Leçon 14 Exercices corrigés

(Une étoile * désignera une question de difficulté supérieure.)

Exercice 1. L'objet de l'exercice est d'obtenir un bon encadrement de la probabilité de queue d'une variable X de loi $\mathcal{N}(0,1)$, à savoir

$$\mathbb{P}(X \ge t) = \frac{1}{\sqrt{2\pi}} \int_{[t,\infty[} e^{-\frac{1}{2}x^2} d\lambda = \frac{1}{\sqrt{2\pi}} \int_t^{\infty} e^{-\frac{1}{2}x^2} dx$$

quand $t \ge 0$ est grand.

a) Démontrer, par une étude de fonction, que pour tout $t \geq 0$,

$$\mathbb{P}(X \ge t) = \frac{1}{\sqrt{2\pi}} \int_{t}^{\infty} e^{-\frac{1}{2}x^{2}} dx \le \frac{1}{2} e^{-\frac{1}{2}t^{2}}.$$

b) Par une intégration par parties, établir que si t > 0,

$$\int_{t}^{\infty} e^{-\frac{1}{2}x^{2}} dx = \frac{1}{t} e^{-\frac{1}{2}t^{2}} - \int_{t}^{\infty} \frac{1}{x^{2}} e^{-\frac{1}{2}x^{2}} dx.$$

c) Vérifier de la même façon que si t > 0,

$$\int_{t}^{\infty} \frac{1}{x^{2}} e^{-\frac{1}{2}x^{2}} dx = \frac{1}{t^{3}} e^{-\frac{1}{2}t^{2}} - \int_{t}^{\infty} \frac{3}{x^{4}} e^{-\frac{1}{2}x^{2}} dx.$$

d) Conclure des questions b) et c) que pour tout t > 0,

$$\left(\frac{1}{t} - \frac{1}{t^3}\right) \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} \le \mathbb{P}(X \ge t) \le \frac{1}{t} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2}.$$

Indication. b) Pour t>0, $\int_t^\infty e^{-\frac{1}{2}x^2}dx=\int_t^\infty \frac{1}{x}\,x\,e^{-\frac{1}{2}x^2}dx$, et $-e^{-\frac{1}{2}x^2}$ est une primitive de $x\,e^{-\frac{1}{2}x^2}$.

Exercice 2. Montrer qu'il existe un vecteur gaussien centré X à valeurs dans \mathbb{R}^3 de matrice de covariance

$$\Sigma = \begin{pmatrix} 5 & 0 & 3 \\ 0 & 6 & 0 \\ 3 & 0 & 2 \end{pmatrix}.$$

Calculer $\mathbb{E}(\langle c, X \rangle^2)$ pour tout $c \in \mathbb{R}^3$.

Indication. D'après les constructions de la leçon, il suffit de vérifier que la matrice Σ est symétrique et (semi-) définie positive.

Exercice 3. Sur $(\Omega, \mathcal{A}, \mathbb{P})$, soit X une variable aléatoire de loi normale $\mathcal{N}(0,1)$ et soit ε , une variable aléatoire indépendante de X, telle que $\mathbb{P}(\varepsilon = -1) = \mathbb{P}(\varepsilon = +1) = \frac{1}{2}$. Démontrer que εX suit la loi $\mathcal{N}(0,1)$. Le couple $(X, \varepsilon X)$ est-il gaussien?

Corrigé. Les divers outils décrivant une loi sont à disposition pour vérifier que εX suit la loi $\mathcal{N}(0,1)$, toutes s'appuyant sur le fait que la densité de la loi $\mathcal{N}(0,1)$ est paire, et en conséquence que la loi de X est symétrique (voir Exercice 5, Leçon 9). Par exemple, avec la fonction de répartition, pour tout $t \in \mathbb{R}$,

$$\mathbb{P}(\varepsilon X \leq t) \, = \, \mathbb{P}(\varepsilon = +1, X \leq t) + \mathbb{P}(\varepsilon = -1, -X \leq t).$$

