

I) Définitions

1) Espace de probabilité

Définition: Soit Ω un univers, on lui adjoint un ensemble $\mathcal{F} \in \mathcal{P}(\Omega)$ tel que $\forall B \in \mathcal{F}, \forall A \in \mathcal{F}, \neg A \in \mathcal{F}, \emptyset \in \mathcal{F}, \Omega \in \mathcal{F}$
 $A \cup B \in \mathcal{F}$ et $A \cap B \in \mathcal{F}$.

Une probabilité sur (Ω, \mathcal{F}) est une application de $\Omega \xrightarrow{P} \mathbb{R}$ mesurable telle que :

$$\begin{cases} \forall E \in \mathcal{F}, P(E) \in [0, 1] \\ P(\Omega) = 1 \\ \text{Soit } (E_i)_{i=1}^{\infty} \text{ disjoints, alors : } P\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} P(E_i). \end{cases}$$

On dit que (Ω, \mathcal{F}, P) est un espace de probabilité.

2) Propriétés des événements

Propriété: Soient deux événements E_1 et $E_2 \in \Omega$.

$$\begin{cases} \bullet E_1 \cap E_2 = \emptyset \Rightarrow P(E_1 \cup E_2) = P(E_1) + P(E_2) \\ \bullet E_1 \text{ et } E_2 \text{ indépendants} \Leftrightarrow P(E_1 \cap E_2) = P(E_1) \times P(E_2) \\ \bullet \text{Si } E_1 \cap E_2 = \emptyset \text{ et sont indépendants, alors soit } P(E_1) = 0 \text{ soit } P(E_2) = 0. \end{cases}$$

Théorème:

$$\begin{cases} \bullet P(A \cup B) = P(A) + P(B) - P(A \cap B) \\ \bullet P\left(\bigcup_{i=1}^n A_i\right) = \sum_{i=1}^n P(A_i) - \sum_{i < j} P(A_i \cap A_j) + \sum_{i < j < k} P(A_i \cap A_j \cap A_k) - \dots + (-1)^{n+1} P\left(\bigcap_{i=1}^n A_i\right) \\ \bullet \text{Soit } \bigcup_{i=1}^n B_i = \Omega \text{ et } B_i \cap B_j = \emptyset \Rightarrow P(A) = \sum_{i=1}^n P(A \cap B_i) = \sum_{i=1}^n P(A) \cdot P(B_i) \\ \bullet \dots P(A|B) = \frac{P(A \cap B)}{P(B)} \end{cases}$$

Corollaire: Soient n événements indépendants et de même probabilité $p \in [0, 1]$.

Alors: $P\left(\bigcup_{i=1}^n E_i\right) < 1$, car $P\left(\bigcup_{i=1}^n E_i\right) = p^n$.

3) Loi de VA

Définition: Une loi de VA discrète X est une ensemble $\{(V_i, p_i) \mid i \in \mathbb{N}\}$ avec V_i les valeurs prises par la VA et $p_i = P(X = V_i)$. De plus $\sum_{i=1}^{\infty} p_i = 1$.

Définition: Une loi de densité pour la VA réelle X est une fonction mesurable $f \geq 0$ telle que $\int_{-\infty}^{+\infty} f = 1$ et $P(X \in [a, b]) = \int_a^b f(x) dx$.

4) Espérance

Définition: Pour une VA discrète $\rightarrow E(X) = \sum_{i=1}^{\infty} V_i \cdot p_i$, lorsque $\sum_{i=1}^{\infty} |V_i| \cdot p_i < \infty$
 Pour une VA réelle $\rightarrow E(X) = \int_{-\infty}^{+\infty} x \cdot f(x) dx$, lorsque $\int_{-\infty}^{+\infty} |x| \cdot f(x) dx < \infty$
 Pour un vecteur aléatoire \rightarrow on définit la moyenne de X : $m_X = (E(X_1), \dots, E(X_n))$
 lorsque $\forall i = 1, \dots, n, \int_{\mathbb{R}^n} |x_i| f(x_1, \dots, x_n) dx_1 \dots dx_n < \infty$

