Petite Classe 4 - MAP 311 - Groupes 1 et 12

Josselin Garnier

15 mai 2017 - salle PC n^o 16

Exercice 1. Soit X une variable aléatoire de loi normale $\mathcal{N}(0,1)$. On pose

$$Y = X\mathbf{1}_{\{|X|<a\}} - X\mathbf{1}_{\{|X|\geq a\}}$$

Montrer que Y est une v.a. normale $\mathcal{N}(0,1)$. Montrer qu'il existe un réel $a > 0$ tel que $\mathrm{Cov}(X,Y) = 0$. Les v.a. X et Y sont elles indépendantes ?

Solution.

$$\begin{aligned}\mathrm{Cov}(X,Y) &= \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = \mathbb{E}\big(X^2\mathbf{1}_{\{|X|<a\}} - X^2\mathbf{1}_{\{|X|\geq a\}}\big) = \mathbb{E}(X^2) - 2\mathbb{E}\big(X^2\mathbf{1}_{\{|X|\geq a\}}\big)\\ &= 1 - \frac{4}{\sqrt{2\pi}}\int_a^\infty e^{-\frac{y^2}{2}}dy\end{aligned}$$

C'est une fonction continue de a, qui vaut -1 en $a=0$ et qui tend vers 1 en $+\infty$, donc il existe $a \in]0,+\infty[$ tel que $\mathrm{Cov}(X,Y) = 0$.

Les v.a. X et Y ne sont pas indépendantes :

$$\mathbb{P}(X+Y=0) = \mathbb{P}(|X| \geq a) \in]0,1[$$

Or si les v.a. étaient indépendantes, alors $X+Y$ serait à densité.

Remarque. Les v.a. X et Y sont gaussiennes, leur covariance est nulle, et pourtant elles ne sont pas indépendantes ! En effet, le vecteur (X,Y) n'est pas gaussien.

Exercice 2. Soient X et Y deux variables aléatoires indépendantes de lois respectives $\Gamma(\alpha,\lambda)$ et $\Gamma(\alpha+1/2,\lambda)$, avec $\alpha>0$ et $\lambda>0$. On pose $(V,W) = (\sqrt{XY},\sqrt{Y})$. Déterminer la loi de (V,W).

On rappelle que la densité de la loi $\Gamma(a,\lambda)$ est (c.f. Section 4.6.3 du poly)

$$\frac{1}{\Gamma(a)}\lambda^a x^{a-1}e^{-\lambda x}\mathbf{1}_{x>0}, \qquad \text{avec} \quad \Gamma(a) = \int_0^\infty z^{a-1}e^{-z}dz.$$

Solution. Comme X et Y sont indépendantes, la loi de (X,Y) a pour densité $(x,y) \mapsto f_X(x)f_Y(y)$, où f_X et f_Y désignent les densités de X et Y. On utilise alors la méthode de la fonction muette. Soit $h : \mathbb{R}^2 \to \mathbb{R}$ une fonction continue bornée. Alors

$$\begin{aligned}\mathbb{E}(h(V,W)) &= \mathbb{E}(h(\sqrt{XY},\sqrt{Y}))\\ &= \int_{\mathbb{R}^2} h(\sqrt{xy},\sqrt{y})f_{(X,Y)}(x,y)dxdy\\ &= \int_{\mathbb{R}^2} h(\sqrt{xy},\sqrt{y})f_X(x)f_Y(y)dxdy\\ &= \frac{\lambda^{2\alpha+1/2}}{\Gamma(\alpha)\Gamma(\alpha+1/2)}\int_{]0,\infty[^2} h(\sqrt{xy},\sqrt{y})(xy)^{\alpha-1/2}e^{-\lambda(x+y)}\frac{dxdy}{\sqrt{x}}.\end{aligned}$$

On considère le changement de variable $v = \sqrt{xy}, w = \sqrt{y}$ qui est un C^1 difféomorphisme de $]0, \infty[^2$ dans lui-même. Le calcul du déterminant de la matrice jacobienne donne

$$dvdw = \frac{dxdy}{4\sqrt{x}}.$$

Ainsi, avec $y = w^2$ et $x = v^2/w^2$,

$$\mathbb{E}(h(V,W)) = \frac{4\lambda^{2\alpha+1/2}}{\Gamma(\alpha)\Gamma(\alpha+1/2)} \int_{]0,\infty[^2} h(v,w) v^{2\alpha-1} e^{-\lambda(w^2+\frac{v^2}{w^2})} dvdw.$$

Ainsi, (V, W) est à densité et sa densité est donnée par

$$f_{(V,W)}(v,w) = \frac{4\lambda^{2\alpha+1/2}}{\Gamma(\alpha)\Gamma(\alpha+1/2)} v^{2\alpha-1} e^{-\lambda(w^2+\frac{v^2}{w^2})} \mathbf{1}_{v>0,w>0}.$$

Exercice 3. Soit (X, Y) un couple de variables aléatoires réelles à densité sur $\mathbb{R}^2$ tel que :

– X suit une loi $\Gamma(2, \lambda)$ (de densité $f_X(x) = \lambda^2 x e^{-\lambda x} \mathbf{1}_{x \geq 0}$),

– la loi conditionnelle de Y sachant X est la loi uniforme sur le segment $[0, X]$ (ou, en d'autres termes, la densité conditionnelle de Y sachant que $X = x$ est $f_{Y|X=x}(y) = \frac{1}{x} \mathbf{1}_{0<y<x}$).

