Statistiques (STA1)

Cours IV – Voyage en pays Gaussien

Luca Ganassali

Laboratoire de Mathématiques d'Orsay, Université Paris-Saclay

Jeudi 16 octobre 2025

Previously in STA1...

ERRATUM: Cas du test de 2 hypothèses simples

On se place dans le cas où $H_0 : \theta = \theta_0$ et $H_1 : \theta = \theta_1$.

Théorème de Neyman-Pearson (cas de deux hyp. simples) On se place dans un modèle $(\mathcal{Z}, (\mathbb{P}_\theta)_{\theta \in \Theta})$ que l'on suppose dominé. On note L sa vraisemblance. Supposons qu'il existe $k_\alpha > 0$ tel que le test de région de rejet

$$\mathcal{R} = \left\{ z \in \mathcal{Z} : \frac{L(\theta_1; z)}{L(\theta_0; z)} > k_\alpha \right\}$$

soit de taille α, i.e. $\mathbb{P}_{\theta_0}(\mathcal{R}) = \alpha$. Alors, ce test est UPP(α) pour tester $H_0 : \theta = \theta_0$ contre $H_1 : \theta = \theta_1$.

En effet, $\mathbb{P}_{\theta_0}(\mathcal{R}) = \alpha$ est utilisé de façon cruciale dans la preuve.

Vecteurs gaussiens : définitions et propriétés

Rappels sur les vecteurs aléatoires

Soit $X = (X_1, \dots, X_d)$ un vecteur aléatoire à valeurs dans $\mathbb{R}^d$, tel que les X_i admettent toutes un moment d'ordre deux fini. On définit son espérance :

$$\mathbb{E}[X] := (\mathbb{E}[X_1], \dots, \mathbb{E}[X_d]) \in \mathbb{R}^d$$

et sa matrice de covariance :

$$\mathrm{Var}(X) := \mathbb{E}[(X-\mu)(X-\mu)^T] \in \mathbb{R}^{d \times d}.$$

Remarque : Par définition, $[\mathrm{Var}(X)]_{i,j} = \mathbb{E}[(X_i - \mu_i)(X_j - \mu_j)] = \mathrm{Cov}(X_i, X_j)$. $\mathrm{Var}(X)$ est donc une matrice symétrique.

Proposition. Soit $X \in \mathbb{R}^d$ de moment d'ordre deux fini, $A \in \mathbb{R}^{m \times d}$ et $b \in \mathbb{R}^m$. Alors $Y = AX + b$ vérifie :

$$\mathbb{E}[Y] = A\mathbb{E}[X] + b, \qquad \mathrm{Var}(Y) = A\mathrm{Var}(X)A^T.$$

Remarque : une matrice de covariance est donc toujours symétrique positive : pour tout $a \in \mathbb{R}^d$, $a^T\mathrm{Var}(X)a = \mathrm{Var}(a^TX) \geq 0$.

Cas unidimensionnel. Une v.a. réelle Z est gaussienne de moyenne $\mu \in \mathbb{R}$ et variance $\sigma^2 > 0$, notée $Z \sim \mathcal{N}(\mu, \sigma^2)$, si sa densité (par rapport à $\mathrm{Leb}_{\mathbb{R}}$) est :

$$f_{\mu,\sigma^2}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) .$$

Propriétés fondamentales des gaussiennes.

(*i*) Fonction caractéristique :

$$\mathbb{E}[e^{itX}] = \exp\left(i\mu t - \frac{1}{2}\sigma^2 t^2\right)$$

(*ii*) Somme de deux gaussiennes indépendantes : Si $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$ et $X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$, alors $X_1 + X_2 \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.

Preuve : à chercher. Pour (i) on définit f telle que $\mathbb{E}[e^{itX}] = e^{i\mu t} f(t)$ et on cherche une équa diff. sur f. Pour (ii), utiliser (i).

Vecteurs gaussiens

Cas multidimensionnel. Un vecteur aléatoire $X = (X_1, \dots, X_d)$ est un vecteur gaussien si toute combinaison linéaire de ses coefficients est une v.a. gaussienne. On note alors $X \sim \mathcal{N}(\mu, \Sigma)$, avec $\mu = \mathbb{E}[X]$ et $\Sigma = \mathrm{Var}(X)$.

