Lois elliptiques

Jérôme Lapuyade-Lahorgue

Lois elliptiques

Jérôme Lapuyade-Lahorgue LITIS - Eq. Quantif - Université de Rouen

1 Hypervolume de parallélogramme

1.1 Aire parallélogramme

1.1.1 Dans $\mathbb{R}^2$

Soient $\vec{u}_1 = \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix}$ et $\vec{u}_2 = \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix}$ deux vecteurs de $\mathbb{R}^2$. Notons $M = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$. L'aire du parallélogramme formé par les deux vecteurs est donnée par $A = \|\vec{u}_1\|\|\vec{u}_2\| \sin\theta$, où θ est l'angle entre les deux vecteurs. On a alors :

$$\begin{aligned} A^2 &= \|\vec{u}_1\|^2\|\vec{u}_2\|^2 - \langle \vec{u}_1, \vec{u}_2 \rangle^2 \\ &= \left(a_{11}^2 + a_{21}^2\right) \times \left(a_{12}^2 + a_{22}^2\right) - \left(a_{11}a_{12} + a_{21}a_{22}\right)^2 \\ &= a_{11}^2 a_{22}^2 + a_{21}^2 a_{12}^2 - 2a_{11}a_{12}a_{21}a_{22} \\ &= \left(\det M\right)^2 . \end{aligned}$$

Ainsi :

$$A = |\det M| . \tag{1}$$

On remarque aussi que $A^2 = \det\left(M^T M\right)$ et :

$$M^T M = \begin{pmatrix} \|\vec{u}_1\|^2 & \langle \vec{u}_1, \vec{u}_2 \rangle \\ \langle \vec{u}_1, \vec{u}_2 \rangle & \|\vec{u}_2\|^2 \end{pmatrix} \tag{2}$$

1.1.2 Dans $\mathbb{R}^d$ avec $d \geq 2$

Soient $\vec{u}_1 = \begin{pmatrix} a_{11} \\ \vdots \\ a_{d1} \end{pmatrix}$ et $\vec{u}_2 = \begin{pmatrix} a_{12} \\ \vdots \\ a_{d2} \end{pmatrix}$ deux vecteurs de $\mathbb{R}^d$. Notons $M = \begin{pmatrix} a_{11} & a_{12} \\ \vdots & \vdots \\ a_{d1} & a_{d2} \end{pmatrix}$. On a également :

$$M^T M = \begin{pmatrix} \|\vec{u}_1\|^2 & \langle \vec{u}_1, \vec{u}_2 \rangle \\ \langle \vec{u}_1, \vec{u}_2 \rangle & \|\vec{u}_2\|^2 \end{pmatrix} \tag{3}$$

$M^T M$ est invariant par rotation, on pose donc :

$$A = \sqrt{\det\left(M^T M\right)}. \tag{4}$$

1.2 Hypervolume parallélogramme de dimension d

Hypothèse de récurrence : On suppose pour tout $k \leq d-1$, l'hypervolume de k vecteurs de $\mathbb{R}^k$ est la racine carrée du déterminant de $M^T M$, où M définie de manière similaire à précédemment.

1.2.1 Dans $\mathbb{R}^d$

Soient $\vec{u}_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{dj} \end{pmatrix}$ avec $1 \leq j \leq d$, d vecteurs de $\mathbb{R}^d$ et soit la matrice $M = (a_{i,j})_{1 \leq i,j \leq d}$. On a :

$$V(\vec{u}_1, \ldots, \vec{u}_d) = V(\vec{u}_1, \ldots, \vec{u}_{d-1}) \times \left|\langle \vec{u}_d, \vec{n} \rangle\right|,$$

où $\vec{n}$ est un vecteur tel que $\|\vec{n}\|^2 = 1$ et $\langle \vec{n}, \vec{u}_j \rangle = 0$ pour $1 \leq j \leq d-1$. Posant $\vec{n} = \begin{pmatrix} n_1 \\ \vdots \\ n_d \end{pmatrix}$, on a alors :

$$\begin{cases} a_{11}n_1 + a_{21}n_2 + \ldots + a_{d-1,1}n_{d-1} & = & -a_{d,1}n_d \\ a_{12}n_1 + a_{22}n_2 + \ldots + a_{d-1,2}n_{d-1} & = & -a_{d,2}n_d \\ \vdots & \vdots & \vdots \\ a_{1,d-1}n_1 + a_{2,d-1}n_2 + \ldots + a_{d-1,d-1}n_{d-1} & = & -a_{d,d-1}n_d \end{cases}$$

Ainsi :

$$n_1 = (-1)^{d-1} n_d \times \frac{\det \begin{pmatrix} a_{21} & a_{22} & \ldots & a_{2,d-1} \\ \vdots & \vdots & \vdots & \vdots \\ a_{d,1} & a_{d,2} & \ldots & a_{d,d-1} \end{pmatrix}}{\det \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1,d-1} \\ \vdots & \vdots & \vdots & \vdots \\ a_{d-1,1} & a_{d-1,2} & \ldots & a_{d-1,d-1} \end{pmatrix}}$$

et pour $2 \leq j \leq d-1$, on a :

$$n_j = (-1)^{d-j} n_d \times \frac{\det \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1,d-1} \\ \vdots & \vdots & \vdots & \vdots \\ a_{j-1,1} & a_{j-1,2} & \dots & a_{j-1,d-1} \\ a_{j+1,1} & a_{j+1,2} & \dots & a_{j+1,d-1} \\ \vdots & \vdots & \vdots & \vdots \\ a_{d,1} & a_{d,2} & \dots & a_{d,d-1} \end{pmatrix}}{\det \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1,d-1} \\ \vdots & \vdots & \vdots & \vdots \\ a_{d-1,1} & a_{d-1,2} & \dots & a_{d-1,d-1} \end{pmatrix}}$$

De $n_1^2 + \ldots + n_d^2 = 1$, on en déduit :

$$n_d^2 = \frac{\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1,d-1} \\ \vdots & \vdots & \vdots & \vdots \\ a_{d-1,1} & a_{d-1,2} & \dots & a_{d-1,d-1} \end{vmatrix}^2}{\sum_{j=1}^{d} \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1,d-1} \\ \vdots & \vdots & \vdots & \vdots \\ a_{j-1,1} & a_{j-1,2} & \dots & a_{j-1,d-1} \\ a_{j+1,1} & a_{j+1,2} & \dots & a_{j+1,d-1} \\ \vdots & \vdots & \vdots & \vdots \\ a_{d,1} & a_{d,2} & \dots & a_{d,d-1} \end{vmatrix}^2}$$

On pose :

$$n_d = (-1)^{d-1} \frac{\det \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1,d-1} \\ \vdots & \vdots & \vdots & \vdots \\ a_{d-1,1} & a_{d-1,2} & \dots & a_{d-1,d-1} \end{pmatrix}}{\sqrt{\sum_{j=1}^{d} \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1,d-1} \\ \vdots & \vdots & \vdots & \vdots \\ a_{j-1,1} & a_{j-1,2} & \dots & a_{j-1,d-1} \\ a_{j+1,1} & a_{j+1,2} & \dots & a_{j+1,d-1} \\ \vdots & \vdots & \vdots & \vdots \\ a_{d,1} & a_{d,2} & \dots & a_{d,d-1} \end{vmatrix}^2}}$$

