

I - La fonction caractéristique.

Nous commençons par introduire ici la fonction caractéristique d'une variable aléatoire. Cette fonction est aussi connue sous le nom de transformée de Fourier dans d'autres domaines (comme le traitement du signal).

① - Notations.

On note $\langle x, y \rangle$ le produit scalaire de deux vecteurs de \mathbb{R}^n . Ainsi, si

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \text{ et } y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \text{ où } x_i, y_i \in \mathbb{R}, \text{ alors } \langle x, y \rangle = x_1 y_1 + \dots + x_n y_n.$$

→ Si $u \in \mathbb{R}^n$, la fonction complexe $x \mapsto e^{i\langle u, x \rangle}$ est continue, de module 1.

→ Si X est une v.a. dans \mathbb{R}^n , on peut considérer $e^{i\langle u, X \rangle}$ comme une v.a. à valeurs complexes, avec comme partie réelle $Y = \cos(\langle u, X \rangle)$ et imaginaire $Z = \sin(\langle u, X \rangle)$ des v.a. réelles.

→ Noter que ces v.a. réelles sont bornées par 1 \Rightarrow elles admettent une espérance :

$$\mathbb{E}[e^{i\langle u, X \rangle}] = \mathbb{E}[Y] + i \mathbb{E}[Z] = \mathbb{E}[\cos(\langle u, X \rangle)] + i \mathbb{E}[\sin(\langle u, X \rangle)].$$
② - Définition.

Définition : On appelle fonction caractéristique d'une v.a. X à valeurs dans \mathbb{R}^n la fonction ϕ_X de \mathbb{R}^n dans \mathbb{C} , définie par :

$$\phi_X(u) = \mathbb{E}[e^{i\langle u, X \rangle}]$$

→ Si X est à valeurs dans \mathbb{R} , ϕ_X est définie de \mathbb{R} dans \mathbb{C} par:

$$\forall u \in \mathbb{R}, \phi_X(u) = \mathbb{E}[e^{iux}].$$

La fonction caractéristique ne dépend que de la loi de X , \mathbb{P}_X : c'est la transformée de Fourier de la loi \mathbb{P}_X . ϕ_X caractérise la loi \mathbb{P}_X .

→ Il existe un lien entre fonction caractéristique et fonction génératrice:

$$\left. \begin{array}{l} \phi_X(u) = \mathbb{E}[e^{iux}] \\ G_X(u) = \mathbb{E}[u^X] \end{array} \right\} \Rightarrow \phi_X(u) = G_X(e^{iu}). \quad (\text{vrai pour des v.a. à valeurs dans } \mathbb{N}!)$$

③ - Propriétés.

Proposition: La fonction ϕ_X est continue, de module inférieur ou égal à 1, et

$$\phi_X(0) = 1$$

$$\phi_X(-u) = \overline{\phi_X(u)}$$

Preuve: notons $|z|$ le module d'un nombre complexe z .

• Comme $\mathbb{E}[Y]^2 \leq \mathbb{E}[Y^2]$ pour toute v.a. réelle Y , on a:

module $|\phi_X(u)|^2 = \mathbb{E}[\cos(\langle u, X \rangle)]^2 + \mathbb{E}[\sin(\langle u, X \rangle)]^2 \leq \mathbb{E}[\underbrace{\cos^2(\langle u, X \rangle) + \sin^2(\langle u, X \rangle)}_{=1}]$

donc $|\phi_X(u)| \leq 1$.

• Pour la continuité, on introduit une suite $u_p \xrightarrow{p \rightarrow +\infty} u$. On a convergence simple de $e^{i\langle u_p, X \rangle}$ vers $e^{i\langle u, X \rangle}$.

Comme ces v.a. sont de module borné par 1, on peut appliquer le théorème de convergence dominée: on obtient ainsi $\phi_X(u_p) \xrightarrow{p \rightarrow +\infty} \phi_X(u)$.

⇒ La fonction ϕ_X est continue.

