Leçon 258 — Couples de variables aléatoires possédant une densité. Couples de variables aléatoires possdant une densité. Covariance. Exemples d'utilisation.

Corrigé partiel des exercices

Exercice 1 (Algorithme de Box–Müller)**.**
Soit $A =]0,1[\times]0,1] \subset \mathbb{R}^2$. On définit une fonction $g : A \to \mathbb{R}^2$ par

$$g(u,v) = \big(\sqrt{-2\ln(u)}\cos(2\pi u), \sqrt{-2\ln(u)}\sin(2\pi u)\big) .$$

On vérifie que g est une bijection de A vers $B = \mathbb{R}^2 \setminus \{(0,0)\}$ (on enlève le point $(0,0)$ et la valeur $u = 1$ car g envoie tous les points $(1,v)$ sur $(0,0)$). Cette bijection est obtenue en composant la bijection $(u,v) \mapsto (r,\varphi) = (\sqrt{-2\ln(u)}, 2\pi v)$ avec une transformation $(r,\varphi) \mapsto (x,y) = (r\cos(\varphi), r\sin(\varphi))$ de coordonnées polaires en coordonnées rectangulaires.

La matrice jacobienne de g est donnée par

$$\begin{pmatrix} \dfrac{\partial g_1}{\partial u} & \dfrac{\partial g_1}{\partial v} \\ \dfrac{\partial g_2}{\partial u} & \dfrac{\partial g_2}{\partial v} \end{pmatrix} = \begin{pmatrix} -\dfrac{\cos(2\pi v)}{u\sqrt{-2\ln(u)}} & -2\pi\sqrt{-2\ln(u)}\sin(2\pi v) \\ -\dfrac{\sin(2\pi v)}{u\sqrt{-2\ln(u)}} & 2\pi\sqrt{-2\ln(u)}\cos(2\pi v) \end{pmatrix} ,$$

dont le déterminant vaut

$$\mathrm{Jac}(g)(u,v) = -\frac{2\pi}{u} .$$

On remarque qu'on peut partiellement inverser g, puisque si $(x,y) = g(u,v)$, alors

$$x^2 + y^2 = -2\ln(u) \quad \Rightarrow \quad u = \mathrm{e}^{-(x^2+y^2)/2} .$$

Ceci nous permet de calculer le Jacobien de la réciproque g^{-1} :

$$\mathrm{Jac}(g^{-1})(x,y) = \frac{1}{\mathrm{Jac}(g)(g^{-1}(x,y))} = -\frac{\mathrm{e}^{-(x^2+y^2)/2}}{2\pi} .$$

Le couple (U,V) admet la densité $f_{U,V}(u,v) = 1_{[0,1]\times[0,1]}(u,v)$. Comme $g^{-1}(x,y) \in A \subset [0,1] \times [0,1]$ pour tout $(x,y) \in B$, la formule de changement de variable montre que la densité du couple (X,Y) est donnée par

$$f_{X,Y}(x,y) = \underbrace{f_{U,V}((g^{-1}(x,y))}_{=1} \big|\mathrm{Jac}(g^{-1})(x,y)\big| = \frac{\mathrm{e}^{-(x^2+y^2)/2}}{2\pi}$$

pour tout $(x,y) \in B = \mathbb{R}^2 \setminus \{(0,0)\}$. On peut prolonger cette fonction par continuité en $(0,0)$. On en conclut que (X,Y) suit une loi normale standard sur $\mathbb{R}^2$, et donc que X et Y sont indépendantes, chacune de loi normale standard.

L'algorithme de Box–Müller permet de simuler des variables aléatoires indépendantes suivant une loi normale à partir de variables aléatoires indépendantes de loi uniforme.