On a alors pour $1 \leq j \leq d-1$:

$$n_j = (-1)^{j+1} \frac{\det \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1,d-1} \\ \vdots & \vdots & \vdots & \vdots \\ a_{j-1,1} & a_{j-1,2} & \dots & a_{j-1,d-1} \\ a_{j+1,1} & a_{j+1,2} & \dots & a_{j+1,d-1} \\ \vdots & \vdots & \vdots & \vdots \\ a_{d,1} & a_{d,2} & \dots & a_{d,d-1} \end{pmatrix}}{\sqrt{\sum_{j=1}^{d} \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1,d-1} \\ \vdots & \vdots & \vdots & \vdots \\ a_{j-1,1} & a_{j-1,2} & \dots & a_{j-1,d-1} \\ a_{j+1,1} & a_{j+1,2} & \dots & a_{j+1,d-1} \\ \vdots & \vdots & \vdots & \vdots \\ a_{d,1} & a_{d,2} & \dots & a_{d,d-1} \end{vmatrix}^2}}.$$

En développant le déterminant, on en déduit :

$$\langle \vec{n}, \vec{u}_d \rangle = \frac{\det M}{\sqrt{\sum_{j=1}^{d} \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1,d-1} \\ \vdots & \vdots & \vdots & \vdots \\ a_{j-1,1} & a_{j-1,2} & \dots & a_{j-1,d-1} \\ a_{j+1,1} & a_{j+1,2} & \dots & a_{j+1,d-1} \\ \vdots & \vdots & \vdots & \vdots \\ a_{d,1} & a_{d,2} & \dots & a_{d,d-1} \end{vmatrix}^2}},$$

ainsi :

$$V(\vec{u}_1, \dots, \vec{u}_d) = V(\vec{u}_1, \dots, \vec{u}_{d-1}) \times \frac{|\det M|}{\sqrt{\sum_{j=1}^{d} \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1,d-1} \\ \vdots & \vdots & \vdots & \vdots \\ a_{j-1,1} & a_{j-1,2} & \dots & a_{j-1,d-1} \\ a_{j+1,1} & a_{j+1,2} & \dots & a_{j+1,d-1} \\ \vdots & \vdots & \vdots & \vdots \\ a_{d,1} & a_{d,2} & \dots & a_{d,d-1} \end{vmatrix}^2}}.$$

Notons :

$$S(\vec{u}_1, \dots, \vec{u}_{d-1}) = \sqrt{\sum_{j=1}^{d} \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1,d-1} \\ \vdots & \vdots & \vdots & \vdots \\ a_{j-1,1} & a_{j-1,2} & \dots & a_{j-1,d-1} \\ a_{j+1,1} & a_{j+1,2} & \dots & a_{j+1,d-1} \\ \vdots & \vdots & \vdots & \vdots \\ a_{d,1} & a_{d,2} & \dots & a_{d,d-1} \end{vmatrix}^2}.$$

On remarque facilement que S satisfait :

1. $S(\vec{e}_1, \ldots, \vec{e}_{d-1}) = 1$.
2. $S(\vec{u}_{\sigma(1)}, \ldots, \vec{u}_{\sigma(d-1)}) = S(\vec{u}_1, \ldots, \vec{u}_{d-1})$, où σ permutation des $d-1$ vecteurs.
3. $S(\vec{u}_1, \ldots, \vec{u}_i, \ldots, \vec{u}_i, \ldots, \vec{u}_{d-1}) = 0$.

De ces propriétés, on déduit que $V(\vec{u}_1, \ldots, \vec{u}_{d-1}) = S(\vec{u}_1, \ldots, \vec{u}_{d-1})$. En effet : $V(\vec{e}_1, \ldots, \vec{e}_{d-2}, \vec{u}_{d-1})^2$ et $S(\vec{e}_1, \ldots, \vec{e}_{d-2}, \vec{u}_{d-1})^2$ sont des polynômes d'indéterminées $a_{j,d-1}$ pour $1 \leq j \leq d-1$ et coïncident pour $\vec{u}_{d-1} = \vec{e}_{d-1}$ (valeur commune 1) et pour $\vec{u}_{d-1} = \vec{e}_j$ avec $1 \leq j \leq d-2$ (valeur commune 0). Ainsi $V(\vec{e}_1, \ldots, \vec{e}_{d-2}, \vec{u}_{d-1})^2 = S(\vec{e}_1, \ldots, \vec{e}_{d-2}, \vec{u}_{d-1})^2$. Par commutativité (propriété 2.) et comme $\vec{u}_{d-1}$ est arbitraire, on montre que $V(\vec{u}_1, \ldots, \vec{e}_{d-1})^2 = S(\vec{u}_1, \ldots, \vec{e}_{d-1})^2$. De même, pour u_1 fixé, $V(\vec{u}_1, \vec{e}_2, \ldots, \vec{e}_{d-2}, \vec{u}_{d-1})^2$ et $S(\vec{u}_1, \vec{e}_2 \ldots, \vec{e}_{d-2}, \vec{u}_{d-1})^2$ sont des polynômes d'indéterminées $a_{j,d-1}$ pour $1 \leq j \leq d-1$ qui coïncident pour $\vec{u}_{d-1} = \vec{e}_{d-1}$, $\vec{u}_{d-1} = \vec{u}_1$ et $\vec{u}_{d-1} = \vec{e}_j$ pour $2 \leq j \leq d-2$. Ainsi $V(\vec{u}_1, \vec{e}_2, \ldots, \vec{e}_{d-2}, \vec{u}_{d-1})^2 = S(\vec{u}_1, \vec{e}_2 \ldots, \vec{e}_{d-2}, \vec{u}_{d-1})^2$. Utilisant de nouveau la propriété 2., on en déduit $V(\vec{u}_1, \vec{u}_2, \vec{e}_3, \ldots, \vec{e}_{d-1})^2 = S(\vec{u}_1, \vec{u}_2, \vec{e}_3 \ldots, \vec{e}_{d-1})^2$. Itérant la méthode, on en déduit le résultat.
On a montré ainsi que :

$$V(\vec{u}_1, \ldots, \vec{u}_d) = |\det M| \,. \tag{5}$$

De plus introduisant la matrice :

$$M^T M = (\langle \vec{u}_i, \vec{u}_j \rangle)_{1 \leq i,j \leq d} \,,$$

on a :

$$V(\vec{u}_1, \ldots, \vec{u}_d) = \sqrt{\det(M^T M)}. \tag{6}$$

La matrice $M^T M$ a la propriété importante d'être invariante par rotation.

1.2.2 Dans $\mathbb{R}^N$ avec $N \geq d$

Soient $\vec{u}_j = \begin{pmatrix} a_{1,j} \\ \vdots \\ a_{N,j} \end{pmatrix}$ avec $1 \leq j \leq d$. Posant $M = (a_{i,j})_{1 \leq i \leq N, 1 \leq j \leq d}$, il est clair que :

$$\begin{aligned} V(\vec{u}_1, \ldots, \vec{u}_d) &= \sqrt{\det(M^T M)} \\ &= \sqrt{\det\left(\langle \vec{u}_i, \vec{u}_j \rangle\right)}. \end{aligned}$$

1.2.3 Théorème de Pythagore pour les hypervolumes

Avec les mêmes notations que précédemment, considérons d vecteurs de $\mathbb{R}^N$ avec $N \geq d$. Par le même raisonnement que pour montrer que $V = S$, on montre

que :

$$V(\vec{u}_1,\ldots,\vec{u}_d)^2 = \sum_{1\leq i_1<i_2<\ldots<i_d\leq N} \begin{vmatrix} a_{i_1,1} & \ldots & a_{i_1,d} \\ \vdots & \vdots & \vdots \\ a_{i_d,1} & \ldots & a_{i_d,d} \end{vmatrix}^2 . \tag{7}$$

C'est le théorème de Pythagore pour les hypervolumes. Ceci nous permet d'introduire les formes volumes alternées définies par :

$$dx_{i_1} \wedge \ldots \wedge dx_{i_d}(\vec{u}_1,\ldots,\vec{u}_d) = \det \begin{pmatrix} a_{i_1,1} & \ldots & a_{i_1,d} \\ \vdots & \vdots & \vdots \\ a_{i_d,1} & \ldots & a_{i_d,d} \end{pmatrix} . \tag{8}$$

Ces formes vérifient $dx_{i_{\sigma(1)}} \wedge \ldots \wedge dx_{i_{\sigma(d)}} = \epsilon(\sigma) dx_1 \wedge \ldots \wedge dx_d$. Selon le choix de l'ordre des facteurs, on a deux résultats au signe près. Le signe détermine l'orientation de l'espace.