Proposition: Si X est une v.a. à valeurs dans \mathbb{R}^n , si $a \in \mathbb{R}^m$ et si A est une matrice de Vailie $m \times n$, on a:

$$\phi_{a+AX}(u) = e^{i\langle u, a \rangle} \phi_X(A^T u), \quad \forall u \in \mathbb{R}^m, \text{ avec } A^T \text{ Transposée de } A.$$

Preuve. On a $e^{i\langle u, a+AX \rangle} = e^{i\langle u, a \rangle} e^{i\langle A^T u, X \rangle}$. En effet, $\langle u, AX \rangle = \langle A^T u, X \rangle$, puis on prend les espérances.

④ - Exemples

On peut faire ces exemples en guise d'exercice.

a) $X \sim \mathcal{B}(n, p)$: $E[e^{iux}] = \sum_{k=0}^n e^{iuk} \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=0}^n \binom{n}{k} (e^{iu} p)^k (1-p)^{n-k} = (1-p + e^{iu} p)^n$ par la formule du binôme de Newton.

b) $X \sim \mathcal{P}(\theta)$: $E[e^{iux}] = \sum_{k=0}^{\infty} e^{iuk} \frac{\theta^k}{k!} e^{-\theta} = \sum_{k=0}^{\infty} e^{-\theta} \frac{(\theta e^{iu})^k}{k!} = e^{-\theta} \sum_{k=0}^{\infty} \frac{(\theta e^{iu})^k}{k!} = e^{-\theta} e^{\theta e^{iu}} = e^{\theta(e^{iu} - 1)}$.

c) $X \sim \mathcal{U}(a, b]$: $E[e^{iux}] = \int_{-\infty}^{\infty} \frac{1}{b-a} \mathbb{1}_{[a, b]}(x) e^{iux} dx = \frac{1}{b-a} \int_a^b e^{iux} dx = \frac{1}{b-a} \frac{1}{iu} [e^{iux}]_a^b = \frac{e^{iub} - e^{iua}}{iu(b-a)}$.

\Rightarrow Cas particulier d'une loi uniforme sur $[-a, a]$: $\phi_X(u) = \frac{e^{iua} - e^{-iua}}{2iua}$

soit par une formule de trigonométrie: $\phi_X(u) = \frac{\sin(ua)}{ua}$.

$$d) \underline{X \sim \text{Exp}(\lambda)} : \lambda > 0 \quad \phi_X(\lambda) = E[e^{i\lambda x}] = \int_{-\infty}^{+\infty} e^{i\lambda x} \lambda e^{-\lambda x} \mathbb{1}_{\{x \geq 0\}}(x) dx$$

$$= \lambda \int_0^{+\infty} e^{-\lambda x} e^{i\lambda x} dx = \lambda \int_0^{+\infty} e^{(i\lambda - \lambda)x} dx = \frac{\lambda}{i\lambda - \lambda} \left[e^{(i\lambda - \lambda)x} \right]_0^{+\infty} = \frac{\lambda}{\lambda - i\lambda}$$

car $\lim_{x \rightarrow +\infty} e^{(i\lambda - \lambda)x} = \lim_{x \rightarrow +\infty} e^{i\lambda x - \lambda x}$, or $e^{i\lambda x}$ a pour module 1 donc

le terme en λx l'emporte. Ainsi, $\lim_{x \rightarrow +\infty} e^{(i\lambda - \lambda)x} = 0$.

e) $X \sim \mathcal{N}(0, 1)$: on note $g(x)$ la densité : $g(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$.

$$\phi_X(\lambda) = E[e^{i\lambda x}] = \int_{-\infty}^{+\infty} e^{i\lambda x} \underbrace{\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}}_{g(x)} dx.$$

En remarquant que $\forall s \in \mathbb{R}, g(x-s) e^{\frac{s^2}{2}} = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-s)^2}{2}} e^{\frac{s^2}{2}} = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2 - 2xs + s^2}{2} + \frac{s^2}{2}}$

$$= \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2} + 2xs} = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} e^{xs} = g(x) e^{xs}, \text{ on peut remplacer}$$

$$\text{dans } \int_{-\infty}^{+\infty} e^{s\lambda x} g(x) dx = \int_{-\infty}^{+\infty} e^{\frac{s^2}{2}} g(x-s) dx = e^{\frac{s^2}{2}} \underbrace{\int_{-\infty}^{+\infty} g(x-s) dx}_{=1 \text{ car densité}} = e^{\frac{s^2}{2}}.$$

On a considéré $s \in \mathbb{R}$: qu'en est-il si $s \in \mathbb{C}$? (car c'est ce que nous avons pr ϕ_X).