Proposition (Densité d'un vecteur gaussien). Soit $X \sim \mathcal{N}(\mu, \Sigma)$ dans $\mathbb{R}^d$.

- si $\Sigma \in \mathbb{R}^{d \times d}$ est inversible, alors X a pour densité (par rapport à $\mathrm{Leb}_{\mathbb{R}^d}$):

$$f_{\mu,\Sigma}(x) = \frac{1}{(2\pi)^{d/2}\sqrt{\det \Sigma}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right) .$$

- si Σ n'est pas inversible, alors $X - \mu$ appartient p.s. à un sous-espace vectoriel de dimension $\mathrm{rg}(\Sigma) < d$.

Exercice : prouver la deuxième affirmation. On pourra prendre $v \in \mathrm{Ker}(\Sigma)$ et calculer $\mathrm{Var}(v^T(X-\mu))$.

Propriétés fondamentales.

(*i*) *Fonction caractéristique.* Si $X \sim \mathcal{N}(\mu, \Sigma)$, alors pour tout $s \in \mathbb{R}^d$,

$$\mathbb{E}[e^{is^T X}] = \exp(is^T \mu - \tfrac{1}{2} s^T \Sigma s)\,.$$

(*ii*) *Somme de vecteurs gaussiens indépendants.* Si $X_1 \sim \mathcal{N}(\mu_1, \Sigma_1)$ et $X_2 \sim \mathcal{N}(\mu_2, \Sigma_2)$ sont indépendants, alors

$$X_1 + X_2 \sim \mathcal{N}(\mu_1 + \mu_2, \Sigma_1 + \Sigma_2)\,.$$

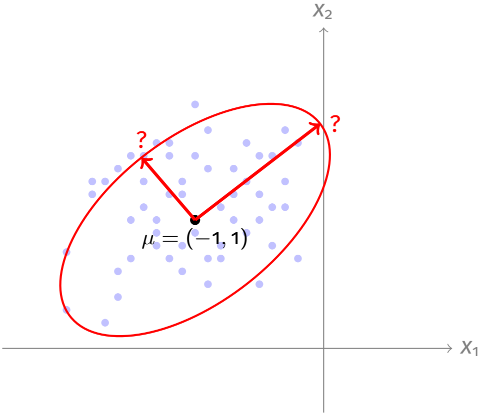
(*iii*) *Stabilité par transformation linéaire.* Si $X \sim \mathcal{N}(\mu, \Sigma)$ alors pour tout A de $\mathbb{R}^{m \times d}$, $b \in \mathbb{R}^m$, $AX + b \sim \mathcal{N}(A\mu + b, A\Sigma A^T)$.

(*iv*) *Indépendances via la matrice de covariance.* Soit $X \sim \mathcal{N}(\mu, \Sigma)$ dans $\mathbb{R}^d$. Pour tout $J \subseteq \{1, \ldots, d\}$,

$$(X_j)_{j \in J} \text{ sont mutuellement indépendants } \iff \Sigma_{J,J} \text{ est diagonale,}$$

où $\Sigma_{J,J}$ = matrice extraite de Σ en ne gardant que les lignes/colonnes d'indices $j \in J$.

Vecteurs gaussiens : illustration géométrique



50 échantillons i.i.d. gaussiens de moyenne $\mu = (-1, 1)$ et de matrice de covariance $\Sigma = \begin{pmatrix} 1 & 0.6 \\ 0.6 & 0.8 \end{pmatrix}$

Quand la géométrie rencontre les probabilités : le Théorème de Cochran

Loi du khi-deux et Théorème de Cochran

Soit $d \geq 1$ et $Z \sim \mathcal{N}(0_d, I_d)$. La loi du khi-deux à d degrés de liberté, notée $\chi^2(d)$, est la loi de la somme des carrés de ses composantes :

$$\|Z\|^2 = Z_1^2 + \ldots + Z_d^2 \sim \chi^2(d)\,.$$

Remarque : Si $Z \sim \chi^2(d)$, on a $\mathbb{E}[Z] = d$ et $\mathrm{Var}(Z) = 2d$.