2 Hypervolume de variétés différentiables. Le cas de l'hypersphère

2.1 Formalisme différentiel

Considérons une variété différentiable de dimension d plongée dans $\mathbb{R}^N$ avec $N \geq d$ paramétrée par :

$$\varphi : (\theta_1,\ldots,\theta_d) \to \begin{pmatrix} x_1 = \varphi_1(\theta_1,\ldots,\theta_d) \\ \vdots \\ x_N = \varphi_N(\theta_1,\ldots,\theta_d) \end{pmatrix} . \tag{9}$$

Les vecteurs tangents sont donnés par :

$$\frac{\vec{\partial}}{\partial\theta_j} = \begin{pmatrix} \frac{\partial\varphi_1}{\partial\theta_j}(\theta_1,\ldots,\theta_d) \\ \vdots \\ \frac{\partial\varphi_N}{\partial\theta_j}(\theta_1,\ldots,\theta_d) \end{pmatrix} , \tag{10}$$

pour $1 \leq j \leq d$.
L'hypervolume entre ces vecteurs tangents est donnée par $\sqrt{\det G}$ où $G = \left(\left\langle \frac{\vec{\partial}}{\partial\theta_i}, \frac{\vec{\partial}}{\partial\theta_j} \right\rangle\right)_{1\leq i,j\leq d}$, appelé tenseur métrique. Ainsi, l'hypervolume d'une portion $\Gamma = \varphi(\Delta)$ de la variété différentiable est défini par :

$$V(\Gamma) = \int_\Delta \sqrt{\det G(\theta)} d\theta_1 \ldots d\theta_d . \tag{11}$$

La forme symétrique volume correspondant à l'image réciproque dans le système de paramétrage de la mesure de Lebesgue de la variété différentiable est par définition $d\omega = \sqrt{\det G(\theta)} d\theta_1 \ldots d\theta_d$.

2.2 Volume de l'hypersphère

2.2.1 Une relation de récurrence

On considère dans cette section l'hypersphère S^{d-1} de dimension $d-1$ plongée dans $\mathbb{R}^d$. Celle-ci est définie par l'ensemble des d-uplets vérifiant $x_1^2 + x_2^2 + \ldots + x_d^2 = 1$. On notera σ_{d-1} son hypervolume et $d\sigma_{d-1}$ sa forme volume symétrique.
Pour $d = 2$, la sphère S^2 est paramétrée par :

$$\left\{ \begin{array}{l} x_1 = \cos(\varphi) \\ x_2 = \sin(\varphi) \end{array} \right. , \tag{12}$$

où $\varphi \in [0, 2\pi]$. Par définition de π, $\sigma_1 = 2\pi$ et $d\sigma_1 = d\varphi$.
Pour $d \geq 3$: en remarquant que $x_d \in [-1, 1]$, on pose $x_d = \cos\theta$ avec $\theta \in [0, \pi]$. On a alors $x_1^2 + x_2^2 + \ldots + x_{d-1}^2 = \sin^2\theta$. Comme $\sin\theta \geq 0$, S^{d-1} peut être paramétré par :

$$\left\{ \begin{array}{l} x_1(\theta, \theta_1, \ldots, \theta_{d-2}) = \sin\theta y_1(\theta_1, \ldots, \theta_{d-2}) \\ \vdots \\ x_{d-1}(\theta, \theta_1, \ldots, \theta_{d-2}) = \sin\theta y_{d-1}(\theta_1, \ldots, \theta_{d-2}) \\ x_d(\theta, \theta_1, \ldots, \theta_{d-2}) = \cos(\theta) \end{array} \right. , \tag{13}$$

où $(y_1, \ldots, y_{d-1}) \in S^{d-2}$, $\theta_1 = \varphi$ et les autres $\theta_j \in [0, \pi]$ construits récursivement de la manière précédente.
Les vecteurs tangents sont donnés par :

$$\frac{\vec{\partial}}{\partial\theta} = \left(\begin{array}{c} \cos\theta y_1 \\ \vdots \\ \cos\theta y_{d-1} \\ -\sin\theta \end{array} \right) \tag{14}$$

et :

$$\frac{\vec{\partial}}{\partial\theta_j} = \left(\begin{array}{c} \sin\theta \frac{\partial y_1}{\partial\theta_j} \\ \vdots \\ \sin\theta \frac{\partial y_{d-1}}{\partial\theta_j} \\ 0 \end{array} \right). \tag{15}$$

Comme $(y_1, \ldots, y_{d-1}) \in S^{d-2}$, alors $\|\frac{\vec{\partial}}{\partial\theta}\|^2 = 1$ et $\left\langle \frac{\vec{\partial}}{\partial\theta}, \frac{\vec{\partial}}{\partial\theta_j} \right\rangle = 0$. On a de plus :

$$\begin{aligned} \left\| \frac{\vec{\partial}}{\partial\theta_j} \right\|^2 &= \sin^2\theta \times \left(\left(\frac{\partial y_1}{\partial\theta_j} \right) + \ldots + \left(\frac{\partial y_{d-1}}{\partial\theta_j} \right) \right) \\ &= \sin^2\theta \left\| \frac{\partial^{(\vec{d-2})}}{\partial\theta_j} \right\|^2 , \end{aligned}$$

où $\frac{\partial^{(\vec{d-2})}}{\partial\theta_j}$ est un vecteur tangent à S^{d-2}. Il s'ensuit que $\det G = \sin^{2(d-2)}\theta \times \det G_{d-2}$, où G_{d-2} est le tenseur métrique de S^{d-2}. On obtient alors les relations de récurrence :

$$d\sigma_{d-1} = \sin^{d-2}\theta d\theta d\sigma_{d-2}, \tag{16}$$

et :

$$\sigma_{d-1} = \sigma_{d-2} \times \int_0^{\pi} \sin^{d-2}\theta d\theta. \tag{17}$$

2.3 La fonction Γ

Pour α réel strictement positif, on définit :

$$\Gamma(\alpha) = \int_0^{+\infty} t^{\alpha-1}\exp(-t)dt. \tag{18}$$

On montre que :

1. $\Gamma(1) = 1$.
2. $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$.
3. $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$.
4. Par Fubini et changement de variable :

$$\Gamma(\alpha)\times\Gamma(\beta) = \Gamma(\alpha+\beta)\times\int_0^1 t^{\alpha-1}(1-t)^{\beta-1}dt.$$

2.4 Calcul volume de l'hypersphère

En effectuant le changement de variable $t = \sin^2\theta$ dans l'intégrale $\int_0^{\pi}\sin^{d-2}\theta d\theta$, on montre que :

$$\begin{aligned}\int_0^{\pi}\sin^{d-2}\theta d\theta &= \int_0^1 \frac{t^{\frac{d-1}{2}-1}}{\sqrt{1-t}}dt \\ &= \sqrt{\pi}\frac{\Gamma\left(\frac{d-1}{2}\right)}{\Gamma\left(\frac{d}{2}\right)}.\end{aligned}$$

En utilisant la formule de récurrence, on montre que :

$$\sigma_{d-1} = 2\times\frac{\pi^{\frac{d}{2}}}{\Gamma\left(\frac{d}{2}\right)}. \tag{19}$$

3 Lois elliptiques et lois normales

Définition 1 (Loi elliptique)**.** *Une variable aléatoire suit une loi elliptique si ses isodensités sont des ellipsoïdes. Elle suit une loi elliptique standard si ses isodensités sont des hypersphères centrées à l'origine.*
Ainsi la densité d'une loi elliptique standard est de la forme :

$$p(x) = h(\|x\|^2). \tag{20}$$

A partir d'une loi elliptique standard, on obtient la densité de la loi elliptique de centre m et de matrice de dispersion Σ (c'est à dire loi de $\Sigma^{\frac{1}{2}}X + m$, où X est standard) par :

$$p(y|m, \Sigma) = \frac{1}{\sqrt{\det \Sigma}} h(\|\Sigma^{-\frac{1}{2}}(y - m)\|^2). \tag{21}$$

Le cas le plus naturel de lois elliptiques est la loi uniforme sur l'hypersphère.

