PC 5 : Calcul de lois & Vecteurs gaussiens

On discutera prioritairement les exercices 1, 2, 5, et 8.
On appelle fonction Gamma, Γ, la fonction définie par

$$\Gamma(\alpha) = \int_0^{+\infty} t^{\alpha-1}e^{-t}\,\mathrm{d}t, \quad \forall\alpha > 0.$$

1 Changements de variables

Exercice 1 (DEUX GAMMA)**.**
Soit X et Y des variables aléatoires indépendantes, X suivant une loi Gamma de paramètres (α, λ), notée $\Gamma(\alpha, \lambda)$, de densité

$$f_{\Gamma(a,\lambda)}(x) = \Gamma(\alpha)^{-1}\,\lambda^\alpha\,x^{\alpha-1}\,\mathrm{e}^{-\lambda x}\,\mathbb{1}_{\mathbb{R}_+^*}(x),$$

et Y suivant une loi Gamma de paramètres (β, λ).

1. Donner une densité conjointe de $U = X + Y$ et $V = \frac{X}{X+Y}$.
2. Quelles sont les lois de U et de V ? Sont-elles indépendantes ?
3. Montrer que $\mathbb{E}[\frac{X}{X+Y}] = \frac{\mathbb{E}[X]}{\mathbb{E}[X]+\mathbb{E}[Y]}$.

Exercice 2 (LOI DE STUDENT)**.**
Soient Z et S des variables indépendantes de lois respectives $\mathcal{N}(0, 1)$ et χ^2_n ; montrer que pour tout $n > 0$ la variable aléatoire

$$T_n = \frac{Z}{\sqrt{S/n}} \qquad \text{a pour densité} \qquad f_{T_n}(t) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{t^2}{n}\right)^{-\frac{n+1}{2}}.$$

On appelle cette loi la *loi de Student* à n degrés de liberté.
On rappelle que $\chi^2_n \sim \Gamma(\alpha = n/2, \lambda = 1/2)$.

Exercice 3 (PALE 2013)**.**
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).

2 Vecteurs gaussiens

Exercice 4 (Méthode de Box-Müller, cf. aussi le cours - exercice corrigé)**.** Soit R une variable aléatoire de loi exponentielle de paramètre $1/2$ et Θ une variable aléatoire de loi uniforme sur $[0, 2\pi]$. On suppose que R et Θ sont indépendantes. Quelle est la loi jointe de $(X, Y) = (\sqrt{R}\cos(\Theta), \sqrt{R}\sin(\Theta))$? En déduire une méthode de simulation d'une variable gaussienne sur $\mathbb{R}^2$ à partir de deux variables i.i.d. uniformément distribuées sur $[0, 1]$.

Solution. On utilise ici la méthode de la fonction muette. Soit f une fonction continue bornée de $\mathbb{R}^2$ dans $\mathbb{R}^2$. En utilisant l'indépendance de Θ et R on voit que la loi du couple (R, Θ) admet une densité par rapport à la mesure de Lebesgue qui est le produit (tensoriel) des deux densités. On obtient alors par le théorème de transfert que

$$\begin{aligned} E\Big[f(\sqrt{R}\cos(\Theta), \sqrt{R}\sin(\Theta))\Big] &= \int_{\mathbb{R}}\int_{\mathbb{R}} f(\sqrt{r}\cos(\theta), \sqrt{r}\sin(\theta))\frac{1}{2}\,\mathrm{e}^{-\frac{1}{2}r}\,\mathbf{1}_{r\in\mathbb{R}^+} \times \frac{1}{2\pi}\mathbf{1}_{\theta\in[0,2\pi]}\mathrm{d}r\mathrm{d}\theta \\ &= \int_0^{+\infty}\int_0^{2\pi} f(\sqrt{r}\cos(\theta), \sqrt{r}\sin(\theta))\frac{1}{2}\,\mathrm{e}^{-\frac{1}{2}r} \times \frac{1}{2\pi}\mathrm{d}r\mathrm{d}\theta. \end{aligned}$$

