Statistique Mathématique 2 (MATH-F-309, Chapitre #2)

Thomas Verdebout

Université Libre de Bruxelles

2015/2016

Plan du cours

1. Vecteurs aléatoires.
2. $\boxed{\text{Loi normale multivariée.}}$
3. Inférence dans les modèles gaussiens.
4. Méthodes classiques de l'analyse multivariée.
5. Données directionelles.
6. Modèle linéaire.

Plan du cours

2. Loi normale multivariée.

    2.1. Définitions, propriétés de base.

    2.2. Indépendance et normalité multivariée.

    2.3. Lois conditionnelles.

    2.4. Loi normale matricielle.

    2.5. Loi de Wishart, lemme de Fisher multivarié.

    2.6. TCL multivarié.

    Appendix: Conditional expectations.

Définitions, propriétés de base

Soit $X = (X_1, \ldots, X_p)'$ un p-v.a.

Définition: *X est de loi normale p-variée centrée réduite (notation: $X \sim \mathcal{N}_p(0, I_p)$)* $\Leftrightarrow$ *les X_i, $i = 1, \ldots, p$ sont i.i.d. $\mathcal{N}(0, 1)$.*

Clairement, cette loi est absolument continue par rapport à la mesure de Lebesgue et a pour densité

$$x \mapsto f^X(x) = \prod_{i=1}^{p} f^{X_i}(x_i) = \prod_{i=1}^{p} \left(\frac{1}{\sqrt{2\pi}} \exp(-x_i^2/2) \right)$$

$$= \frac{1}{(2\pi)^{p/2}} \exp(-\sum_{i=1}^{p} x_i^2/2) = \left(\frac{1}{2\pi} \right)^{\frac{p}{2}} \exp(-\|x\|^2/2).$$

Aussi,

$$\mathrm{E}[X] = \begin{pmatrix} \mathrm{E}[X_1] \\ \vdots \\ \mathrm{E}[X_p] \end{pmatrix} = 0 \quad \text{et} \quad \mathrm{Var}[X] = \Big(\mathrm{Cov}[X_i, X_j] \Big)_{i,j=1,\ldots,p} = I_p.$$

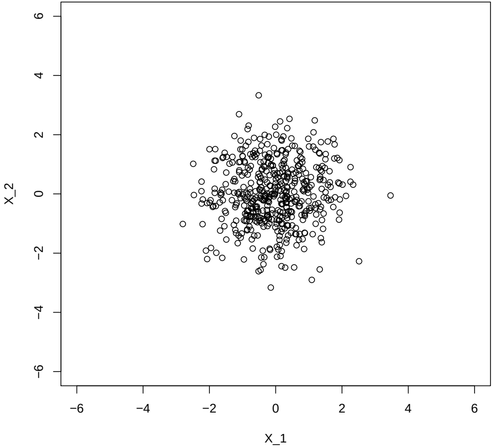
Définitions, propriétés de base

Remarque: si $X \sim \mathcal{N}_p(0, I_p)$, X a des moments finis de tout ordre. En effet, pour tout $s > 0$, on a

$$E[\|X\|^s] \leq C_{p,s} \sum_{i=1}^{p} E|X_i|^s = pC_{p,s} E|X_1|^s < \infty.$$

Définitions, propriétés de base

500 observations i.i.d. de loi $\mathcal{N}_2(0, I_2)$:



Définitions, propriétés de base

Définition: *X est de loi normale p-variée $\Leftrightarrow$ il y a un vecteur $\mu \in \mathbb{R}^p$ et une matrice A $(p \times q)$ tels que $X = AZ + \mu$, où $Z \sim \mathcal{N}_q(0, I_q)$.*

Remarques:

(i) $\mathrm{E}[X] = A\mathrm{E}[Z] + \mu = \mu$ et $\mathrm{Var}[X] = A\mathrm{Var}[Z]A' = AA'$.

(ii) $\varphi_X(t) = \mathrm{E}[e^{it'X}] = \mathrm{E}[e^{it'(AZ+\mu)}] = \mathrm{E}[e^{i(A't)'Z}]e^{it'\mu}$. Donc

$$\varphi_X(t) = \exp(it'\mu)\exp\left(-\frac{1}{2}t'AA't\right).$$

Ça implique que la loi de X dépend que de μ et $\mathrm{Var}(X) = AA' =: \Sigma$.

On écrit: $X \sim \mathcal{N}_p(\mu, \Sigma)$.

(iii) Si $X \sim \mathcal{N}_p(\mu, \Sigma)$, alors $Y = BX + \nu$ ($\nu \in \mathbb{R}^d$ et $B \in \mathbb{R}^{d \times p}$) est de loi normale d-variée, $Y \sim \mathcal{N}_p(\nu + B\mu, B\Sigma B')$.

En particulier, en prenant $B = \left(I_{p_1} \,|\, 0_{p_1 \times (p-p_1)}\right)$ et $\nu = 0$, il découle de cette proposition le résultat suivant.