3.1 Loi uniforme sur l'hypersphère

D'après la forme volume de l'hypersphère, le vecteur aléatoire $(U_1, \ldots, U_d)$ suit une loi uniforme sur la sphère S^{d-1} si et seulement si :

$$\begin{cases} U_1 = \sin\Theta V_1 \\ \vdots \\ U_{d-1} = \sin\Theta V_{d-1} \\ U_d = \cos\Theta \end{cases}, \tag{22}$$

où Θ et $(V_1, \ldots, V_{d-1})$ sont des variables aléatoires indépendantes, $(V_1, \ldots, V_{d-1})$ suit une loi uniforme sur la sphère S^{d-2} et Θ suit une loi de densité de support $[0, \pi]$ définie par :

$$\theta \to \frac{\Gamma\left(\frac{d}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{d-1}{2}\right)} \sin^{d-2}\theta. \tag{23}$$

Lois marginales :
Par symétrie sphérique, toutes les marginales U_j suivent la même loi. Par changement de variable, la loi de $U_d^2 = \cos^2\theta$ est de densité :

$$t \in [0, 1] \to \frac{\Gamma\left(\frac{d}{2}\right)}{\sqrt{\pi}\Gamma\left(\frac{d-1}{2}\right)} \frac{(1-t)^{\frac{d-1}{2}-1}}{\sqrt{t}}. \tag{24}$$

Toujours par symétrie sphérique, U_d prend autant de valeurs positives que négatives, ainsi $U_d = S \times \sqrt{T}$, où S et T sont deux variables indépendantes de lois respectives : uniforme sur $\{-1, 1\}$ et T de densité donnée par la formule précédente.
Moments :
On montre facilement que :

1. $\mathbb{E}(U_j) = 0$.
2. $\sigma_{i,j} = 0$ si $i \neq j$.
3. $\sigma_i^2 = \frac{1}{d}$,

où σ_i^2 désigne la variance de U_i et $\sigma_{i,j}$ est la covariance entre U_i et U_j.
On remarque notamment que la corrélation entre les marginales est nulle sans que celles-ci soient indépendantes.

3.2 Loi elliptique et loi uniforme sur la sphère

On montre facilement la proposition suivante :

Proposition 1. *Un vecteur aléatoire X suit une loi elliptique standard dans $\mathbb{R}^d$ de densité donnée par $p(x) = h\left(\|x\|^2\right)$ si et seulement si les variables aléatoires $R = \|X\|$ et $U = \dfrac{X}{\|X\|}$ sont indépendantes de lois respectives : R de densité :*

$$r \in [0, +\infty[\to \frac{2\pi^{\frac{d}{2}}}{\Gamma\left(\frac{d}{2}\right)} r^{d-1} h(r^2), \tag{25}$$

et U uniforme sur la sphère S^{d-1}.

On montre facilement qu'un vecteur aléatoire X suivant une loi elliptique standard vérifie :

$$\mathbb{E}(X) = 0 \tag{26}$$

$$\mathbb{E}(XX^T) = \frac{\mathbb{E}(R^2)}{d} I_d, \tag{27}$$

où I_d est la matrice identité.

3.3 Loi normale dans $\mathbb{R}^d$

Définition 2 (Loi normale standard)**.** *Un vecteur aléatoire X de $\mathbb{R}^d$ suit une loi normale standard si c'est un vecteur aléatoire elliptique standard satisfaisant les conditions :*

1. *$\mathbb{E}(XX^T) = I_d$.*
2. *Les marginales de X sont indépendantes.*

La proposition suivante montre l'unicité de la densité de la loi normale standard.

Proposition 2. *Un vecteur aléatoire X de $\mathbb{R}^d$ suit une loi normale standard si et seulement si sa densité est :*

$$x \in \mathbb{R}^d \to \frac{1}{(2\pi)^{\frac{d}{2}}} \exp\left(-\frac{\|x\|^2}{2}\right) \tag{28}$$

Démonstration. Comme toute loi elliptique standard, la densité de X est de la forme $p(x) = h\left(\|x\|^2\right)$. La condition 2. d'indépendance des marginales est satisfaite si et seulement si $h\left(\|x\|^2\right) = C\exp\left(-\alpha\|x\|^2\right)$, où C et α sont des constantes positives réelles. Comme p est une densité de probabilité, on montre que $C = \left(\frac{\alpha}{\pi}\right)^{\frac{d}{2}}$. Ainsi $h\left(\|x\|^2\right) = \left(\frac{\alpha}{\pi}\right)^{\frac{d}{2}}\exp\left(-\alpha\|x\|^2\right)$ et R a pour densité :

$$r \in [0, +\infty[\to \frac{2\alpha^{\frac{d}{2}}}{\Gamma\left(\frac{d}{2}\right)} r^{d-1}\exp\left(-\alpha r^2\right).$$

Calculant $\mathbb{E}(R^2)$ en fonction de α, on montre que la condition 1. est satisfaite si et seulement si $\alpha = \frac{1}{2}$. □

Définition 3 (Loi normale)**.** *Un vecteur aléatoire Y suit une loi normale de moyenne $m \in \mathbb{R}^d$ et de matrice de dispersion (covariance) $\Sigma \in \mathcal{M}_{d,d}$ si et seulement si $Y = \Sigma^{\frac{1}{2}}X + m$, où X suit une loi normale standard.*
La densité de Y est donnée par :

$$y \in \mathbb{R}^d \to \frac{1}{(2\pi)^{\frac{d}{2}}\sqrt{\det\Sigma}}\exp\left(-\frac{y^T\Sigma^{-1}y}{2}\right). \tag{29}$$

4 Inférence de la loi normale sur $\mathbb{R}$

On rappelle la définition suivante :

Définition 4 (Fonction caractéristique)**.** *La fonction caractéristique d'un vecteur aléatoire Y de dimension d de densité p est donnée par :*

$$\mathbb{E}\left[\exp\left(i\left\langle t, Y\right\rangle\right)\right] = \int_{\mathbb{R}^d} p(y)\exp\left(i\left\langle t, y\right\rangle\right)dy, \tag{30}$$

où $\langle .,. \rangle$ est le produit scalaire de $\mathbb{R}^d$.

