PC 3 : Variables aléatoires réelles - Vecteurs aléatoires

1 Espérance,Variance et loi d'une variable aléatoire réelle

Exercice 1. *Soit V une variable aléatoire de loi uniforme sur $[0, \pi]$. Déterminer la loi de $\sin(V)$.*

Solution. *Soit $f : \mathbb{R} \to \mathbb{R}$ continue bornée. Comme $f(\sin(V))$ est bornée, l'espérance est bien définie. Par la Proposition 4.5.1 on a :*

$$\mathbb{E}[f(\sin(V))] = \frac{1}{\pi} \int_0^{\pi} f(\sin(x))dx.$$

Changement de variable $u = \sin(x)$ mais attention : sin *n'est pas injective sur $[0, \pi]$. On regarde $[0, \pi/2]$ et $[\pi/2, \pi]$. On remarque :*

$$\frac{1}{\pi} \int_{\pi/2}^{\pi} f(\sin(x))dx = \frac{1}{\pi} \int_0^{\pi/2} f(\sin(\pi - x))dx = \frac{1}{\pi} \int_0^{\pi/2} f(\sin(x))dx$$

On se ramène à

$$\begin{aligned} \frac{1}{\pi} \int_0^{\pi} f(\sin(x))dx &= \frac{2}{\pi} \int_0^{\pi/2} f(\sin(x))dx \\ &= \frac{2}{\pi} \int_0^{1} \frac{f(u)}{\sqrt{1-u^2}}du, \end{aligned}$$

en faisant le changement de variable $u = \sin(x)$. Donc la densité de $\sin(V)$ notée ϕ est donnée par

$$\phi(x) = \frac{2}{\pi\sqrt{1-x^2}} \mathbf{1}_{0 \le x < 1}.$$

Exercice 2 (Loi de Cauchy)**.** *Soit X une variable aléatoire de Cauchy, de densité $(\pi(1+x^2))^{-1}$. Reconnaître la loi de $1/X$ en utilisant la méthode de la fonction muette.*

Solution. *Soit $f : \mathbb{R} \longrightarrow \mathbb{R}$ continue bornée. On a*

$$\mathbb{E}[f(\frac{1}{X})] = \int_{\mathbb{R}} f(\frac{1}{x}) \frac{1}{\pi(1+x^2)}dx.$$

On a envie de faire le changement $u = 1/x$ mais pas bijectif sur $\mathbb{R}$! on scinde en deux

$$\mathbb{E}[f(\frac{1}{X})] = \int_0^{+\infty} f(\frac{1}{x}) \frac{1}{\pi(1+x^2)}dx + \int_{-\infty}^{0} f(\frac{1}{x}) \frac{1}{\pi(1+x^2)}dx.$$

On pose la variable $u = 1/x$ donc $du = -u^2 dx$ ainsi

$$\begin{aligned} \int_0^{+\infty} f(\frac{1}{x}) \frac{1}{\pi(1+x^2)}dx &= \int_0^{+\infty} f(u) \frac{1}{u^2} \frac{1}{\pi(1+u^{-2})}du \\ &= \int_0^{+\infty} f(u) \frac{1}{\pi(1+u^2)}du. \end{aligned}$$

De plus en faisant $y=-x$ dans l'integrale sur $\mathbb{R}^-$ on a

$$\int_{-\infty}^{0} f(\frac{1}{x})\frac{1}{\pi(1+x^2)}dx = \int_{0}^{+\infty} f(-\frac{1}{y})\frac{1}{\pi(1+y^2)}dy.$$

On pose alors $u=-\frac{1}{y}$ et on obtient par calculs similaires

$$\int_{-\infty}^{0} f(\frac{1}{x})\frac{1}{\pi(1+x^2)}dx = \int_{0}^{+\infty} f(u)\frac{1}{\pi(1+u^2)}du.$$

Donc $\frac{1}{X}$ a même loi que X... comme vu dans la PC2.