On développe en série entière :

$$e^{s^2/2} = \sum_{n \geq 0} \frac{s^{2n}}{2^n n!}$$

$$\int_{-\infty}^{+\infty} e^{s\lambda x} g(x) dx = \sum_{n \geq 0} \frac{s^n}{n!} \int_{-\infty}^{+\infty} x^n g(x) dx$$

\Rightarrow En identifiant les coefficients de s^n , on en déduit les moments de g :

$$\underbrace{\int_{-\infty}^{+\infty} x^{2n+1} g(x) dx}_{\text{car impair } E[X^{2n+1}]} = 0 \quad ; \quad \underbrace{\int_{-\infty}^{+\infty} x^{2n} g(x) dx}_{\text{car pair } E[X^{2n}]} = \frac{(2n)!}{2^n n!} \text{ pour } X \sim \mathcal{N}(0, 1)$$

Le rayon de convergence de la série entière est infini.

On déduit de ces résultats que si s est complexe, les développements de Taylor des 2 membres $\left(e^{\frac{s^2}{2}} \int_{-\infty}^{+\infty} e^{s^2 x} g(x) dx \right)$ sont encore égaux. Ainsi, $\phi_X(u) = e^{-\frac{u^2}{2}}$.

g) $X \sim \mathcal{N}(m, \sigma^2)$: on utilise que $X = m + \sigma Y$ avec $Y \sim \mathcal{N}(0, 1)$, et on trouve le résultat avec la proposition du cours.

$$\text{Ainsi, } \phi_X(u) = \phi_{m+\sigma Y}(u) = e^{i u m} \phi_Y(\sigma u) = e^{i u m} e^{-\frac{\sigma^2 u^2}{2}} = e^{i u m - \frac{u^2 \sigma^2}{2}}$$

⑤ - Résultats fondamentaux.

Théorème : La fonction caractéristique ϕ_X caractérise la loi du vecteur aléatoire X .
Ainsi, si deux vecteurs aléatoires X et Y ont même fonction caractéristique, alors ils ont même loi.

Corollaire : Soit $X = (X_1, \dots, X_n)$ un vecteur aléatoire à valeurs dans \mathbb{R}^n .

Les composantes X_i sont $\perp \Leftrightarrow \forall (u_1, \dots, u_n) \in \mathbb{R}^n, \phi_X(u_1, \dots, u_n) = \prod_{j=1}^n \phi_{X_j}(u_j)$,
où ϕ_{X_j} désigne la fct caractéristique de la composante X_j et ϕ_X celle du vecteur.

Preuve : Supposons que les X_i sont \perp . On sait par ailleurs que $\langle u, X \rangle = \sum_{j=1}^n u_j X_j$,
donc $E[e^{i \langle u, X \rangle}] = E[e^{i \sum_{j=1}^n u_j X_j}] = E[e^{i u_1 X_1} e^{i u_2 X_2} \dots e^{i u_n X_n}] \stackrel{\perp}{=} \phi_{X_1}(u_1) \dots \phi_{X_n}(u_n)$.

• Supposons que $\forall (u_1, \dots, u_n) \in \mathbb{R}^n, \phi_X(u_1, \dots, u_n) = \prod_{j=1}^n \phi_{X_j}(u_j)$. On peut construire des v.a. $X'_j \perp$, telles que X'_j et X_j aient même loi $\forall j$. Donc $\phi_{X_j} = \phi_{X'_j}$. Si $X' = (X'_1, \dots, X'_n)$, alors $\phi_{X'} = \phi_X$ en utilisant la condition nécessaire.

Du coup, X et X' ont même loi, ce qui entraîne que pour tout borélien A_j ,

$$P\left(\bigcap_j \{X_j \in A_j\}\right) = P\left(\bigcap_j \{X'_j \in A_j\}\right) \stackrel{X'_j \perp}{=} \prod_{j=1}^n P(\{X'_j \in A_j\}) = \prod_{j=1}^n P(\{X_j \in A_j\}),$$

d'où l'indépendance entre les X_j .

⑥ - Transformée de Laplace.

→ Utile pour une variable aléatoire à valeurs dans \mathbb{R}^+ .

