Igor Kortchenski

Cours 7 : Convergence en loi et fonctions caractéristiques

Rappel (Cours 6) Si $X = (X_1, \dots, X_d)$ v.a. dans $\mathbb{R}^d$,

$$\varphi_X(u) = \mathbb{E}\left[e^{i\langle X, u\rangle}\right] = \mathbb{E}\left[\exp\left(i\left(X_1 u_1 + \dots + X_d u_d\right)\right)\right]$$

avec $u = (u_1, \dots, u_d) \in \mathbb{R}^d$.

On a vu que

$$\varphi_X = \varphi_Y \iff X \overset{(d)}{=} Y$$

1) Théorème de Lévy

2) Théorème central limite

3) Vecteurs gaussiens

1) Théorème de Lévy

Théorème (Lévy) Soit X_n, X v.a. dans $\mathbb{R}^d$. Alors $X_n \xrightarrow{(d)} X$ ssi $\forall u \in \mathbb{R}^d$, $\varphi_{X_n}(u) \underset{n\to\infty}{\longrightarrow} \varphi_X(u)$

On fait la preuve pour $d=1$ ($\mathbb{R}$) pour simplifier

Preuve

$\boxed{\Rightarrow}$ Pour $u \in \mathbb{R}$, $x \mapsto e^{iux}$ est continue bornée. Donc par définition de la convergence en loi,

$$\mathbb{E}[e^{iuX_n}] \underset{n\to\infty}{\longrightarrow} \mathbb{E}[e^{iuX}]$$

$\boxed{\Leftarrow}$ On suppose $\forall u \in \mathbb{R}^d$, $\varphi_{X_n}(u) \xrightarrow[n\to\infty]{} \varphi_X(u)$

On utilise un peu la même idée que pour le théorème $\varphi_X = \varphi_Y \Leftrightarrow X \overset{(d)}{=} Y$

idée : ajouter une petite perturbation gaussienne

$$Z_k \overset{(d)}{=} \mathcal{N}\left(0, \frac{1}{k^2}\right) \perp\!\!\!\perp X_n, X$$

On montre : (¥)

$$\boxed{\forall k \geq 1, \; X_n + Z_k \xrightarrow{(d)} X + Z_k}$$

En effet, supposons (¥)

Montrons que $X_n \xrightarrow{(d)} X$.

On utilise porte-manteau :
Soit $F : \mathbb{R} \to \mathbb{R}$ lipschitzienne bornée, $\forall x, y \in \mathbb{R}$, $|F(x) - F(y)| \leq L|x-y|$,
et on montre que $\mathbb{E}[F(X_n)] \xrightarrow[n\to\infty]{} \mathbb{E}[F(X)]$
Soit $k \geq 1$

$$|\mathbb{E}[F(X_n)] - \mathbb{E}[F(X)]|$$
$$\leq \mathbb{E}[|F(X_n) - F(X_n + Z_k)|]$$
$$+ |\mathbb{E}[F(X_n + Z_k)] - \mathbb{E}[F(X + Z_k)]|$$
$$+ \mathbb{E}[|F(X + Z_k) - F(X)|]$$
$$\leq 2L\,\mathbb{E}[|Z_k|] = \frac{2L\,\mathbb{E}[|Z_1|]}{k}$$
$$+ \underbrace{|\mathbb{E}[F(X_n + Z_k)] - \mathbb{E}[F(X + Z_k)]|}_{\xrightarrow[n\to\infty]{} 0}$$

<u>Conclusion</u> : $\limsup_{n\to\infty} |\mathbb{E}[F(X_n)] - \mathbb{E}[F(X)]|$

$\leq \dfrac{2L\, \mathbb{E}[|Z_1|]}{k} \quad \forall k \geq 1.$

On conclut en faisant $k \to \infty$.

