PC 5 – Calcul de lois & Vecteurs gaussiens

Exercice 1. Soient X et Y deux variables aléatoires indépendantes gaussiennes centrées réduites.

1. Déterminer la loi de $\left(\frac{X+Y}{\sqrt{2}}, \frac{X-Y}{\sqrt{2}}\right)$.
2. Déterminer la loi de X/Y.

Solution. Comme X et Y sont indépendantes, la loi de (X,Y) a une densité $\frac{1}{2\pi}e^{-\frac{x^2+y^2}{2}}$ sur $\mathbb{R}^2$. Soit $g : \mathbb{R}^2 \to \mathbb{R}$ une fonction continue bornée. On applique la méthode de la fonction muette en calculant $\mathbb{E}\left[g\left(\frac{X+Y}{\sqrt{2}}, \frac{X-Y}{\sqrt{2}}\right)\right]$:

$$\mathbb{E}\left[g\left(\frac{X+Y}{\sqrt{2}}, \frac{X-Y}{\sqrt{2}}\right)\right] = \frac{1}{2\pi}\int_{\mathbb{R}^2} g\left(\frac{x+y}{\sqrt{2}}, \frac{x-y}{\sqrt{2}}\right) e^{-\frac{x^2+y^2}{2}}\, dx\, dy.$$

Or $(x,y) \in \mathbb{R}\times\mathbb{R} \mapsto (\frac{x+y}{\sqrt{2}}, \frac{x-y}{\sqrt{2}}) \in \mathbb{R}\times\mathbb{R}$ est un C^1-difféormorphisme de jacobien 1. Le changement de variable $u = \frac{x+y}{\sqrt{2}}$ et $v = \frac{x-y}{\sqrt{2}}$ donne $x = \frac{u+v}{\sqrt{2}}$ et $y = \frac{u-v}{\sqrt{2}}$, de sorte que :

$$\begin{aligned}\mathbb{E}\left[g\left(\frac{X+Y}{\sqrt{2}}, \frac{X-Y}{\sqrt{2}}\right)\right] &= \frac{1}{2\pi}\int_{\mathbb{R}^2} g(u,v)\, e^{-\frac{(u+v)^2+(u-v)^2}{4}}\, du\, dv \\ &= \frac{1}{2\pi}\int_{\mathbb{R}^2} g(u,v)\, e^{-\frac{u^2+v^2}{2}}\, du\, dv.\end{aligned}$$

On en déduit que $(\frac{X+Y}{\sqrt{2}}, \frac{X-Y}{\sqrt{2}})$ est à densité, de densité donnée par $(u,v) \mapsto \frac{1}{2\pi}e^{-\frac{u^2+v^2}{2}}$. Ainsi, $(\frac{X+Y}{\sqrt{2}}, \frac{X-Y}{\sqrt{2}})$ a la même loi qu'un couple de deux variables aléatoires gaussiennes centrées réduites indépendantes.

Comme X et Y sont indépendantes, la loi de (X,Y) a une densité $\frac{1}{2\pi}e^{-\frac{x^2+y^2}{2}}$ sur $\mathbb{R}^2$. Soit $g : \mathbb{R} \to \mathbb{R}$ une fonction continue bornée. On applique la méthode de la fonction muette en calculant $\mathbb{E}\left[g(X/Y)\right]$:

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

Or $(x,y) \in \mathbb{R}\times\mathbb{R}^* \mapsto (x/y, y) \in \mathbb{R}\times\mathbb{R}^*$ est un C^1-difféormorphisme de jacobien y^{-1}. En faisant le changement de variable $u = x/y$ et $v = y$, de sorte que $x = uv$ et $y = v$, on a

$$\begin{aligned}\int_{\mathbb{R}^2} g\left(\frac{x}{y}\right) e^{-\frac{x^2+y^2}{2}}\, dx\, dy &= \int_{\mathbb{R}^2} g\left(\frac{x}{y}\right) |y|\, e^{-\frac{y^2}{2}(\frac{x^2}{y^2}+1)} |y|^{-1}\, dx\, dy \\ &= \int_{\mathbb{R}^2} g(u)\, |v|\, e^{-\frac{v^2}{2}(u^2+1)}\, du\, dv \\ &= \int_{\mathbb{R}} g(u) \left(\int_{\mathbb{R}} |v|\, e^{-\frac{v^2}{2}(u^2+1)}\, dv\right) du \\ &= 2\int_{\mathbb{R}} g(u) \frac{1}{u^2+1}\, du.\end{aligned}$$

Donc

$$\mathbb{E}\left(g\left(\frac{X}{Y}\right)\right) = \frac{1}{\pi}\int_{\mathbb{R}} g(u)\frac{1}{u^2+1}\,du,$$

ce qui signifie que la loi de X/Y est la loi de Cauchy, c'est-à-dire la loi de densité $(\pi(1+x^2))^{-1}$ par rapport à la mesure de Lebesgue.