Remarque: $E\left(\frac{1}{[a, b]}\right) = \int_a^b f(x) dx = P(X \in [a, b])$. Cauchy-Schwarz: $E(Z_1 Z_2) \leq E(Z_1^2) E(Z_2^2)$ pour $Z_1, Z_2 \in L^2$

5) Variance

Définition: Soit une VA réelle X de densité f telle que $X^2 f(X) \in L^1(\mathbb{R})$.
 On définit alors $E(X^2) = \int_{-\infty}^{+\infty} x^2 f(x) dx$, et: $\text{Var}(X) = E(X^2) - E(X)^2 = E((X - E(X))^2)$

Définition: On note l'écart type $\sigma_X = \sqrt{\text{Var}(X)}$.

Propriétés: Soient X_1, \dots, X_n deux à deux indépendantes, alors: $E(X_1 \cdot X_2 \cdot \dots \cdot X_n) = E(X_1) \cdot E(X_2) \cdot \dots \cdot E(X_n)$
 $\text{Var}(X_1 + \dots + X_n) = \sum_{i=1}^n \text{Var}(X_i)$

6) Densité conditionnelle, loi marginale

Théorème: Densité conditionnelle $f_{X|Y}(x, y) = \frac{f_{X,Y}(x, y)}{f_Y(y)}$ avec $f_Y(y) = \int_{-\infty}^{+\infty} f_{X,Y}(x, y) dx$
 Si $f_Y(y) = 0$ alors $f_{X|Y}(x, y) = 0$ par convention.

Propriété: La loi marginale de (X_2, \dots, X_n) avec $X_1 \in A$ est $g(x_2, \dots, x_n) = \frac{\int_A f(x_1, \dots, x_n) dx_1}{P(X_1 \in A)} = \frac{\int_A f(x_1, \dots, x_n) dx_1}{\int_A \int_{\mathbb{R}^{n-1}} f(x_1, \dots, x_n) dx_2 \dots dx_n}$

Propriété: Soient X_1 et X_2 de lois f et g . X_1, X_2 indépendantes \Rightarrow loi $(X_1 + X_2)$: $(f * g)(x) = \int_{\mathbb{R}} f(y) g(x - y) dy$.

Préciser: "Fubini" quand on fait $\int_{\mathbb{R}^2} f(x+y) g(x-y) dx dy$

7) Définitions

Définition : la distribution de X est $F(x) = P(X \leq x)$.

Définition : la densité de X , notée $f_X(x)$ est telle que $P(X \in [a, b]) = \int_a^b f_X(x) dx$. Donc, $F(x) = \int_{-\infty}^x f_X(t) dt$.

8) Lois usuelles discrètes

- Bernoulli : pour une v.a. $x \in \{0, 1\}$, $P(X=0) = p$ et $P(X=1) = 1-p$.
 $E(X) = 1-p$ $Var(X) = p(1-p)$
 - Bernoulli : pour une v.a. $x \in \{0, 1, \dots, r\}$, $P(X=i) = p_i$ avec $\sum_{i=0}^r p_i = 1$ et $P(X=i) = 0$ pour $i \geq r+1$.
 - Bernoulli : pour une v.a. $x \in \mathbb{N}$, $P(X=i) = p_i$ avec $\sum_{i=0}^{\infty} p_i = 1$.
 - Poisson : $P(X=k) = \frac{\lambda^k}{k!} e^{-\lambda}$ avec $\lambda > 0$, notée $P(\lambda)$, maximum atteint en $X=E(X)$, car $\frac{d}{d\lambda} \frac{\lambda^k}{k!} e^{-\lambda} = \frac{\lambda^{k-1}}{(k-1)!} e^{-\lambda} - \frac{\lambda^k}{k!} e^{-\lambda} = 0 \Leftrightarrow \lambda = k$.
 - Binomiale : $P(X=k) = C_n^k p^k (1-p)^{n-k}$ pour $k \in [0, n]$, notée $B(n, p)$, $E(X) = np$, $Var(X) = np(1-p)$.
 $P(X=k) = P(X \leq k) = P(X \geq k)$
 - Géométrique : $P(X=k) = p(1-p)^{k-1}$ c'est la proba d'avoir pile au k -ième tirage. $k \in \mathbb{N}^*$.
- Remarque : Soit n v.a. indép de Bernoulli $P(X_i=1)=p$ et $P(X_i=0)=1-p$, alors $S_n = \sum_{i=1}^n X_i$ suit une loi binomiale de paramètres (n, p) .