Proposition: *soit $X = (X_1', X_2')' \sim \mathcal{N}_p(\mu, \Sigma)$, où X_i est un p_i-v.a. ($i = 1, 2$) et où*

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \quad et \quad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$$

Alors $X_1 \sim \mathcal{N}_{p_1}(\mu_1, \Sigma_{11})$.

Ceci montre donc que tous les v.a. extraits d'un v.a. de loi normale multivariée sont également de loi normale.

(iv) Soit $\mu \in \mathbb{R}^p$ et $\Sigma \in \mathbb{R}^{p \times p}$ symétrique et semi-défini positive (on écrit: $\Sigma \geq 0$), alors il y a un vecteur aléatoire $X \sim \mathcal{N}_p(\mu, \Sigma)$.

Preuve: Le *théorème spectral* implique, que

$$\Sigma = O \Lambda O',$$

où $\Lambda \geq 0$ $(p \times p)$ est une matrice diagonale et O $(p \times p)$ est une matrice *orthogonale* $(O'O = I_p)$. Donc si on définit $A = O\Lambda^{1/2}O'$ et $Z \sim \mathcal{N}_p(0, I_p)$, alors $\mu + AZ \stackrel{\mathcal{D}}{=} \mathcal{N}_p(\mu, \Sigma)$. $\square$

(v) Comme dans le cas standard (i.e., $\mu = 0$ et $\Sigma = I_p$), $X \sim \mathcal{N}_p(\mu, \Sigma)$ a des moments finis de tout ordre.

(vi) Si Σ est une matrice $(p \times p)$ symétrique et définie positive (on écrit: $\Sigma > 0$), et $X \sim \mathcal{N}(\mu, \Sigma)$, alors la loi d'un tel X est absolument continue par rapport à la mesure de Lebesgue et a pour densité

$$\begin{aligned} x \mapsto f^X(x) &= |A^{-1}| f^Z(A^{-1}(x-\mu)) \\ &= \left(\frac{1}{2\pi}\right)^{\frac{p}{2}} \frac{1}{|A|} \exp(-\|A^{-1}(x-\mu)\|^2/2) \\ &= \left(\frac{1}{2\pi}\right)^{\frac{p}{2}} \frac{1}{|\Sigma|^{\frac{1}{2}}} \exp(-(x-\mu)'\Sigma^{-1}(x-\mu)/2), \end{aligned}$$

(nous avons pris $A = \Sigma^{1/2}$ de (iv)).

(vii) Soit $X \sim \mathcal{N}(\mu, \Sigma)$. Dans le cas $\Sigma > 0$, on a $\Sigma = O\Lambda O'$ avec $\Lambda > 0$. Donc, on peut définir $\Sigma^{-1/2} = O\Lambda^{-1/2}O'$. Alors

$$\Sigma^{-1/2}(X - \mu) \sim \mathcal{N}(0, I_p).$$

Ça ressemble à

$$\frac{1}{\sigma}(X - \mu) \sim N(0, 1),$$

dans le cas où X a de loi normale univariée $N(\mu, \sigma)$.

Définitions, propriétés de base

Proposition: $X \sim \mathcal{N}_p(\mu, \Sigma) \Leftrightarrow \forall a \in \mathbb{R}^p \backslash \{0\},\ a'X \sim \mathcal{N}_1(a'\mu, a'\Sigma a)$.

Preuve:

($\Rightarrow$) $\forall t \in \mathbb{R}$, on a
$$\varphi_{a'X}(t) = \varphi_X(ta) = e^{i(ta)'\mu} e^{-(ta)'\Sigma(ta)/2} = e^{it(a'\mu)} e^{-(a'\Sigma a)t^2/2} = \varphi_Y(t),$$
où $Y \sim \mathcal{N}_1(a'\mu, a'\Sigma a)$, de sorte que $a'X \sim \mathcal{N}_1(a'\mu, a'\Sigma a)$.

($\Leftarrow$) $\forall t \in \mathbb{R}^p \backslash \{0\}$, on a
$$\varphi_X(t) = \varphi_{t'X}(1) = e^{i \times 1 \times (t'\mu)} e^{-(t'\Sigma t) \times 1^2/2} = e^{it'\mu} e^{-t'\Sigma t/2} = \varphi_Y(t),$$
où $Y \sim \mathcal{N}_p(\mu, \Sigma)$. Bien entendu, on a aussi $\varphi_X(0) = 1 = \varphi_Y(0)$. Donc $X \sim \mathcal{N}_p(\mu, \Sigma)$. $\square$

Définitions, propriétés de base

Il ne suffit pas que toutes les marges multivariées soient de loi normale pour que le vecteur lui-même soit de loi normale!

Exemple:

Soit $X = (X_1, X_2)'$, où $X_1 \sim \mathcal{N}(0,1)$ et $X_2 = \delta X_1$, où δ est independante de X_1 est $P(\delta = \pm 1) = \frac{1}{2}$.

