cours 7

Central limite $(X_n)_{n\geq 1}$ v.a.r. indépendantes de même loi que X v.a.t.
$\sigma^2 = var(X)$ t.q. $\mathbb{E}[|X|^2] < \infty$

$$\frac{S_n - n\mathbb{E}[X]}{\sqrt{n}} \xrightarrow{loi} \mathcal{N}(0,\sigma^2)$$

cas particulier : Théorème de de Moivre-Laplace

$(X_n)_{n\geq 1}$ v.a.i de loi de Bernoulli de paramètre $\frac{1}{2}$

i.e. $P(X_n=0) = P(X_n=1) = \frac{1}{2}$

(S_n suit une loi binomiale $B(n,\frac{1}{2})$)

$$P(S_n=k) = \binom{n}{k}\frac{1}{2^n}$$

$X = X_1$

$$\sigma^2(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = 0\,\frac{1}{2} + 1\cdot\frac{1}{2} - \left(\frac{1}{2}\right)^2 = \frac{1}{4}$$

Théorème C.-L.

$$\Rightarrow \quad \frac{S_n - n\mathbb{E}[X]}{\sqrt{n}} \xrightarrow{loi} \mathcal{N}\left(0,\frac{1}{4}\right)$$

Remarque : $\dfrac{S_n - n\mathbb{E}[X]}{\sigma\sqrt{n}} \xrightarrow{loi} \mathcal{N}(0,\sigma^2)$ 1

$$var(\alpha X) = \alpha^2 var(X)$$

Cas particulier : $(X_n)_{n \geq 1}$ v.a.i. de même loi que ~~ce~~ de loi binômiale de paramètre $p \in]0,1[$

$$\sigma^2 = p(1-p)$$

$$\text{T.C.L} \Rightarrow \frac{S_n - n\mathbb{E}[X]}{\sqrt{p(1-p)}\sqrt{n}} \xrightarrow{\text{loi}} \mathcal{N}(0,1)$$

Théorème central limite vectoriel

rappel : loi gaussienne de moyenne m et variance $\sigma^2 > 0$ est de densité

$$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}$$

Définition : Une v.a. $X = (X_1, \dots, X_n)$ à valeurs dans $\mathbb{R}^d$ est un <u>vecteur gaussien</u> si toute combinaison linéaire de ses coordonnées est une v.a. gaussienne.

- $\lambda_1 X_1 + \dots + \lambda_n X_n$ $\quad \lambda_i \in \mathbb{R}$ $\quad 1 \leq i \leq n$ est une v.a gaussienne

remarque : Si $(X_i)_{1 \leq i \leq n}$ est une famille de v.a. gaussiennes indépendantes

alors $X = (X \dots X)$ est ...

alors $X = (X_1, \ldots, X_n)$ est un vecteur gaussien (somme de v.a. gaussienne ind. est v.a. gaussienne)

• X est vecteur gaussien alors chaque composante est gaussienne.

Exemple : N v.a. gaussienne

ε v.a. de loi de Bernouilli

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

indépendantes

• εN est aussi gaussienne.

car $P(\varepsilon N < t) = P(\varepsilon = 1) P(N < t) + P(\varepsilon = -1) P(N > -t)$

$\Big(\{\varepsilon N < t\} = \{1 . N < t\} \cup \{(-1) N < t\}$

$$= \frac{1}{2} P(N < t) + \frac{1}{2} P(N > -t)$$

$$= P(N < t) \quad \left(\text{car } \underset{\displaystyle P(N > -t)}{\overset{\displaystyle P(N < t)}{\shortparallel}} \right)$$

• Mais $(N, \varepsilon N)$ n'est pas un vecteur gaussien :

$N + \varepsilon N$ n'est pas gaussien :

$$P(N + \varepsilon N = 0) = P((1+\varepsilon) N = 0)$$

$$= P(1 + \varepsilon = 0) = \frac{1}{2}$$

Définition : Soit $X = (X_1 \dots X_n)$ un vecteur gaussien dans $\mathbb{R}^d$. On définit