Théorème de Cochran. Soit $Z \sim \mathcal{N}(\mu, \sigma^2 I_n)$ un vecteur gaussien de $\mathbb{R}^n$. Soient $r \geq 1$ et $E_1, \ldots, E_r$ des sous-espaces vectoriels orthogonaux tels que $\bigoplus_{i=1}^r E_i = \mathbb{R}^n$. Pour $1 \leq j \leq n$, notons Π_j le projecteur orthogonal sur E_j et d_j sa dimension. Alors,

- (i) les vecteurs aléatoires $\Pi_1 Z, \ldots, \Pi_r Z$ sont gaussiens mutuellement indépendants et de lois respectives $\mathcal{N}(\Pi_1\mu, \sigma^2\Pi_1), \ldots, \mathcal{N}(\Pi_r\mu, \sigma^2\Pi_r)$;
- (ii) Les variables $\frac{\|\Pi_1(Z-\mu)\|^2}{\sigma^2}, \ldots, \frac{\|\Pi_r(Z-\mu)\|^2}{\sigma^2}$ sont mutuellement indépendantes et de lois respectives $\chi^2(d_1), \ldots, \chi^2(d_r)$.

Sans perte de généralité, on peut supposer $\mu = 0$ et $\sigma^2 = 1$, i.e. $Z \sim \mathcal{N}(0, I_n)$. Considérons le vecteur concaténé

$$U := (\Pi_1 Z, \ldots, \Pi_r Z) \in \mathbb{R}^{nr},$$

et définissons la matrice de projection en blocs

$$\Pi := \begin{pmatrix} \Pi_1 \\ \vdots \\ \Pi_r \end{pmatrix} \in \mathbb{R}^{nr \times n},$$

de sorte que $U = \Pi Z$. Pour étudier la covariance de U, on calcule

$$\mathrm{Var}(U) = \Pi \mathrm{Var}(Z) \Pi^T = \Pi\Pi^T = \begin{pmatrix} \Pi_1 \\ \vdots \\ \Pi_r \end{pmatrix} (\Pi_1^T, \ldots, \Pi_r^T) = \begin{pmatrix} \Pi_1\Pi_1^T & \Pi_1\Pi_2^T & \cdots & \Pi_1\Pi_r^T \\ \Pi_2\Pi_1^T & \Pi_2\Pi_2^T & \cdots & \Pi_2\Pi_r^T \\ \vdots & & \ddots & \vdots \\ \Pi_r\Pi_1^T & \cdots & \cdots & \Pi_r\Pi_r^T \end{pmatrix}.$$

Les E_j sont 2 à 2 orthogonaux $\longrightarrow \Pi_j \Pi_{j'}^T = \Pi_j \Pi_{j'} = 0$ pour $j \neq j'$. Donc :

$$\mathrm{Var}(U) = \Pi\Pi^T = \begin{pmatrix} \Pi_1\Pi_1^T & 0 & \cdots & 0 \\ 0 & \Pi_2\Pi_2^T & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \Pi_r\Pi_r^T \end{pmatrix} = \begin{pmatrix} \Pi_1 & 0 & \cdots & 0 \\ 0 & \Pi_2 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & \Pi_r \end{pmatrix}.$$

Les propriétés fondamentales des vecteurs gaussiens impliquent que les vecteurs projetés $\Pi_j Z$ sont mutuellement indépendants, et on lit $\Pi_j Z \sim \mathcal{N}(0, \Pi_j)$, ce qui conclut (i).

Pour (ii), pour chaque j choisissons une base orthonormée $(e_{j,1}, \ldots, e_{j,d_j})$ de E_j. On a

$$\|\Pi_j Z\|^2 = \sum_{k=1}^{d_j} (e_{j,k}^T Z)^2 \sim \chi^2(d_j),$$

car les $e_{j,k}^T Z \sim \mathcal{N}(0,1)$ sont indépendants. De plus, les $\Pi_j Z$ sont mutuellement indépendants donc les $\|\Pi_j Z\|^2$ le sont également, ce qui conclut (ii).