Alors on vérifie facilement que

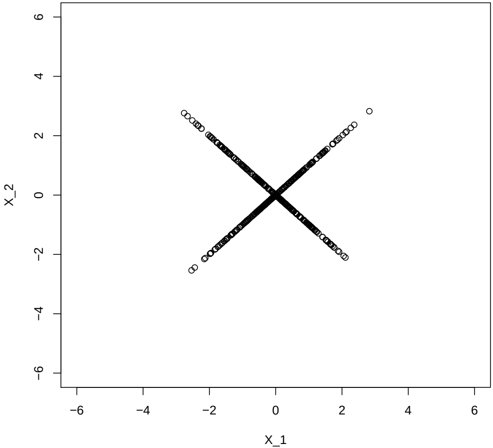
- $X_2 \sim \mathcal{N}(0,1)$ (exercice), mais que
- X n'est pas de loi normale bivariée.

Ce second point est obtenu en observant que

$$P(X_1 + X_2 = 0) = \frac{1}{2},$$

et par consequence $a'X$, avec $a = (1,1)'$, n'est pas de loi normal.

500 observations i.i.d. qui ont la loi bivariée ci-dessus:



Définitions, propriétés de base

Définition: la distance de Mahalanobis entre x et y dans la métrique associée à Σ (notation: $d_\Sigma(x, y)$) est la quantité $\sqrt{(x-y)'\Sigma^{-1}(x-y)}$.

La densité d'une loi normale p-variée est alors

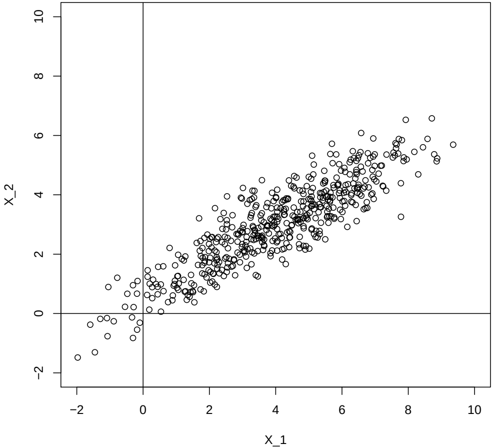
$$x \mapsto f^X(x) = \left(\frac{1}{2\pi}\right)^{\frac{p}{2}} \frac{1}{|\Sigma|^{\frac{1}{2}}} \exp\left(-d^2_\Sigma(x, \mu)/2\right);$$

les courbes de niveau de f^X sont donc des hyper-ellipsoïdes (dans $\mathbb{R}^p$) de centre μ et dont la forme et l'orientation sont déterminées par Σ.

Définitions, propriétés de base

500 observations i.i.d. de loi $\mathcal{N}_2(\mu, \Sigma)$, où

$$\mu = \begin{pmatrix} 4 \\ 3 \end{pmatrix} \quad \text{et} \quad \Sigma = \begin{pmatrix} 5 & 3 \\ 3 & 2.25 \end{pmatrix}.$$



Définitions, propriétés de base

Proposition: si $X \sim \mathcal{N}_p(\mu, \Sigma)$ où $\Sigma > 0$, $d^2_\Sigma(X, \mu) \sim \chi^2_p$.

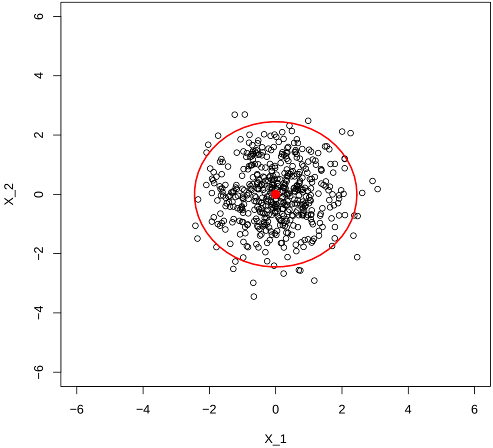
Preuve: en utilisant le fait que $X \stackrel{\mathcal{D}}{=} AZ + \mu$, où $Z \sim \mathcal{N}_p(0, I_p)$ et $AA' = \Sigma$, on obtient

$$\begin{aligned} d^2_\Sigma(X, \mu) = (X - \mu)'\Sigma^{-1}(X - \mu) &\stackrel{\mathcal{D}}{=} ((AZ + \mu) - \mu)'\Sigma^{-1}(AZ + \mu) - \mu) \\ &= (AZ)'\Sigma^{-1}(AZ) = Z'A'(AA')^{-1}AZ = Z'Z = \sum_{i=1}^{p} Z_i^2, \end{aligned}$$

qui est bien de loi χ^2_p, puisque les Z_i sont i.i.d. $\mathcal{N}(0, 1)$ par définition. $\square$

Par conséquent, l'ellipsoïde $E_{1-\alpha} := \{y \in \mathbb{R}^p \mid d^2_\Sigma(y, \mu) \leq \chi^2_{p;1-\alpha}\}$ (où $\chi^2_{p;1-\alpha}$ désigne le quantile d'ordre $1 - \alpha$ de la loi χ^2_p) contient une masse de probabilité d'exactement $1 - \alpha$. On parlera de zone de tolérance (à $(1 - \alpha) \times 100\%$) .