Exercice 2. (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).

On rappelle que la densité de la loi $\Gamma(a,\lambda)$ est

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

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 :

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

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 $x=v^2/w^2$,

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

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.

1. Soit (X,Y) un couple de variables indépendantes de lois respectives $\Gamma(a,\lambda)$ et $\Gamma(b,\lambda)$. Déterminer la loi jointe du vecteur aléatoire (U,V) où $U=X/Y$ et $V=X+Y$.
2. Soient Z et S des variables indépendantes de lois respectives $\mathcal{N}(0,1)$ et χ^2_n. On appelle *loi de Student* à n degrés de liberté la loi de la variable $T=\frac{Z}{\sqrt{S/n}}$. Montrer que la densité de T est donnée sur $\mathbb{R}$ par

$$t\mapsto \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{t^2}{n}\right)^{-\frac{n+1}{2}}.$$

Solution. 1. D'après le cours, V suit la loi Gamma $\Gamma(a+b,\lambda)$. En revanche, pour déterminer la densité jointe de (U,V) on utilisera la méthode de la fonction muette. Soit $h:\mathbb{R}^2\to\mathbb{R}^2$ une fonction continue, bornée. Notons que par indépendance $f_{(X,Y)}=f_Xf_Y$. On a

$$\begin{aligned}\mathbb{E}\left[h(U,V)\right] &= \int_{\mathbb{R}^2} h\left(\frac{x}{y},x+y\right) f_{(X,Y)}(x,y)d(x,y)\\ &= \frac{\lambda^{a+b}}{\Gamma(a)\Gamma(b)}\int_{\mathbb{R}_+^2} h\left(\frac{x}{y},x+y\right) x^{a-1}y^{b-1}e^{-\lambda(x+y)}d(x,y).\end{aligned}$$

Faisons le changement de variables $(u,v)=\phi(x,y):=(x/y,x+y)$. La réciproque est $\phi^{-1}(u,v)=\left(\frac{uv}{1+u},\frac{v}{1+u}\right)$. Le Jacobien J de $\phi^{-1}(u,v)$ vaut

$$J=\det\begin{pmatrix}\frac{v}{(1+u)^2} & \frac{u}{1+u}\\ -\frac{v}{(1+u)^2} & \frac{1}{1+u}\end{pmatrix}=\frac{v}{(1+u)^3}+\frac{uv}{(1+u)^3}=\frac{v}{(1+u)^2}.$$

On en déduit que

$$\begin{aligned}\mathbb{E}\left[h(U,V)\right] &= \frac{\lambda^{a+b}}{\Gamma(a)\Gamma(b)}\int_{\mathbb{R}^2} h(u,v)\left(\frac{uv}{1+u}\right)^{a-1}\left(\frac{v}{1+u}\right)^{b-1}e^{-\lambda v}\frac{|v|}{(1+u)^2}\mathbb{1}_{\{\frac{uv}{1+u}>0,\frac{v}{1+u}>0\}}d(u,v)\\ &= \int_{\mathbb{R}^2} h(u,v)\underbrace{\frac{\lambda^{a+b}}{\Gamma(a+b)}v^{a+b-1}e^{-\lambda v}\mathbb{1}_{v>0}}_{=f_V(v)}\underbrace{\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\frac{u^{a-1}}{(1+u)^{a+b}}\mathbb{1}_{u>0}}_{=f_U(u)}d(u,v).\end{aligned}$$

On observe que V suit bien la loi Gamma $\Gamma(a+b,\lambda)$. La loi de U est dite loi beta prime de paramètres a et b. En plus, U et V sont indépendantes, car la densité jointe se factorise.