1. Déterminer la densité de (X, Y) ainsi que la loi de Y.
2. Calculer la densité conditionnelle de X sachant Y.
3. Calculer les quantités suivantes:
   (a) $\mathbb{E}(Y|X)$,
   (b) $\mathbb{E}(X|Y)$,
   (c) $\mathbb{E}(X + XY|Y)$,
   (d) $\mathbb{E}(\mathbb{E}(Y|X))$,
   (e) $\mathbb{E}(XY)$ (on pourra utiliser le fait que $\mathbb{E}(X^2) = \frac{6}{\lambda^2}$).

Solution.

1. Par définition,

$$f_{Y|X=x}(y) = \frac{f_{(X,Y)}(x,y)}{f_X(x)},$$

   où $f_{(X,Y)}$ est la densité de (X, Y). On en déduit que

$$f_{(X,Y)}(x,y) = f_{Y|X=x}(y) f_X(x) = \lambda^2 e^{-\lambda x} \mathbf{1}_{0<y<x}.$$

   La loi de Y est alors à densité, de densité f_Y donnée par

$$f_Y(y) = \int_{\mathbb{R}} f_{(X,Y)}(x,y) dx = \lambda^2 \int_y^{\infty} e^{-\lambda x} dx \mathbf{1}_{y>0} = \lambda e^{-\lambda y} \mathbf{1}_{y>0}.$$

   Ainsi, Y suit une loi exponentielle de paramètre λ.

2. Par définition, $f_{X|Y=y}(x)$ est donné par

$$f_{X|Y=y}(x) = \frac{f_{(X,Y)}(x,y)}{f_Y(y)} = \lambda e^{-\lambda x + \lambda y}\mathbf{1}_{0<y<x}.$$

3. (a) On a $\mathbb{E}(Y|X) = \Psi_1(X)$, avec

$$\Psi_1(x) = \mathbb{E}(Y|X=x) = \int_{\mathbb{R}} y f_{Y|X=x}(y)dy = \int_0^x \frac{y}{x}dy\mathbf{1}_{x>0} = \frac{x}{2}\mathbf{1}_{x>0}.$$

Ainsi, $\mathbb{E}(Y|X) = \frac{X}{2}$.

(b) On a $\mathbb{E}(X|Y) = \Psi_2(Y)$ avec

$$\Psi_2(y) = \mathbb{E}(X|Y=y) = \int_{\mathbb{R}} x f_{X|Y=y}(x)dx = \lambda e^{\lambda y}\mathbf{1}_{y>0}\int_y^{\infty} xe^{-\lambda x}dx.$$

Pour calculer cette intégrale, on fait le changement de variable $u = x - y$, de sorte que

$$\Psi_2(y) = \mathbf{1}_{y>0}\int_0^{\infty}(u+y)\lambda e^{-\lambda u}du = \left(\frac{1}{\lambda} + y\right)\mathbf{1}_{y>0}.$$

Ainsi, $\mathbb{E}(X|Y) = \frac{1}{\lambda} + Y$.

(c) On a

$$\mathbb{E}(X + XY|Y) = \mathbb{E}(X|Y) + \mathbb{E}(XY|Y) = \mathbb{E}(X|Y) + Y\mathbb{E}(X|Y) = (Y+1)(\frac{1}{\lambda} + Y).$$

(d) On sait que $\mathbb{E}(\mathbb{E}(Y|X)) = \mathbb{E}(Y)$. Comme Y suit une loi exponentielle de paramètre λ, on a donc $\mathbb{E}(\mathbb{E}(Y|X)) = 1/\lambda$.

(e) Pour simplifier le calcul, l'idée est d'écrire $\mathbb{E}(XY) = \mathbb{E}(\mathbb{E}(XY|X))$. En effet, $\mathbb{E}(XY|X) = X\mathbb{E}(Y|X) = \frac{X^2}{2}$, donc

$$\mathbb{E}(XY) = \mathbb{E}(\frac{X^2}{2}) = \frac{\mathbb{E}(X^2)}{2} = \frac{3}{\lambda^2}.$$

Exercice 4. Soit (X,Y) un couple de variables aléatoires réelles à densité sur $\mathbb{R}^2$. On suppose que X et Y sont indépendantes.

1. Montrer que
$$\mathbb{E}(X|Y) = \mathbb{E}(X).$$

2. Plus généralement, montrer que $\mathbb{E}(h(X,Y)|X) = \Phi(X)$, avec
$$\Phi(x) = \mathbb{E}(h(x,Y)).$$

Solution.

1. Par définition, $\mathbb{E}(X|Y) = \Psi(Y)$, avec
$$\Psi(y) = \int_{\mathbb{R}} x \frac{f_{(X,Y)}(x,y)}{f_Y(y)} dx.$$
Or X et Y sont indépendantes, donc $f_{(X,Y)}(x,y) = f_X(x) f_Y(y)$. On en déduit que
$$\Psi(y) = \int_{\mathbb{R}} x f_X(x) dx = \mathbb{E}(X).$$
Donc $\mathbb{E}(X|Y)$ est une variable aléatoire constante qui vaut $\mathbb{E}(X)$.