Maintenant on montre

$$\boxed{\forall k \geq 1, \quad X_n + Z_k \xrightarrow{(d)} X + Z_k} \ (*)$$

Pour cela, par restriction des fonctions test, il suffit de montrer que $\forall f : \mathbb{R} \to \mathbb{R}$ continue à support compact

$$\mathbb{E}[f(X_n + Z_k)] \xrightarrow[n\to\infty]{} \mathbb{E}[f(X + Z_k)]$$

<u>Rappel</u> :

$\mathbb{E}[g(X_n + Z_k)]$

$$= \int_{\mathbb{R}} dz\, g(z) \left(\frac{k}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{iuz} g_k(u) \varphi_{X_n}(-u) du \right)$$

avec $g_k(u) = \frac{1}{\sqrt{2\pi k}} e^{-\frac{u^2}{2k}}$

<u>Rappel</u> : $\varphi_{X_n}(-u) \xrightarrow[n\to\infty]{} \varphi_X(-u)$

$\forall u \in \mathbb{R}$ par hypothèse.

On va utiliser convergence dominée <u>deux</u> fois.

① $\cdot\ e^{iuz} g_k(u) \varphi_{X_n}(-u) \xrightarrow[n\to\infty]{} e^{iuz} g_k(u) \varphi_X(u)$

$\cdot\ |e^{iuz} g_k(u) \varphi_{X_n}(-u)| \leq g_k(u)$, intégrable.

Donc

$$\frac{R}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{iuz} g_R(u)\, \varphi_{X_n}(-u)\, du$$

$$\xrightarrow[n \to \infty]{} \frac{R}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{iuz} g_R(u)\, \varphi_X(-u)\, du$$

(2)

$$\left| f(z) \frac{R}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{iuz} g_R(u)\, \varphi_{X_n}(-u)\, du \right|$$

$$\leq |f(z)| \frac{R}{\sqrt{2\pi}} \int_{\mathbb{R}} g_R(u)\, du$$

$$= |f(z)|$$

Or f est à support compact, donc $\int_{\mathbb{R}} |f(z)|\, dz < \infty$.

On conclut par convergence

dominée

Application 1 Soit $X^n = (X^n_1, \dots, X^n_d)$ v.a dans $\mathbb{R}$, $X = (X_1, \dots, X_d)$ v.a dans $\mathbb{R}^d$. Alors $X^n \xrightarrow{(d)} X$ ssi $\forall u_1, \dots, u_d \in \mathbb{R}$,

$$u_1 X^n_1 + \dots + u_d X^n_d \xrightarrow{(loi)} u_1 X_1 + \dots + u_d X_d$$

(principe de Cramer-Wold)

Preuve

$\boxed{\Rightarrow}$ Clair par principe de composition : la fonction

$$\mathbb{R}^d \longrightarrow \mathbb{R}$$

$$(x_1, \dots, x_d) \longmapsto u_1 x_1 + \dots + u_d x_d$$

est continue.

$\boxed{\Leftarrow}$ Posons $u = (u_1, \dots, u_d) \in \mathbb{R}^d$

On remarque que

$$\varphi_{X^n}(u) = \mathbb{E}\left[e^{i(u_1 X_1^n + \dots + u_d X_d^n)}\right]$$

$$= \varphi_{u_1 X_1^n + \dots + u_d X_d^n}(1)$$

$$\xrightarrow[n \to \infty]{\text{Lévy}} \varphi_{u_1 X_1 + \dots + u_d X_d}(1)$$

$$= \varphi_X(u).$$

Application 2 On suppose

$X_n \xrightarrow{(d)} X$ dans $\mathbb{R}^d$

$Y_n \xrightarrow{(d)} Y$ dans $\mathbb{R}^{d'}$

et $\boxed{X_n \perp\!\!\!\perp Y_n.}$

Alors (X_n, Y_n) converge en loi vers une v.a. de même loi que (X', Y') avec

$X' \overset{loi}{=} X$

$Y' \overset{loi}{=} Y$ et $X' \perp\!\!\!\perp Y'$.