500 observations i.i.d. de loi $\mathcal{N}_2(0, I_2)$ et la zone de tolérance $E_{.95}$.

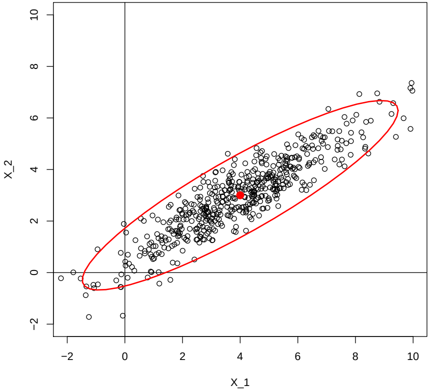


Définitions, propriétés de base

500 observations i.i.d. de loi $\mathcal{N}_2(\mu, \Sigma)$, où

$$\mu = \begin{pmatrix} 4 \\ 3 \end{pmatrix} \quad \text{et} \quad \Sigma = \begin{pmatrix} 5 & 3 \\ 3 & 2.25 \end{pmatrix}.$$

et la zone de tolérance $E_{.95}$.



Plan du cours

Indépendance et normalité multivariée

Soit X_1 un p_1-v.a. et X_2 un p_2-v.a. On sait que

- $X_1 \perp\!\!\!\perp X_2 \Rightarrow \mathrm{Cov}[X_1, X_2] = 0$, mais que
- la réciproque n'est pas vraie en général (exemple: $p_1 = p_2 = 1$, $X_1 \sim \mathcal{N}_1(0,1)$ et $X_2 = (X_1)^2$).

Le résultat suivant montre que la réciproque tient dans le cas où $X = (X_1', X_2')'$ est de loi normale p-variée ($p = p_1 + p_2$).

Proposition: *soit $X = (X_1', X_2')' \sim \mathcal{N}_p(\mu, \Sigma)$, où X_i est un p_i-v.a. ($i = 1, 2$) et où*

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \quad et \quad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$$

Alors $X_1 \perp\!\!\!\perp X_2$ ssi $\Sigma_{12} = 0$.

Remarque: $\Sigma_{12} = \mathrm{Cov}[X_1, X_2]$.

Preuve:

($\Rightarrow$) Prouvé au 1er chapitre (sans hypothèse de normalité).

($\Leftarrow$) Par définition,

$$X = AZ + \mu,$$

où $Z \sim \mathcal{N}_p(0, I_p)$ et où A est une matrice $(p \times p)$ quelconque telle que $AA' = \Sigma$. Clairement, puisque $\Sigma_{12} = 0$, on peut prendre

$$A = \begin{pmatrix} A_{11} & 0 \\ 0 & A_{22} \end{pmatrix},$$

où A_{ii} est une matrice $(p_i \times q_i)$ telle que $A_{ii}A'_{ii} = \Sigma_{ii}$ $(i = 1, 2)$. On obtient alors

$$\begin{pmatrix} X_1 \\ X_2 \end{pmatrix} = X = AZ + \mu = \begin{pmatrix} A_{11}Z_1 + \mu_1 \\ A_{22}Z_2 + \mu_2 \end{pmatrix} = \begin{pmatrix} h_1(Z_1) \\ h_2(Z_2) \end{pmatrix},$$

ce qui montre que $X_1 \perp\!\!\!\perp X_2$ (puisque $Z_1 \perp\!\!\!\perp Z_2$). $\square$

Indépendance et normalité multivariée

Vice versa, il convient la suivante:

Proposition: *soit* $X = (X_1', X_2')'$*, où* $X_i \sim \mathcal{N}_{p_i}(\mu_i, \Sigma_i)$ $(i = 1, 2)$ *sont indépendants. Alors* $X \sim \mathcal{N}_{p_1+p_2}(\mu, \Sigma)$*, où*

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \quad et \quad \Sigma = \begin{pmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{pmatrix}.$$

En utilisant la proposition, on obtient pour n p-v.a. independants et de la loi normal, que $X = (X_1', \ldots, X_n')' \sim \mathcal{N}_{np}(\mu, \Sigma)$, où

$$\mu = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_n \end{pmatrix} \quad \text{et} \quad \Sigma = \begin{pmatrix} \Sigma_1 & & 0 \\ & \ddots & \\ 0 & & \Sigma_n \end{pmatrix}.$$

Corollaire: *soient X_i, $i = 1, \ldots, n$ des p-v.a. indépendants tels que $X_i \sim \mathcal{N}_p(\mu_i, \Sigma_i)$. Soient c_i, d_i, $i = 1, \ldots, n$ des constantes réelles. Alors*