Par indépendance entre ε et X,

$$\mathbb{P}(\varepsilon X \leq t) \, = \, \frac{1}{2} \, \mathbb{P}(X \leq t) + \frac{1}{2} \, \mathbb{P}(-X \leq t) \, = \, \mathbb{P}(X \leq t)$$

car

$$\mathbb{P}(-X \le t) = \mathbb{P}(X \ge -t) = \frac{1}{\sqrt{2\pi}} \int_{[-t, +\infty[} e^{-\frac{1}{2}x^2} d\lambda$$
$$= \frac{1}{\sqrt{2\pi}} \int_{]-\infty, t]} e^{-\frac{1}{2}x^2} d\lambda = \mathbb{P}(X \le t)$$

après le changement de variable $x \mapsto -x$ (il n'est pas inutile non plus de représenter graphiquement ces intégrales). Ainsi εX et X ont même fonction de répartition, donc même loi. Si le couple $(X, \varepsilon X)$ était gaussien, par exemple la combinaison linéaire $X + \varepsilon X$ devrait suivre une loi normale en tant que variable aléatoire réelle. À ce titre, $\mathbb{P}(X + \varepsilon X = 0) = 0$ car les lois gaussiennes ont une densité par rapport à la mesure de Lebesgue sur \mathbb{R} . Mais

$$\mathbb{P}(X + \varepsilon X = 0) \ge \mathbb{P}(\varepsilon = -1) = \frac{1}{2}$$

ce qui exprime la contradiction.

Exercice 4. Soient X et Y deux variables aléatoires normales centrées réduites $\mathcal{N}(0,1)$ indépendantes. Quelle est la loi du couple (X,Y)? Déterminer la loi de $\frac{X}{Y}$.

Corrigé. Par indépendance, la loi du couple (X,Y) a pour densité $\frac{1}{2\pi}e^{-\frac{1}{2}(x^2+y^2)}$, $(x,y)\in\mathbb{R}^2$, par rapport à la mesure de Lebesgue λ^2 sur \mathbb{R}^2 . Comme $\mathbb{P}(Y=0)=0$, la variable aléatoire $\frac{X}{Y}$ est bien définie presque sûrement. Si $\phi:\mathbb{R}\to\mathbb{R}$ est borélienne, positive ou bornée, par le théorème de transport pour la loi du couple (X,Y),

$$\mathbb{E}\left(\phi\left(\frac{X}{Y}\right)\right) = \frac{1}{2\pi} \int_{\mathbb{R}^2} \phi\left(\frac{x}{y}\right) e^{-\frac{1}{2}(x^2 + y^2)} d\lambda^2(x, y).$$

Après le changement de variable $(u, v) = (\frac{x}{y}, y)$ de jacobien $\frac{1}{y} = \frac{1}{v}$,

$$\mathbb{E}\left(\phi\left(\frac{X}{Y}\right)\right) = \frac{1}{2\pi} \int_{\mathbb{R}^2} \phi(u) \, e^{-\frac{1}{2}(1+u^2)v^2} |v| d\lambda^2(u,v).$$

(Rigoureusement, le changement de variable prendra place pour $(x, y) \in \mathbb{R} \times \mathbb{R}^*$.) Comme $\lambda^2 = \lambda \otimes \lambda$, en vertu du théorème de Fubini-Tonelli,

$$\mathbb{E}\left(\phi\left(\frac{X}{Y}\right)\right) = \frac{1}{2\pi} \int_{\mathbb{R}} \phi(u) \left(\int_{\mathbb{R}} e^{-\frac{1}{2}(1+u^2)v^2} |v| d\lambda(v)\right) d\lambda(u)$$
$$= \int_{\mathbb{R}} \phi(u) \frac{1}{\pi(1+u^2)} d\lambda(u)$$

après une intégration par calcul de primitive en la variable v (séparer v > 0 et v < 0). La loi de $\frac{X}{Y}$ est donc la loi de Cauchy de densité $\frac{1}{\pi(1+u^2)}$, $u \in \mathbb{R}$, par rapport à la mesure de Lebesgue sur \mathbb{R} .

Exercice 5*. Soit (X,Y) un couple aléatoire de loi $\mathcal{N}(0,\mathrm{Id})$ sur \mathbb{R}^2 ; rappeler les lois marginales. Démontrer que XY a même loi que $\frac{1}{2}(X^2-Y^2)$. (Indication: utiliser la formule de polarisation $4XY=(X+Y)^2-(X-Y)^2$.)

Corrigé. Comme X et Y sont indépendantes de même loi $\mathcal{N}(0,1)$, $\frac{1}{\sqrt{2}}(X+Y)$ et $\frac{1}{\sqrt{2}}(X-Y)$ sont toutes deux de loi $\mathcal{N}(0,1)$ après vérification de leur moyenne et de leur variance. La formule de polarisation indiquée se transcrit sous la forme

$$XY = \frac{1}{2} \left(\left[\frac{1}{\sqrt{2}} (X + Y) \right]^2 - \left[\frac{1}{\sqrt{2}} (X - Y) \right]^2 \right).$$