Exercice 3 (Loi uniforme)**.** *Soit X une v.a. uniforme sur $[0,1]$. On définit $Y=\min(X,1-X)$ et $Z=\max(X,1-X)$. Trouver les lois de Y et Z. Calculer $\mathbb{E}[YZ]$.*

Solution. *La variable aléatoire Y prend ses valeurs dans $[0,1/2]$ et pour tout $t\in[0,1/2]$,*

$$F_Y(t)=\mathbb{P}(Y\le t, X\le 1/2)+\mathbb{P}(Y\le t, X\ge 1/2)=\mathbb{P}(X\le t)+\mathbb{P}(X\ge 1-t)=2t$$

donc Y suit la loi uniforme sur $[0,1/2]$. De même, on montre que Z suit la loi uniforme sur $[1/2,1]$. Pour calculer $\mathbb{E}[YZ]$, on remarque que $YZ=X(1-X)$ et donc

$$\mathbb{E}[YZ]=\mathbb{E}[X(1-X)]=\int_0^1 (t-t^2)\,dt=[t^2/2-t^3/3]_0^1=1/2-1/3=1/6.$$

Remarque : on peut aussi calculer la densité de YZ de la manière suivante. Comme $Y+Z=1$, la variable aléatoire $YZ=Z(1-Z)$ prend ses valeurs dans $[0,1/4]$, et pour tout $t\in[0,1]$,

$$\mathbb{P}(YZ>t)=\mathbb{P}((1-Z)Z>t)=\mathbb{P}\Big(Z\in\Big[\frac{1-\sqrt{1-4t}}{2},\frac{1+\sqrt{1-4t}}{2}\Big]\Big)=\mathbb{P}(2Z-1<\sqrt{1-4t})=\sqrt{1-4t}.$$

La densité de YZ est

$$f_{YZ}(t)=F'_{YZ}(t)=\frac{2}{\sqrt{1-4t}}\mathbf{1}_{[0,1/4]}(t).$$

On peut utiliser cette densité pour retrouver la valeur de $\mathbb{E}[YZ]$ mais les calculs sont plus longs qu'avec la méthode précédente.

$$\mathbb{E}[YZ]=\int_0^{1/4}\frac{2t}{\sqrt{1-4t}}\,dt=\left[\frac{\sqrt{1-4t}\,(1+2t)}{6}\right]_0^{1/4}=\frac{1}{6}.$$

Exercice 4. *Soit X une variable aléatoire de densité $f(x)=\dfrac{\mathbf{1}_{[0,1]}(x)}{(\ln 2)(1+x)}$. Montrer que $Y:=\dfrac{1}{X}-\left[\dfrac{1}{X}\right]$ a même loi que X, où $[x]$ désigne la partie entière de x.*

Solution. *Soit h une fonction de $\mathbb{R}$ dans $\mathbb{R}$ continue bornée. On calcule*

$$\begin{aligned}\mathbb{E}[h(Y)] &= \int_0^1 \frac{1}{(\ln 2)(1+x)}h\Big(\frac{1}{x}-\Big[\frac{1}{x}\Big]\Big)dx\\ &= \sum_{n=1}^{+\infty}\int_{\frac{1}{n+1}}^{\frac{1}{n}}\frac{1}{(\ln 2)(1+x)}h\Big(\frac{1}{x}-n\Big)dx\end{aligned}$$

changement de variable : $y = \frac{1}{x} - n$ donc $dx = -\frac{dy}{(y+n)^2}$. On obtient

$$\begin{aligned}\mathbb{E}[h(Y)] &= \sum_{n=1}^{+\infty} \int_0^1 \frac{1}{(\ln 2)(1+\frac{1}{y+n})(y+n)^2} h(y)dy \\ &= \sum_{n=1}^{+\infty} \int_0^1 \frac{1}{(\ln 2)(y+n)(y+n+1)} h(y)dy \\ &= \frac{1}{\ln 2} \sum_{n=1}^{+\infty} \int_0^1 \left(\frac{1}{y+n} - \frac{1}{y+n+1} \right) h(y)dy \\ &= \frac{1}{\ln 2} \int_0^1 \frac{1}{y+1} h(y)dy.\end{aligned}$$

Exercice 5. *Soient X, Y et Z des variables aléatoires réelles définies sur un même espace de probabilité.*

(1) On suppose que $\mathbb{P}(X = Y) = 1$. Montrer que X et Y ont la même loi. Montrer que la réciproque est fausse.