Exercice 4 (Loi Gamma)**.**
Si X_1 suit une loi exponentielle de paramètre 1, alors sa densité est donnée par

$$f(x) = \mathrm{e}^{-x}\, 1_{\{x>0\}} \; .$$

Soit $S_2 = X_1 + X_2$ la somme de deux variables aléatoires réelles indépendantes de loi exponentielle de paramètre 1. Sa densité est donnée par la convolution

$$f_{S_2}(y) = \int_{-\infty}^{\infty} f(y-x)f(x)\,\mathrm{d}x = \int_0^y \mathrm{e}^{-y}\, 1_{\{y>0\}}\,\mathrm{d}x = y\,\mathrm{e}^{-y}\, 1_{\{y>0\}} \; .$$

C'est la densité d'une loi Gamma $\gamma(2,1)$. Montrons alors par récurrence que S_n suit une loi $\gamma(n,1)$. L'initialisation vient d'être montrée. Supposons que S_n suit une loi $\gamma(n,1)$, de densité $\frac{1}{(n-1)!}x^{n-1}\,\mathrm{e}^{-x}\, 1_{\{x>0\}}$, et montrons que $S_{n+1} = S_n + X_{n+1}$ suit une loi $\gamma(n+1,1)$. La densité de S_{n+1} est donnée par la convolution

$$\begin{aligned} f_{S_{n+1}}(y) &= \int_{-\infty}^{\infty} f_{S_n}(y-x)f(x)\,\mathrm{d}x \\ &= \frac{1}{(n-1)!}\int_0^y (y-x)^{n-1}\,\mathrm{e}^{-y}\, 1_{\{y>0\}}\,\mathrm{d}x \\ &= \frac{1}{n!}y^n\,\mathrm{e}^{-y}\, 1_{\{y>0\}} \; . \end{aligned}$$

C'est bien la densité de la loi $\gamma(n+1,1)$.

Exercice 5.

1. La fonction f est manifestement non-négative. De plus on a
$$\int_{\mathbb{R}^2} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_0^{\infty}\int_0^y \mathrm{d}x\,\mathrm{e}^{-y}\,\mathrm{d}y = \int_0^{\infty} y\,\mathrm{e}^{-y}\,\mathrm{d}y = -y\,\mathrm{e}^{-y}\Big|_0^{\infty} + \int_0^{\infty}\mathrm{e}^{-y}\,\mathrm{d}y = 1 \; ,$$
ce qui montre que f est bien une densité.

2. Les densités marginales valent respectivement
$$\begin{aligned} f_1(x) &= \int_{\mathbb{R}} f(x,y)\,\mathrm{d}y = \int_x^{\infty} \mathrm{e}^{-y}\,\mathrm{d}y\, 1_{\{x>0\}} = \mathrm{e}^{-x}\, 1_{\{x>0\}} \; , \\ f_2(y) &= \int_{\mathbb{R}} f(x,y)\,\mathrm{d}x = \int_0^y \mathrm{e}^{-y}\,\mathrm{d}x\, 1_{\{y>0\}} = y\,\mathrm{e}^{-y}\, 1_{\{y>0\}} \; . \end{aligned}$$
X suit donc une loi exponentielle de paramètre 1, alors que Y suit une loi Gamma de paramètres $(1,1)$.

3. Comme $f_1(x)f_2(y) = y\,\mathrm{e}^{-(x+y)}\, 1_{\{x>0,y>0\}} \neq f(x,y)$, les variables X et Y ne sont pas indépendantes.

4. Pour $0 \leqslant z \leqslant 1$, et $Y > 0$, la condition $X/Y \leqslant z$ équivaut à $X \leqslant zY$. Il suit
$$\mathbb{P}\biggl\{\frac{X}{Y} \leqslant z, Y \leqslant t\biggr\} = \int_0^t\int_0^{zy} \mathrm{e}^{-y}\,\mathrm{d}x\,\mathrm{d}y = \int_0^t zy\,\mathrm{e}^{-y}\,\mathrm{d}y = z\bigl[1 + (t-1)\,\mathrm{e}^{-t}\bigr] \; .$$
On en déduit
$$\mathbb{P}\biggl\{\frac{X}{Y} \leqslant z\biggr\} = \lim_{t\to\infty}\mathbb{P}\biggl\{\frac{X}{Y} \leqslant z, Y \leqslant t\biggr\} = z$$
pour $z \in [0,1]$, donc que X/Y suit la loi uniforme sur $[0,1]$.