2. Par définition, $\mathbb{E}(h(X,Y)|X) = \Phi(X)$ avec
$$\Phi(x) = \int_{\mathbb{R}} h(x,y) \frac{f_{(X,Y)}(x,y)}{f_X(x)} dy = \int_{\mathbb{R}} h(x,y) f_Y(y) dy = \mathbb{E}(h(x,Y)).$$

Exercice 5. Soient $X_1, \dots, X_n$ des variables aléatoires exponentielles indépendantes de même paramètre $\lambda > 0$ avec $n \geq 2$. On pose $S_n = X_1 + \cdots + X_n$.

1. Identifier la loi de S_n comme une loi classique.
2. Trouver la densité conditionnelle de X_1 sachant que $S_n = t$.
3. En déduire la valeur de l'espérance conditionnelle de X_1 sachant S_n. Pouvait-on prévoir ce résultat ?

Solution.

1. S_n suit la loi $\Gamma(n, \lambda)$ comme somme de n variables aléatoires exponentielles indépendantes de même paramètre $\lambda > 0$.

2. On commence par déterminer la loi jointe (X_1, S_n) par la méthode de la fonction muette. Soit $h : \mathbb{R}^2 \to \mathbb{R}$ une fonction continue bornée et on calcule $\mathbb{E}(h(X_1, S_n))$. Pour simplifier les calculs, on remarque que $S_n = X_1 + Y$ avec Y indépendante de X_1 et qui suit une loi $\Gamma(n-1, \lambda)$. Donc
$$\begin{aligned} \mathbb{E}(h(X_1, S_n)) &= \mathbb{E}(h(X_1, X_1 + Y) \\ &= \int_{]0,\infty[^2} h(x, x+y) \lambda e^{-\lambda x} \frac{\lambda^{n-1}}{(n-2)!} y^{n-2} e^{-\lambda y} dx dy. \end{aligned}$$
L'application $(x,y) \mapsto (x, x+y)$ est un C^1 difféomorphisme de $]0,\infty[^2$ sur $\{(u,v) \in]0,\infty[^2 \colon u < v\}$, de jacobien 1. On fait alors changement de variable $u = x$ et $v = x+y$:
$$\mathbb{E}(h(X_1, S_n)) = \int_{\mathbb{R}^2} h(u,v) \frac{\lambda^n}{(n-2)!} (v-u)^{n-2} e^{-\lambda v} \mathbf{1}_{v>u>0} du dv.$$
Donc (X_1, S_n) est à densité, de densité donnée par
$$f(u,v) = \frac{\lambda^n}{(n-2)!} (v-u)^{n-2} e^{-\lambda v} \mathbf{1}_{v>u>0}.$$

La densité conditionnelle de X_1 sachant que $S_n = t$ vaut donc

$$f_{X_1|S_n=t}(u) = \frac{\frac{\lambda^n}{(n-2)!}(t-u)^{n-2}e^{-\lambda t}\mathbf{1}_{t>u>0}}{\frac{\lambda^n}{(n-1)!}t^{n-1}e^{-\lambda t}} = (n-1)\frac{(t-u)^{n-2}}{t^{n-1}}\mathbf{1}_{t>u>0}$$

3. On a donc

$$\begin{aligned}
\mathbb{E}(X_1|S_n = t) &= \int_{\mathbb{R}} u f_{X_1|S_n=t}(u)du = \frac{n-1}{t^{n-1}}\int_0^t u(t-u)^{n-2}du = \frac{n-1}{t^{n-1}}\int_0^t (t-u)u^{n-2}du \\
&= \frac{n-1}{t^{n-1}}\left(t\frac{t^{n-1}}{n-1} - \frac{t^n}{n}\right) = \frac{t}{n}
\end{aligned}$$

Ainsi,

$$\mathbb{E}(X_1|S_n) = \frac{S_n}{n}.$$

Retrouvons ce résultat en utilisant la symétrie du problème. Par symétrie, on a $\mathbb{E}(X_1|S_n) = \mathbb{E}(X_2|S_n) = \cdots = \mathbb{E}(X_n|S_n) = \Psi(S_n)$. Alors

$$n\Psi(S_n) = \mathbb{E}(X_1|S_n) + \mathbb{E}(X_2|S_n) + \cdots + \mathbb{E}(X_n|S_n) = \mathbb{E}(X_1 + \cdots + X_n|S_n) = \mathbb{E}(S_n|S_n) = S_n.$$

Donc $\mathbb{E}(X_1|S_n) = \frac{S_n}{n}$.

Exercice 6. Soit $X = (X_1, X_2, X_3)$ un vecteur gaussien centré de matrice de covariance

$$\Gamma = \begin{pmatrix} 3 & 1 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

1. Que peut-on dire de X_3 et de (X_1, X_2) ?
2. Quelle est la loi de (X_1, X_2) ?
3. Montrer que pour tout $a \in \mathbb{R}$ le vecteur $(X_2, X_2 + aX_1)$ est un vecteur gaussien.
4. En choisissant a de sorte que X_2 et $X_2 + aX_1$ soient indépendants, calculer $\mathbb{E}(X_1|X_2)$.

Solution.