- *(i)* $\sum_{i=1}^n c_i X_i \sim \mathcal{N}_p(\sum_{i=1}^n c_i \mu_i, \sum_{i=1}^n c_i^2 \Sigma_i)$;
- *(ii)*

$$\begin{pmatrix} \sum_{i=1}^n c_i X_i \\ \sum_{i=1}^n d_i X_i \end{pmatrix} \sim \mathcal{N}_{2p}\left(\begin{pmatrix} \sum_{i=1}^n c_i \mu_i \\ \sum_{i=1}^n d_i \mu_i \end{pmatrix}, \begin{pmatrix} \sum_{i=1}^n c_i^2 \Sigma_i & \sum_{i=1}^n c_i d_i \Sigma_i \\ \sum_{i=1}^n c_i d_i \Sigma_i & \sum_{i=1}^n d_i^2 \Sigma_i \end{pmatrix} \right);$$

- *(iii) si* $\Sigma_i = \Sigma_1 \ \forall i$, $(\sum_{i=1}^n c_i X_i) \perp\!\!\!\perp (\sum_{i=1}^n d_i X_i) \Leftrightarrow \sum_{i=1}^n c_i d_i = 0$.

Preuve: (i) et (ii) découlent de $BX \sim \mathcal{N}_p(B\mu, B\Sigma B')$, où

$$B = (c_1 I_p \ldots c_n I_p) \quad \text{et} \quad B = \begin{pmatrix} c_1 I_p \ldots c_n I_p \\ d_1 I_p \ldots d_n I_p \end{pmatrix}.$$

(iii) est une conséquence directe du fait que la non-corrélation des marges équivaut à leur indépendance pour les v.a. normaux. $\square$

Corollaire: *soient X_i, $i = 1, \ldots, n$ des p-v.a. i.i.d. de loi $\mathcal{N}_p(\mu, \Sigma)$. Alors $\bar{X} := \frac{1}{n}\sum_{i=1}^{n} X_i \sim \mathcal{N}_p(\mu, \frac{1}{n}\Sigma)$.*

Plan du cours

Lois conditionnelles

Proposition: *soit* $X = (X_1', X_2')' \sim \mathcal{N}_p(\mu, \Sigma)$*, où* X_i *est un* p_i*-v.a.* $(i = 1, 2)$*,* $\Sigma > 0$ *et où*

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} \quad et \quad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}.$$

Alors $X_2 \,|\, X_1 = x_1 \sim \mathcal{N}_{p_2}(\mu_2 + \Sigma_{21}\Sigma_{11}^{-1}(x_1 - \mu_1), \Sigma_{22.1})$*, où*
$\Sigma_{22.1} := \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}$.

Remarques:

- la variance de $X_2 \,|\, X_1 = x_1$ ne dépend pas de x_1. Ce phénomène est connu sous le nom d'homoscédasticité.
- La variance des lois conditionnelles est plus petite que celle des lois originales. En effet, $\mathrm{Var}[X_2] - \mathrm{Var}[X_2 | X_1 = x_1]$
$= \Sigma_{22} - \Sigma_{22.1} = \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12} \geq 0$.

Lois conditionnelles

Preuve: Σ_{11} est définie positive (puisque c'est le cas de Σ), et est donc aussi inversible. Bien entendu, $X - \mu \sim \mathcal{N}_p(0, \Sigma)$. Donc, en posant

$$B = \begin{pmatrix} I_{p_1} & 0 \\ -\Sigma_{21}\Sigma_{11}^{-1} & I_{p_2} \end{pmatrix}$$

on obtient $B(X - \mu) \sim \mathcal{N}_p(0, B\Sigma B')$, où

$$B(X - \mu) = \begin{pmatrix} X_1 - \mu_1 \\ -\Sigma_{21}\Sigma_{11}^{-1}(X_1 - \mu_1) + (X_2 - \mu_2) \end{pmatrix}$$

et

$$B\Sigma B' = \begin{pmatrix} \Sigma_{11} & 0 \\ 0 & \Sigma_{22.1} \end{pmatrix}.$$

Nous posons $Y := -\Sigma_{21}\Sigma_{11}^{-1}(X_1 - \mu_1) + (X_2 - \mu_2)$.

Alors

$$\begin{aligned}
&Y \sim \mathcal{N}_{p_2}(0, \Sigma_{22.1}),\\
&X_2 = \mu_2 + Y + \Sigma_{21}\Sigma_{11}^{-1}(X_1 - \mu_1),\\
&X_1 \perp\!\!\!\perp Y,
\end{aligned}$$

et donc

$$\begin{aligned}
&E\left[\exp(it'X_2)|X_1 = x_1\right]\\
&\quad = E\left[\exp(it'[\mu_2 + Y + \Sigma_{21}\Sigma_{11}^{-1}(X_1 - \mu_1)])|X_1 = x_1\right]\\
&\quad = \exp(it'[\Sigma_{21}\Sigma_{11}^{-1}(x_1 - \mu_1)]) \times \underbrace{E\left[\exp(it'[\mu_2 + Y])|X_1 = x_1\right]}_{=E[\exp(it'[\mu_2+Y])] \quad \text{puisque } X_1 \perp\!\!\!\perp Y}\\
&\quad = \exp(it'[\Sigma_{21}\Sigma_{11}^{-1}(x_1 - \mu_1) + \mu_2]) \times E\left[\exp(it'Y)\right]
\end{aligned}$$

$\square$

Plan du cours

$\boxed{\text{2. Loi normale multivariée.}}$

2.1. Définitions, propriétés de base.

2.2. Indépendance et normalité multivariée.

2.3. Lois conditionnelles.

$\boxed{\text{2.4. Loi normale matricielle.}}$

2.5. Loi de Wishart, lemme de Fisher multivarié.

2.6. TCL multivarié.

Loi normale matricielle

Pour définir la loi normale matricielle, nous aurons besoin des deux notations suivantes, qui sont classiques en analyse multivariée.