4.1 Lois des moments empiriques

Définition 5 (Loi du Chi-deux)**.** *Une variable S suit une loi du Chi-deux à N degrés de liberté si et seulement si $S = X_1^2 + \ldots + X_N^2$, où $X_1, \ldots, X_N$ est un échantillon iid suivant la loi normale $\mathcal{N}_{\mathbb{R}}(0,1)$.*
La densité de S est donnée par :

$$s \in [0, +\infty[\to \frac{1}{\Gamma\left(\frac{N}{2}\right)2^{\frac{N}{2}}} s^{\frac{N}{2}-1}\exp\left(-\frac{s}{2}\right), \tag{31}$$

et sa fonction caractéristique est donnée par :

$$t \in \mathbb{R} \to \frac{1}{(1-2it)^{\frac{N}{2}}}. \tag{32}$$

Proposition 3 (Moments empiriques d'une loi normale centrée-réduite)**.** *Soit* $(X_1, \ldots, X_N)$ *un échantillon iid d'une loi normale* $\mathcal{N}_{\mathbb{R}}(0,1)$*, alors les variables aléatoires :*

$$\begin{aligned} \overline{X} &= \frac{1}{N}\sum_{n=1}^{N} X_n \\ S &= \sum_{n=1}^{N}(X_n - \overline{X})^2 \end{aligned}$$

sont indépendantes et suivent respectivement les lois $\mathcal{N}_{\mathbb{R}}\left(0, \frac{1}{N}\right)$ *et* $\mathcal{X}^2_{N-1}$ *(Chi-deux à* $N-1$ *degrés de liberté).*
L'estimateur sans biais de la variance est alors donné par :

$$\hat{\sigma}^2 = \frac{1}{N-1}\sum_{n=1}^{N}(X_n - \overline{X})^2. \tag{33}$$

Démonstration. Soit le vecteur aléatoire $(Z_1; \ldots; Z_N)$ construit par :

$$\begin{pmatrix} Z_1 \\ \vdots \\ Z_N \end{pmatrix} = \underbrace{\begin{pmatrix} \cdots & \cdots & \cdots \\ \vdots & \vdots & \vdots \\ \frac{1}{\sqrt{N}} & \cdots & \frac{1}{\sqrt{N}} \end{pmatrix}}_{P} \times \begin{pmatrix} X_1 \\ \vdots \\ X_N \end{pmatrix},$$

où P est une matrice orthogonale. On en déduit que $(Z_1, \ldots, Z_N)$ est également un échantillon iid de $\mathcal{N}_{\mathbb{R}}(0,1)$ et que $Z_1^2 + \ldots + Z_N^2 = X_1^2 + \ldots + X_N^2$. De plus, on a $Z_N = \sqrt{N}\overline{X}$, ainsi :

$$\begin{aligned} \sum_{n=1}^{N}(X_n - \overline{X})^2 &= \sum_{n=1}^{N} X_n^2 - N\overline{X}^2 \\ &= \sum_{n=1}^{N-1} Z_n^2, \end{aligned} \tag{34}$$

d'où le résultat. □

Corollaire 1. *Soit* $(Y_1, \ldots, Y_N)$ *un échantillon iid d'une loi normale* $\mathcal{N}_{\mathbb{R}}(m, \sigma^2)$*, alors les variables aléatoires :*

$$\begin{aligned} \overline{Y} &= \frac{1}{N}\sum_{n=1}^{N} Y_n \\ S &= \sum_{n=1}^{N}(Y_n - \overline{Y})^2 \end{aligned}$$

sont indépendantes et suivent respectivement les lois $\mathcal{N}_{\mathbb{R}}\left(m, \frac{\sigma^2}{N}\right)$ *et* $\sigma^2 \times \mathcal{X}^2_{N-1}$.

4.2 Intervalles de confiance : point de vue fréquentiel

Rappelons les définitions :

Définition 6 (Loi de Student signée à N degrés de liberté)**.** *Une variable aléatoire réelle T suit une loi de Student (signée) à N degrés de liberté si elle s'écrit :*

$$T = \sqrt{N} \times \frac{X}{\sqrt{S}},$$

où X et S sont indépendantes de lois respectives $\mathcal{N}_{\mathbb{R}}(0,1)$ et $\mathcal{X}_N^2$.
La densité de T est :

$$t \in \mathbb{R} \to \frac{N^{\frac{N}{2}} \Gamma\left(\frac{N+1}{2}\right)}{\Gamma\left(\frac{N}{2}\right)\sqrt{\pi}} \times \frac{1}{(N+t^2)^{\frac{N+1}{2}}}. \tag{35}$$

Définition 7 (Loi de Student non signée à N degrés de liberté)**.** *Une variable aléatoire réelle T suit une loi de Student (non signée) à N degrés de liberté si elle est la valeur absolue d'une variable aléatoire de Student signée.*
La densité d'une loi aléatoire de Student non signée est donnée par :

$$t \in \mathbb{R}^+ \to \frac{2N^{\frac{N}{2}} \Gamma\left(\frac{N+1}{2}\right)}{\Gamma\left(\frac{N}{2}\right)\sqrt{\pi}} \times \frac{1}{(N+t^2)^{\frac{N+1}{2}}}. \tag{36}$$

D'après le paragraphe précédent sur les lois des moments empiriques, on a :

Proposition 4. *Soit $(Y_1, \ldots, Y_N)$ un échantillon iid de $\mathcal{N}_{\mathbb{R}}(m, \sigma^2)$, alors :*

$$T = \sqrt{N} \frac{\left|\overline{Y} - m\right|}{\hat{\sigma}}, \tag{37}$$

suit une loi de Student non signée à $N-1$ degrés de liberté.

Intervalle de confiance de la moyenne au risque p : On a :

$$\mathbb{P}\left[\sqrt{N} \frac{\left|\overline{Y} - m\right|}{\hat{\sigma}} < \alpha\right] = 1 - p,$$

où α est le $(1-p)$-quantile de la loi de Student T_{N-1}. Soit $(\overline{y}, \hat{s})$ réalisation de $(\overline{Y}, \hat{\sigma})$. Il y a une probabilité de $1-p$ que cette réalisation satisfasse :

$$\sqrt{N} \frac{\left|\overline{y} - m\right|}{\hat{s}} < \alpha.$$

On considère alors qu'il y a une probabilité de $1-p$ (et donc un risque égal à p) que la moyenne m appartienne à l'intervalle :

$$I_p(m) = \left[\overline{y} - \frac{\alpha \hat{s}}{\sqrt{N}}, \overline{y} + \frac{\alpha \hat{s}}{\sqrt{N}}\right], \tag{38}$$

où α est le $(1-p)$-quantile de la loi T_{N-1}. Cet interval s'appelle **l'intervalle de confiance de la moyenne au risque** p.
Intervalle de confiance de la variance au risque p **:** De même, on a :

$$\mathbb{P}\left[\beta < \frac{(N-1)\hat{\sigma}^2}{\sigma^2} < \alpha\right] = 1-p,$$

où α est le $(1-\frac{p}{2})$-quantile de la loi $\mathcal{X}^2_{N-1}$ et β le $\frac{p}{2}$-quantile de la loi $\mathcal{X}^2_{N-1}$. Si $\hat{s}$ est une réalisation de $\hat{\sigma}$, on a une probabilité de $1-p$ que cette réalisation satisfasse :

$$\beta < \frac{(N-1)\hat{s}^2}{\sigma^2} < \alpha.$$

Ainsi, on considère qu'il y a une probabilité de $1-p$ (et donc un risque égal à p) que σ^2 appartienne à l'intervalle :

$$I_p(\sigma^2) = \left[\frac{(N-1)\hat{s}^2}{\alpha}, \frac{(N-1)\hat{s}^2}{\beta}\right]. \tag{39}$$

C'est **l'intervalle de confiance de la variance au risque p**.