2. Soit $h:\mathbb{R}\to\mathbb{R}$ une fonction continue, bornée. On a

$$\begin{aligned}\mathbb{E}\left[h(T)\right] &= \int_{\mathbb{R}^2} h\left(\sqrt{\frac{n}{s}}z\right) f_Z(z)f_S(s)d(z,s)\\ &= \int_{\mathbb{R}_+}\int_{\mathbb{R}} h\left(\sqrt{\frac{n}{s}}z\right)\frac{1}{\sqrt{2\pi}}e^{-\frac{z^2}{2}}\frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}s^{\frac{n}{2}-1}e^{-\frac{s}{2}}dzds.\end{aligned}$$

Par le changement de variable $t=\sqrt{\frac{n}{s}}z$ avec $dz=\sqrt{\frac{s}{n}}dt$, on obtient

$$\begin{aligned}\mathbb{E}\left[h(T)\right] &= \frac{1}{2^{\frac{n+1}{2}}\Gamma\left(\frac{n}{2}\right)\sqrt{\pi}}\int_{\mathbb{R}_+}\int_{\mathbb{R}} h(t)e^{-\frac{st^2}{2n}}s^{\frac{n}{2}-1}e^{-\frac{s}{2}}\sqrt{\frac{s}{n}}dtds\\ &= \frac{1}{2^{\frac{n+1}{2}}\Gamma\left(\frac{n}{2}\right)\sqrt{n\pi}}\int_{\mathbb{R}} h(t)\int_{\mathbb{R}_+}\underbrace{s^{\frac{n+1}{2}-1}\exp\left\{-\frac{s}{2}\left(\frac{t^2}{n}+1\right)\right\}}_{=\text{ densité de la loi }\Gamma\left(\frac{n+1}{2},\frac{1}{2}\left(\frac{t^2}{n}+1\right)\right)\text{ à une constante près}}dsdt\\ &= \frac{\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\sqrt{n\pi}}\int_{\mathbb{R}} h(t)\left(\frac{t^2}{n}+1\right)^{-\frac{n+1}{2}}dt.\end{aligned}$$

On trouve bien que T a la densité $\frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{t^2}{n}\right)^{-\frac{n+1}{2}}$.

Exercice 4. 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. On va montrer que (X_1, X_2) est un vecteur gaussien centré, de matrice de covariance $\Gamma_{12} = \begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix}$. On calcul la densité de (X_1, X_2) par la formule

$$\begin{aligned} f_{(X_1,X_2)}(x_1, x_2) &= \int_{\mathbb{R}} \frac{1}{(2\pi)^{3/2}\sqrt{|\Gamma_{12}|}} e^{-\frac{1}{2}(\mathbf{x}\Gamma^{-1}\mathbf{x})} dx_3 \\ &= \frac{1}{2\pi\sqrt{|\Gamma_{12}|}} \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(\mathbf{x}\Gamma^{-1}\mathbf{x})} dx_3, \end{aligned}$$

puisque $|\Gamma| = |\Gamma_{12}|$. On a de plus $\Gamma^{-1} = \begin{pmatrix} \Gamma_{12}^{-1} & 0 \\ 0 & 1 \end{pmatrix}$ donc

$$\begin{aligned} f_{(X_1,X_2)}(x_1, x_2) &= \frac{1}{2\pi\sqrt{|\Gamma_{12}|}} e^{-\frac{1}{2}((x_1,x_2)\Gamma_{12}^{-1}(x_1,x_2)^t)} \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}|x_3|^2} dx_3 \\ &= \frac{1}{2\pi\sqrt{|\Gamma_{12}|}} e^{-\frac{1}{2}((x_1,x_2)\Gamma_{12}^{-1}(x_1,x_2)^t)}. \end{aligned}$$

3. On a $(X_2, X_2 + aX_1)^t = A(X_1, X_2)^t$ avec $A = \begin{pmatrix} 0 & 1 \\ a & 1 \end{pmatrix}$. Donc $(X_2, X_2 + aX_1)$ est un vecteur gaussien centré de matrice $A\Gamma_{12}A^t$.
4. Comme le vecteur (X_2, X_2+aX_1) est gaussien (centré), X_2 et X_2+aX_1 sont indépendants si et seulement si $\mathrm{Cov}(X_2, X_2 + aX_1) = 0$. Or

$$\mathrm{Cov}(X_2, X_2 + aX_1) = \mathrm{Var}(X_2) + a\mathrm{Cov}(X_2, X_1) = 2 + a.$$

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 :

$$\mathbb{E}[X_1|X_2] = -\frac{1}{2}\mathbb{E}[X_2 - 2X_1|X_2] + \mathbb{E}\left[\frac{1}{2}X_2|X_2\right] = -\frac{1}{2}\mathbb{E}[X_2 - 2X_1] + \frac{1}{2}X_2 = \frac{1}{2}X_2.$$

Exercice 5. Soient X, Y, Z trois variables aléatoires indépendantes, de même loi $\mathcal{N}(0, 1)$. Montrer que la variable aléatoire $(X - Y)^2 + (X - Z)^2 + (Y - Z)^2$ est indépendante de la variable aléatoire $X + Y + Z$.