9) Lois usuelles continues

- Uniforme : sur $[a, b]$, $f_X(x) = \frac{1}{b-a} \mathbb{1}_{[a, b]}$.
- Exponentielle : $f_X(x) = \beta e^{-\beta x} \mathbb{1}_{[0, +\infty[}$, avec $\beta > 0$. Elle est sans mémoire.
- Weibull : $f_X(x) = \alpha \beta x^{\beta-1} e^{-(\beta x)^\alpha} \mathbb{1}_{[0, +\infty[}$, avec $\alpha, \beta > 0$.
- Normale : $f_X(x) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$, $x \in \mathbb{R}$, $\sigma > 0$, $\mu \in \mathbb{R}$. Notée $N(\mu, \sigma^2)$.
- Cauchy : $f_X(x) = \frac{1}{\pi} \frac{1}{1+x^2}$, $x \in \mathbb{R}$. Δ n'admet pas de moyenne et encore moins de variance!

Définition : Une loi est dite "sans mémoire" lorsque $P(X \in [t, t+h] | X \geq t) = P(X \in [0, h])$

Proposition : Soit X une v.a. de loi sans mémoire, alors c'est la loi exponentielle.

Proposition : Soit X suivant $N(\mu, \sigma^2)$ alors $E(X) = \mu$ et $Var(X) = \sigma^2$.

$$\begin{aligned} CV \text{ PS} &\Rightarrow P(\lim_{n \rightarrow \infty} X_n = X) = 1 \\ CV \text{ proba} &\Rightarrow \lim_{n \rightarrow \infty} P(X_n = X) = 1 \\ CV \text{ LP} &\Rightarrow \lim_{n \rightarrow \infty} E(|X_n - X|^p) = 0 \end{aligned}$$

II) Suites de VA et convergence

1) CV presque sûre

Définition : Soit une suite de VA $(X_n)_n \in (\Omega, \mathcal{F}, P)$. La suite (X_n) CV presque sûrement vers une VA X définie sur (Ω, \mathcal{F}, P) si $N = \{\omega \in \Omega / \lim_{n \rightarrow \infty} X_n(\omega) \neq X(\omega)\}$ satisfait à $P(N) = 0$.

Lemme : (Borel Cantelli) Soit (A_n) une suite d'événements de (Ω, \mathcal{F}, P) .

a/ Si $\sum_{n=1}^{\infty} P(A_n) < +\infty \Rightarrow P(\{A_n \text{ inf. souvent}\}) = 0$.

b/ Si $\sum_{n=1}^{\infty} P(A_n) = +\infty$ et les A_i indépendants $\Rightarrow P(\{A_n \text{ inf. souvent}\}) = 1$.

Définition : $\{A_n \text{ inf. souvent}\} = \{\omega / \text{qui appartient à une infinité de } A_n\} = \bigcap_{m \in \mathbb{N}} \bigcup_{l=m}^{\infty} A_l$

Corollaire : Soit (A_n) une suite d'événements indépendants, alors il y a deux possibilités : $P(\{A_n \text{ i.s.}\}) = 0$ ou $P(\{A_n \text{ i.s.}\}) = 1$

Application : Soit (X_n) une suite de v.a. indépendantes, alors l'événement $\{\omega / \lim_{n \rightarrow \infty} X_n(\omega) \text{ existe}\} = E$ satisfait à $P(E) = 0$ ou $P(E) = 1$.