4.3 Intervalles de confiance : point de vue bayésien

Le point de vue fréquentiel des intervalles de confiance, outre un manque de rigueur souffre des inconvénients suivants :

1. Les paramètres à estimer sont implicitement considérés comme réalisation de variables aléatoires sans en donner explicitement le formalisme.
2. Pour l'intervalle de confiance de la variance, on considère un risque égal d'avoir des valeurs plus petites que la borne inférieure que d'avoir des valeurs plus grandes que la borne supérieure. Or, par l'intuition, on peut se convaincre qu'il est plus probable d'avoir une petite variance que d'en avoir une grande.
3. L'intervalle de confiance de la moyenne ne prend pas en compte les variabilités vis-à-vis de la valeur de la variance. Plus la variance est grande et moins la différence entre deux valeurs de la moyenne est discriminante.
4. L'inconvénient le plus majeur réside dans le changement de paramétrage. Par la technique fréquentielle, l'intervalle de confiance de l'écart-type s'obtient en prenant la racine carrée des bornes de l'intervalle de confiance. Or prendre une fonction sur une variable aléatoire en change sa densité ; ceci se traduit par un changement d'incertitude sur le paramètre.

Dans le cadre bayésien, les paramètres sont considérés comme des réalisations de variable aléatoire. Appelons désormais $v = \sigma^2$ la variance de Y et considérons $(y_1, \ldots, y_N)$ la réalisation d'un échantillon iid de Y. En l'absence totale d'information sur le couple (m, v) (cad. avant l'observation de l'échantillon), on considère que le couple (m, v) suit la loi impropre de Jeffreys. En effet, si on considère (m, v) comme l'entrée d'un canal bruité dont la sortie est $(y_1, \ldots, y_N)$,

une telle loi atteint la capacité du canal. D'un point de vue statistique, choisir (m, v) comme suivant la loi de Jeffreys permet de garantir le maximum de dépendance entre $(y_1, \ldots, y_N)$ et les paramètres et donc l'estimation la plus optimale. La loi de Jeffreys est obtenue en prenant la racine carrée du déterminant de la matrice d'information de Fisher. Dans le cas des lois normales, celle-ci vaut :

$$p(m, v) \propto \frac{1}{v\sqrt{v}}.$$

La loi a posteriori conditionnellement à l'échantillon $(y_1, \ldots, y_N)$ est alors donnée par :

$$p(m, v|y_1, \ldots, y_N) \propto \frac{1}{v^{\frac{N+1}{2}+1}} \exp\left(-\frac{\sum_{n=1}^{N}(y_n - m)^2}{2v} \right). \tag{40}$$

Remarque : De la formule précédente, on peut obtenir une zone de confiance conjointe du couple (m, v). On remarque de plus, que contrairement aux estimateurs ponctuels qui sont des variables aléatoires indépendantes, les estimateurs bayésiens et donc les intervalles de confiance respectifs ne sont pas indépendants. En marginalisant dans la formule (40), on obtient :

$$\begin{aligned} p(m|y_1, \ldots, y_N) &\propto \int_0^{+\infty} \frac{1}{v^{\frac{N+1}{2}+1}} \exp\left(-\frac{\sum_{n=1}^{N}(y_n - m)^2}{2v} \right) dv \\ &\propto \frac{1}{\left[\sum_{n=1}^{N}(y_n - m)^2\right]^{\frac{N-1}{2}}}. \end{aligned} \tag{41}$$

Remarque : Contrairement au résultat obtenu avec le point de vue fréquentiel, cette fois-ci, c'est :

$$\sqrt{N(N-2)}\frac{|m - \overline{y}|}{\hat{s}},$$

qui suit une loi de Student à $N - 2$ **degrés de liberté**. La perte d'un degré de liberté étant due à l'ajout d'information sur la valeur de la variance (pénaliser les variances trop grande).
De même, posant :

$$t = \frac{\sum_{n=1}^{N}(y_n - m)^2}{2v}, \tag{42}$$

la loi de t sachant $(m, y_1, \ldots, y_N)$ est une loi $\Gamma\left(\frac{N+1}{2}\right)$.
On a également :

$$\begin{aligned} p(v|y_1, \ldots, y_N) &\propto \int_0^{+\infty} \frac{1}{v^{\frac{N+1}{2}+1}} \exp\left(-\frac{\sum_{n=1}^{N}(y_n - m)^2}{2v} \right) dm \\ &\propto \frac{1}{v^{\frac{N}{2}+1}} \exp\left(-\frac{(N-1)\hat{s}^2}{2v} \right). \end{aligned} \tag{43}$$

On en déduit que $\dfrac{(N-1)\hat{s}^2}{v}$ suit une loi du Chi-deux à N **degré de liberté**.

5 Inférence de la loi normale sur $\mathbb{R}^d$

On va généraliser ce que l'on a vu précédemment sur un échantillon iid $(Y_1, \ldots, Y_N)$ de $\mathcal{N}_{\mathbb{R}^d}(m, \Sigma)$, où Σ est la matrice de covariance.

5.1 Loi Γ multivariée

Notons E l'ensemble des matrices symétriques réelles définies positives de dimension $d \times d$.

Proposition 5. *Soit $B \in E$ et $\alpha > 0$, considérons l'intégrale :*

$$I = \int_E (\det M)^{\alpha-1} \exp\left(-Tr(BM)\right) dM,$$

alors on a :

$$I = \frac{\pi^{\frac{d(d-1)}{4}} \prod_{i=1}^{d} \Gamma\left(\alpha + \frac{d-i}{2}\right)}{(\det B)^{\alpha + \frac{d-1}{2}}}.$$

Démonstration. B est symétrique réelle et définie positive donc diagonalisable. Soit $B = P\Delta P^T$ sa diagonalisation. On notera δ_i les coefficients diagonaux de Δ. On effectue le changement de variable $M = PNP^T$. On montre facilement que :

$$I = \int_E (\det N)^{\alpha-1} \exp\left(-\mathrm{Tr}(\Delta N)\right) dN.$$

Comme les matrices sont symétriques, on fera attention que $dN = \prod_{1 \le i \le j \le d} dn_{i,j}$, où $n_{i,j}$ sont les coefficients de N.
On effectue ensuite le changement de variable $N = TT^T$ par la transformée de Cholesky de N, où T est triangulaire inférieure (cad $t_{i,j} = 0$ si $i < j$). Par des

calculs simples mais fastidieux, on montre que $dN = 2^d t_{11}^d t_{22}^{d-1} \dots t_{dd} dT$. On a également $\det N = t_{11}^2 t_{22}^2 \dots t_{dd}^2$ et

$$\mathrm{Tr}(\Delta N) = \sum_{i=1}^{d} \delta_i t_{ii}^2 + \sum_{i=2}^{d} \sum_{j=1}^{i-1} \delta_i t_{ij}^2.$$

On en déduit que :

$$\begin{aligned} I &= 2^d \prod_{i=1}^{d} \int_0^{+\infty} t^{2\alpha+d-i-1} \exp(-\delta_i t^2) dt \\ &\times \prod_{i=2}^{d} \prod_{j=1}^{i-1} \int_{\mathbb{R}} \exp(-\delta_i t^2) dt. \end{aligned}$$

Comme :

$$\begin{aligned} \int_{\mathbb{R}} \exp(-\delta_i t^2) dt &= \frac{\sqrt{\pi}}{\sqrt{\delta_i}} \\ \int_0^{+\infty} t^{2\alpha+d-i-1} \exp(-\delta_i t^2) dt &= \frac{\Gamma\left(\alpha + \frac{d-i}{2}\right)}{2\delta_i^{\alpha+\frac{d-i}{2}}}, \end{aligned}$$

on en déduit le résultat. □

Définition 8 (Loi Γ multivariée)**.** *La loi Γ en dimension d de paramètre de forme $\alpha > 0$ et de paramètre d'échelle $B \in E$ est la loi de densité :*

$$M \in E \to \frac{(\det B)^{\alpha+\frac{d-1}{2}}}{\pi^{\frac{d(d-1)}{4}} \prod_{i=1}^{d} \Gamma\left(\alpha + \frac{d-i}{2}\right)} \times (\det M)^{\alpha-1} \exp\left(-Tr(BM)\right).$$

Proposition 6 (Fonction caractéristique loi Γ)**.** *La fonction caractéristique de la loi Γ est donnée par :*

$$\varphi_M : t \in E \to \frac{1}{\left[\det\left(I_d - iB^{-1}t^*\right)\right]^{\alpha+\frac{d-1}{2}}},$$

où $t_{ij}^* = \begin{cases} t_{ii} & si\ i = j \\ \frac{t_{ij}}{2} & sinon. \end{cases}$.