Solution. Comme X, Y, Z sont i.i.d. de loi normale standard, le vecteur (X, Y, Z) est gaussien centrée de matrice de variance l'identité (produit des densités). On a $(X, Y, Z) \sim \mathcal{N}_3(0, I_3)$. Considérons le vecteur

$$\mathbf{V} := \begin{pmatrix} X - Y \\ X - Z \\ Y - Z \\ X + Y + Z \end{pmatrix} = A \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}, \qquad \text{avec } A = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & -1 \\ 1 & 2 & 3 \end{pmatrix}.$$

Donc, $\mathbf{V}$ est vecteur gaussien en tant que transformation linéaire d'un vecteur gaussien. Sa matrice de variance-covariance est donnée par

$$\mathrm{Var}(\mathbf{U}) = \mathrm{Var}\left(A\begin{pmatrix}X\\Y\\Z\end{pmatrix}\right) = A\mathrm{Var}\left(\begin{pmatrix}X\\Y\\Z\end{pmatrix}\right)A^T = AA^T = \begin{pmatrix}2 & 1 & -1 & 0\\1 & 2 & 1 & 0\\-1 & 1 & 2 & 0\\0 & 0 & 0 & 3\end{pmatrix}.$$

Les covariances nulles implique l'indépendance du sous-vecteur $\mathbf{W} := \begin{pmatrix}X-Y\\X-Z\\Y-Z\end{pmatrix}$ de la variable $X+Y+Z$. Or, $(X-Y)^2+(X-Z)^2+(Y-Z)^2 = \|\mathbf{W}\|^2$ est fonction mesurable de $\mathbf{W}$, donc l'indépendance de $X+Y+Z$ est préservée.

Exercice 6 (Théorème de Cochran)**.** Soit Z un vecteur gaussien de $\mathbb{R}^n$ d'espérance nulle et de matrice de covariance I_n où I_n est la matrice identité de dimension n. Supposons que $\mathbb{R}^n$ s'écrit comme la somme directe de J sous-espaces vectoriels orthogonaux $V_1, \cdots, V_J$ de dimensions respectives $p_1, \cdots, p_J$. On désigne par Π_{V_j} la matrice de projection orthogonale sur V_j.

1. Montrer que $\Pi_{V_1}Z, \cdots, \Pi_{V_k}Z$ sont des vecteurs aléatoires indépendants. Déterminer leurs lois.
2. Montrer que $\|\Pi_{V_j}Z\|^2$ suit la loi $\chi^2(p_j)$ pour tout $1 \leq j \leq J$.
3. Application. Soient $X_i, i = 1, \ldots, n$ des variables aléatoires indépendantes de loi normale $\mathcal{N}(\mu, \sigma)$ avec $\mu \in \mathbb{R}$ et $\sigma > 0$. On pose $\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$ et $S_n^2 = \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X}_n)^2$. Déterminer la loi jointe du vecteur aléatoire $(\bar{X}_n, S_n^2)$.

Solution. 1. Formons d'abord une grande matrice A avec toutes les matrices de projections :

$$A = \begin{pmatrix}\Pi_{V_1}\\\Pi_{V_2}\\\vdots\\\Pi_{V_1}\end{pmatrix}.$$

Puisque Z est un vecteur gaussien, $\begin{pmatrix}\Pi_{V_1}Z\\\cdots\\\Pi_{V_J}Z\end{pmatrix} = AZ$ l'est aussi comme transformation affine d'un vecteur gaussien. Sa moyenne est $\mathbb{E}[AZ] = A\mathbb{E}[Z] = 0$ et sa matrice de variance-covariance est $\mathrm{Var}(AZ) = A\mathrm{Var}(Z)A^T = AA^T$.