(2) On suppose que X et Y ont la même loi.

(a) Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction mesurable. Montrer que les variables aléatoires $f(X)$ et $f(Y)$ ont la même loi.

(b) Montrer que les variables aléatoires XZ et YZ n'ont pas nécessairement la même loi.

Solution. *1. Si $\mathbb{P}(X = Y) = 1$, alors pour toute fonction continue bornée $h : \mathbb{R} \to \mathbb{R}$ on a l'égalité $\mathbb{P}(h(X) - h(Y) = 0) = 1$ et donc $\mathbb{E}[h(X) - h(Y)] = 0$, de sorte que $\mathbb{E}[h(X)] = \mathbb{E}[h(Y)]$, ce qui montre que X et Y ont la même loi.*

La réciproque est fausse. Considérons une variable aléatoire X de loi normale $\mathcal{N}(0,1)$ (c'est-à-dire de densité $\sqrt{2\pi}^{-1} e^{-x^2/2}$ par rapport à la mesure de Lebesgue). Posons $Y = -X$. Alors Y est une variable aléatoire X de loi normale $\mathcal{N}(0,1)$. En effet soit $g : \mathbb{R} \to \mathbb{R}$ une fonction continue bornée. Alors

$$\mathbb{E}[g(Y)] = \int_{-\infty}^{+\infty} g(-x)e^{-x^2/2}\, dx = \int_{-\infty}^{+\infty} g(x)e^{-x^2/2}\, dx.$$

Donc X et Y ont la même loi mais ne sont pas égales avec probabilité 1.

2. (a) Pour toute fonction continue bornée $h : \mathbb{R} \to \mathbb{R}$, la fonction $h \circ f$ est mesurable bornée. Comme X et Y ont la même loi, on a

$$\mathbb{E}[h \circ f(X)] = \int_R h(f(x))\mathbb{P}_X(dx) = \int_R h(f(x))\mathbb{P}_Y(dx) = \mathbb{E}[h \circ f(Y)],$$

ce qui montre que $f(X)$ et $f(Y)$ ont la même loi.

(b) On reprend les variables X et Y de la question 1. Soit $Z = X$. Alors $XZ = X^2$ et $YZ = -X^2$. La loi de X^2 est une mesure de probabilité sur $\mathbb{R}_+$ (différente de δ_0 la mesure de Dirac en 0) et la loi de $-X^2$ est une mesure de probabilité sur $\mathbb{R}_-$ donc XZ et YZ n'ont pas la même loi.

2 Inégalités

Exercice 6.

Soit X une variable aléatoire réelle de carré intégrable centrée (i.e. telle que $\mathbb{E}(X) = 0$) et $a > 0$.

1. *Montrer que* $a \leq \mathbb{E}\left((a-X)\mathbf{1}_{\{X<a\}}\right) \leq \sqrt{\mathbb{P}(X<a)} \times \sqrt{\mathrm{Var}(X)+a^2}$.
2. *En déduire que* $\mathbb{P}(X \geq a) \leq \frac{\mathrm{Var}(X)}{\mathrm{Var}(X)+a^2}$ *et comparer avec la majoration obtenue par l'inégalité de Bienaymé Chebychev.*

Solution. 1. *Puisque* X *est centrée :*

$$\begin{aligned} a &= \mathbb{E}[a-X] \\ &= \mathbb{E}[(a-X)\mathbf{1}_{X<a}] + \mathbb{E}[(a-X)\mathbf{1}_{X\geq a}] \\ &\leq \mathbb{E}[(a-X)\mathbf{1}_{X<a}] \\ &\leq \sqrt{\mathbb{P}(X<a)} \times \sqrt{\mathbb{E}[(a-X)^2]} \\ &\leq \sqrt{\mathbb{P}(X<a)} \times \sqrt{a^2+\mathrm{Var}(X)}, \end{aligned}$$

en utilisant l'inégalité de Cauchy Schwarz pour l'avant dernière ligne et le fait que X *est centrée entraine que* $\mathrm{Var}(X) = \mathbb{E}[X^2]$.