Démonstration. Il suffit de reconnaître que :

$$\begin{aligned} \langle t, M \rangle &= \sum_{1 \le i \le j \le d} t_{ij} M_{ij} \\ &= \mathrm{Tr}(t^* M). \end{aligned}$$

□

5.2 Lois des moments d'une normale multivariée

Définition 9 (Loi de Wishart à N degrés de liberté en dimension d)**.** *On dit qu'une matrice aléatoire W suit une loi de Wishart à N degrés de liberté en dimension d (notée $\mathcal{W}(N,d)$) si*

$$W = \sum_{n=1}^{N} X_n X_n^T,$$

où $(X_1, \ldots, X_N)$ est un échantillon iid de $\mathcal{N}_{\mathbb{R}^d}(0, I_d)$.

Proposition 7 (Fonction caractéristique loi de Wishart)**.** *La fonction caractéristique de la loi de Wishart est donnée par :*

$$\varphi_W : t \in E \to \frac{1}{\left[\det\left(I_d - 2it^*\right)\right]^{\frac{N}{2}}}.$$

Démonstration. Les X_n étant indépendants on a :

$$\begin{aligned}
\varphi_W(t) &= \prod_{n=1}^{N} \varphi_{X_n X_n^T}(t) \\
&= \prod_{n=1}^{N} \mathbb{E}\left[\exp\left(i\mathrm{Tr}(t^* X_n X_n^T)\right)\right] \\
&= \prod_{n=1}^{N} \mathbb{E}\left[\exp(i X_n^T t^* X_n)\right].
\end{aligned}$$

La matrice t^* est également symétrique réelle et définie positive donc diagonalisable avec une base de vecteurs propres orthogonaux. Soit $t^* = P\Delta P^T$ sa diagonalisation, où les δ_j sont les coefficients diagonaux de Δ. On pose $Z_n = P^T X_n$. On montre facilement que $(Z_1, \ldots, Z_N)$ est un échantillon iid de $\mathcal{N}_{\mathbb{R}^d}(0, Id)$ et que :

$$\begin{aligned}
\mathbb{E}\left[\exp(i X_n^T t^* X_n)\right] &= \mathbb{E}\left[\exp(i Z_n^T \Delta Z_n)\right] \\
&= \prod_{j=1}^{d} \varphi_{\chi_1^2}(\delta_j) \\
&= \prod_{j=1}^{d} \frac{1}{\sqrt{1 - 2i\delta_j}} \\
&= \frac{1}{\sqrt{\det\left(Id - 2it^*\right)}}.
\end{aligned}$$

On en déduit le résultat. □

Corollaire 2 (Densité loi de Wishart)**.** *La loi de Wishart admet une densité pour $N \geq d$ donnée par :*

$$W \in E \to \frac{1}{2^{\frac{Nd}{2}} \pi^{\frac{d(d-1)}{4}} \prod_{i=1}^{d} \Gamma\left(\frac{N-i+1}{2}\right)} \times (\det W)^{\frac{N-d-1}{2}} \exp\left(-\frac{1}{2} Tr(W)\right).$$

Proposition 8. *Soit $(X_1, \ldots, X_N)$ un échantillon iid de $\mathcal{N}_{\mathbb{R}^d}(0, Id)$. Les variables aléatoires :*

$$\begin{aligned} \overline{X} &= \frac{1}{N} \sum_{n=1}^{N} X_n \\ S &= \sum_{n=1}^{N} (X_n - \overline{X})(X_n - \overline{X})^T \end{aligned}$$

sont indépendantes et suivent respectivement une loi $\mathcal{N}_{\mathbb{R}^d}\left(0, \frac{1}{N} Id\right)$ et une loi de Wishart $\mathcal{W}(N-1, d)$.

Démonstration. Même raisonnement que pour le cas monovarié. □

Définition 10 (Estimateurs sans biais)**.** *Soit $(Y_1, \ldots, Y_N)$ un échantillon iid de $\mathcal{N}_{\mathbb{R}^d}(m, \Sigma)$, les estimateurs sans biais de m et de Σ sont donnés par :*

$$\begin{aligned} \overline{Y} &= \frac{1}{N} \sum_{n=1}^{N} Y_n \\ \hat{\Sigma} &= \frac{1}{N-1} \sum_{n=1}^{N} (Y_n - \overline{Y})(Y_n - \overline{Y})^T. \end{aligned}$$

Ils sont indépendants et suivent respectivement les lois $\mathcal{N}_{\mathbb{R}^d}\left(m, \frac{1}{N}\Sigma\right)$ et la même loi que $\frac{1}{N-1}\Sigma^{\frac{1}{2}} W \Sigma^{\frac{1}{2}}$, où W suit une Wishart $\mathcal{W}(N-1, d)$.

5.3 Zones de confiance : point de vue fréquentiel

5.3.1 Pré-requis

Définition 11 (Loi de Student-Hotelling vectorielle)**.** *Un vecteur aléatoire V suit une loi de Student-Hotelling vectorielle en dimension d et à N degrés de liberté si il s'écrit $V = \sqrt{N} W^{-\frac{1}{2}} Z$, où W et Z sont des variables aléatoires indépendantes telle que W suit une loi $\mathcal{W}(N, d)$ et Z suit une loi $\mathcal{N}_{\mathbb{R}^d}(0, Id)$.*

Proposition 9 (Densité de la loi de Student-Hotelling vectorielle)**.** *La densité de la loi de Student-Hotelling vectorielle est donnée par :*

$$v \in \mathbb{R}^d \to \frac{\Gamma\left(\frac{N+1}{2}\right)}{(N\pi)^{\frac{d}{2}} \Gamma\left(\frac{N-d+1}{2}\right)} \times \frac{1}{\left(1 + \frac{v^T v}{N}\right)^{\frac{N+1}{2}}} \tag{44}$$

Démonstration. Soit $V = \sqrt{N} W^{-\frac{1}{2}} Z$ et ϕ une fonction intégrable par rapport à la densité de V, on a :

$$\begin{aligned}
\mathbb{E}\left[\phi(V)\right] &= \frac{1}{(2\pi)^{\frac{d}{2}} 2^{\frac{Nd}{2}} \pi^{\frac{d(d-1)}{4}} \prod_{j=1}^{d} \Gamma\left(\frac{N-j+1}{2}\right)} \\
&\times \int_{E\times\mathbb{R}^d} \phi\left(\sqrt{N} W^{-\frac{1}{2}} z\right) (\det W)^{\frac{N-d-1}{2}} \exp\left(-\frac{1}{2}\mathrm{Tr}(W)\right) \exp\left(-\frac{z^T z}{2}\right) dW dz.
\end{aligned}$$

Dans l'intégrale :

$$\int_{\mathbb{R}^d} \phi\left(\sqrt{N} W^{-\frac{1}{2}} z\right) \exp\left(-\frac{z^T z}{2}\right) dz,$$

on effectue le changement de variables $v = \sqrt{N} W^{-\frac{1}{2}} z$, on obtient alors :

$$\begin{aligned}
\mathbb{E}\left[\phi(V)\right] &= \frac{1}{(2\pi)^{\frac{d}{2}} N^{\frac{d}{2}} 2^{\frac{Nd}{2}} \pi^{\frac{d(d-1)}{4}} \prod_{j=1}^{d} \Gamma\left(\frac{N-j+1}{2}\right)} \\
&\times \int_{\mathbb{R}^d} \phi(v) \times \left(\int_E (\det W)^{\frac{N-d}{2}} \exp\left(-\frac{1}{2N}\mathrm{Tr}((NId + vv^T)W)\right) dW\right) dv.
\end{aligned}$$