Rappel sur les projecteurs orthogonaux : on a pour tout j, $\Pi_{V_j} = \Pi_{V_j}^T$ (symétrie) et $\Pi_{V_j}^2 = \Pi_{V_j}$, et par orthogonalité des sous-espaces vectoriels V_j on a $\Pi_{V_j}\Pi_{V_l} = 0$ pour tout $j \neq l$.

Donc, la matrice de variance-covariance de AZ est diagonale par block :

$$\mathrm{Var}(AZ) = \begin{pmatrix}\Pi_{V_1} & 0 & \cdots & 0\\0 & \Pi_{V_2} & \ddots & \vdots\\\vdots & & \ddots & \\0 & \cdots & \cdots & \Pi_{V_k}\end{pmatrix}.$$

La structure de la matrice $\mathrm{Var}(AZ)$ implique que les sous-vecteurs $\Pi_{V_j}Z$ sont des vecteurs gaussiens indépendants. Et $\Pi_{V_j}Z$ est de moyenne nulle et $\mathrm{Var}(\Pi_{V_j}Z) = \Pi_{V_j}$ pour tout j. Donc, $\Pi_{V_j}Z \sim \mathcal{N}_{p_j}(0, \Pi_{V_j})$.

2. Comme Π_{V_j} est symétrique, il existe une matrice Γ orthogonale telle que $\Pi_{V_j} = \Gamma\Lambda\Gamma^T$, où $\Lambda = \mathrm{Diag}(\lambda_1, \ldots, \lambda_p)$ est la matrice diagonale des valeurs propres de Π_{V_j}. Alors,

$$\|\Pi_{V_j} Z\|^2 = Z^T \Pi_{V_j}^T \Pi_{V_j} Z = Z^T \Pi_{V_j} Z = (Z^T\Gamma)\Lambda(\Gamma^T Z) = U^T \Lambda U = \sum_{i=1}^{k} \lambda_i U_i^2 \ ,$$

où $U = \Gamma^T Z = (U_1, \ldots, U_n)^T$. En utilisant l'orthogonalité de Γ, on vérifie que U est un vecteur normal de loi $\mathcal{N}_k(0, I_k)$. En effet,

$$\mathbb{E}\left[U\right] = \Gamma^T \mathbb{E}\left[Z\right] = 0 \quad \text{ et } \quad \mathrm{Var}(U) = \Gamma^T \mathrm{Var}(Z)\Gamma = \Gamma^T\Gamma = I_k \ .$$

Or, Π_{V_j} est un projecteur orthogonal, donc $\lambda_j \in \{0, 1\}$ et $\mathrm{Card}\{j : \lambda_j = 1\} = \mathrm{Rang}(\Pi_{V_j}) = p_j$. Donc,

$$\|\Pi_{V_j} Z\|^2 = \sum_{i:\lambda_i = 1} U_i^2 \sim \chi^2_{p_j}.$$

3. Le vecteur aléatoire $X = (X_1, \ldots, X_n)^T$ est de loi normale de moyenne $(\mu, \ldots, \mu)^T$ et de matrice de variance $\sigma^2 I_n$. Posons $Z = \frac{X-\mu}{\sigma}$. Alors Z est un vecteur gaussien centré de variance I_n. Notons $V_1 = \mathrm{Vect}(\mathbb{1}_n)$ le sous-espace vectoriel engendré par le vecteur $\mathbb{1}_n = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \in \mathbb{R}^n$. La projection orthogonale $\Pi_{V_1} Z$ de Z sur V_1 est donnée par

$$\langle \frac{1}{\sqrt{n}} \mathbb{1}_n, Z \rangle \frac{1}{\sqrt{n}} \mathbb{1}_n = \frac{1}{n} \mathbb{1}_n^T Z \mathbb{1}_n = \bar{Z}_n \mathbb{1}_n,$$

où $\bar{Z}_n = \bar{X}_n - \mu$. On en déduit que la projection orthogonale de Z sur $V_2 = V_1^\perp$ est donnée par $Z - \Pi_{V_1} Z = Z - \bar{Z}_n \mathbb{1}_n$. Par le théorème de Cochran, $\bar{Z}_n \mathbb{1}_n$ et $Z - \bar{Z}_n \mathbb{1}_n$ sont des vecteurs gaussiens indépendants. De plus, $\|Z - \bar{Z}_n \mathbb{1}_n\|^2 = \frac{nS_n^2}{\sigma^2}$ suit la loi de khi-deux dont le nombre de degrés de liberté est $\dim(V_2) = n - \dim(V_1) = n - 1$. On en déduit l'indépendance de $\bar{X}_n$ et S_n^2 avec $\bar{X}_n \sim \mathcal{N}(\mu, \sigma^2)$ et $\frac{nS_n^2}{\sigma^2} \sim \chi^2_{n-1}$.