1. On voit que pour $i = 1$ et $i = 2$ on a $\mathrm{Cov}(X_3, X_i) = 0$. Donc X_3 est indépendant du vecteur gaussien (X_1, X_2).
2. Le vecteur gaussien (X_1, X_2) est centré, de matrice de covariance $\Gamma_{12} = \begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix}$.
3. Tout combinaison de linéaire de X_2 et de $X_2 + aX_1$ est une combinaison linéaire de X_1 et de X_2, donc est une variable gaussienne car (X_1, X_2) est un vecteur gaussien. Donc le vecteur $(X_2, X_2 + aX_1)$ est gaussien (centré).

4. Comme le vecteur $(X_2, X_2 + aX_1)$ est gaussien, X_2 et $X_2 + aX_1$ sont indépendants si et seulement si $\text{Cov}(X_2, X_2 + aX_1) = 0$. Or

$$\begin{aligned}\text{Cov}(X_2, X_2 + aX_1) &= \mathbb{E}(X_2(X_2 + aX_1)) - \mathbb{E}(X_2)\mathbb{E}(X_2 + aX_1) = \mathbb{E}(X_2^2) + a\mathbb{E}(X_1X_2) \\ &= 2 + a.\end{aligned}$$

En prenant $a = -2$, on en déduit que $X_2 - 2X_1$ et X_2 sont indépendants.

Pour calculer $\mathbb{E}(X_1|X_2)$, l'idée est de choisir α, β pour avoir $X_1 = \alpha(X_2 - 2X_1) + \beta X_2$: en écrivant $X_1 = -\frac{1}{2}(X_2 - 2X_1) + \frac{1}{2}X_2$, on obtient donc:

$$\mathbb{E}(X_1|X_2) = -\frac{1}{2}\mathbb{E}(X_2 - 2X_1|X_2) + \mathbb{E}(\frac{1}{2}X_2|X_2) = -\frac{1}{2}\mathbb{E}(X_2 - 2X_1) + \frac{1}{2}X_2 = \frac{1}{2}X_2.$$

Exercice 7. (Simulation d'une loi de Poisson) Soit $(U_n)_{n\geq 1}$ une suite de variables aléatoires indépendantes de loi uniforme, et soit $\lambda > 0$. On pose $E_n = -\ln(U_n)/\lambda$ pour tout $n \geq 1$ et

$$T_n = E_1 + \cdots + E_n.$$

1. Identifier la loi de $(E_i)_{1\leq i\leq n}$, puis la loi de T_n comme une loi classique.

2. On pose $N = \min\{n \geq 0 : U_1U_2 \cdots U_{n+1} < e^{-\lambda}\}$. Montrer que N suit une loi de Poisson de paramètre λ.

Solution.

1. Soit $1 \leq i \leq n$. Si $x \leq 0$, on a $\mathbb{P}(E_i \leq x) = 0$, et pour $x \geq 0$ on a

$$\mathbb{P}(E_i \leq x) = \mathbb{P}(U_n \geq e^{-\lambda x}) = 1 - e^{-\lambda x}.$$

Donc E_i suit une loi exponentielle de paramètre λ. Par ailleurs, d'après le principe de composition, les variables aléatoires $(E_i)_{1\leq i\leq n}$ sont indépendantes.

Ainsi, T_n suit la loi $\Gamma(n, \lambda)$ comme somme de n variables aléatoires exponentielles indépendantes de même paramètre $\lambda > 0$.

2. Soit $n \geq 0$ un entier. Comme les événements $\{N = n\}$ et $\{T_1 \leq 1, T_2 \leq 1, \ldots, T_{n+1} > 1\}$ sont égaux, on a, en utilisant le théorème de transfert:

$$\begin{aligned}\mathbb{P}(N = n) &= \mathbb{P}\big(E_1 \leq 1, E_1 + E_2 \leq 1, \ldots, E_1 + \cdots + E_n \leq 1, E_1 + \cdots + E_{n+1} > 1\big) \\ &= \int_{]0,\infty[^{n+1}} \mathbf{1}_{x_1\leq 1, x_1+x_2\leq 1, \ldots, x_1+\cdots+x_n\leq 1, x_1+\cdots+x_{n+1}>1} \lambda^{n+1} e^{-\lambda(x_1+\cdots+x_{n+1})} dx_1 \cdots dx_n\end{aligned}$$

L'application $(x_1, \ldots, x_{n+1}) \mapsto (x_1, x_1+x_2, \ldots, x_1+\cdots+x_{n+1})$ est un C^1 difféomorphisme de $]0,\infty[^{n+1}$ sur $\{(y_1, \ldots, y_{n+1}) \in]0,\infty[^{n+1}: y_1 < y_2 < \cdots < y_{n+1}\}$ de jacobien 1. Le changement de variable $y_i = x_1 + \cdots + x_i$ donne

$$\begin{aligned}
\mathbb{P}(N=n) &= \int_{0<y_1<\cdots<y_n\leq 1, y_{n+1}>1} \lambda^{n+1} e^{-\lambda y_{n+1}} dy_1 \cdots dy_{n+1} \\
&= e^{-\lambda}\lambda^n \int_{0<y_1<\cdots<y_n\leq 1} dy_1 \cdots dy_n \\
&= e^{-\lambda}\lambda^n \int_{0<y_2<\cdots<y_n\leq 1} y_2 dy_2 \cdots dy_n \\
&= e^{-\lambda}\lambda^n \int_{0<y_3<\cdots<y_n\leq 1} \frac{y_3^2}{2} dy_3 \cdots dy_n \\
&= e^{-\lambda}\lambda^n \int_{0<y_n\leq 1} \frac{y_n^{n-1}}{(n-1)!} dy_n \qquad \text{(par récurrence)} \\
&= e^{-\lambda}\frac{\lambda^n}{n!},
\end{aligned}$$

ce qui permet de conclure.