De manière équivalente si

$\mathbb{P}_{X_n} \longrightarrow \mathbb{P}_X$ étroitement

$\mathbb{P}_{Y_n} \longrightarrow \mathbb{P}_Y$ étroitement

alors

$\mathbb{P}_{X_n} \otimes \mathbb{P}_{Y_n} \longrightarrow \mathbb{P}_X \otimes \mathbb{P}_Y$ étroitement

(⚠ ⚠ en général

$X_n \xrightarrow{(d)} X$, $Y_n \xrightarrow{(d)} Y \not\Rightarrow (X_n, Y_n) \xrightarrow{(d)} (X,Y)$

<u>Preuve</u> On va appliquer le théorème de Lévy. Soit $u, v \in \mathbb{R}$. On écrit

$$\mathbb{E}[e^{iuX_n + ivY_n}]$$

$$= \mathbb{E}[e^{iuX_n}]\,\mathbb{E}[e^{ivY_n}] \text{ par } \perp\!\!\!\perp$$

$$\xrightarrow[n\to\infty]{} \mathbb{E}[e^{iuX}]\,\mathbb{E}[e^{ivY}]$$

par hypothèse et théorème de Lévy

Ainsi

$$\varphi_{(X_n, Y_n)}(u,v)$$

$$\xrightarrow[n\to\infty]{} \varphi_X(u)\,\varphi_Y(v)$$

$$= \varphi_{(X', Y')}(u,v)$$

Donc $(X_n, Y_n) \xrightarrow{(d)} (X', Y')$ par théorème de Lévy

2) Théorème central limite

Motivations : soit $(X_n)_{n\geq 1}$ des v.a. iid réelles dans L^1.

LGN : $\frac{1}{n} \sum_{k=1}^{n} X_k \xrightarrow[n\to\infty]{p.s} \mathbb{E}[X_1]$

On cherche à savoir à quelle vitesse cette convergence a lieu, c'est-à-dire quel est l'ordre de grandeur de

$$\frac{1}{n} \sum_{k=1}^{n} X_k - \mathbb{E}[X_1]$$

quand n est grand.

Si on suppose de plus X_1 dans L^2 on devine la réponse en calculant

$$\mathbb{E}[(X_1+\dots+X_n - n\mathbb{E}[X_1])^2]$$
$$= \text{Var}\, X_1 + \dots + \text{Var}\, X_n$$
$$= n \,\text{Var}\, X_1 \quad \text{par} \perp\!\!\!\perp.$$

Ceci suggère que

$X_1+\dots+X_n - n\mathbb{E}[X_1]$ est d'ordre $\sqrt{n}$ et donc

$\frac{X_1+\dots+X_n}{n} - \mathbb{E}[X_1]$ est d'ordre $\frac{1}{\sqrt{n}}$.

Théorème (Théorème central limite TCL ; CLT en anglais)

Soit $(X_n)_{n\geq 1}$ des v.a

- iid
- dans L^2

Soit $\sigma^2 = \mathrm{Var}(X_1) = \mathbb{E}[(X_1 - \mathbb{E}[X_1])^2]$

On suppose $\sigma^2 > 0$. Alors

$$\frac{1}{\sqrt{n}}(X_1 + \dots + X_n - n\mathbb{E}[X_1]) \xrightarrow[n\to\infty]{(d)} N(0, \sigma^2).$$

De manière équivalente,

$\forall a < b$:

$$\mathbb{P}\left(a < \frac{X_1 + \dots + X_n - n\mathbb{E}[X_1]}{\sigma\sqrt{n}} < b\right) \xrightarrow[n\to\infty]{} \int_a^b \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\,dx.$$

<u>Preuve</u> : la deuxième partie est une conséquence de la première (convergence en loi et fonctions de répartition)

- Pour la première partie, quitte à considérer

$$X_i - \mathbb{E}[X_i]$$

on peut supposer $\mathbb{E}[X_1] = 0$

Posons alors $Z_n = \dfrac{X_1 + \cdots + X_n}{\sigma\sqrt{n}}$

<u>But</u> $Z_n \xrightarrow{(d)} N(0,1)$

Alors

$$\varphi_{Z_n}(u) = \mathbb{E}\left[e^{iu\left(\frac{X_1+\cdots+X_n}{\sigma\sqrt{n}}\right)}\right]$$

$$= \mathbb{E}\left[e^{i\frac{u}{\sigma\sqrt{n}}X_1}\right]^n \quad \text{car iid}$$