- $m = (m^1, \dots, m^d)$ $\quad m^i = \mathbb{E}[X_i]$
  la moyenne
- $D = (D_{ij})$ $\quad D_{ij} = \mathbb{E}[X_i X_j] - \mathbb{E}[X_i]\mathbb{E}[X_j]$
  la covariance (ou matrice de dispersion)

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

Exemple : si $X = (X_1, \dots, X_n)$ et $Y = (Y_1, \dots Y_n)$ sont deux vecteurs gaussien indépendants alors $X + Y$ est un v.g. de moyenne $m_X + m_Y$ et covariance $D_X + D_Y$

Proposition : X vecteur gaussien
$\forall \lambda \in \mathbb{R}^d$ la v.a. $\langle \lambda, X \rangle$
a pour moyenne
$$m_\lambda = \langle \lambda, m \rangle \text{ et}$$
variance

$$\sigma_\lambda^2 = \langle \lambda, D\lambda \rangle$$ $\hookrightarrow$ covariance de X :

$\Big($ comme $\sigma_\lambda^2 \geq 0 \implies \underset{\text{notation}}{D \geq 0}\Big)$

$\langle \lambda, D\lambda \rangle \geq 0 \quad \forall \lambda \in \mathbb{R}^d$

preuve : • $m_\lambda = \mathbb{E}[\langle \lambda, X \rangle]$

$$= \sum_{i=1}^d \mathbb{E}[\lambda_i X_i]$$

$$= \sum_i \lambda_i \underbrace{\mathbb{E}[X_i]}_{m_i} = \sum \lambda_i m_i = \langle \lambda, m \rangle$$

• $\mathbb{E}[\langle \lambda, X \rangle^2] - \mathbb{E}[\langle \lambda, X \rangle]^2$

$$= \sum_{i,j} \mathbb{E}[\lambda_i \lambda_j X_i X_j] - \sum_{i,j} \mathbb{E}[\lambda_i X_i]\mathbb{E}[\lambda_j X_j]$$

$$= \sum_{i,j} \lambda_i \lambda_j \mathbb{E}[X_i X_j] - \sum_{i,j} \lambda_i \lambda_j \mathbb{E}[X_i]\mathbb{E}[X_j]$$

$$= \sum_{i,j} \lambda_i \lambda_j \big(\mathbb{E}[X_i X_j] - \mathbb{E}[X_i]\mathbb{E}[X_j]\big)$$

$$= \langle \lambda, D\lambda \rangle$$

∎

Proposition : La fonction caractéristique d'un vecteur gaussien X est

$\lambda \in \mathbb{R}^d \qquad \Phi_X(\lambda) = e^{i\langle \lambda, m \rangle - \frac{\langle \lambda, D\lambda \rangle}{2}}$

Remarque : Un vecteur gaussien est déterminé par sa moyenne et sa covariance.

preuve : On sait que

$\langle \lambda, X \rangle$ est v.a. gaussienne

avec $m_\lambda = \langle \lambda, m \rangle$ et $\sigma_\lambda^2 = \langle \lambda, D\lambda \rangle$

$$\Phi_{\langle \lambda, X \rangle}(t) = \mathbb{E}\left[e^{it\langle \lambda, X \rangle}\right] = e^{itm_\lambda - \frac{t^2\sigma_\lambda^2}{2}}$$

on fait $t=1$:

$$\underbrace{\mathbb{E}\left[e^{i\langle \lambda, X \rangle}\right]}_{\substack{\text{définition} \\ \Phi_X(\lambda)}} = e^{im_\lambda - \frac{\sigma_\lambda^2}{2}}$$

$\boxtimes$

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

preuve : i) Soit $N_1, \dots, N_d$ v.a. gaussienne normal standard indépendantes

$m_i = 0$, $\sigma_i^2 = 1$

$N = (N_1, \dots, N_d)$ vecteur gaussien

ii) Soit C matrice symétrique

positive t.q. $\boxed{C^2 = D}$

On diagonalise $D = P \begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{pmatrix} P^T$

rappel $d_i \geqslant 0$ car $D \geqslant 0$: $C = P \begin{pmatrix} \sqrt{d_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{d_n} \end{pmatrix} P^T$

iii) On définit $X = CN + m$

Il faut vérifier que X est vecteur gaussien de moyenne m et covariance D.