Exercice 8. (Pale 2014) On considère le jeu de la courte paille entre deux joueurs. Une troisième personne choisit un bâton puis elle le coupe en deux. Elle présente alors les deux morceaux aux joueurs en cachant la différence de longueur dans sa main. Le premier joueur choisit un morceau et le second prend lautre. Celui qui a le morceau le plus long gagne. La longueur U du bâton est supposée suivre la loi uniforme sur $[0,1]$ et les deux morceaux finaux ont pour longueur $X = UV$ et $Y = U(1-V)$ avec V de loi uniforme sur $[0,1]$ indépendante de U.

1. Calculer $\mathbb{E}(X)$ et $\mathrm{Var}(X)$.
2. Donner la loi de $-\ln(U)$. En déduire sans calcul que $\xi = -\ln(X)$ suit la loi $\Gamma(2,1)$. Déterminer la loi de X.
3. Quelle est la loi de (X,Y) (faire attention au domaine image du changement de variables) ? Les longueurs des morceaux sont-elles indépendantes ? Comparer les lois de (Y,X) et de (X,Y). Le jeu est-il équitable ? Retrouver la densité de X.

Solution.

1. Puisque U et V sont indépendantes, $\mathbb{E}(X) = \mathbb{E}(UV) = \mathbb{E}(U)\mathbb{E}(V) = 1/4$. De même, $\mathbb{E}(X^2) = \mathbb{E}(U^2)\mathbb{E}(V^2) = 1/9$ et $\mathrm{Var}(X) = 1/9 - 1/16 = 7/144$.
2. On sait que $-\ln(U)$ suit la loi exponentielle de paramètre 1. La variable aléatoire ξ est donc la somme de deux variables exponentielles de paramètre 1 et suit donc une loi $\Gamma(2,1)$, de densité $xe^{-x}\mathbf{1}_{x>0}$. Pour déterminer la loi de X, on utilise la méthode de la fonction muette. Soit $h : \mathbb{R} \to \mathbb{R}$ une fonction continue bornée. On calcule $\mathbb{E}(h(X))$ avec le théorème de transfert puis avec le changement de variable $x = e^{-z}$:

$$\mathbb{E}(h(X)) = \mathbb{E}(h(e^{-\xi})) = \int_0^\infty h(e^{-z}) z e^{-z} dz = -\int_0^1 \phi(x)\ln(x)dx.$$

Donc X est à densité, de densité $-\mathbf{1}_{0<x<1}\ln(x)$.

3. On utilise la méthode de la fonction muette. Soit $h : \mathbb{R}^2 \to \mathbb{R}$ une fonction conitbue bornée. D'après le théorème de transfert:

$$\mathbb{E}(h(X,Y)) = \mathbb{E}(h(UV, U(1-V)) = \int_0^1 \int_0^1 h(uv, u(1-v))dudv.$$

On a donc envie de faire le changement de variable $x = uv$ et $y = u(1-v)$. Par ailleurs $0 < u, v < 1$ si et seulement si $0 < x + y < 1$ et $x, y > 0$. Ainsi, l'application $(u, v) \mapsto (uv, u(1-v))$ est un C^1 difféomorphisme de $]0,1[^2$ sur $\{(x,y) \in]0,1[^2 \colon x + y < 1\}$, de matrice jacobienne

$$\begin{pmatrix} v & u \\ 1-v & -u \end{pmatrix}$$

et donc de jacobien $|u|$. On a donc

$$\begin{aligned} \mathbb{E}(h(X,Y)) &= \int_0^1 \int_0^1 h(uv, u(1-v))u \cdot \frac{dudv}{u} \\ &= \int_{\mathbb{R}^2} h(x,y) \frac{\mathbf{1}_{x>0,y>0,x+y<1}}{x+y} dxdy. \end{aligned}$$

Le couple (X,Y) est donc à densité, de densité donnée par $f_{(X,Y)}(x,y) = \frac{\mathbf{1}_{x>0,y>0,x+y<1}}{x+y}$. Comme cette quantité ne peut pas se mettre sous forme du produit dune fonction de x par une fonction de y. Les variables aléatoires X et Y ne sont donc pas indépendantes. Cependant, la fonction $f_{(X,Y)}$ est symétrique en ces variables, de sorte que (X,Y) et (Y,X) ont la même loi, ce qui entraîne que

$$\mathbb{P}(X < Y) = \mathbb{P}(Y < X) = \frac{1 - \mathbb{P}(X = Y)}{2} = \frac{1-0}{2}$$

si bien que le jeu est équitable. On retrouve la densité de X en intégrant par rapport à y:

$$f_X(x) = \int_{\mathbb{R}} f_{(X,Y)}(x,y)dy = \mathbf{1}_{x>0} \int_{\mathbb{R}} \mathbf{1}_{0<y<1-x} \frac{dy}{x+y} = \mathbf{1}_{0<x<1} \int_0^{1-x} \frac{dy}{x+y} = -\mathbf{1}_{0<x<1} \ln(x)$$

Pour aller plus loin

Exercice 9. (**Réarrangement croissant de lois exponentielles**) Soient $\lambda_1, \ldots, \lambda_n > 0$. On considère des variables aléatoires $(E_i)_{1 \le i \le n}$ indépendantes telles que E_k suit une loi exponentielle de paramètre λ_k. On note

$$E_{(1)} \le E_{(2)} \le \cdots \le E_{(n)}$$

les variables aléatoires $(E_i)_{1 \le i \le n}$ réarrangées dans l'ordre croissant.