$$= \varphi_{X_1}\left(\frac{u}{\sigma\sqrt{n}}\right)$$

Mais $X_1 \in L^2$ et on a vu que

$$\phi_{X_1}(u) = 1 + iu\,\mathbb{E}[X_1] - \frac{1}{2}u^2\,\mathbb{E}[X_1^2] + o(u^2)$$

$$= 1 - \frac{1}{2}u^2\sigma^2 + o(u^2) \qquad (u \to 0)$$

Donc

$$\varphi_{X_1}\left(\frac{u}{\sigma\sqrt{n}}\right) = 1 - \frac{u^2}{2n} + o\left(\frac{1}{n}\right)$$

Ainsi

$$\varphi_{Z_n}(u) = \left(1 - \frac{u^2}{2n} + o\left(\frac{1}{n}\right)\right)^n$$

$$\xrightarrow[n\to\infty]{} e^{-u^2/2} = \varphi_{N(0,1)}(u)$$

Ainsi $Z_n \xrightarrow{(d)} \mathcal{N}(0,1)$ par le théorème de Lévy

~

<u>But</u> : généraliser le TCL à $\mathbb{R}^d$, $d > 1$.

3) Vecteurs gaussiens

Définition Une v.a. $X = (X_1, \dots, X_d)$ à valeurs dans $\mathbb{R}^d$ est un vecteur gaussien si toutes ses combinaisons linéaires $\lambda_1 X_1 + \dots + \lambda_d X_d$ avec $\lambda_1, \dots, \lambda_d \in \mathbb{R}$ sont des v.a. (réelles) gaussiennes

Par convention, une v.a. constante m est une $\mathcal{N}(m, 0)$

Rappel Si $X \sim \mathcal{N}(m, \sigma^2)$, on a $\mathbb{E}[e^{iuX}] = \exp(imu - \frac{\sigma^2}{2}u^2)$

Exemple : Si $X_1, \dots, X_d$ sont des v.a. gaussiennes $\perp\!\!\!\perp$, $(X_1, \dots, X_d)$ est un vecteur gaussien

⚠ Faux sans $\perp\!\!\!\perp$:

Si $X \sim \mathcal{N}(0,1)$

$\mathbb{P}(\varepsilon = 1) = \mathbb{P}(\varepsilon = -1) = \frac{1}{2}$

avec $X \perp\!\!\!\perp \varepsilon$, alors

$X, \varepsilon X$ sont des $\mathcal{N}(0,1)$ mais $(X, \varepsilon X)$ n'est pas un vecteur gaussien car $\mathbb{P}(X + \varepsilon X = 0) = \frac{1}{2}$

Remarque $(X_1, \dots, X_d)$ vecteur gaussien $\Rightarrow$ $X_1, \dots, X_d$ sont gaussiennes (réciproque fausse)

Définition Soit $X = (X_1, \dots, X_d)$ un vecteur gaussien. Le vecteur
$$m_X = (\mathbb{E}[X_1], \dots, \mathbb{E}[X_d])$$
s'appelle la moyenne de X

On dit que X est centré si $m_X = (0, 0, \dots, 0)$

Clairement, $X - m_X$ est un vecteur gaussien centré et si X, Y sont deux vecteurs gaussiens $\perp\!\!\!\perp$, $X+Y$ est un vecteur gaussien

$$m_{X+Y} = m_X + m_Y$$

- La matrice $d \times d$

$$K_X = \left(\mathbb{E}[X_i X_j] - \mathbb{E}[X_i]\mathbb{E}[X_j]\right)_{1 \le i,j \le d}$$

s'appelle la covariance de X.

On a $K_{X+Y} = K_X + K_Y$

<u>Rappel</u> On note $\langle \cdot , \cdot \rangle$ le produit scalaire sur $\mathbb{R}^d$.