2. *On en déduit*

$$\mathbb{P}(X<a) \geq \frac{a^2}{a^2+\mathrm{Var}(X)}$$

et donc

$$\mathbb{P}(X \geq a) = 1 - \mathbb{P}(X<a) \leq 1 - \frac{a^2}{a^2+\mathrm{Var}(X)} = \frac{\mathrm{Var}(X)}{a^2+\mathrm{Var}(X)}$$

Avec Bienaymé Chebychev on aurait

$$\mathbb{P}(X \geq a) \leq \mathbb{P}(|X| \geq a) \leq \frac{\mathrm{Var}(X)}{a^2},$$

donc on a une meilleure estimation qu'avec Bienaymé Chebychev.

Exercice 7. *(Inégalité de Paley-Zygmund) Soit* X *une variable aléatoire réelle intégrable telle que* $\mathbb{E}[X] \geq 0$.

1. *Montrer que pour tout* $\lambda > 0$, $X \leq \lambda\mathbb{E}[X] + X\mathbf{1}_{\{X>\lambda\mathbb{E}[X]\}}$.
2. *On suppose que, de plus,* $0 < \mathbb{E}[X^2] < +\infty$. *Montrer que pour tout* $\lambda \in]0,1[$ *on a*

$$\mathbb{P}(X > \lambda\mathbb{E}[X]) \geq (1-\lambda)^2 \frac{\mathbb{E}[X]^2}{\mathbb{E}[X^2]}.$$

Solution. 1. *On a*

$$\begin{aligned} X &= X\mathbf{1}_{\{X>\lambda\mathbb{E}[X]\}} + X\mathbf{1}_{\{X\leq\lambda\mathbb{E}[X]\}} \\ &\leq X\mathbf{1}_{\{X>\lambda\mathbb{E}[X]\}} + \lambda\mathbb{E}[X], \end{aligned}$$

en utilisant que $\mathbb{E}[X] \geq 0$ *et* $\lambda > 0$.

2. *On prend l'espérance de l'inégalité de la question précédente :*

$$\mathbb{E}[X] \leq \lambda\mathbb{E}[X] + \mathbb{E}[X\mathbf{1}_{\{X>\lambda\mathbb{E}[X]\}}].$$

D'après l'inégalité de Cauchy–Schwarz,

$$\mathbb{E}[X\mathbf{1}_{\{X>\lambda\mathbb{E}[X]\}}] \leq \mathbb{E}[X^2]^{1/2}\mathbb{E}[\mathbf{1}^2_{\{X>\lambda\mathbb{E}[X]\}}]^{1/2} = \mathbb{E}[X^2]^{1/2}\mathbb{P}(X > \lambda\mathbb{E}[X])^{1/2}.$$

Donc

$$(1-\lambda)^2\mathbb{E}[X]^2 \leq \mathbb{E}[X^2]\mathbb{P}(X > \lambda\mathbb{E}[X])$$

et le résultat voulu en découle immédiatement.

Remarque. *L'inégalité de Markov permet de majorer* $\mathbb{P}(X > \lambda\mathbb{E}[X])$, *alors que l'inégalité de Paley-Zygmund permet de minorer cette quantité.*

3 Loi jointe-loi marginale-vecteurs aléatoires

Exercice 8 (Lois jointes, lois marginales et lois conditionnelles)**.** *Soit (X,Y) et (X',Y') des couples de variables aléatoires de densités*

$$f_{(X,Y)}(x,y) = \frac{1}{4}(1+xy)\mathbf{1}_{[-1,1]^2}(x,y) \quad et \quad f_{(X',Y')}(x',y') = \frac{1}{4}\mathbf{1}_{[-1,1]^2}(x',y').$$

1. *Vérifier qu'il s'agit bien de densités.*
2. *Montrer que (X,Y) et (X',Y') ne suivent pas la même loi.*
3. *Montrer que (X,Y) et (X',Y') ont les mêmes lois marginales, c'est-à-dire que X et X' sont de même loi, et que Y et Y' sont de même loi (en fait X, X', Y, Y' sont de même loi !).*