- $\langle \lambda, X \rangle = \langle \lambda, CN \rangle + \langle \lambda, m \rangle$

  est une somme de v.a. gauss. indépen.

  $\Rightarrow \langle \lambda, X \rangle$ est gaussien

  $\Rightarrow X$ est gaussien

- $m_\lambda = \langle \lambda, m \rangle$ évident

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

$$= \mathbb{E}[\langle \lambda, CN + m \rangle^2] - \mathbb{E}^2[\langle \lambda, CN \rangle + \langle \lambda, m \rangle]$$

$$= \mathbb{E}[(\langle \lambda, CN \rangle + \langle \lambda, m \rangle)^2] - \Big(\mathbb{E}[\langle \lambda, CN \rangle] + \underbrace{\mathbb{E}[\langle \lambda, m \rangle]}_{= \langle \lambda, m \rangle}\Big)^2$$

$$= \mathbb{E}[\langle \lambda, CN \rangle^2] + 2\langle \lambda, m \rangle \cancel{\mathbb{E}[\langle \lambda, CN \rangle]} + \cancel{\langle \lambda, m \rangle^2}$$

$$- \Big(\mathbb{E}[\cancel{\langle \lambda, CN \rangle}]^2 + 2 \langle \lambda, m \rangle \cancel{\mathbb{E}[\langle \lambda, CN \rangle]} + \cancel{\langle \lambda, m \rangle^2} \Big)$$

~~0 car N_i est standard normal~~ $(\mathbb{E}[N_i]=0)$

$$= \mathbb{E}[\langle \lambda, CN\rangle^2]$$

symétrique

$$= \mathbb{E}\Big[\big(\sum_{i,j} \lambda_i C_{ij} N_j\big)^2\Big]$$

$$= \mathbb{E}\Big[\big(\sum_{i,j} \lambda_j C_{ij} N_i\big)^2\Big]$$

$$= \mathbb{E}\Big[\big(\sum_{i,j} \lambda_j C_{ij} \underline{N_i}\big)\big(\sum_{k,l} \lambda_l C_{kl} \underline{N_k}\big)\Big]$$

$$= \mathbb{E}\Big[\sum_{i,j,l} C_{ij}\lambda_j \underline{N_i}\, C_{il}\lambda_l \underline{N_i}\Big]$$ par indépendance

$\big(\mathbb{E}[N_i N_j] = \mathbb{E}[N_i]\mathbb{E}[N_j]\big)$ pour $i \neq j$, $\mathbb{E}[N_i] = 0$, $\mathbb{E}[N_j] = 0$

$$= \sum_{i,j} C_{ij}\lambda_j C_{il}\lambda_l$$ $\left(\sigma_{N_i}^2 = 1\right) \Longrightarrow \mathbb{E}[N_i^2] = 1$

symétrique

$$= \sum_{i,j} \lambda_j C_{ji} C_{il} \lambda_l$$

$$= \langle \lambda, C^2\lambda\rangle = \langle \lambda, D\lambda\rangle$$

∎

Corollaire : $m \in \mathbb{R}^d$, $D \geq 0$ symétrique $\det(D) > 0$. Alors la loi gaussienne $\mathcal{N}(m, D)$ (obtenu dans la proposition) est de densité

$$-\tfrac{1}{2}\langle x-m, D^{-1}(x-m)\rangle$$

$$\frac{1}{(2\pi)^{\frac{d}{2}}} \frac{1}{(\det D)^{\frac{1}{2}}} e$$

preuve : $\mathbb{E}[f(X)] \overset{?}{=} \int f(x_1, \dots, x_n) p(x_1, \dots x_n) dx_1 \dots dx_n$

Par la proposition précédente :

$$X = CN + m \quad \text{où} \quad C^2 = D$$

N est de densité $f_N(z) = \prod_{i=1}^{d} \frac{1}{\sqrt{2\pi}} e^{-\frac{z_i^2}{2}}$