<u>Proposition</u> Soit X un vecteur gaussien dans $\mathbb{R}^d$ et $\lambda \in \mathbb{R}^d$. Alors $\langle \lambda, X \rangle$ $(= \lambda_1 X_1 + \dots + \lambda_d X_d)$ suit une loi $\mathcal{N}(m_\lambda, \sigma_\lambda^2)$ avec

- $m_\lambda = \langle \lambda, m_X \rangle$
- $\sigma_\lambda^2 = \langle \lambda, K_X \lambda \rangle$

En particulier, K_X est une matrice symétrique positive

<u>Preuve</u> :

- $m_\lambda = \langle \lambda, X \rangle$ provient de la linéarité de l'espérance

• Ensuite

$$\mathbb{E}[\langle \lambda | X \rangle^2] = \sum_{i,j} \lambda_i \lambda_j \mathbb{E}[X_i X_j]$$

et

$$\langle \lambda | m \rangle^2 = \sum_{i,j} \lambda_i \lambda_j \mathbb{E}[X_i]\mathbb{E}[X_j]$$

$$\parallel$$

$$\mathbb{E}[\langle \lambda | X \rangle]^2$$

Donc $\mathrm{Var}(\langle \lambda, X \rangle)$

$$= \sum_{i,j} \lambda_i \lambda_j (K_X)_{i,j}$$

$$= \langle \lambda, K_X \lambda \rangle$$

Corollaire La fonction caractéristique d'un vecteur gaussien X est donnée par

$$\boxed{\underbrace{\varphi_X(\lambda)}_{\mathbb{E}[e^{i\langle\lambda,X\rangle}]} = \exp\left(i\langle\lambda, m_X\lambda\rangle - \frac{1}{2}\langle\lambda, K_X\lambda\rangle\right)}$$

En particulier la loi d'un vecteur gaussien est complètement déterminée par sa moyenne et sa covariance

Preuve

$$\mathbb{E}[e^{iu\langle X,\lambda\rangle}]$$

$$= \varphi_{\langle X,\lambda\rangle}(u)$$

Mais $\langle X,\lambda\rangle \overset{(\text{loi})}{=} \mathcal{N}(\langle\lambda,m_X\rangle,\langle\lambda,K_X\lambda\rangle)$

Donc

$\varphi_{\langle X,\lambda\rangle}(u)$ $\quad (\ast)$

$$= \exp\left(iu\langle\lambda,m_X\rangle-\frac{u^2}{2}\langle\lambda,K_X\lambda\rangle\right)$$

(fonction caractéristique d'une gaussienne)

Mais $\varphi_X(\lambda) = \mathbb{E}[e^{i\langle X,\lambda\rangle}]$.

On conclut en prenant $u=1$ dans $(\ast)$

Proposition Soit $m \in \mathbb{R}^d$ et K une matrice $d \times d$ symétrique positive. Alors il existe un vecteur gaussien de $\mathbb{R}^d$ de moyenne m et de matrice de covariance K.

Preuve : **Idée** : écrire $K = C^2$ avec C symétrique

En effet, on diagonalise K en base orthonormée :

$$K = P D P^{-1}$$ avec P orthogonale ($P P^T = P^T P = I$)

et $D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_d \end{pmatrix}$ avec $\lambda_1, \dots, \lambda_d \geq 0$

On définit $C = P\sqrt{D}\,P^{-1}$.

Soit maintenant

$N = (N_1, \ldots, N_d)$ avec

$N_1, \ldots, N_d$ des $\mathcal{N}(0,1)$ iid

On pose

$$X = C \cdot N + m$$

On vérifie que X convient

- Pour $\lambda \in \mathbb{R}^d$,

$\langle \lambda | X \rangle = \langle \lambda | CN \rangle + \langle \lambda | m \rangle$

$(\star) = \langle C\lambda | N \rangle + \langle \lambda | m \rangle$

car C symétrique

Comme N est un vecteur gaussien, $\langle \lambda | X \rangle$ est gaussienne, donc X est un

vecteur gaussien.
Il reste à calculer la moyenne et la variance de $\langle\lambda|X\rangle$