Définition: *soient $A = (A_1 \dots A_n) = (a_{ij})$ une matrice $m \times n$ et B une matrice $p \times q$. Alors*

$$\mathrm{vec}\, A = \begin{pmatrix} A_1 \\ \vdots \\ A_n \end{pmatrix} \quad \textit{et} \quad A \otimes B = \begin{pmatrix} a_{11}B & \dots & a_{1n}B \\ \vdots & & \vdots \\ a_{m1}B & \dots & a_{mn}B \end{pmatrix}.$$

Remarques:

- $A \otimes B$ est appelé le produit de Kronecker de A et B.
- $A \otimes B$ est de taille $mp \times nq$.
- En général, $A \otimes B \neq B \otimes A$.

Loi normale matricielle

Propriétés:

Pour toute matrice $A, A_1, A_2, B, B_1, B_2, C$ et pour tout réel λ, μ,

- $(A_1 + A_2) \otimes B = A_1 \otimes B + A_2 \otimes B$ et $A \otimes (B_1 + B_2) = A \otimes B_1 + A \otimes B_2$.
- $(A \otimes B) \otimes C = A \otimes (B \otimes C)$.
- $(A_1 \otimes B_1)(A_2 \otimes B_2) = (A_1 A_2) \otimes (B_1 B_2)$, $(A \otimes B)' = A' \otimes B'$ et $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$.
- $\mathrm{tr}[A \otimes B] = (\mathrm{tr}\, A)(\mathrm{tr}\, B)$ et, si A $(m \times m)$ et B $(n \times n)$, $\det[A \otimes B] = (\det\, A)^n (\det\, B)^m$.
- $A, B > 0 \Rightarrow A \otimes B > 0$.

On a aussi (et surtout) les liens suivants entre “vec" et “$\otimes$":

- $(\mathrm{vec}\, A)'(\mathrm{vec}\, B) = \mathrm{tr}[A'B]$ et $\mathrm{vec}\,(ABC) = (C' \otimes A)(\mathrm{vec}\, B)$.

Loi normale matricielle

Soit X une matrice aléatoire de dimension $n \times p$.

Définition: *X est de loi normale de moyenne M et de variance-covariance Ω (notation: $X \sim \mathcal{N}_{n,p}(M, \Omega)$) $\Leftrightarrow \operatorname{vec} X' \sim \mathcal{N}_{np}(\operatorname{vec} M', \Omega)$.*

Exemple: posons $1_n = (1, 1, \ldots, 1)' \in \mathbb{R}^n$. Alors, si $X_1, \ldots, X_n$ sont i.i.d. $\mathcal{N}_p(\mu, \Sigma)$, la matrice échantillon

$$X = \begin{pmatrix} X_1' \\ \vdots \\ X_n' \end{pmatrix} \sim \mathcal{N}_{n,p}\Big(1_n \otimes \mu', I_n \otimes \Sigma\Big)$$

puisque $\operatorname{vec} X' = (X_1', \ldots, X_n')' \sim \mathcal{N}_{np}(1_n \otimes \mu, I_n \otimes \Sigma)$, où $1_n \otimes \mu = \operatorname{vec}(1_n \otimes \mu')'$.

Loi normale matricielle

Le résultat suivant indique comment une matrice aléatoire de loi normale se comporte sous l'application de "transformations linéaires".

Proposition: *soit $X \sim \mathcal{N}_{n,p}(M, \Omega)$ et soient A, B des matrices de dimensions respectives $(r \times n)$ et $(p \times s)$. Alors*

$$AXB \sim \mathcal{N}_{r,s}(AMB, (A \otimes B')\Omega(A \otimes B')').$$

Preuve: On a vu que
$\mathrm{vec}\,(AXB)' = (A \otimes B')(\mathrm{vec}\,X') \sim \mathcal{N}_{rs}((A \otimes B')(\mathrm{vec}\,M'), (A \otimes B')\Omega(A \otimes B')').$ $\square$

Exemple: Si X désigne la matrice échantillon dans la situation ci-dessus, on a que $\bar{X}' = n^{-1}1_n' X I_p \sim \mathcal{N}_{1,p}(\mu', \frac{1}{n}\Sigma)$, ce qui signifie bien que $\bar{X} = \mathrm{vec}\,(\bar{X}')' \sim \mathcal{N}_p(\mu, \frac{1}{n}\Sigma)$.