$z = (z_1, \dots, z_d)$

$$= \boxed{\frac{1}{(2\pi)^{d/2}} e^{-\frac{1}{2}\|z\|^2}}$$

$$\mathbb{E}[f(X)] = \mathbb{E}[f(CN+m)] \quad (\text{transfer})$$

$$= \int_{\mathbb{R}^d} f(Cz+m) f_N(z)\, dz$$

changement de variable :

$$x = Cz + m \qquad \text{Jacobien} : \det C^{-1} > 0$$

$$z = C^{-1}(x-m)$$

$$\Rightarrow \mathbb{E}[f(X)] = \int_{\mathbb{R}^d} f(x) \boxed{f_N(C^{-1}(x-m))} \det C^{-1}\, dx$$

$$= \frac{1}{(2\pi)^{d/2}} \frac{1}{\sqrt{\det D}} \int_{\mathbb{R}^d} f(x)\, e^{-\frac{1}{2}\|C^{-1}(x-m)\|^2} dx$$

On obtient que

$$\|C^{-1}(x-m)\|^2 = \langle C^{-1}(x-m), C^{-1}(x-m) \rangle$$

$$= \langle x-m, C^{-2}(x-m)\rangle$$

$$= \langle x-m, D^{-1}(x-m)\rangle$$

densité $\boxed{\dfrac{1}{(2\pi)^{d/2}} \dfrac{1}{\sqrt{\det D}} e^{-\frac{1}{2}\langle x-m, D^{-1}(x-m)\rangle}}$

$\square$

Proposition Soit $X = (X_1, \dots, X_n)$ un vecteur gaussien. Alors

$(X_i)_{1\le i\le n}$ est une famille indépendante

ssi

la covariance est diagonale.

preuve : $\Rightarrow$ $(X_i)_{1\le i\le n}$ est indépendante

alors $D_{ij} = 0$ si $i \neq j$ (par définition)

$$D_{ij} = \mathbb{E}[X_i X_j] - \mathbb{E}[X_i]\mathbb{E}[X_j] = 0$$

$\Leftarrow$ si D est diagonale

$$\begin{pmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_n^2 \end{pmatrix}$$

Alors $\forall \lambda \in \mathbb{R}^n$

$$e^{i\langle \lambda, m\rangle - \frac{1}{2}\langle \lambda, D\lambda\rangle}$$

$$\Phi_X(\lambda) = e$$
$$= e^{\sum i\lambda_k m_k - \frac{1}{2}\sum \lambda_k^2 \sigma_k^2}$$
$$= \prod_{k=1}^{n} \Phi_{X_k}(\lambda_k)$$

$\Rightarrow$ critère d'indépendance $(X_i)_{1\leq i\leq n}$ est une famille indépendante. ■

Théorème central limite multidimensionnelle.

$\boxed{X = (X^{(1)}, \dots, X^{(d)})}$ v.a. à valeurs dans $\mathbb{R}^d$

$\mathbb{E}[\|X\|^2] < \infty$ covariance$(X) = D$

Soit $(X_i)_{i\geq 1}$ suite de vecteurs aléatoires indépendants de même loi que X.

Théorème : C.L.M

$$\frac{S_n - n.m}{\sqrt{n}} \xrightarrow[n\to\infty]{(loi)} \mathcal{N}(0, D)$$

preuve : On suppose $m = 0$ (sans perte de généralité)

$$D_{ij} = \mathbb{E}[X^{(i)} X^{(j)}]$$

On utilise les fonctions caractéristiques :

$$\Phi_X(\lambda) = \int_{\mathbb{R}^d} e^{i\langle\lambda, x\rangle} P_X(dx)$$

$$\frac{\partial^2 \Phi_X(\lambda)}{\partial\lambda^{(i)}\partial\lambda^{(j)}} = -\int_{\mathbb{R}^d} x^{(i)} x^{(j)} e^{i\langle\lambda, x\rangle} P_X(dx)$$