Solution. 1. *Clairement, $f_{(X,Y)} \geq 0$ et $f_{(X',Y')} \geq 0$ sur $\mathbb{R}^2$. On remarque que $\int_{-1}^1 \int_{-1}^1 xy\, dxdy = 0$ et $\int_{-1}^1 \int_{-1}^1 \frac{1}{4} dxdy = 1$. Ainsi*

$$\int_{-1}^1 \int_{-1}^1 f_{(X,Y)}(x,y)\, dxdy = \int_{-1}^1 \int_{-1}^1 f_{(X',Y')}(x,y)\, dxdy = 1.$$

2. *On a*

$$\begin{aligned}
\mathbb{P}(X \geq 0, Y \geq 0) &= \frac{1}{4} \int_0^1 \int_0^1 1 + xy\, dxdy = \frac{1}{4}\left(\int_0^1 dx \int_0^1 dy + \int_0^1 x\, dx \int_0^1 y\, dy\right) \\
&= \frac{1}{4}\left(1 + \frac{1}{2} \times \frac{1}{2}\right) = \frac{5}{16} \\
\mathbb{P}(X' \geq 0, Y' \geq 0) &= \frac{1}{4} \int_0^1 \int_0^1 dxdy = \frac{1}{4}.
\end{aligned}$$

Donc (X,Y) et (X',Y') n'ont pas la même loi.

3. $f_{X'}(x) = \frac{1}{2}\mathbf{1}_{[-1,1]}(x) = \mathbf{1}_{[-1,1]}(x) \int_{-1}^1 \frac{1}{4} du$ *et de même* $f_{Y'}(y) = \frac{1}{2}\mathbf{1}_{[-1,1]}(y)$*. De plus*

$$f_X(x) = \frac{1}{2}\mathbf{1}_{[-1,1]}(x) \int_{-1}^1 \left(\frac{1}{4} + xy\right) dy = \frac{1}{2}\mathbf{1}_{[-1,1]}(x),$$

de même $f_Y(y) = \frac{1}{2}\mathbf{1}_{[-1,1]}(y)$.

Exercice 9. *Soit (X_1, X_2, X_3) un vecteur aléatoire centré de matrice de variance covariance*

$$A = \begin{bmatrix} 2 & 1 & 3 \\ 1 & 5 & 6 \\ 3 & 6 & 9 \end{bmatrix}$$

1. *Calculer la variance de $X_3 - \alpha_1 X_1 - \alpha_2 X_2$ pour $\alpha_1, \alpha_2 \in \mathbb{R}$.*
2. *En déduire que X_3 est combinaison linéaire de X_1 et X_2 p.s.*
3. *Plus généralement, pour un vecteur aléatoire Y de matrice de variance covariance Γ, donner une condition nécessaire et suffisante sur Γ pour que l'une des composantes de Y soit fonction affine des autres composantes de Y p.s.*
4. *Soit maintenant Z un vecteur aléatoire à valeurs dans $\mathbb{R}^d$, $d \geq 1$. Supposons que Z a une densité f par rapport à la mesure de Lebesgue sur $\mathbb{R}^d$. Soit $x \in \mathbb{R}^d$ un vecteur non-nul. Montrer qu'alors la v.a. $U = x^t Z$ a une densité sur $\mathbb{R}$.*

5. *Si Y est un vecteur aléatoire de matrice de variance covariance non-inversible, peut-il avoir une densité ? Le vecteur (X_1, X_2, X_3) a-t-il une densité ?*

Solution. 1. *On obtient*

$$\begin{aligned}Var(X_3 - \alpha_1 X_1 - \alpha_2 X_2) &= \alpha_1^2 Var(X_1) + \alpha_2^2 Var(X_2) + Var(X_3) \\ &\quad - 2\alpha_1 Cov(X_1, X_3) - 2\alpha_2 Cov(X_2, X_3) + 2\alpha_1\alpha_2 Cov(X_1, X_2).\end{aligned}$$