1. Montrer que $E_{(1)}$ suit une loi exponentielle de paramètre $\lambda_1 + \cdots + \lambda_n$.

2. Montrer que $\mathbb{P}(E_{(1)} < E_{(2)}) = 1$ (ainsi le minimum des variables est atteint une unique fois presque sûrement).

3. On note $N = \min\{1 \leq i \leq n : E_i = E_{(1)}\}$. Montrer que N et $E_{(1)}$ sont indépendants, et que $\mathbb{P}(N = k) = \frac{\lambda_k}{\lambda_1 + \cdots + \lambda_n}$ pour tout $1 \leq k \leq n$.

Solution.

1. Si $x < 0$, on a $\mathbb{P}(E_{(1)} \geq x) = 1$. Si $x \geq 0$, on a

$$\mathbb{P}(E_{(1)} \geq x) = \mathbb{P}(E_1 \geq x, \ldots, E_n \geq x) = \prod_{i=1}^{n} \mathbb{P}(E_i \geq x) = e^{-\sum_{i=1}^{n} \lambda_i}.$$

Donc $E_{(1)}$ suit une loi exponentielle de paramètre $\lambda_1 + \cdots + \lambda_n$.

2. On remarque que l'événement $\{E_{(1)} < E_{(2)}\}$ contient l'événement

$$\{\forall i, j \in \{1, 2, \ldots, n\} \text{ tel que } i \neq j \text{ on a } X_i \neq X_j\}$$

Il suffit donc de montrer que ce dernier événement est de probabilité 1 ou, de manière équivalente, que son complémentaire est de probabilité nulle. Pour cela, on adapte l'argument de l'exercice 14 (1) de la PC 3: on a

$$\mathbb{P}\big(\exists i, j \in \{1, 2, \ldots, n\} : i \neq j \text{ et } E_i = E_j\big) \leq \sum_{i,j \in \{1,2,\ldots,n\}, i \neq j} \mathbb{P}(E_i = E_j).$$

Or, d'après le théorème de transfert, en notant f_k la densité de E_k, pour $i \neq j$,

$$\begin{aligned}
\mathbb{P}(E_i = E_j) &= \mathbb{E}(\mathbf{1}_{E_i = E_j}) \\
&= \int_{\mathbb{R}^n} dx_1 dx_2 \cdots dx_n f_1(x_1) f_2(x_2) \cdots f_n(x_n) \mathbf{1}_{x_i = x_j} \\
&= \int_{\mathbb{R}^2} dx dy f_i(x) f_j(y) \mathbf{1}_{x=y} \\
&= \int_{\mathbb{R}} dx f_i(x) \int_x^x f_j(y) dy \\
&= 0.
\end{aligned}$$

D'où le résultat.

3. Calculons pour un entier $1 \leq k \leq n$ et $t \geq 0$ la quantité $\mathbb{P}(E_{(1)} > t, N = k)$ en remarquant les événements $\{E_{(1)} > t, N = k\}$ et $\{E_k > t, E_j > E_k \text{ pour tout } j \neq k\}$ sont égaux:

$$\begin{aligned}
\mathbb{P}(E_{(1)} > t, N = k) &= \mathbb{P}\big(E_k > t, E_j > E_k \text{ pour tout } j \neq k\big) \\
&= \int_{]0,\infty[^n} \mathbf{1}_{x_k > t, x_j > x_k \text{ pour tout } i \neq k} \prod_{i=1}^{n} \lambda_i e^{-\lambda_i x_i} dx_i \\
&= \int_{]0,\infty[} \mathbf{1}_{x_k > t} \lambda_k e^{-(\lambda_1 + \cdots + \lambda_k) x_k} dx_k \\
&= e^{-t(\lambda_1 + \cdots + \lambda_n)} \frac{\lambda_k}{\lambda_1 + \cdots + \lambda_n}.
\end{aligned}$$

En prenant $t = 0$, on trouve que $\mathbb{P}(N = k) = \frac{\lambda_k}{\lambda_1 + \cdots + \lambda_n}$, ce qui permet d'identifier la loi de N, et en réinjectant dans l'égalité précédente on voit que

$$\mathbb{P}(E_{(1)} > t, N = k) = \mathbb{P}(E_{(1)} > t)\mathbb{P}(N = k).$$

Ceci étant valable pour tout $t \in \mathbb{R}$ (il n'y a rien à faire si $t < 0$) et $1 \leq k \leq n$, on en déduit que $E_{(1)}$ et N sont indépendants.