→ Si X est une v.a. à valeurs dans \mathbb{R}^+ , sa Transformée de Laplace est donnée par:

$$\Psi_X(\lambda) = E[e^{-\lambda X}] \quad \text{avec } \lambda \in \mathbb{R}^+.$$

→ Cette fonction est:

- définie sur \mathbb{R}^+ ,
- indéfiniment dérivable sur $]0, +\infty[$,
- qui satisfait $\Psi_X(\lambda) = \phi_X(i\lambda) \Rightarrow \Psi_X$ caractérise la loi P_X .

→ Si X est à valeurs dans \mathbb{N} , de fonction génératrice G_X , alors $\Psi_X(\lambda) = G_X(e^{-\lambda})$.

⑦ - Loi d'une somme de vecteurs aléatoires \perp .

Proposition: Si X et Y sont 2 vecteurs aléatoires indépendants à valeurs dans \mathbb{R}^n , la fonction caractéristique de la somme $X+Y$ vaut

$$\phi_{X+Y} = \phi_X \phi_Y$$

Preuve: Immédiate. En effet, $e^{i\langle u, X+Y \rangle} = e^{i\langle u, X \rangle} e^{i\langle u, Y \rangle}$, puis on applique le fait que si $X \perp Y$, alors $g(X)$ et $h(Y)$ le sont aussi. De plus, si $g(X)$ et $h(Y) \in L^1$, alors $g(X)h(Y)$ aussi, et $E[g(X)h(Y)] = E[g(X)]E[h(Y)]$.

→ Exemples:

- $X \sim \mathcal{P}(\lambda_1)$
 $Y \sim \mathcal{P}(\lambda_2)$ $\left\{ \Rightarrow Z = X+Y \sim \mathcal{P}(\lambda_1+\lambda_2) \right.$
- $X \sim \mathcal{N}(m_1, \sigma_1^2)$
 $Y \sim \mathcal{N}(m_2, \sigma_2^2)$ $\left\{ \Rightarrow Z = X+Y \sim \mathcal{N}(m_1+m_2, \sigma_1^2+\sigma_2^2) \right.$
- $X \sim \mathcal{B}(n_1, p)$
 $Y \sim \mathcal{B}(n_2, p)$ $\left\{ \Rightarrow Z = X+Y \sim \mathcal{B}(n_1+n_2, p) \right.$

⑧ - Lien entre fonction caractéristique et moments.

Proposition: Soit X un vecteur aléatoire de \mathbb{R}^n . Si la variable aléatoire $|X|^m$ (où $|X|$ désigne la norme euclidienne de X) $\in L^1$ pour un entier m , alors la fonction ϕ_X est m fois continûment différentiable sur \mathbb{R}^n , et pour tout choix des indices i_1, \dots, i_m , on a

$$\frac{\partial^m}{\partial u_{i_1} \partial u_{i_2} \dots \partial u_{i_m}} \phi_X(u) = i^{i_1} \dots i^{i_m} \mathbb{E}[e^{i\langle u, X \rangle} X_{i_1} \dots X_{i_m}]$$

→ Remarque : en prenant $u=0$, cela permet de calculer $\mathbb{E}[X_{i_1} \dots X_{i_m}]$: ce sont les dérivées à l'origine de ϕ_X .

→ Exemple : si X est à valeurs réelles, intégrable (ou de carré intégrable); on a

$$\mathbb{E}[X] = +i \phi_X'(0) \quad \text{et} \quad \mathbb{E}[X^2] = -\phi_X''(0).$$

⑨ - Cas des vecteurs gaussiens

Rappel: Un vecteur aléatoire $X = (X_1, \dots, X_n)$ à valeurs dans \mathbb{R}^n est un vecteur gaussien si toute combinaison linéaire $\sum_{j=1}^n a_j X_j = \langle a, X \rangle$ pour $a = (a_1, \dots, a_n) \in \mathbb{R}^n$ suit une loi normale.

Bien entendu, cela implique que chaque $X_j \sim \mathcal{N}(m, \sigma^2)$, puisque le vecteur x peut avoir une seule composante non-nulle!

→ Si les X_i sont des v.e. normales indépendantes, le vecteur X est gaussien (avec matrice de variance-covariance diagonale).

