquel la copule C est associée.