Par $(\forall)$ sa moyenne vaut clairement $\langle\lambda|m\rangle$

La variance de $\langle\lambda|X\rangle$ vaut

$$\mathbb{E}[\langle C\lambda|N\rangle^2]$$

$$= \mathbb{E}\Big[\sum_{i,j,k,l} C_{ij}\lambda_i N_j C_{kl}\lambda_k N_l\Big]$$

$$= \sum_{i,j,k} C_{ij}\lambda_i C_{kj}\lambda_k$$

car $\mathbb{E}[N_j N_l] = 1$ si $j=l$
$= 0$ sinon

$$= \langle C\lambda|C\lambda\rangle = \langle\lambda|K\lambda\rangle$$

Notation

$\mathcal{N}(m, K)$: v.a. dont la loi est celle d'un vecteur gaussien de moyenne m et de matrice de covariance K.

Corollaire $\mathcal{N}(m, K)$ admet une densité sur $\mathbb{R}^d$ ssi $\det(K) > 0$, et alors sa densité est

$$\frac{1}{(2\pi)^{d/2}} \frac{1}{\sqrt{\det K}} e^{-\frac{1}{2}\langle x-m, K^{-1}(x-m)\rangle} dx$$

idée de la preuve : principe

de la fonction muette avec
$X = CN + m$.

Théorème (caractérisation de l'$\perp\!\!\!\perp$ pour des vecteurs gaussiens)

① Soit $(X_1, \dots, X_d)$ un vecteur gaussien de $\mathbb{R}^d$. Alors
$X_1, \dots, X_d$ sont $\perp\!\!\!\perp$ ssi
K_X est diagonale ($\mathbb{E}[X_i X_j]$
$= \mathbb{E}[X_i]\,\mathbb{E}[X_j]$ pour $i \neq j$)

② Soit $Z = (X_1, \dots, X_p, Y_1, \dots, Y_q)$ un vecteur gaussien de
$\mathbb{R}^{p+q}$

Alors $(X_1, \dots, X_p) \perp\!\!\!\perp (Y_1, \dots, Y_q)$ ssi $\mathrm{Cov}(X_i, Y_j) = 0$ pour tout $1 \leq i \leq p$, $1 \leq j \leq q$

$Z = (X_1, \dots, X_p, Y_1, \dots, Y_q)$ un vecteur gaussien de $\mathbb{R}^{p+q}$

<u>Preuve</u> : Sans perte de généralité on peut supposer les vecteurs gaussiens centrés.

(1) $\boxed{\Rightarrow}$ si $X \perp\!\!\!\perp Y$ on sait que $\mathbb{E}[XY] = \mathbb{E}[X]\,\mathbb{E}[Y]$

$\boxed{\Leftarrow}$ Soit $u = (u_1, \dots, u_d) \in \mathbb{R}^d$

En notant $K_X = \begin{pmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_d^2 \end{pmatrix}$ on a

$\mathbb{E}[\exp(i \langle u | X \rangle)]$

$= \exp\left(-\frac{1}{2}\langle u | K_X u\rangle\right)$ (proposition)

$$= \prod_{i=1}^{d} \exp\left(-u_i^2 \cdot \frac{\sigma_i^2}{2}\right)$$

Donc $(X_1, \dots, X_d)$ sont $\perp\!\!\!\perp$ et $X_i \overset{d}{=} N(0, \sigma_i^2)$.

(2) Preuve similaire

$\boxed{\Rightarrow}$ Pour $1 \leq i \leq p$, $1 \leq j \leq q$

$X_i \perp\!\!\!\perp Y_j \Rightarrow \mathbb{E}[X_i Y_j] = \mathbb{E}[X_i]\,\mathbb{E}[Y_j]$