→ ⚠: Si ces composantes sont de loi normale mais pas indépendantes, parfois X n'est pas un vecteur gaussien!

Ex: $X_1 \sim \mathcal{N}(0,1)$

- Alors $X_2 \sim \mathcal{N}(0,1)$

$$X_2 = \begin{cases} X_1 & \text{si } |X_1| \leq 1 \\ -X_1 & \text{sinon} \end{cases}$$

Mais $X = (X_1, X_2)$ n'est pas un vecteur gaussien, puisque $P(X_1 + X_2 = 0) > 0$ (donc $X_1 + X_2$ ne suit pas une loi normale, sinon cette proba serait nulle...).

Théorème: X est un vecteur gaussien \Leftrightarrow Sa fonction caractéristique $\phi_X(u) = e^{i\langle u, m \rangle - \frac{1}{2}\langle u, C u \rangle}$ où $m \in \mathbb{R}^n$ et C est une matrice $n \times n$ symétrique positive. $\langle u, C u \rangle$

Dans ce cas, $m = E[X] = \begin{pmatrix} E[X_1] \\ E[X_2] \\ \vdots \\ E[X_n] \end{pmatrix}$ et C est la matrice de covariance de X .

Preuve: • CS: On suppose que $\phi_X(u) = e^{i\langle u, m \rangle - \frac{1}{2}\langle u, C u \rangle}$.

Pour toute combinaison linéaire $Y = \sum_j a_j X_j = \langle a, X \rangle$, et pour $v \in \mathbb{R}$, on a

$$\phi_Y(v) = \phi_X(va) = e^{i v \langle m, a \rangle - \frac{v^2}{2} \langle a, C a \rangle}, \text{ donc } Y \sim \mathcal{N}(\langle a, m \rangle, \langle a, C a \rangle).$$

• CN: Soit C la matrice de covariance de X , et m le vecteur moyenne de X .

Ces quantités existent puisque chaque $X_j \sim \mathcal{N}$. Si $Y = \langle a, X \rangle = \sum_j a_j X_j$ avec $a \in \mathbb{R}^n$, on a immédiatement $E[Y] = \langle a, m \rangle$ et $\text{Var}(Y) = \langle a, C a \rangle$.

Par hyp., $Y \sim \mathcal{N}$, donc $\phi_Y(v) = e^{i v \langle a, m \rangle - \frac{v^2}{2} \langle a, C a \rangle}$.

Or $\phi_Y(1) = \phi_{\langle a, X \rangle}(1) = E[e^{i\langle a, X \rangle}] = \phi_X(a)$, d'où le résultat.

(5)

Proposition: Soit X un vecteur gaussien, à valeurs dans \mathbb{R}^n , de moyenne m .

Il existe des v.a.r. \perp Y_1, \dots, Y_n de lois normales $\mathcal{N}(0, \lambda_j)$ avec $\lambda_j \geq 0$ (si $\lambda_j = 0$ on admet que $Y_j = 0$) et une matrice orthogonale A telles que :

$$X = m + AY, \text{ où } Y = (Y_1, \dots, Y_n).$$

Preuve: Comme C est une matrice symétrique positive, il existe une matrice orthogonale A et une matrice diagonale Λ dont les éléments diagonaux vérifient $\lambda_j \geq 0$, et telle que la matrice de covariance de X s'écrit $C = A\Lambda A^t$.

Soit $Y = A^t(X - m)$, alors Y est un vecteur gaussien de covariance $C' = A^t C A = \Lambda$ et de moyenne nulle. Les composantes Y_j de Y répondent à la prescription.

II - Extension vers les statistiques : autres résultats fondamentaux.

① - La loi du Chi-2 (χ^2)

Définition: On appelle variable aléatoire du Chi-2 (χ^2) à d degrés de liberté toute variable aléatoire positive qui a même loi que la somme de d carrés de variables aléatoires de loi normale $\mathcal{N}(0, 1)$ indépendantes.

→ En pratique, cette loi est très utile dans le domaine de la statistique, plus particulièrement des tests. En effet, elle a 2 applications majeures :

- Test d'adéquation à une loi donnée,
- Test d'indépendance entre deux variables aléatoires.

② - Indicateurs statistiques fondamentaux.