(Φ_X est de classe C^2)

$$\Rightarrow \frac{\partial^2 \Phi_X(0)}{\partial\lambda^{(i)}\partial\lambda^{(j)}} = -\int_{\mathbb{R}^d} x^{(i)} x^{(j)} P_X(dx) = -D_{ij}$$

La formule de Taylor :

$$\Phi_X(\lambda) = 1 - \frac{1}{2}\langle\lambda, D\lambda\rangle + o(|\lambda|^2)$$

$$\Phi_{S_n}(\lambda) = (\Phi_X(\lambda))^n \quad (\text{indépendance})$$

$$\Rightarrow \Phi_{\frac{S_n}{\sqrt{n}}}(\lambda) = \left(\Phi_X\left(\frac{\lambda}{\sqrt{n}}\right)\right)^n$$

$$= \left(1 - \frac{1}{2n}\langle\lambda, D\lambda\rangle + o\left(\frac{|\lambda|^2}{n}\right)\right)^n$$

$$\Rightarrow \Phi_{\frac{S_n}{n}}(\lambda) \longrightarrow e^{-\frac{1}{2}\langle\lambda, D\lambda\rangle} \quad \forall \lambda \in \mathbb{R}^d$$

Théorème de Levy on obtient la convergence en loi

∎

Chapitre 11 Conditionnement

11.1 Conditionnement discret

$(\Omega, \mathcal{A}, P)$ un espace de probabilité
Soit $B \in \mathcal{A}$ t.q. $P(B) > 0$

Définition : Probabilité sur $(\Omega, \mathcal{A})$ sachant B

$$P(A|B) = P(A \cap B) / P(B)$$

Définit : Pour toute v.a.

$$X \geq 0 \quad \text{ou} \quad X \in L^1(\Omega, \mathcal{A}, P)$$

On définit l'espérance conditionnelle de X sachant B par

$$\boxed{\mathbb{E}[X|B] = \underline{\mathbb{E}[X \cdot \mathbb{1}_B]} / P(B)}$$

i.e. $\mathbb{E}[X|B]$ est la moyenne de X quand B est réalisé.

Remarque : $P(\cdot|B)$ est une nouvelle mesure sur Ω

Proposition : si $\mathbb{E}[X] < \infty$ alors X

et $P(\cdot|B)$ ...

est $P(\cdot|B)$-intégrable et

$$\mathbb{E}[X|B] = \int x\, P(\cdot|B)(dx)$$

espérance de X par rapport à la mesure $P(\cdot|B)$.

<u>preuve</u> : i) on suppose $X = 1\!\!1_A$ $A \in \mathcal{A}$.

$$\int_\Omega x\, P(\cdot|B)dx = \int_\Omega 1\!\!1_A(x) P(\cdot|B)dx$$

$$= P(A|B) = P(A\cap B)/P(B).$$

$$= \frac{1}{P(B)} \int_\Omega 1\!\!1_{A\cap B}(x)\, P(dx)$$

$$= \frac{1}{P(B)} \int_\Omega 1\!\!1_A \cdot 1\!\!1_B\, P(dx)$$

$$\overset{\text{def}}{=} \mathbb{E}[X|B]$$

ii) fonctions étagées

$$X = \sum \lambda_i 1\!\!1_{A_i}$$

on obtient l'égalité par linéarité

iii) X mesurable positive

On écrit $X_n \longrightarrow X$ X_n étagées positives

$$\int_\Omega X_n P(\cdot|B) = \frac{1}{P(B)} \int_\Omega X_n 1\!\!1_B P(dx)$$

$$\int x\, P(\cdot \mid B) = \frac{1}{P(B)} \int_\Omega x \, 1\!\!1_B \, P(dx)$$

iv) X intégrable

$X = X^+ - X^-$ $\quad X^+, X^- \geq 0$ intégrables

par iii) $\displaystyle\int_\Omega X^\pm(x) P(\cdot \mid B)(dx) = \frac{1}{P(B)} \int X^\pm(x) 1\!\!1_B P(dx)$

$\Rightarrow |X| = X^+ + X^-$ est intégrable pour $P(\cdot \mid B)$ aussi.

$\blacksquare$

On veut définir l'espérance cond. sachant <u>une variable aléatoire</u>.