L'affirmation demandée s'ensuivra donc sous réserve de démontrer que $\frac{1}{\sqrt{2}}(X+Y)$ et $\frac{1}{\sqrt{2}}(X-Y)$ sont indépendantes. (Autrement dit, le couple

$$\left(\frac{1}{\sqrt{2}}(X+Y), \frac{1}{\sqrt{2}}(X-Y)\right)$$

a même loi que (X,Y).) Il revient au même de démontrer que X+Y et X-Y sont indépendantes. Mais, le couple (X+Y,X-Y) étant gaussien (toute combinaison linéaire des coordonnées l'est car (X,Y) lui-même est gaussien), il suffit de démontrer que X+Y et X-Y ne sont pas corrélées. Or $\mathbb{E}(X+Y)=\mathbb{E}(X-Y)=0$ et

$$\mathbb{E}((X+Y)(X-Y)) = \mathbb{E}(X^2) - \mathbb{E}(Y^2) = 0.$$

Le résultat est établi.

Exercice 6. Soit $X = (X_1, ..., X_d)$ un vecteur aléatoire de loi gaussienne standard $\mathcal{N}(0, \mathrm{Id})$ sur \mathbb{R}^d ; démontrer que la loi de $\frac{X}{|X|}$ est invariante par transformation orthogonale (pour rappel $|X| = \left(\sum_{k=1}^d X_k^2\right)^{\frac{1}{2}}$). Quelle est cette loi?

Corrigé. Une matrice $d \times d$ orthogonale O est telle que $O^{\top}O = {}^{\top}O O = {\rm Id}$; en particulier, pour tout $x \in \mathbb{R}^d$, |Ox| = |x|. Par suite, pour tout borélien B de \mathbb{R}^d ,

$$\mathbb{P}(OX \in B) = \mathbb{P}(X \in {}^{\top}OB)$$

$$= \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{{}^{\top}OB} e^{-\frac{1}{2}|x|^2} d\lambda$$

$$= \frac{1}{(2\pi)^{\frac{d}{2}}} \int_{B} e^{-\frac{1}{2}|x|^2} d\lambda = \mathbb{P}(X \in B)$$

où l'avant-dernière égalité résulte du changement de variable y = Ox et de l'invariance par transformation orthogonale de la mesure de Lebesgue. Donc la loi de X est invariante par transformation orthogonale, et comme |OX| = |X|, il en va de même de la loi de $\frac{X}{|X|}$. Il est clair que la variable aléatoire $\frac{X}{|X|}$ est concentrée sur la sphère unité de \mathbb{R}^d . Donc sa loi est la mesure uniforme sur celle-ci (décrite dans le Théorème 3, Leçon 4).

Exercice 7*. Soit $X = (X_1, X_2, X_3, X_4)$ un vecteur gaussien centré de dimension 4; établir l'identité

$$\mathbb{E}(X_1 X_2 X_3 X_4) = \mathbb{E}(X_1 X_2) \mathbb{E}(X_3 X_4) + \mathbb{E}(X_1 X_3) \mathbb{E}(X_2 X_4) + \mathbb{E}(X_1 X_4) \mathbb{E}(X_2 X_3).$$

(Indications: Plusieurs arguments sont possibles. Par exemple, si $Y = (Y_1, Y_2, Y_3, Y_4)$ est un vecteur indépendant de même loi que X, un contrôle des covariances assure que X + Y a même loi que $\sqrt{2}X$. Développer alors l'identité qui en résulte

$$4\mathbb{E}(X_1X_2X_3X_4) = \mathbb{E}((X_1+Y_1)(X_2+Y_2)(X_3+Y_3)(X_4+Y_4)).$$

Sinon, si $u=(u_1,u_2,u_3,u_4)\in\mathbb{R}^4$, l'expression de la fonction caractéristique est $\mathbb{E}(e^{i\langle u,X\rangle})=e^{-\frac{1}{2}\mathbb{E}(\langle u,X\rangle^2)}$; développer les deux expressions à l'ordre 4, et identifier les termes correspondants aux u_k tous différents.)

Exercice 8. Soient $0 = t_0 < t_1 < t_2 < \cdots < t_d$ des nombres reéls, et soit (X_0, X_1, \ldots, X_d) un vecteur aléatoire gaussien centré de matrice de covariance

$$\mathbb{E}(X_k X_\ell) = \min(t_k, t_\ell), \quad k, \ell = 0, 1, \dots, d.$$

(Comme $\mathbb{E}(X_0^2) = 0$, il sera convenu que $X_0 = 0$ presque sûrement.) Déterminer les lois marginales. Poser $Y_k = \frac{X_k - X_{k-1}}{\sqrt{t_k - t_{k-1}}}, k = 1, \dots, d$. Démontrer que le vecteur aléatoire (Y_1, \dots, Y_d) est de loi $\mathcal{N}(0, \mathrm{Id})$ dans \mathbb{R}^d .