On rappelle ici (cf fin chap. 2) le résultat suivant:

Proposition: Soient n v.a. $\perp (X_1, X_2, \dots, X_n)$, iid, de loi normale $\mathcal{N}(m, \sigma^2)$,

Considérons $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ et $V = \frac{1}{n} ((X_1 - \bar{X})^2 + \dots + (X_n - \bar{X})^2)$.

Alors:

$$\rightarrow \bar{X} \sim \mathcal{N}(m, \frac{\sigma^2}{n})$$

$$\rightarrow \frac{n}{\sigma^2} ((X_1 - \bar{X})^2 + \dots + (X_n - \bar{X})^2) \sim \chi^2(n-1)$$

$$\rightarrow \bar{X} \perp V.$$

III - Convergence en loi

Après avoir vu les convergences en probabilité, en moyenne, presque sûre, nous introduisons ici la notion de convergence en loi.

Cette notion est radicalement différente des précédentes, et concerne les lois des variables aléatoires. Elle représente une notion où les lois sont asymptotiquement "proches" (mais ce n'est pas forcément le cas des v.a.).

① - Définition

On considère une suite de v.a. (X_n) , et X . Les v.a. sont toutes à valeurs dans le même espace \mathbb{R}^d (mais peuvent être définies sur des espaces d'états $\neq \mathbb{R}^d$!).

Déf: La suite (X_n) converge en loi vers X si pour toute fonction f continue et bornée sur \mathbb{R}^d , on a: $\mathbb{E}[f(X_n)] \xrightarrow{n \rightarrow +\infty} \mathbb{E}[f(X)]$.

On note alors $X_n \xrightarrow{L} X$.

→ Exemple 1: si l'ordre les v.a. X_n ont un nombre fini de valeurs, ⑥
 $\{x_i, 1 \leq i \leq N\}$. Alors $X_n \xrightarrow{\mathcal{L}} X \Leftrightarrow \lim_{n \rightarrow \infty} P(X_n = x_i) = P(X = x_i) \forall i$.
 En effet, il suffit de voir que $E[f(X_n)] = \sum_{i=1}^N f(x_i) P(X_n = x_i)$.

→ Exemple 2: Soient $(X_n)_n$ des v.a. de loi $X_n \sim \mathcal{N}(0, \sigma_n^2)$.
 X une v.a. de loi $X \sim \mathcal{N}(0, \sigma^2)$.

On suppose que la suite de réels positifs $(\sigma_n)_n$ converge vers $\sigma > 0$ pd $n \rightarrow \infty$.
 Alors $X_n \xrightarrow{\mathcal{L}} X$.

En effet, avec f continue bornée sur \mathbb{R} ,

$$E[f(X_n)] = \int_{\mathbb{R}} f(y) \frac{1}{\sqrt{2\pi} \sigma_n} e^{-\frac{y^2}{2\sigma_n^2}} dy \xrightarrow[n \rightarrow \infty]{\substack{\uparrow \\ \text{théorème de} \\ \text{CV dominée}}} \int_{\mathbb{R}} f(y) \frac{1}{\sqrt{2\pi} \sigma} e^{-\frac{y^2}{2\sigma^2}} dy = E[f(X)]$$

② - Propriétés.

La convergence en loi est plus faible que la convergence en probabilité.

Proposition: Si $X_n \xrightarrow[n \rightarrow \infty]{P} X$ alors $X_n \xrightarrow[n \rightarrow \infty]{\mathcal{L}} X$.

⇒ Les convergences en moyenne et presque-sûre entraînent aussi la CV en loi.

Preuve: Supposons que la suite (X_n) CV en proba. vers X . Soit f une fonction continue et bornée sur \mathbb{R}^d . On sait que la suite $(f(X_n))_n$ CV aussi en proba vers $f(X)$. Comme f est bornée, $f(X_n)$ CV aussi en moyenne vers $f(X)$.
 Et donc en particulier, $E[f(X_n)]$ CV vers $E[f(X)]$.

Proposition: Soient X_n et X des v.a.r. de fonctions de répartition F_n et F .