Lemme : Soit (X_n) v.a. indépendantes. $X_n \xrightarrow{p.s.} X \Leftrightarrow \forall \varepsilon > 0, \sum_{n=1}^{\infty} P(|X_n - X| > \varepsilon) < +\infty$

2) CV en proba

Définition : Une suite (X_n) de v.a. dans (Ω, \mathcal{F}, P) converge en proba vers X $\Leftrightarrow \forall \varepsilon > 0, P(\{\omega / |X_n(\omega) - X(\omega)| > \varepsilon\}) \xrightarrow{n \rightarrow \infty} 0$ ou la notation $X_n \xrightarrow{p} X$
 $\Leftrightarrow \lim_{n \rightarrow \infty} P(|X_n - X| > \varepsilon) = 0$

Proposition : Inégalité de Markov. Si X est une v.a. réelle d'espérance finie $E(|X|) < +\infty$, alors $P(|X| > \varepsilon) \leq \frac{E(|X|)}{\varepsilon}$

Proposition : Inégalité de Chebychev. Si X est une v.a. de carré intbl $E(X^2) < +\infty$, alors $P(|X - E(X)| > \varepsilon) \leq \frac{Var(X)}{\varepsilon^2}$

Corollaire: Soit X_i une suite de v.a. indep. suivant $B(p)$ Bernoulli. Alors $S_n = \sum_{i=1}^n X_i$ vérifie $\frac{S_n}{n} \xrightarrow{P} p$

3) CV en moyenne L^p Dans L^2 , raisonner avec des produits scalaires et des Normes! Comme dans l'exercice 1
 Définition: X est $L^p \Leftrightarrow E(|X|^p) < +\infty \Leftrightarrow \int_{-\infty}^{+\infty} |x|^p f_X(x) dx < +\infty$
 Définition: Une suite $(X_n) \in L^p$ CV vers une v.a. $X \in L^p \Leftrightarrow E(|X - X_n|^p) \xrightarrow{n \rightarrow \infty} 0$, noté $X_n \xrightarrow{L^p} X$

4) Relation entre les CV

Théorème: $X_n \xrightarrow{L^p} X \Leftrightarrow X_n \xrightarrow{P} X$

Lemme: $X_n \xrightarrow{L^1} X \Leftrightarrow E\left(\frac{|X_n - X|}{1 + |X_n - X|}\right) \xrightarrow{n \rightarrow \infty} 0$

Théorème: $X_n \xrightarrow{P} X \Leftrightarrow X_n \xrightarrow{P} X$, avec $X_n, X \in (\Omega, \mathcal{F}, P)$

Théorème: $X_n \xrightarrow{P} X \Leftrightarrow \exists m \uparrow \text{ tq } X_{n_k} \xrightarrow{P} X$

Théorème: $X_n \xrightarrow{P} X$ et $\forall n, |X_n| \leq Y$ avec $Y \in L^1 \Rightarrow X \in L^1$ et $X_n \xrightarrow{L^1} X$