Théorème de Cochran : illustration géométrique

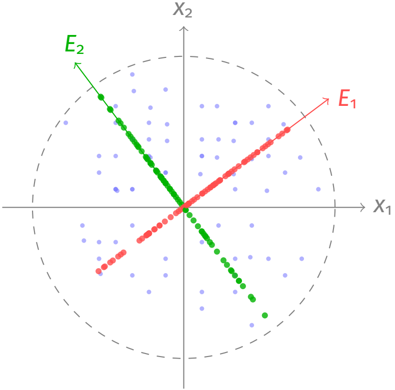


Illustration du théorème de Cochran en dimension 2, avec $d_1 = d_2 = 1$.

Autres lois du monde Gaussien

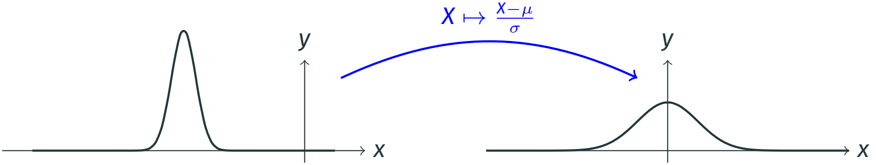
Lois de Student et de Fisher

Loi de Student. Soit $Z \sim \mathcal{N}(0, 1)$ et $K \sim \chi^2(d)$ indépendantes. La loi de Student à d degrés de liberté, notée $\mathcal{T}(d)$, est la loi de la variable

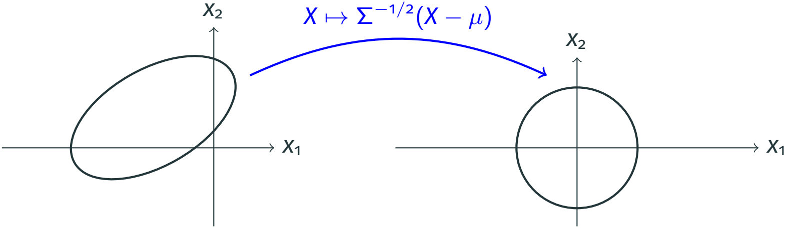
$$T = \frac{Z}{\sqrt{K/d}} \sim \mathcal{T}(d)\,.$$

Loi de Fisher. Soient $K_1 \sim \chi^2(n_1)$ et $K_2 \sim \chi^2(n_2)$ indépendantes. La loi de Fisher à (n_1, n_2) degrés de liberté, notée $\mathcal{F}(n_1, n_2)$, est la loi de la variable

$$F = \frac{K_1/n_1}{K_2/n_2} \sim \mathcal{F}(n_1, n_2)\,.$$



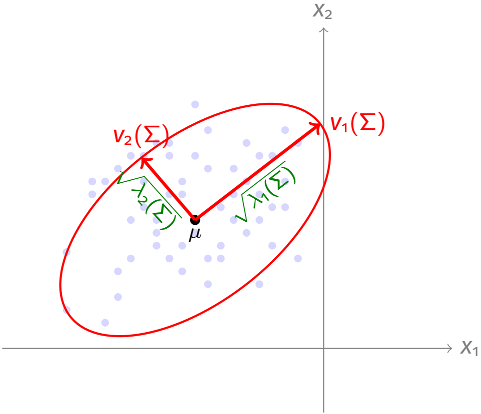
$X \sim \mathcal{N}(\mu, \sigma^2)$ dans $\mathbb{R}$ $\qquad$ $X \sim \mathcal{N}(0, 1)$ dans $\mathbb{R}$



$X \sim \mathcal{N}(\mu, \Sigma)$ dans $\mathbb{R}^d$ $\qquad$ $X \sim \mathcal{N}(0, I_2)$ dans $\mathbb{R}^d$

Bonus : géométrie du monde Gaussien

Bonus : retour sur l'illustration géométrique



50 échantillons i.i.d. gaussiens de moyenne $\mu = (-1, 1)$ et de matrice de covariance $\Sigma = \begin{pmatrix} 1 & 0.6 \\ 0.6 & 0.8 \end{pmatrix}$

Merci !

Rdv en TD pour les questions et la pratique de ces notions.

(contenu du cours disponible sur ma page web : `lganassali.github.io`)