Plan du cours

Lemme de Fisher multivarié

Nous savons que l'inférence statistique dans de nombreux modèles (gaussiens) univariés est fondée sur le lemme de Fisher:

Proposition: *soient* $X_1, \ldots, X_n$ *i.i.d.* $\mathcal{N}_1(\mu, \sigma^2)$. *Alors, en notant* $\bar{X} := \frac{1}{n}\sum_{i=1}^n X_i$ *et* $S := \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X})^2$, *on a*

- *(i)* $\bar{X} \sim \mathcal{N}_1(\mu, \frac{\sigma^2}{n})$,
- *(ii)* $(n-1)S \sim \sigma^2 \chi^2_{n-1}$, *et*
- *(iii)* $\bar{X} \perp\!\!\!\perp S$.

Notre but est ici d'étendre ce résultat au cas multivarié :

Proposition: *soient* $X_1, \ldots, X_n$ *i.i.d.* $\mathcal{N}_p(\mu, \Sigma)$. *Alors*

- *(i)* $\bar{X} \sim \mathcal{N}_p(\mu, \frac{1}{n}\Sigma)$,
- *(ii)* $S \sim ?$, *et*
- *(iii)* $\bar{X} \perp\!\!\!\perp S$.

Lemme de Fisher multivarié

Nous savons déjà que $\bar{X} \sim \mathcal{N}_p(\mu, \frac{1}{n}\Sigma)$. Pour montrer que $\bar{X} \perp\!\!\!\perp S$, nous aurons besoin du lemme suivant.

Lemme: *soient P, Q deux matrices de projection sur $\mathbb{R}^n$ (i.e., deux matrices symétriques et idempotentes de dimension $n \times n$). Supposons que $PQ = 0$. Si $X \sim \mathcal{N}_{n,p}(0, I_{np})$, alors $PX \perp\!\!\!\perp QX$.*

Preuve: Tous les composants de X sont indépendantes. Alors $PX_i \perp\!\!\!\perp QX_j$ si $i \neq j$ (où X_i est la i-ème colonne de X). Dans le cas $i = j$, on obtient que

$$\mathrm{Cov}(PX_i, QX_i) = E[PX_i(QX_i)'] = PE[X_iX_i']Q = PQ = 0.$$

$\square$

Lemme de Fisher multivarié

Proposition: *soient $X_1, \ldots, X_n$ i.i.d. $\mathcal{N}_p(\mu, \Sigma)$. Alors $\bar{X} \perp\!\!\!\perp S$.*

Preuve:

En écrivant, $X_i \stackrel{\mathcal{D}}{=} AZ_i + \mu$ (où $\Sigma = AA'$ et les Z_i sont i.i.d. $\mathcal{N}_p(0, I_p)$), on a $\bar{X} = A\bar{Z} + \mu$ et $S = AS_zA'$, de sorte qu'il suffit de montrer que $\bar{Z} \perp\!\!\!\perp S_z$.

Pour ce faire, posons $P = \frac{1}{n}1_n1_n'$, $Q = I_n - P$ et $Z = (Z_1, \ldots, Z_n)'$. On vérifie alors très facilement (exercice) que

- $Z \sim \mathcal{N}_{n,p}(0, I_{np})$
- $\bar{Z} = \big(PZ\big)'1_n$.
- $S_z = (n-1)^{-1}(QZ)'(QZ)$.

La proposition découle donc du lemme précédent. $\square$

Lemme de Fisher multivarié

Il ne nous reste donc qu'à préciser/établir le point (ii) du lemme de Fisher multivarié:

Proposition (lemme de Fisher): *soient* $X_1, \ldots, X_n$ *i.i.d.* $\mathcal{N}_p(\mu, \Sigma)$. *Alors*

- *(i)* $\bar{X} \sim \mathcal{N}_p(\mu, \frac{1}{n}\Sigma)$.
- *(ii)* $S \sim$?
- *(iii)* $\bar{X} \perp\!\!\!\perp S$.

Que faut-il écrire à la place de “ ? "

Une extension multivariée de la loi χ^2_{n-1}...

Loi de Wishart

Définition: *soit V une matrice aléatoire $p \times p$. Alors*

- *V est de loi de Wishart à m degrés de liberté (notation: $V \sim W_p(m)$) $\Leftrightarrow$ $V \stackrel{\mathcal{D}}{=} \sum_{i=1}^{m} Z_i Z_i'$, où les Z_i sont i.i.d. $\mathcal{N}_p(0, I_p)$.*
- *V est de loi de Wishart de paramètre Σ (une matrice $p \times p$ symétrique et définie positive) à m degrés de liberté (notation: $V \sim W_p(m, \Sigma)$) $\Leftrightarrow$ $V \stackrel{\mathcal{D}}{=} \sum_{i=1}^{m} X_i X_i'$, où les X_i sont i.i.d. $\mathcal{N}_p(0, \Sigma)$.*

Loi de Wishart

Remarques:

- $V \sim W_1(m) \Leftrightarrow V \sim \chi^2_m$ (et $V \sim W_1(m, \sigma^2) \Leftrightarrow V/\sigma^2 \sim \chi^2_m$); la loi de Wishart généralise donc bien la loi χ^2 dans le cas multivarié.
- En écrivant comme d'habitude $X_i \stackrel{\mathcal{D}}{=} AZ_i$ (où $\Sigma = AA'$ et les Z_i sont i.i.d. $\mathcal{N}_p(0, I_p)$), on voit que $V \sim W_p(m, \Sigma) \Leftrightarrow V \stackrel{\mathcal{D}}{=} AV_0A'$, où $V_0 \sim W_p(m)$.