Exercice 10. (**Processus de Poisson**) Soit $(X_n, n \geq 1)$ une suite de variables aléatoires indépendantes définies sur $(\Omega, \mathcal{A}, \mathbb{P})$, de même loi exponentielle de paramètre 1. On pose $T_0 = 0$ et pour tout $n \geq 1$,

$$T_n = X_1 + \cdots + X_n.$$

Pour tout $t \geq 0$, on pose

$$N_t = \max\{n \geq 0 : T_n \leq t\}.$$

1. Soit $n \geq 1$. Calculer la loi du n-uplet $(T_1, \ldots, T_n)$.
2. En déduire la loi de N_t pour tout $t > 0$.
3. Pour $n \geq 1$ et $t > 0$, on définit sur Ω une nouvelle mesure de probabilité $\mathbb{Q}^{n,t}$ par la formule

$$\mathbb{Q}^{n,t}(A) = \frac{\mathbb{P}(A \cap \{N_t = n\})}{\mathbb{P}(N_t = n)} \quad \text{pour tout} \quad A \in \mathcal{A}.$$

Calculer la loi du n-uplet $(T_1, \ldots, T_n)$ sous la mesure de probabilité $\mathbb{Q}^{n,t}$.

Solution.

1. Soit $f : \mathbb{R}_+^n \to \mathbb{R}$, une fonction continue bornée. On a

$$\mathbb{E}(f(T_1, \ldots, T_n)) = \int_{\mathbb{R}_+^n} f(x_1, x_1 + x_2, \ldots, x_1 + x_2 + \cdots + x_n) e^{-(x_1 + x_2 + \cdots + x_n)}\, dx_1 \ldots dx_n.$$

Or $\phi : (x_1, \ldots, x_n) \in (R_+^*)^n \mapsto (x_1, x_1 + x_2, \ldots, x_1 + x_2 + \cdots + x_n) \in \{0 < t_1 < \ldots < t_n\}$ est un $\mathcal{C}^1$-difféomorphisme de jacobien égal à 1 donc d'après la formule du changement de variables,

$$\mathbb{E}(f(T_1, \ldots, T_n)) = \int_{\mathbb{R}_+^n} f(t_1, t_2, \ldots, t_n) e^{-t_n} \mathbf{1}_{\{0 < t_1 < \ldots < t_n\}}\, dt_1 \ldots dt_n,$$

ce qui signfie que la loi de $(T_1, \ldots, T_n)$ a pour densité $e^{-t_n}\mathbf{1}_{\{0 < t_1 < \ldots < t_n\}}$ par rapport à la mesure de Lebesgue sur $\mathbb{R}^n$.

2. Soit $n \in \mathbb{N}$. On remarque que $\mathbb{P}(N_t = n) = \mathbb{P}(T_n \leq t, T_{n+1} > t)$ (car avec probabilité 1 la suite $(T_n)_{n \geq 1}$ est croissante). On en déduit d'après la question (1) que pour $n \geq 1$,

$$\begin{aligned}
\mathbb{P}(N_t = n) &= \int_{\mathbb{R}_+^{n+1}} e^{-t_{n+1}} \mathbf{1}_{\{0 < t_1 < \ldots < t_n < t_{n+1}\}} \mathbf{1}_{\{t_n \leq t\}} \mathbf{1}_{\{t_{n+1} > t\}}\, dt_1 \ldots dt_n dt_{n+1} \\
&= \int_{\mathbb{R}_+^n} \mathbf{1}_{\{0 < t_1 < \ldots < t_n \leq t\}}\, dt_1 \ldots dt_n \int_t^\infty e^{-t_{n+1}} dt_{n+1} \\
&= \frac{t^n}{n!}\, e^{-t},
\end{aligned}$$

où la deuxième égalité est une conséquence du théorème de Fubini-Tonelli. Et l'on a

$$\mathbb{P}(N_t = 0) = \mathbb{P}(T_1 > t) = \mathbb{P}(X_1 > t) = \int_t^{\infty} e^{-x}\, dx = e^{-t}.$$

On voit que N_t suit la loi de Poisson de paramètre t (c'est en fait ce qu'on avait déjà démontré à la question (2) de l'exercice 7).

3. Soit $f : \mathbb{R}_+^n \to \mathbb{R}^+$, une fonction mesurable. On a

$$\begin{aligned}
&\mathbb{E}(f(T_1, \ldots, T_n)\mathbf{1}_{\{N_t=n\}}) \\
&= \mathbb{E}(f(T_1, \ldots, T_n)\mathbf{1}_{\{T_n \le t, T_{n+1} > t\}}) \\
&= \int_{\mathbb{R}_+^{n+1}} f(t_1, \ldots, t_n)\mathbf{1}_{\{t_n \le t, t_{n+1} > t\}})e^{-t_{n+1}}\mathbf{1}_{\{0<t_1<\ldots<t_n<t_{n+1}\}}\, dt_1 \ldots dt_n dt_{n+1} \\
&= \int_{\mathbb{R}_+^n} f(t_1, \ldots, t_n)\mathbf{1}_{\{0<t_1<\ldots<t_n<t\}}\, dt_1 \ldots dt_n \int_t^{\infty} e^{-t_{n+1}} dt_{n+1} dt_{n+1} \\
&= e^{-t} \int_{\mathbb{R}_+^n} f(t_1, \ldots, t_n)\mathbf{1}_{\{0<t_1<\ldots<t_n<t\}}\, dt_1 \ldots dt_n,
\end{aligned}$$

où la troisième égalité est une conséquence du théorème de Fubini-Tonelli.