Pour mener à bien la méthode de la fonction muette jusqu'à sa fin, on doit/veut effectuer le changement de variables suivant : $(x, y) = (\sqrt{r}\cos(\theta), \sqrt{r}\sin(\theta))$ dans la dernière intégrale. Pour cela, on va appliquer le théorème de changement de variables qui s'applique bien quand la fonction de changement de variables est un $\mathcal{C}^1$-difféomorphisme. On considère

$$\varphi : \left\{ \begin{array}{ccc} U & \to & V \\ (r, \theta) & \mapsto & (\sqrt{r}\cos(\theta), \sqrt{r}\sin(\theta)) \end{array} \right.$$

où $U :=]0, +\infty[\times]0, 2\pi[$ et $V = \mathbb{R}^2 \setminus \{0\} \times \mathbb{R}^+$ sont deux ouverts de $\mathbb{R}^2$. On voit que φ est une bijection de U sur V en tant que composition de deux bijections : $(r', \theta) \longrightarrow (r'\cos(\theta), r'\sin(\theta))$ est le changement en coordonnées polaires qui est une bijection de U sur V et $(r, \theta) \to (\sqrt{r}, \theta)$ est une bijection de U sur U. Par ailleurs, φ est $\mathcal{C}^1$ en tant que composée de fonctions de classe $\mathcal{C}^1$ et sa matrice jacobienne est donnée pour tout $(r, \theta) \in U$ par

$$J(\varphi)(r, \theta) = \begin{pmatrix} \frac{\cos(\theta)}{2\sqrt{r}} & -\sqrt{r}\sin(\theta) \\ \frac{\sin(\theta)}{2\sqrt{r}} & \sqrt{r}\cos(\theta) \end{pmatrix}$$

qui est inversible car de déterminant égale à 1/2. La fonction φ est bien un $\mathcal{C}^1$-difféomorphisme (on utilise ici que φ est inversible, $\mathcal{C}^1$ et de différentielle inversible – car de matrice Jacobienne inversible – comme caractérisation d'un $\mathcal{C}^1$-difféomorphisme) et le théorème de changement de variable donne

$$\int_V h(x, y)\mathrm{d}x\mathrm{d}y = \int_U h(\varphi(r, \theta))|\det(J(\varphi)(r, \theta))|\mathrm{d}r\mathrm{d}\theta$$

(de manière informelle $(x, y) = (\sqrt{r}\cos(\theta), \sqrt{r}\sin(\theta))$ et $\mathrm{d}x\mathrm{d}y = |\det(J(\varphi)(r, \theta))|\mathrm{d}r\mathrm{d}\theta$. On a aussi la formule d'inversion $r = x^2 + y^2$; on ne donne pas celle en θ car on ne s'en servira pas et elle fait intervenir plusieurs cas en fonction du signe de x et y).
On obtient

$$\begin{aligned} E\Big[f(\sqrt{R}\cos(\Theta), \sqrt{R}\sin(\Theta))\Big] &= \int_{\mathbb{R}}\int_{\mathbb{R}} f(x, y)\frac{1}{2}\,\mathrm{e}^{-\frac{1}{2}(x^2+y^2)} \times \frac{1}{2\pi}2\mathrm{d}x\mathrm{d}y \\ &= \int_{\mathbb{R}}\int_{\mathbb{R}} f(x, y)\,\mathrm{e}^{-\frac{1}{2}(x^2+y^2)} \times \frac{1}{2\pi}\mathrm{d}x\mathrm{d}y. \end{aligned}$$

On en déduit, par la méthode de la fonction muette, que (X, Y) suit une loi normale standard (centrée et de matrice de covariance identité I_2) sur $\mathbb{R}^2$.