$\boxed{\Leftarrow}$ La matrice de covariance de Z s'écrit :

$$K_Z = \begin{pmatrix} K_X & 0 \\ 0 & K_Y \end{pmatrix}$$

Donc pour

$\xi = (u_1, \dots, u_p, v_1, \dots, v_q) \in \mathbb{R}^{p+q}$

on a

$$\mathbb{E}\left[\exp\left(i \langle \xi \mid Z \rangle\right)\right]$$

$$= \exp\left(-\frac{1}{2} \langle \xi \mid K_Z \xi \rangle\right)$$

$$= \exp\left(-\frac{1}{2} \langle u \mid K_X u \rangle\right) \cdot \exp\left(-\frac{1}{2} \langle v \mid K_Y v \rangle\right)$$

avec $u = (u_1, \dots, u_p)$

$v = (v_1, \dots, v_q)$

Donc

$$\mathbb{E}\left[e^{i(\langle u \mid X \rangle + \langle v \mid Y \rangle)}\right]$$

$$= \mathbb{E}\left[e^{i\langle u \mid X \rangle}\right] \mathbb{E}\left[e^{i \langle v \mid Y \rangle}\right]$$

Donc $X \perp\!\!\!\perp Y$

⚠ Dans ② il est important que $(X_1, \dots, X_p, Y_1, \dots, Y_q)$ soit un vecteur gaussien. « $(X_1, \dots, X_p)$ gaussien et $(Y_1, \dots, Y_q)$ gaussien » ne suffit pas.

3) Théorème central limite vectoriel

Théorème Soit $(X_i)_{i \geq 1}$ des v.a dans $\mathbb{R}^d$

- iid
- $\mathbb{E}[|X_1|^2] < \infty$

(Soit $(X_i)_{i\geq 1}$ des v.a dans $\mathbb{R}^d$ • iid • $\mathbb{E}[|X_1|^2]<\infty$)

Alors (ici $|\cdot|$ est une norme sur $\mathbb{R}^d$)

$$\frac{1}{\sqrt{n}}\left(X_1+\dots+X_n - n\,\mathbb{E}[X_1]\right) \xrightarrow[n\to\infty]{(d)} \mathcal{N}(0, K_{X_1})$$

où $K_{X_1} = \left(\mathbb{E}[(X_1)_i (X_1)_j] - \mathbb{E}[(X_1)_i]\,\mathbb{E}[(X_1)_j]\right)_{1\leq i,j\leq d}$

est la matrice de covariance de X_1

Ceci explique l'importance des vecteurs gaussiens en probabilités et en statistiques.

Preuve C'est essentiellement la même que pour $d=1$:

- On peut supposer $\mathbb{E}[X_1]=0$
- Pour $u \in \mathbb{R}^d$,

$$\mathbb{E}\left[\exp\left(i\left\langle u \,\middle|\, \frac{X_1+\dots+X_n}{\sqrt{n}}\right\rangle\right)\right]$$

$$= \mathbb{E}\left[\exp\left(i\left\langle \frac{u}{\sqrt{n}}, X_1\right\rangle\right)\right]^n \qquad \text{(iid)}$$

$$= \varphi_{X_1}\left(\frac{u}{\sqrt{n}}\right)^n$$

D'autre part, comme $X_1 \in L^2$ on a la formule de Taylor

$$\Phi_{X_1}\left(\frac{u}{\sqrt{n}}\right) = 1 - \frac{1}{2n} \langle u \mid K_{X_1} u \rangle + o\left(\frac{1}{n}\right)$$

Ainsi

$$\varphi_{\frac{X_1 + \dots + X_n}{\sqrt{n}}}(u)$$

$$= \left(1 - \frac{1}{2n} \langle u \mid K_{X_1} u \rangle + o\left(\frac{1}{n}\right)\right)^n$$

$$\xrightarrow[n \to \infty]{} \exp\left(-\frac{1}{2} \langle u \mid K_{X_1} u \rangle\right)$$

qui est la fonction caractéristique d'une v.a. $N(0, K_{X_1})$. On conclut par théorème de Lévy

Remarque

Pourquoi peut-on supposer $\mathbb{E}[X_i]=0$?

Réponse : Si le TCL est vrai pour $(Y_i)_{i\geq 1}$ il sera aussi vrai pour $(Y'_i)_{i\geq 1}$ avec $Y'_i = Y_i + c$

car

$$\frac{Y_1 + \cdots + Y_n - n\,\mathbb{E}[Y_1]}{\sqrt{n}}$$

$$= \frac{Y'_1 + \cdots + Y'_n - n\,\mathbb{E}[Y'_1]}{\sqrt{n}}$$