On a :

$$\int_E (\det W)^{\frac{N-d}{2}} \exp\left(-\frac{1}{2N}\mathrm{Tr}((NId + vv^T)W)\right) dW = \frac{\pi^{\frac{d(d-1)}{4}} \prod_{j=1}^{d} \Gamma\left(\frac{N-j+2}{2}\right) 2^{\frac{(N+1)d}{2}}}{\left[\det\left(Id + \frac{vv^T}{N}\right)\right]}.$$

En remarquant que $\det\left(Id + \frac{vv^T}{N}\right) = \left(1 + \frac{v^T v}{N}\right)$ et que :

$$\frac{\prod_{j=1}^{d} \Gamma\left(\frac{N-j+2}{2}\right)}{\prod_{j=1}^{d} \Gamma\left(\frac{N-j+1}{2}\right)} = \frac{\Gamma\left(\frac{N+1}{2}\right)}{\Gamma\left(\frac{N-d+1}{2}\right)},$$

on en déduit le résultat. □

Définition 12 (Loi β)**.** *Une variable aléatoire B à valeurs dans $[0,1]$ suit une loi β de paramètres de forme $a_1 > 0$ et $a_2 > 0$, notée $\beta(a_1, a_2)$ si sa densité est :*

$$b \in [0,1] \to \frac{\Gamma(a_1 + a_2)}{\Gamma(a_1)\Gamma(a_2)} b^{a_1 - 1} (1-b)^{a_2 - 1}.$$

Proposition 10 (Moments de la loi β)**.** *Si B suit une loi $\beta(a_1, a_2)$ alors pour tout entier $k \geq 0$, on a :*

$$\mathbb{E}\left[B^k\right] = \frac{\Gamma(a_1 + a_2)\Gamma(a_1 + k)}{\Gamma(a_1)\Gamma(a_1 + a_2 + k)}.$$

Définition 13 (Loi de Student-Hotelling positive)**.** *Une variable aléatoire réelle positive T suit une loi de Student-Hotelling positive à (N, d)-degrés de liberté, notée $T(N, d)$, si elle s'écrit $T = \sqrt{NZ^T W^{-1} Z}$, où Z et W sont des variables aléatoires indépendantes de lois respectives $\mathcal{N}_{\mathbb{R}^d}(0, Id)$ et $\mathcal{W}(N, d)$.*

Proposition 11 (Densité loi $T(N, d)$)**.** *La densité de la loi de Student-Hotelling positive est donnée par :*

$$t \in \mathbb{R}^+ \to \frac{2\Gamma\left(\frac{N+1}{2}\right)}{N^{\frac{d}{2}}\Gamma\left(\frac{N-d+1}{2}\right)\Gamma\left(\frac{d}{2}\right)} \times \frac{t^{d-1}}{\left(1 + \frac{t^2}{N}\right)^{\frac{N+1}{2}}}.$$

Démonstration. On remarque que $T = \sqrt{V^T V}$, où V suit une loi de Student-Hotelling vectorielle à N degrés de liberté en dimension d.
Première étape : Loi de :

$$\tilde{V} = \frac{V}{\sqrt{N}}.$$

Par changement de variable, on montre que la densité de $\tilde{V}$ est :

$$\tilde{v} \in \mathbb{R}^d \to \frac{\Gamma\left(\frac{N+1}{2}\right)}{\pi^{\frac{d}{2}}\Gamma\left(\frac{N-d+1}{2}\right)} \times \frac{1}{(1 + \tilde{v}^T v)^{\frac{N+1}{2}}}.$$

Deuxième étape : Loi de :

$$B = \frac{1}{1 + \tilde{V}^T \tilde{V}}.$$

On montre facilement que pour tout entier $k \geq 0$, on a :

$$\mathbb{E}\left[B^k\right] = \frac{\Gamma\left(\frac{N+1}{2}\right)\Gamma\left(\frac{N-d+1}{2} + k\right)}{\Gamma\left(\frac{N-d+1}{2}\right)\Gamma\left(\frac{N+1}{2} + k\right)}.$$

On en déduit que B suit une loi $\beta\left(\frac{N-d+1}{2}, \frac{d}{2}\right)$.
Dernière étape : Loi de T. Comme :

$$T = \sqrt{N \times \frac{(1 - B)}{B}},$$

on en déduit la loi de T par changement de variable. □

Pour finir, nous aurons besoin du résultat suivant.

Proposition 12. *Soit W une matrice aléatoire suivant la loi de Wishart $\mathcal{W}(N, d)$, alors les valeurs propres de W sont des variables aléatoires indépendantes suivant la loi du Chi-deux $\mathcal{X}^2_{N-d+1}$.*

5.3.2 Intervalles de confiance

Intervalle de confiance de la moyenne au risque p :

Proposition 13. *Soit $(Y_1, \ldots, Y_N)$ un échantillon iid de $\mathcal{N}_{\mathbb{R}^d}(m, \Sigma)$ avec $N \geq d+1$, alors :*

$$T(m) = \sqrt{N} \times \sqrt{(\overline{Y} - m)^T \hat{\Sigma}^{-1} (\overline{Y} - m)},$$

suit une loi de Student-Hotelling $T(N-1, d)$.

Démonstration. Preuve facile. □

On a alors :

$$\mathbb{P}\left[T(m) \leq \alpha\right] = 1 - p,$$

où α est le $(1-p)$-quantile de la loi $T(N-1, d)$. Soit $(\overline{y}, \hat{s})$ réalisation de $(\overline{Y}, \hat{\Sigma})$, on peut considérer l'intervalle de confiance de la moyenne au risque p comme étant l'intervalle :

$$I_p(m) = \left\{ m : \sqrt{N} \times \sqrt{(\overline{y} - m)^T \hat{s}^{-1} (\overline{y} - m)} < \alpha \right\}. \tag{45}$$

Intervalle de confiance de la covariance au risque p :

Proposition 14. *Soit $(Y_1, \ldots, Y_N)$ un échantillon iid de $\mathcal{N}_{\mathbb{R}^d}(m, \Sigma)$ avec $N \geq d+1$, notons $\mu_1, \ldots, \mu_d$ les valeurs propres de $(N-1)\Sigma^{-\frac{1}{2}} \hat{\Sigma} \Sigma^{-\frac{1}{2}}$, alors :*

$$D(\Sigma, \hat{\Sigma}) = \sum_{j=1}^{d} \left|\log(\mu_j)\right|,$$

est une fonction positive et nulle si et seulement si $\Sigma = \hat{\Sigma}$; de plus elle suit la même loi que :

$$lW = \sum_{j=1}^{d} \left|\log(S_j)\right|,$$

où les S_j sont des variables indépendantes de loi $\mathcal{X}^2_{N-d}$. Nous appelerons cette loi "log-Wishart" de degrés de liberté $(N-1, d)$ et nous la noterons $lW(N-1, d)$.

On a alors :

$$\mathbb{P}\left[D(\Sigma, \hat{\Sigma}) < \alpha\right] = 1 - p,$$

où α est le $(1-p)$-quantile de $lW(N-1, d)$. Soit $\hat{s}$ une réalisation de $\hat{\Sigma}$, l'intervalle de confiance de la covariance au risque p est alors :

$$I_p(\Sigma) = \left\{ \Sigma : D(\Sigma, \hat{s}) < \alpha \right\}. \tag{46}$$