5. Les variables X/Y et Y sont indépendantes, puisque nous venons d'établir que leur fonction de répartition conjointe se factorise en produit d'une fonction de z et d'une fonction de t.

Exercice 6 (Aiguille de Buffon)**.**

On suppose que Y et Φ sont indépendantes, avec Y de loi uniforme sur $[0,1[$ et Φ de loi uniforme sur $[0,2\pi[$ (on suppose que les deux extrémités de l'aiguille sont distinguables, par exemple de couleur bleue et rouge, et que Φ mesure l'angle par rapport au Nord de l'extrémité bleue). Cette hypothèse est justifiée si les jets sont suffisamment "aléatoires" : aucune direction et aucune position n'est privilégiée. Le couple (Φ, Y) admet donc la densité

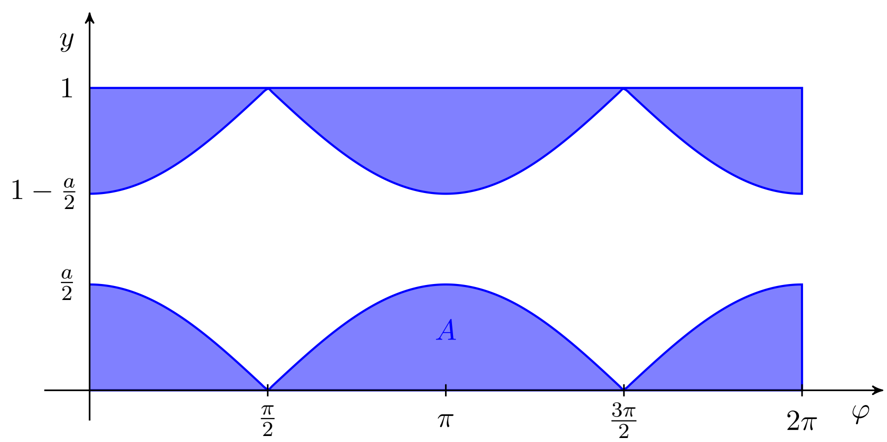
$$f_{\Phi,Y}(\varphi, y) = \frac{1}{2\pi} 1_{[0,2\pi[\times[0,1[}(\varphi, y) .$$

Les extrémités de l'aiguille ont comme coordonnées y

$$Y \pm \frac{a}{2}\big|\cos\Phi\big| .$$

Par conséquent, l'aiguille intersecte une ligne si

- soit $Y - \frac{a}{2}|\cos\Phi| < 0$,
- soit $Y + \frac{a}{2}|\cos\Phi| > 1$.



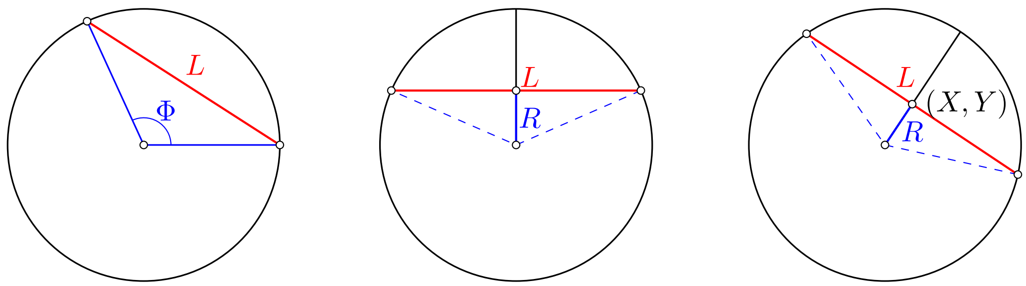
Soit A le domaine de $[0,2\pi[\times[0,1[$ défini par ces conditions. La probabilité cherchée est donnée par

$$P = \int_A f_{\Phi,Y}(\varphi, y)\,\mathrm{d}\varphi\,\mathrm{d}y = \frac{4}{2\pi}\int_{\pi/2}^{3\pi/2}\left(-\frac{a}{2}\cos\varphi\right)\mathrm{d}\varphi = \frac{2a}{\pi} .$$