On a: $X_n \xrightarrow[n \rightarrow \infty]{\mathcal{L}} X \iff F_n(x) \xrightarrow[n \rightarrow \infty]{} F(x) \quad \forall x \text{ en lequel } F \text{ est continue.}$

Preuve: admis - L'intuition est que les FdR caractérisent les lois.

Corollaire: Si la suite (X_n) de v.a.r. converge en loi vers X , et si la loi de X admet une densité, alors $\forall a < b$,

$$P(X_n \in]a, b]) \xrightarrow[n \rightarrow \infty]{} P(X \in]a, b]).$$

Proposition: Soient X_n et X des vecteurs aléatoires.

$$X_n \xrightarrow[n \rightarrow \infty]{\mathcal{L}} X \iff E[f(X_n)] \xrightarrow[n \rightarrow \infty]{} E[f(X)] \quad \forall f \text{ lipschitzienne bornée.}$$

③ - Résultats fondamentaux.

a) Lemme de Slutsky.

Le lemme de Slutsky est régulièrement utilisé en statistique pour justifier en statistique de la convergence en loi d'un estimateur.

Lemme: Soit $(X_n)_n$ et $(Y_n)_n$ deux suites de vecteurs aléatoires. Supposons que

$$X_n \xrightarrow[n \rightarrow \infty]{\mathcal{L}} X \text{ et que } X_n - Y_n \xrightarrow[n \rightarrow \infty]{P} 0. \text{ Alors } Y_n \xrightarrow[n \rightarrow \infty]{\mathcal{L}} X.$$

Preuve: Il suffit de montrer que $\lim_{n \rightarrow \infty} E[f(Y_n)] = E[f(X)]$ pr lte fonction bornée lipschitzienne f . On a alors $|f(x)| \leq k$ et $|f(x) - f(y)| \leq C|x - y|$ pour des constantes k et C .

Soit $\varepsilon > 0$ fixé. On peut écrire:

$$|E[f(Y_n)] - E[f(X_n)]| \leq E[|f(Y_n) - f(X_n)|] \leq C\varepsilon + 2h \underbrace{P(|X_n - Y_n| > \varepsilon)}_{\xrightarrow{n \rightarrow \infty} 0 \text{ car } X_n - Y_n \xrightarrow{P} 0}.$$

Or ε est arbitrairement petit,

$$\text{donc } \lim_{n \rightarrow \infty} |E[f(Y_n)] - E[f(X_n)]| = 0 \quad \text{CQFD.}$$

→ Il existe plusieurs versions du théorème de Slutsky. Par exemple,

$$\bullet \text{ Si } \begin{cases} X_n \xrightarrow{L} X \\ Y_n \xrightarrow{P} c \end{cases}, \text{ alors } (X_n, Y_n) \xrightarrow{L} (X, c)$$

$$\text{Si par exemple } X_n \text{ et } Y_n \text{ peuvent \u00eatre ajout\u00e9s, on a } \begin{cases} X_n + Y_n \xrightarrow{L} X + c \\ Y_n X_n \xrightarrow{L} cX \end{cases}$$

b) Th\u00e9or\u00e8me de L\u00e9vy:

Ce th\u00e9or\u00e8me caract\u00e9rise la convergence en loi gr\u00e2ce aux fonctions caract\u00e9ristiques. C'est donc un outil central dans la pratique. Il intervient par exemple dans la d\u00e9monstration du th\u00e9or\u00e8me central limit\u00e9.

Th\u00e9or\u00e8me: Soit X_n des variables al\u00e9atoires \u00e0 valeurs dans \mathbb{R}^d .

a) Si $X_n \xrightarrow[n \rightarrow \infty]{L} X$ alors $\phi_{X_n} \rightarrow \phi_X$ (au sens de la convergence simple des fct.)

b) Si les ϕ_{X_n} convergent simplement vers une fonction (complexe) ϕ sur \mathbb{R}^d , et si cette fonction est continue en 0, alors c'est la fonction caract\u00e9ristique d'une variable al\u00e9atoire X et on a $X_n \xrightarrow[n \rightarrow \infty]{L} X$.

Preuve: admis.

→ Exemple: en utilisant les fonctions caractéristiques, on déduit immédiatement par exemple que: si
$$\left. \begin{array}{l} X_n \sim B(m, p_n) \\ X \sim B(m, p) \\ p_n \xrightarrow{n \rightarrow \infty} p \end{array} \right\} \Rightarrow \text{alors } X_n \xrightarrow{\mathcal{L}} X.$$

IV - Théorème de la limite centrale.