Exercice 7 (Tomber dans le cercle)**.** Soit X, Y, Z trois vecteurs aléatoires indépendants à valeurs dans $\mathbb{R}^2$ de loi gaussienne standard. Montrer que la probabilité que Z tombe dans le cercle de diamètre $Y - X$ qui passe par X et Y vaut $1/4$.

Solution. L'événement considéré se produit si et seulement si

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

en notant $\|\cdot\|$ la norme euclidienne standard sur $\mathbb{R}^2$. En élevant au carré et en développant, on trouve la condition

$$\left(Z_1 - \frac{(X_1+Y_1)}{2}\right)^2 + \left(Z_2 - \frac{(X_2+Y_2)}{2}\right)^2 \leq \left(\frac{X_1 - Y_1}{2}\right)^2 + \left(\frac{X_2 - Y_2}{2}\right)^2 .$$

Posons $U_1 = Z_1 - \frac{(X_1+Y_1)}{2}$, $U_2 = Z_2 - \frac{(X_2+Y_2)}{2}$, $U_3 = \frac{X_1-Y_1}{2}$ et $U_4 = \frac{X_2-Y_2}{2}$ et déterminons la loi du vecteur $U = (U_1, U_2, U_3, U_4)$. Tout d'abord U est un vecteur Gaussien. En effet, on peut

écrire $U = AV$, où $V = (X_1, X_2, Y_1, Y_2, Z_1, Z_2)$ est un vecteur Gaussien standard sur $\mathbb{R}^6$ et A une matrice de taille 4×6 donnée par

$$A = \begin{pmatrix} -\frac{1}{2} & 0 & -\frac{1}{2} & 0 & 1 & 0 \\ 0 & -\frac{1}{2} & 0 & -\frac{1}{2} & 0 & 1 \\ \frac{1}{2} & 0 & -\frac{1}{2} & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & -\frac{1}{2} & 0 & 0 \end{pmatrix}.$$

On a $\mathbb{E}[U] = A\mathbb{E}[V] = 0$ et pour la matrice de variance

$$\mathrm{Var}(U) = A\mathrm{Var}(V)A^T = AA^T = \begin{pmatrix} \frac{3}{2} & 0 & 0 & 0 \\ 0 & \frac{3}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & \frac{1}{2} \end{pmatrix}.$$

La structure diagonale de $\mathrm{Var}(U)$ implique que les composantes du vecteur U sont indépendantes. Par ailleurs, U_1 et U_2 sont de loi normale $\mathcal{N}(0, \frac{3}{2})$ et U_3 et U_4 sont de loi normale $\mathcal{N}(0, \frac{1}{2})$.

Pour résumer, on a

$$U_1 = \sqrt{3/2}\varepsilon_1, \quad U_2 = \sqrt{3/2}\varepsilon_2, \quad U_3 = (1/\sqrt{2})\varepsilon_3, \quad U_4 = (1/\sqrt{2})\varepsilon_4,$$

avec $(\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4)$ un vecteur gaussien standard de $\mathbb{R}^4$. La probabilité p recherchée vaut donc

$$p := \mathbb{P}\left(\left\| Z - \frac{(X+Y)}{2} \right\| \leq \frac{1}{2}\|X - Y\|\right) = \mathbb{P}\left(3(\varepsilon_1^2 + \varepsilon_2^2) \leq \varepsilon_3^2 + \varepsilon_4^2\right).$$

Par la définition de la loi du khi-deux, les variables alétaoires $R_1^2 := \varepsilon_1^2 + \varepsilon_2^2$ et $R_2^2 := \varepsilon_3^2 + \varepsilon_4^2$ suivent la loi $\chi_2^2 = \Gamma(\frac{2}{2}, \frac{1}{2}) = \mathrm{Exp}(\frac{1}{2})$. Par conséquent,

$$p = \frac{1}{4}\int_0^{+\infty}\int_0^{+\infty} \mathbf{1}_{3r_1 \leq r_2} e^{-r_1/2} e^{-r_2/2}\, dr_1 dr_2 = \frac{1}{2}\int_0^{+\infty} e^{-2r_1}\, dr_1 = \frac{1}{4}.$$