Donc

$$Var(X_3 - \alpha_1 X_1 - \alpha_2 X_2) = 2\alpha_1^2 + 5\alpha_2^2 + 9 - 6\alpha_1 - 12\alpha_2 + 2\alpha_1\alpha_2.$$

2. *En prenant $\alpha_1 = \alpha_2 = 1$ on obtient*

$$Var(X_3 - X_1 - X_2) = 0$$

donc il existe une constante c telle que

$$X_3 = X_1 + X_2 + c.$$

3. *Soit Y un vecteur de taille n de matrice de variance-covariance Γ. On remarque que*

$$\begin{aligned}Ker(\Gamma) \neq 0 &\iff \exists\, u \neq 0_{\mathbb{R}^n},\ \Gamma u = 0 \\ &\iff \exists\, u \neq 0_{\mathbb{R}^n},\ u^\top \Gamma u = 0 \\ &\iff \exists\, u \neq 0_{\mathbb{R}^n},\ \sum_{i,j} u_j Cov(Y_i, Y_j) u_i = 0 \\ &\iff \exists\, u \neq 0_{\mathbb{R}^n},\ Var\Big(\sum_{i=1}^n u_i Y_i\Big) = 0.\end{aligned}$$

Comme Γ est symétrique, on en déduit que Γ est non-inversible si et seulement si l'une des composantes de Y est fonction affine des autres composantes de Y p.s.

4. *Soit $h : \mathbb{R} \to \mathbb{R}$ une fonction continue bornée. Puisque x est non nul, il existe un indice k tel que $x_k \neq 0$. On pose le changement de variable $(z_1, z_2, \dots, x^t(z_1, \dots, z_d), \dots, z_{d-1}, z_d) = (u_1, u_2, \dots, u_k, \dots, u_{d-1}, u_d)$ dont l'inverse est donné par $(z_1, z_2, \dots, z_k, \dots, z_{d-1}, z_d) = (u_1, u_2, \dots, \frac{u_k - \sum_{i\neq k} x_i u_i}{x_k}, \dots, u_{d-1}, u_d)$. Le Jacobien est x_k^{-1} d'où*

$$\begin{aligned}\mathbb{E}[h(U)] = \mathbb{E}[h(x^t Z)] &= \int_{\mathbb{R}^d} h(x^t(z_1, \dots, z_d)) f(z_1, \dots, z_d)\, dz_1 \cdots dz_d \\ &= \frac{1}{x_k} \int_{\mathbb{R}^d} h(x^t(z_1, \dots, z_d)) f(u_1, \dots, \frac{u_k - \sum_{i\neq k} x_i u_i}{x_k}, \dots, u_d)\, du_1 \cdots du_d \\ &= \int_{\mathbb{R}} h(u_k) \Big(\int_{\mathbb{R}^{d-1}} \frac{1}{x_k} f(u_1, \dots, \frac{u_k - \sum_{i\neq k} x_i u_i}{x_k}, \dots, u_d)\, d\hat{u}^k \Big) du_k\end{aligned}$$

avec la notation $d\hat{u}^k$ pour la mesure de Lebesgue $du_1 \cdots du_d$ en excluant du_k, donc U possède une densité donnée par la fonction

$$f_U(u) = \int_{\mathbb{R}^{d-1}} \frac{1}{x_k} f(u_1, \dots, \frac{u - \sum_{i\neq k} x_i u_i}{x_k}, \dots, u_d)\, d\hat{u}^k.$$

5. *Si la matrice Γ est non-inversible, alors il existe $u \in \mathbb{R}^n \setminus \{0\}$ tel que $\langle u, Y \rangle = c$ où c est une constante, par la question 2. Donc Y ne peut pas avoir de densité sinon on contredit la question 4. On en déduit que X n'a pas de densité car Γ est non-inversible.*

Exercice 10. *Soit (X,Y) un couple de variables aléatoires à valeurs dans $\mathbb{R}^2$ dont la loi a pour densité*

$$f_{(X,Y)}(x,y) = \frac{2}{\pi} e^{-x(1+y^2)} \mathbf{1}_{x,y\geq 0}.$$

1. *Vérifier que $f_{(X,Y)}$ est bien une densité.*
2. *Déterminer les lois de X et de Y.*
3. *Les variables aléatoires X et Y sont-elles indépendantes ?*