Lemme de Fisher multivarié

Nous pouvons maintenant compléter le lemme de Fisher multivarié:

Proposition (lemme de Fisher): *soient $X_1, \dots, X_n$ i.i.d. $\mathcal{N}_p(\mu, \Sigma)$. Alors*

- *(i) $\bar{X} \sim \mathcal{N}_p(\mu, \frac{1}{n}\Sigma)$.*
- *(ii) $(n-1)S \sim W_p(n-1, \Sigma)$.*
- *(iii) $\bar{X} \perp\!\!\!\perp S$.*

Preuve: il ne reste plus qu'à prouver (ii).

Pour ce faire, écrivons une fois de plus $X_i \stackrel{\mathcal{D}}{=} AZ_i$ (où $\Sigma = AA'$ et les Z_i sont i.i.d. $\mathcal{N}_p(0, I_p)$). Alors

$$(n-1)S = A[(n-1)S_z]A',$$

de sorte qu'il suffit de montrer que $(n-1)S_z \sim W_p(n-1)$.

Lemme de Fisher multivarié

Posons comme plus haut $Q = I_n - \frac{1}{n}1_n 1_n'$. Q admet une décomposition spectrale de la forme $Q = O\Lambda O'$.

En posant $Z = (Z_1, \ldots, Z_n)'$ $(\sim \mathcal{N}_{n,p}(0, I_{np}))$, on a

$$(n-1)S_z = (QZ)'(QZ) = Z'QZ = (O'Z)'\Lambda(O'Z) = Y'\Lambda Y$$

où $Y = (Y_1, \ldots, Y_n)' = O'Z \sim \mathcal{N}_{n,p}(0, I_{np})$ (de sorte que les Y_i sont aussi i.i.d. $\mathcal{N}_p(0, I_p)$).

Nous determinons la décomposition spectrale Q:

1. Pour un matrice de projection les valeurs propres λ_i sont égales à 1 ou 0. [$Q = Q^2$ implique que $\Lambda = \Lambda^2$]
2. $\mathrm{tr}(Q) = \sum_{i=1}^n \lambda_i$. [$\mathrm{tr}(Q) = \mathrm{tr}(O\Lambda O') = \mathrm{tr}(\Lambda O'O)$]
3. $\mathrm{tr}(Q) = n - 1$. [par la définition de Q]

Il découle donc que $\lambda_1 = \ldots = \lambda_{n-1} = 1$ et $\lambda_n = 0$.

Lemme de Fisher multivarié

Alors Q admet une décomposition spectrale de la forme

$$Q = O\Lambda O', \quad \text{où} \quad \Lambda := \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & 0 \end{pmatrix}.$$

Ceci fournit le résultat, puisque

$$(n-1)S_z = Y'\Lambda Y = \sum_{i=1}^{n-1} Y_i Y_i' \sim W_p(n-1). \qquad \square$$

Plan du cours

TCL multivarié

Nous terminons ce chapitre en étendant au cas multivarié le théorème central limite.

Proposition: *soient $X_1, X_2, \ldots$ des p-v.a. i.i.d., avec des moments finis d'ordre 2. Notons $\mu = \mathrm{E}[X_1]$, $\Sigma = \mathrm{Var}[X_1]$ et $\bar{X}^{(n)} = \frac{1}{n}\sum_{i=1}^{n} X_i$. Alors, si $n \to \infty$, $\sqrt{n}\big(\bar{X}^{(n)} - \mu\big) \xrightarrow{\mathcal{L}} \mathcal{N}_p(0, \Sigma)$.*

Preuve:

Fixons $u \in \mathbb{R}^p$ tel que $\|u\| = 1$, et posons $Y_i = u'X_i$ $(i = 1, 2, \ldots)$
Clairement, les Y_i sont i.i.d. et leur loi commune, qui admet des moments finis d'ordre 2, a pour moyenne $\mathrm{E}[Y_1] = u'\mu$ et pour variance $\mathrm{Var}[Y_1] = u'\Sigma u (> 0)$. Le TCL univarié livre donc que

$$u'\left[\sqrt{n}\left(\bar{X}^{(n)} - \mu\right)\right] = \sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} u'X_i - u'\mu\right) \xrightarrow{\mathcal{L}} \mathcal{N}_1(0, u'\Sigma u).$$

Le théorème de Cramér-Wold permet donc de conclure. $\square$