Soit Y v.a. à valeurs dans E (dénombrable)

$E' = \{ y \in E \mid P(Y = y) > 0 \}$

si $y \in E'$ et $X \in L^1(\Omega, \mathcal{A}, P)$

on définit

$$\mathbb{E}[X \mid Y = y] = \mathbb{E}[X \cdot 1\!\!1_{\{Y=y\}}] / P(Y=y)$$

Définition : Soit $X \in L^1(\Omega, \mathcal{A}, P)$.

L'espérance conditionnell de X sachant Y est la v.a. définie sur Ω

$$\mathbb{E}[X \mid Y] = \varphi(Y)$$

où $\varphi : E \longrightarrow \mathbb{R}$ est :

$$\varphi(y) = \begin{cases} \mathbb{E}[X \mid Y = y] & \text{si } y \in E' \\ 0 & \text{sinon} \end{cases}$$

__Remarque__ $P(Y \in E \setminus E') = \sum_{y \in E \setminus E'} \overbrace{P(Y=y)}^{0} = 0$

$\Rightarrow E \setminus E'$ est de mesure nulle.

On aurait pu définir $\mathbb{E}[X \mid Y] \Big|_{E \setminus E'}$ differement.

__remarque__

On définit une v.a. __p.s.__

$$\underline{\mathbb{E}[X \mid Y](\omega)} = \underline{\mathbb{E}[X \mid Y = y]}$$
si $Y(\omega) = y$

__remarque__ : $\mathbb{E}[X \mid Y]$ est une v.a. $\sigma(Y)$-mesurable. C'est la meilleure approximation de X par une fonction de Y.

__Exemple__ : Lancer d'un dé

$\Omega = \{1, \dots, 6\}$ $P(\{\omega\}) = \frac{1}{}$

6

$\forall \omega \in \Omega$

$$Y(\omega) = \begin{cases} 1 & \text{si } \omega \text{ impaire} \quad 1,3,5 \\ 0 & \text{'' pair} \quad 2,4,6 \end{cases}$$

$$X(\omega) = \omega$$

$$\mathbb{E}[X \mid Y](\omega) \quad ?$$

$$\mathbb{E}[X \mid Y](1) = \mathbb{E}[X \mid \underbrace{Y=1}_{\text{impaire}}] = \frac{1}{P(Y=1)} \mathbb{E}[X \, \mathbb{1}_{\{Y=1\}}]$$

$\underline{Y(1)=1}$

$$= \frac{1}{\left(\frac{3}{6}\right)} \left(\frac{1}{6} + \frac{3}{6} + \frac{5}{6} \right)$$

$$= 3$$

$$\mathbb{E}[X \mid Y](2) = \mathbb{E}[X \mid Y=0] = \frac{1}{P(Y=0)} \mathbb{E}[X \, \mathbb{1}_{\{Y=0\}}]$$

$Y(2)=0$

$$= \frac{1}{\left(\frac{3}{6}\right)} \left(\frac{2}{6} + \frac{4}{6} + \frac{6}{6} \right)$$

$$= 4$$

$$\mathbb{E}[X \mid Y](3) = \mathbb{E}[X \mid Y=1] = 3$$

$Y(3)=1$

$$\mathbb{E}[X \mid Y](4) = \mathbb{E}[X \mid Y=0] = 4$$

$Y(4)=0$

$$\mathbb{E}[X,Y](\omega) = \begin{cases} 3 & \omega \text{ est impaire} \\ 4 & \text{paire} \end{cases}$$

Proposition : Propriétés :

i) $\mathbb{E}\left[\,|\mathbb{E}[X|Y]|\,\right] \leq \mathbb{E}[|X|]$

en particulier $E[X|Y] \in L^1(\Omega, \mathcal{A}, P)$

(car $X \in L^1(\Omega, \mathcal{A}, P)$)

ii) Si Z est v.a.r $\sigma(Y)$-mesurable bornée.

Alors

$$\mathbb{E}[ZX] = \mathbb{E}[Z\,\mathbb{E}[X|Y]]$$

preuve :