On appelle souvent ce théorème le théorème central limite. (TCL).

On se place dans le même contexte que pour la loi des grands nombres: on considère une suite de v.a. X_n indépendantes, de même loi et de carré intégrable. (en notant m et σ^2 la moyenne et la variance communes):

On s'intéresse à $\boxed{S_n = X_1 + \dots + X_n}$ (ainsi que $M_n = \frac{1}{n} S_n$).

Nous avons déjà démontré que:

→ $\frac{1}{n} S_n$ converge en moyenne vers m .

→ $\frac{1}{n} S_n$ converge presque sûrement vers m .

On aimerait savoir à quelle vitesse cette convergence a lieu... Ceci est lié à la variance de l'estimateur.

Pour évaluer cette vitesse, pour trouver un équivalent de $\frac{S_n}{n} - m$, on est amené à étudier la limite (si elle existe!) de $n^\alpha \left(\frac{S_n}{n} - m \right)$ pour \neq valeurs de α .

- Si α est petit, la suite va encore tendre vers 0.
- Si α est grand, elle va exploser.

→ On peut espérer que cette suite CV vers une limite (ni ∞ , ni nulle) pour une valeur de α .

(8)

En l'occurrence, la suite $n^\alpha \left(\frac{S_n}{n} - m \right)$ ne CV pas presque-sûrement, ni même en probabilité, pour aucune valeur de α .

Mais elle converge en loi pour la même valeur $\alpha = \frac{1}{2}$ quelle que soit la loi des X_n , toujours vers une loi normale!

TCL: Si les X_n sont des v.a. réelles indépendantes et de même loi, de carré intégrable, de moyenne m et de variance σ^2 avec $\sigma^2 > 0$, alors les variables

$$\frac{S_n - nm}{\sigma\sqrt{n}} \underset{n \rightarrow +\infty}{\overset{\mathcal{L}}{\sim}} \mathcal{N}(0,1), \text{ où } \mathcal{N}(0,1) \text{ désigne la loi normale centrée réduite.}$$

Preuve: cette preuve s'appuie sur le théorème de Lévy. Il existe d'autres preuves.
Soit ϕ la fonction caractéristique de $X_n - m$.

Notons $Y_n = \frac{S_n - nm}{\sigma\sqrt{n}}$.

Puisque les X_n sont indépendantes, la fonction caractéristique de Y_n vaut:

$$\phi_n(u) = \left(\phi\left(\frac{u}{\sigma\sqrt{n}}\right) \right)^n.$$

Comme $E[X_n - m] = 0$ et $E[(X_n - m)^2] = \sigma^2$, la relation qui lie la fonction caractéristique aux moments implique que:

$$\phi(u) = 1 - \frac{u^2 \sigma^2}{2} + u^2 o(|u|) \text{ quand } u \rightarrow 0.$$

Comme $\phi(0) = 1$ et que ϕ est continue en 0, on en déduit que pour u fixé et n assez grand,

$$\left| \phi\left(\frac{u}{\sigma\sqrt{n}}\right) - 1 \right| \leq \frac{1}{2}.$$

Il est possible de généraliser la notion de logarithme aux complexes z tels que $|z-1| \leq \frac{1}{2}$. (cf cours d'analyse complexe).

La fonction $\ln(z)$ définie sur le disque $\{z \in \mathbb{C} : |z-1| \leq \frac{1}{2}\}$ admet le même développement limité au voisinage de $z=1$ que le logarithme réel. Donc :

$$\phi_n(u) = \left(\phi\left(\frac{u}{\sqrt{n}}\right) \right)^n = \exp \left[n \ln \left(1 - \frac{u^2}{2n} + \frac{1}{n} \varepsilon_n(u) \right) \right], \text{ où } \varepsilon_n(u) \xrightarrow{n \rightarrow \infty} 0.$$

On en déduit immédiatement que $\phi_n(u)$ converge vers $e^{-\frac{u^2}{2}}$, qui est la fonction caractéristique d'une loi $\mathcal{N}(0,1)$.

On obtient le résultat en appliquant le théorème de Lévy.