Exercice 9. Soient X_1, \ldots, X_n des variables aléatoires indépendantes de loi $\mathcal{N}(0,1)$, et soient $a_1, \ldots, a_n, b_1, \ldots, b_n$ des réels. Démontrer que les variables aléatoires $\sum_{k=1}^n a_k X_k$ et $\sum_{k=1}^n b_k X_k$ sont indépendantes si et seulement si $\sum_{k=1}^n a_k b_k = 0$.

Indication. Utiliser l'équivalence entre indépendance et orthogonalité pour des variables gaussiennes.

Exercice 10. Soient $X = (X_1, ..., X_d)$ et $Y = (Y_1, ..., Y_d)$ deux vecteurs aléatoires gaussiens centrés sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, supposés indépendants et de même loi. Pour tout réel θ , soient $X(\theta) = X \sin(\theta) + Y \cos(\theta)$ et $X'(\theta) = X \cos(\theta) - Y \sin(\theta)$; démontrer que pour tout θ , $X(\theta)$ et $X'(\theta)$ sont des vecteurs aléatoires gaussiens indépendants de même loi que X.

Corrigé. Comme X et Y sont indépendants, le couple Z=(X,Y) est un vecteur gaussien (centré) de \mathbb{R}^{2d} . Pour tout $\theta \in \mathbb{R}$, il en va de même du couple $Z(\theta)=(X(\theta),X'(\theta))$ car toute combinaison linéaire des coordonnées de $Z(\theta)$ est une combinaison linéaire des coordonnées de (X,Y). Par indépendance, la matrice de covariance de Z dans \mathbb{R}^{2d} est une matrice à blocs de la forme

$$\begin{pmatrix} \Sigma & 0 \\ 0 & \Sigma \end{pmatrix}$$

où Σ est la matrice de covariance $d \times d$ commune à X et Y (et 0 la matrice $d \times d$ nulle). La suite de la démonstration va consister à montrer que, pour

tout $\theta \in \mathbb{R}$, la matrice de covariance de $Z(\theta)$ est la même que celle-ci, ce qui démontrera à la fois l'indépendance de $X(\theta)$ et $X'(\theta)$ et le fait qu'ils ont tous deux même loi que X et Y. Par centrage, il revient à examiner $\mathbb{E}(Z(\theta)_k Z(\theta)_\ell)$ pour $k, \ell = 1, \ldots, 2d$. Trois cas sont à étudier. Pour $k, \ell = 1, \ldots, d$,

$$\mathbb{E}(Z(\theta)_k Z(\theta)_\ell) = \mathbb{E}(X(\theta)_k X(\theta)_\ell)$$

$$= \mathbb{E}([X_k \sin(\theta) + Y_k \cos(\theta)] [X_\ell \sin(\theta) + Y_\ell \cos(\theta)])$$

$$= \sin^2(\theta) \mathbb{E}(X_k X_\ell) + \cos^2(\theta) \mathbb{E}(Y_k Y_\ell))$$

où il a été utilisé que, par indépendance de X et Y, $\mathbb{E}(X_kY_\ell)=0$ et $\mathbb{E}(X_\ell Y_k)=0$. Comme X et Y ont la même loi, il s'ensuit que $\mathbb{E}(Z(\theta)_k Z(\theta)_\ell)=\Sigma_{k,\ell}$ pour tous $k,\ell=1,\ldots,d$. Un raisonnement similaire indique que $\mathbb{E}(Z(\theta)_{d+k}Z(\theta)_{d+\ell})=\Sigma_{k,\ell}$ pour tous $k,\ell=1,\ldots,d$. En revanche, pour $k,\ell=1,\ldots,d$,

$$\mathbb{E}(Z(\theta)_k Z(\theta)_{d+\ell}) = \mathbb{E}(X(\theta)_k X'(\theta)_\ell)$$

$$= \mathbb{E}([X_k \sin(\theta) + Y_k \cos(\theta)][X_\ell \cos(\theta) - Y_\ell \sin(\theta)])$$

$$= \sin(\theta) \cos(\theta) \left(\mathbb{E}(X_k X_\ell) - \mathbb{E}(Y_k Y_\ell)\right)$$

qui est toujours nul car X et Y sont de même loi. Il en va de même pour $\mathbb{E}(Z(\theta)_{d+k}Z(\theta)_{\ell})$. Ainsi $Z(\theta)$ a même matrice de covariance que Z, ce qui conclut la démonstration.