On s'est servi ici des symétries du problème, qui impliquent que l'intégrale sur A est donnée par 4 fois l'intégrale sur un "lobe".

L'expérience de l'aiguille de Buffon permet d'obtenir une approximation du nombre π : en jetant un grand nombre d'aiguilles, la loi des grands nombres implique que la proportion p d'aiguilles coupant une ligne a une grande probabilité d'être proche de P, et donc que π a une grande probabilité d'être proche de $2a/p$.

Exercice 7 ("Paradoxe" de Bertrand)**.**



1. On choisit deux points sur le cercle, formant un angle Φ avec le centre du cercle, où Φ suit une loi uniforme sur $[0, 2\pi[$. La longueur de la corde entre ces points vaut

$$L = 2\sin\Big(\frac{\Phi}{2}\Big) .$$

Il suit que

$$\mathbb{P}\{L > \sqrt{3}\} = \mathbb{P}\Big\{\sin\Big(\frac{\Phi}{2}\Big) > \frac{\sqrt{3}}{2}\Big\} = \mathbb{P}\Big\{\frac{2\pi}{3} < \Phi < \frac{4\pi}{3}\Big\} = \int_{2\pi/3}^{4\pi/3} \frac{1}{2\pi}\,\mathrm{d}\varphi = \frac{1}{3} .$$

Remarque: On obtient le même résultat en partant d'un couple (Φ_1, Φ_2) de v.a. de loi uniforme sur $[0, 2\pi[\times[0, 2\pi[$, et considérer la variable aléatoire $L = 2\sin(\frac{\Phi_2-\Phi_1}{2})$.

2. On choisit un point sur le rayon, dont la distance R au centre du cercle suit une loi uniforme sur $[0, 1]$. La corde perpendiculaire au rayon en ce point est la base d'un triangle isocèle de hauteur R, dont la longueur vaut

$$L = 2\sqrt{1 - R^2} .$$

Il suit que

$$\mathbb{P}\{L > \sqrt{3}\} = \mathbb{P}\{4(1 - R^2) > 3\} = \mathbb{P}\Big\{R^2 < \frac{1}{4}\Big\} = \mathbb{P}\Big\{R < \frac{1}{2}\Big\} = \frac{1}{2} .$$

3. Le couple (X, Y) d'un point choisi de manière uniforme dans le cercle admet la densité

$$f_{X,Y}(x, y) = \frac{1}{\pi} 1_{\{x^2+y^2<1\}} .$$

Ce point se trouve à distance $R = \sqrt{X^2 + Y^2}$ du centre, et la corde admettant ce point pour milieu a la longueur

$$L = 2\sqrt{1 - R^2} = 2\sqrt{1 - X^2 - Y^2} .$$

Par conséquent,

$$\mathbb{P}\{L > \sqrt{3}\} = \mathbb{P}\Big\{1 - X^2 - Y^2 > \frac{3}{4}\Big\} = \mathbb{P}\Big\{X^2 + Y^2 < \frac{1}{4}\Big\} = \int_{x^2+y^2<\frac{1}{4}} \frac{1}{\pi}\,\mathrm{d}x\,\mathrm{d}y = \frac{1}{4} ,$$

où la dernière intégrale se calcule en passant en coordonnées polaires.

Cet exemple fut introduit par le mathématicien Joseph Bertrand dans son ouvrage de 1889 "Calcul des probabilités", afin de montrer qu'une probabilité peut ne pas être bien définie si le choix de l'aléa n'est pas spécifié.