Nous allons appliquer la méthode de la fonction muette pour trouver la loi de $(T_1, \ldots, T_n)$ sous la nouvelle mesure de probabilité $\mathbb{Q}^{n,t}$. Pour cela, notons $\mathbb{E}^{\mathbb{Q}^{n,t}}$ l'espérance – c'est-à-dire l'intégrale au sens de Lebegue – par rapport à la mesure de probabilité $\mathbb{Q}^{n,t}$, et démontrons d'abord que

$$\mathbb{E}^{\mathbb{Q}^{n,t}}[F] = \frac{\mathbb{E}[F\mathbf{1}_{\{N_t=n\}}]}{\mathbb{P}(N_t = n)}$$

pour toute fonction mesurable positive $F : \Omega \to \mathbb{R}^+$. C'est vrai pour toute fonction étagée F par définition de $\mathbb{Q}^{n,t}$. En effet, si $F = \sum_{i=1}^k a_i \mathbf{1}_{B_i}$, alors

$$\begin{aligned}
\mathbb{E}^{\mathbb{Q}^{n,t}}[F] &= \sum_{i=1}^k a_i \mathbb{E}^{\mathbb{Q}^{n,t}}[\mathbf{1}_{B_i}] \quad \text{(linéarité de l'espérance)} \\
&= \sum_{i=1}^k a_i \mathbb{Q}^{n,t}(B_i) \quad \text{(l'espérance de } \mathbf{1}_{B_i} \text{ est la probabilité de } B_i\text{)} \\
&= \sum_{i=1}^k a_i \frac{\mathbb{P}(B_i \cap \{N_t = n\})}{\mathbb{P}(N_t = n)} \quad \text{(par définition de } \mathbb{Q}^{n,t}\text{)} \\
&= \sum_{i=1}^k a_i \frac{\mathbb{E}[\mathbf{1}_{B_i \cap \{N_t=n\}}]}{\mathbb{P}(N_t = n)} \quad \text{(la probabilité de } B_i \cap \{N_t = n\} \text{ est l'espérance de } \mathbf{1}_{B_i \cap \{N_t=n\}}\text{)} \\
&= \frac{\mathbb{E}[(\sum_{i=1}^k a_i \mathbf{1}_{B_i})\mathbf{1}_{\{N_t=n\}}]}{\mathbb{P}(N_t = n)} \quad \text{(linéarité de l'espérance)} \\
&= \frac{\mathbb{E}[F\mathbf{1}_{\{N_t=n\}}]}{\mathbb{P}(N_t = n)} \quad \text{(par définition de } F\text{).}
\end{aligned}$$

Pour passer des fonctions étagées aux fonctions mesurables positives, on utilise un procédé d'approximation: il existe une suite de fonctions F_n étagées positives (c'est-à-dire de la forme $F_n = \sum_{i=1}^k a_i \mathbf{1}_{B_i}$ avec $a_i \geq 0$ et B_i des événements, de sorte que $\mathbb{E}(F_n) = \sum_{i=1}^n a_i \mathbb{P}(B_i)$) telles que pour tout $\omega \in \Omega$ la suite $(F_n(\omega))_{n\geq 1}$ est croissante et converge simplement vers F (voir l'équation 4.3.16 du poly pour un raisonnement similaire). D'après un théorème de théorie de la mesure (théorème de convergence monotone), on a alors $\mathbb{E}(F_n) \to \mathbb{E}(F)$. Ainsi,

$$\mathbb{E}^{\mathbb{Q}^{n,t}}[F] = \frac{\mathbb{E}[F\mathbf{1}_{\{N_t=n\}}]}{\mathbb{P}(N_t = n)}$$

pour toute fonction mesurable positive $F : \Omega \to \mathbb{R}^+$.

On trouve alors que

$$\mathbb{E}^{\mathbb{Q}^{n,t}}[f(T_1,\dots,T_n)] = \frac{\mathbb{E}[f(T_1,\dots,T_n)\mathbf{1}_{\{N_t=n\}}]}{\mathbb{P}(N_t = n)}$$

en prenant pour F la fonction $F = f(T_1,\dots,T_n)$.

Ainsi,

$$\begin{aligned}
\mathbb{E}^{\mathbb{Q}^{n,t}}(f(T_1,\dots,T_n)) &= \frac{\mathbb{E}(f(T_1,\dots,T_n)\mathbf{1}_{\{N_t=n\}})}{\mathbb{P}(N_t = n)} \\
&= \frac{n!}{t^n} \int_{\mathbb{R}_+^n} f(t_1,\dots,t_n)\mathbf{1}_{\{0<t_1<\dots<t_n<t\}}\, dt_1 \dots dt_n,
\end{aligned}$$

ce qui signfie que la loi de $(T_1,\dots,T_n)$ sous $\mathbb{Q}^{n,t}$ a pour densité $n!\, t^{-n}\ \mathbf{1}_{\{0<t_1<\cdots<t_n<t\}}$ par rapport à la mesure de Lebesgue sur $\mathbb{R}^n$.

Remarque: une adaptation de l'exercice 9 (en remplaçant les variables uniformes sur $[0,1]$ par des variables uniformes sur $[0,t]$) montre que la loi de $(T_1,\dots,T_n)$ sous $\mathbb{Q}^{n,t}$ a la même loi que le réarrangement croissant de n variables aléatoires indépendantes et uniformes sur $[0,t]$.