Soit $\mu \in \mathbb{R}^2$ et $\Sigma \in \mathbb{R}^{2\times 2}$ une matrice symétrique semi-définie positive. On souhaite simuler une v.a.r. distribuée selon $\mathcal{N}(\mu, \Sigma)$. On commence par observer que si $X \sim \mathcal{N}(0, I_2)$ alors $\mu + \Sigma^{1/2}X$ suit une $\mathcal{N}(\mu, \Sigma)$. Il suffit donc de simuler une $X \sim \mathcal{N}(0, I_2)$. Pour cela, on utilise la question précédente et la méthode d'inversion de la fonction de répartition. Soit U_1 et U_2 deux variables i.i.d. uniformément distribuées sur $[0, 1]$. On a vu en PC2 que $F^{(-1)} : p \in]0, 1[\mapsto -2\ln(1-p)$ est l'inverse généralisée de la fonction de répartition d'une Exp(1/2). On en déduit que $R := F^{(-1)}(U_1)$ est une variable aléatoire de loi exponentielle de paramètre 1/2 et $\Theta := 2\pi U_2$ est une variable aléatoire de loi uniforme sur $[0, 2\pi]$. De plus, R et Θ sont indépendantes vu que U_1 et U_2 le sont. On conclut avec la première question.

Remarque : du fait des fonctions trigonométriques, certains argumentent qu'une façon plus rapide de simuler cette loi consiste à tirer un point uniformément au hasard dans le disque unité via la méthode du rejet, cela donne l'angle Θ, et on simule la norme R^2 comme précédemment.

Exercice 5.
Soit (X, Y) un couple gaussien suivant la loi $\mathcal{N}(0, \mathrm{Id}_2)$, c'est-à-dire que X et Y sont indépendantes et de même loi $\mathcal{N}(0, 1)$.

1. Trouver la loi du couple (U, V) où $U = \frac{X+Y}{\sqrt{2}}$ et $V = \frac{Y-X}{\sqrt{2}}$.
2. Comment interprétez-vous la transformation qui fait passer de (X, Y) à (U, V) ?

Exercice 6. Soit X et Y deux variables aléatoires indépendantes telles que X suit la loi normale standard et $\mathbb{P}(Y = 1) = 1 - \mathbb{P}(Y = -1) = p \in]0, 1[$. Montrer que :

1. $Z := XY$ suit la loi normale standard.

2. $\operatorname{Cov}(X, Z) = 2p - 1$.
3. X et Z ne sont pas indépendants.
4. Le vecteur aléatoire (X, Z) n'est pas un vecteur gaussien.

Exercice 7. 1. Soit $\theta \in]0, \pi[$; montrer l'identité

$$\int_{]0,\infty[^2} \mathrm{e}^{-(x^2+y^2+2xy\cos\theta)} \,\mathrm{d}x\mathrm{d}y = \frac{\theta}{2\sin\theta}.$$

2. En déduire que si (X, Y) est un vecteur gaussien tel que chaque variable est centrée réduite et $\operatorname{Cov}(X, Y) = \rho \in]-1, 1[$, alors

$$\mathbb{P}(X > 0, Y > 0) = \frac{1}{4} + \frac{\arcsin(\rho)}{2\pi}.$$

Commenter pour $\rho = 0$, $\rho \uparrow 1$ et $\rho \downarrow -1$.

3. Calculer pour ces v.a. $\mathbb{E}[X \mid Y]$ et $\operatorname{Var}(X \mid Y) := \mathbb{E}[(X - \mathbb{E}[X \mid Y])^2 \mid Y]$.

3 Somme de variables aléatoires (exponentielles)

Exercice 8 (Somme d'exponentielles).
Soient $X_1, \ldots, X_n$ 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 une densité conditionnelle de X_1 sachant que $S_n = t$ pour tout $t > 0$.
3. En déduire la valeur de $\mathbb{E}[X_1 \mid S_n]$. Pouvait-on prévoir ce résultat ?

Exercice 9 (Max d'exponentielles).
Soit $X_1, \ldots, X_n$ des variables aléatoires indépendantes qui suivent toutes une loi exponentielle de paramètre $\lambda > 0$. Montrer que les variables aléatoires

$$Y_n = \max(X_1, \ldots, X_n) \qquad \text{et} \qquad Z_n = X_1 + \frac{X_2}{2} + \cdots + \frac{X_n}{n}$$

ont la même loi.

Exercice 10 (D'autres exponentielles).
Soit $X_1, \ldots, X_n$ des variables aléatoires indépendantes, telles que X_i suit la loi exponentielle de paramètre $\lambda_i > 0$ pour chaque $i \in \{1, \ldots, n\}$. On suppose que ces λ_i sont tous distincts ; donner la loi de la somme $X_1 + \cdots + X_n$.