Solution. 1. *$f_{(X,Y)}$ est bien positive, vérifions que son intégrale sur $\mathbb{R}^2$ vaut bien 1.*

$$\begin{aligned}
\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} dxdy\ \frac{2}{\pi} e^{-x(1+y^2)} \mathbf{1}_{x,y\geq 0} &= \int_0^{\infty} dy \frac{2}{\pi} \int_0^{\infty} e^{-x(1+y^2)} dx \\
&= \int_0^{\infty} dy \frac{2}{\pi} \cdot \frac{1}{1+y^2} \\
&= \frac{2}{\pi} \left[\arctan(y)\right]_0^{\infty} \\
&= 1.
\end{aligned}$$

2. *Soit $h : \mathbb{R} \longrightarrow \mathbb{R}$ une fonction continue bornée. On calcule $\mathbb{E}[h(X)]$ et $\mathbb{E}[h(Y)]$:*

$$\begin{aligned}
\mathbb{E}[h(X)] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} dxdy\ h(x) \frac{2}{\pi} e^{-x(1+y^2)} \mathbf{1}_{x,y\geq 0} \\
&= \int_0^{\infty} dx\ h(x) \frac{2e^{-x}}{\pi} \int_0^{\infty} e^{-xy^2} dy \\
&= \int_0^{\infty} dx\ h(x) \frac{2e^{-x}}{\pi\sqrt{2x}} \int_0^{\infty} e^{-z^2/2} dz \qquad (\textit{en posant } z = \sqrt{2x}\cdot y) \\
&= \int_0^{\infty} dx\ h(x) \frac{2e^{-x}}{\pi\sqrt{2x}} \cdot \frac{\sqrt{2\pi}}{2}.
\end{aligned}$$

Donc X est à densité, et $f_X(x) = \frac{e^{-x}}{\sqrt{\pi x}} \mathbf{1}_{x\geq 0}$ est une densité de X.
Aussi,

$$\begin{aligned}
\mathbb{E}[h(X)] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} dxdy\ h(y) \frac{2}{\pi} e^{-x(1+y^2)} \mathbf{1}_{x,y\geq 0} \\
&= \int_0^{\infty} dy \frac{2}{\pi} h(y) \int_0^{\infty} e^{-x(1+y^2)} \\
&= \int_0^{\infty} dy \frac{2}{\pi} h(y) \cdot \frac{1}{1+y^2}.
\end{aligned}$$

Donc Y est à densité, et $f_Y(y) = \frac{2}{\pi}\cdot\frac{1}{1+y^2}\mathbf{1}_{y\geq 0}$ est une densité de Y.

3. *X et Y sont indépendantes si et seulement si pour presque tous $(x,y) \in \mathbb{R}^2$ on a $f_{(X,Y)}(x,y) = f_X(x) f_Y(y)$, ou encore, pour tous $x, y \geq 0$,*

$$e^{-x(1+y^2)} = \frac{e^{-x}}{\sqrt{\pi x}} \cdot \frac{1}{1+y^2},$$

ce qui n'est clairement pas le cas. Donc X et Y ne sont pas indépendantes.

Remarque. *Le* clairement pas *met des choses sous le tapis... Justifions cela. Tout d'abord, si $g : \mathbb{R}^2 \to \mathbb{R}$ est une fonction continue nulle presque partout (pour la mesure de Lebesgue sur $\mathbb{R}^2$), alors en fait g est nulle partout. En effet, raisonnons par l'absurde et choisissons $x \in \mathbb{R}^2$*

tel que $g(x) \neq 0$. Alors il existe une boule ouverte B_x centrée en x telle que $g(y) \neq 0$ pour tout $y \in B_x$. Or B_x est de mesure de Lebesgue strictement positive, absurde.

Dans notre cas, supposons par l'absurde que pour presque tous $(x,y) \in \mathbb{R}_+^2$ on a

$$e^{-x(1+y^2)} = \frac{e^{-x}}{\sqrt{\pi x}} \cdot \frac{1}{1+y^2}.$$

La fonction continue $g(x,y) = e^{-x(1+y^2)} - \frac{e^{-x}}{\sqrt{\pi x}} \cdot \frac{1}{1+y^2}$ est donc presque partout nulle, et donc partout nulle. Or $g(1,0) \neq 0$, absurde.

Exercice 11. *On considère un gâteau circulaire avec une cerise sur le bord. On découpe le gâteau en deux parts en coupant suivant deux rayons choisis au hasard.*

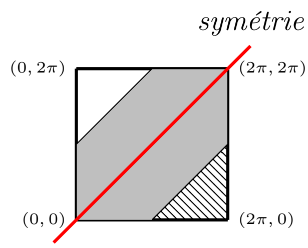
1. *Avec quelle probabilité la part contenant la cerise est-elle plus petite que la part ne contenant pas la cerise ?*
2. *Quelle est la longueur angulaire moyenne de la part contenant la cerise ?*

Solution. *On note Θ_1 et Θ_2 les angles formés par les deux rayons et le rayon qui passe par la cerise. L'énoncé du problème indique que Θ_1 et Θ_2 sont indépendants et suivent la loi uniforme sur $[0, 2\pi]$. La longueur angulaire de la part contenant la cerise est égale à $2\pi - |\Theta_1 - \Theta_2|$.*

1. *La probabilité pour que la part contenant la cerise soit la plus petite est égale à $\mathbb{P}(2\pi - |\Theta_1 - \Theta_2| < |\Theta_1 - \Theta_2|)$. On en déduit que*

$$\begin{aligned} \mathbb{P}(2\pi - |\Theta_1 - \Theta_2| < |\Theta_1 - \Theta_2|) &= \mathbb{E}[\mathbf{1}_{2\pi - |\Theta_1 - \Theta_2| < |\Theta_1 - \Theta_2|}] \\ &= \frac{1}{(2\pi)^2} \int_0^{2\pi} \int_0^{2\pi} d\theta_1 d\theta_2 \mathbf{1}_{|\theta_1 - \theta_2| > \pi} \\ &= \frac{1}{(2\pi)^2} 2 \int_0^{2\pi} d\theta_1 \int_0^{\theta_1} d\theta_2 \mathbf{1}_{\theta_1 - \theta_2 > \pi} \\ &= \frac{1}{4} \end{aligned}$$

où pour l'avant dernière égalité on a utilisé la symétrie entre θ_1 et θ_2 pour se restreindre au cas où $\theta_2 < \theta_1$. Géométriquement la dernière égalité est l'aire du petit triangle hachuré. La probabilité pour que la part contenant la cerise soit la plus petite est donc $1/4$.



2. *La longueur moyenne de la part contenant la cerise est égale à $2\pi - \mathbb{E}[|\Theta_1 - \Theta_2|]$, qu'on calcule avec le théorème de transfert et la même astuce de symétrie :*

$$\begin{aligned} \mathbb{E}[|\Theta_1 - \Theta_2|] &= \frac{1}{(2\pi)^2} \int_0^{2\pi} \int_0^{2\pi} d\theta_1 d\theta_2 |\theta_1 - \theta_2| \\ &= \frac{1}{(2\pi)^2} 2 \int_0^{2\pi} d\theta_1 \int_0^{\theta_1} (\theta_1 - \theta_2) d\theta_2 \\ &= \frac{2\pi}{3} \end{aligned}$$

La longueur moyenne de la part contenant la cerise est donc $4\pi/3$.

La part qui contient la cerise est plus grande en moyenne et elle est également plus grande dans 75% des cas. Pour voir que ces résultats ne contredisent pas l'intuition il faut inverser les *opérations. On découpe d'abord au hasard deux rayons dans le gâteau, puis on jette au hasard la cerise sur le bord. Celle-ci a intuitivement plus de chance de tomber sur la part la plus grosse ! Il reste à se convaincre que jeter la cerise sur le bord puis couper le gâteau au hasard, ou couper le gâteau au hasard puis jeter la cerise sur le bord donne bien le même résultat*