Théorème: Si f est une fonction continue, $\begin{cases} X_n \xrightarrow{P} X \Rightarrow f(X_n) \xrightarrow{P} f(X) \\ X_n \xrightarrow{L^1} X \Rightarrow f(X_n) \xrightarrow{L^1} f(X) \end{cases}$



~~~~~ = loi limite  
 - - - - = majoration, CV dominée

### III Fonctions caractéristiques et CV en loi

Définition: Soit une v.a.  $X$  à valeurs dans  $\mathbb{R}^d$ . la fonction caractéristique est  $\phi_X(t) = E(e^{-i2\pi \langle t, X \rangle})$ ,  $t \in \mathbb{R}^d$ .  
 et  $\langle t, x \rangle = \sum t_i x_i$

Théorème:  $\phi_X$  détermine de façon unique la loi de  $X$ , i.e.  $\phi_X = \phi_Y \Rightarrow X \sim Y$ .

Lemme: Si  $X$  est de carré intégrable, alors  $\phi_X$  est  $C^1$  et:  $\phi_X(t) = 1 + i2\pi t E(X) + \frac{(t \cdot t)}{2} \text{Var}(X) + o(t^2)$

Fonctions caractéristiques des lois usuelles:

- Bernoulli:  $X \sim B(p)$   $\phi_X(t) = (1-p) + pe^{-i2\pi t}$
- Binomiale:  $X \sim B(n, p)$   $\phi_X(t) = ((1-p) + pe^{-i2\pi t})^n$
- Poisson:  $X \sim P(\lambda)$   $\phi_X(t) = e^{\lambda(-1 + e^{-i2\pi t})}$
- Exponentielle:  $X \sim \Sigma(\lambda)$   $\phi_X(t) = \frac{\lambda}{\lambda - i2\pi t}$
- Normale:  $X \sim N(m, \sigma^2)$   $\phi_X(t) = e^{-i2\pi t m} e^{-\frac{\sigma^2}{2} (2\pi t)^2}$

#### 2) CV en loi

Définition:  $\begin{cases} \bullet \text{ Soit une suite de mesures } (\mu_n) \text{ sur } \mathbb{R}^d. \mu_n \xrightarrow{w} \mu \Leftrightarrow \forall \varphi \text{ bornée } C^0 \text{ de } \mathbb{R}^d \rightarrow \mathbb{R}, \int_{\mathbb{R}^d} \varphi(x) \mu_n(x) dx \xrightarrow{n \rightarrow \infty} \int_{\mathbb{R}^d} \varphi(x) \mu(x) dx \\ \bullet \text{ Soit une suite de v.a. } (X_n) \in \mathbb{R}^d. X_n \xrightarrow{w} X \Leftrightarrow \forall \varphi \text{ bornée } C^0, E(\varphi(X_n)) \rightarrow E(\varphi(X)). \end{cases}$

Théorème: Soit  $(\mu_n)$  suite de probas sur  $\mathbb{R}^d$ . Soit  $B \in \mathcal{B}(\mathbb{R}^d)$ .

$\mu_n \xrightarrow{w} \mu$  et  $\mu(B) = 0 \Rightarrow \mu_n(B) \rightarrow 0$  Remarque:  $\mu(\{0\}) > 0 \Rightarrow \mu(\{0\}) = \mu(\{0\}) = \mu(\{0\})$ .

Corollaire:  $X_n \xrightarrow{w} X$  et  $F_X(x) = P(X \leq x)$  est  $C^1 \Rightarrow F_{X_n}(x) \xrightarrow{n \rightarrow \infty} F_X(x)$ . Discrete:  $X_n \xrightarrow{w} X \Leftrightarrow P(X_n = k) \xrightarrow{n \rightarrow \infty} P(X = k)$

Lemme:  $\forall \varphi$  continue sur  $\mathbb{R}^d$  à support compact,  $\int_{\mathbb{R}^d} \varphi(x) \mu_n(x) dx \xrightarrow{n \rightarrow \infty} \int_{\mathbb{R}^d} \varphi(x) \mu(x) dx \Rightarrow \mu_n \xrightarrow{w} \mu$

Corollaire:  $X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{w} X$ .

Lemme: Une mesure de proba est déterminée de manière unique par sa transformée de Fourier:  $\hat{\mu}(t) = \int_{\mathbb{R}^d} e^{-i2\pi \langle t, x \rangle} \mu(x) dx$

Théorème:  $\begin{cases} \mu_n \xrightarrow{w} \mu \Leftrightarrow \hat{\mu}_n(t) \xrightarrow{n \rightarrow \infty} \hat{\mu}(t), \forall t \in \mathbb{R}^d \\ X_n \xrightarrow{w} X \Leftrightarrow \phi_{X_n}(t) \xrightarrow{n \rightarrow \infty} \phi_X(t), \forall t \in \mathbb{R}^d \end{cases}$

#### 3) loi forte des gds nombres

- $X_n$  indépendants
  - $X_n$  identiquement distrib.
  - $E(X_n) < +\infty, \forall n$
- $\Rightarrow \frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{n \rightarrow \infty} E(X_1)$

ou:  $\begin{cases} \forall i \neq j, E((X_i - E(X_i))(X_j - E(X_j))) = 0 \\ X_n \text{ identiquement distrib.} \\ \text{Var}(X_n) < +\infty, \forall n \end{cases} \Rightarrow \frac{1}{n} \sum_{i=1}^n X_i \xrightarrow{n \rightarrow \infty} E(X_1)$

#### 4) Théorème limite centrale

Soit  $X_n$  une suite de v.a. iid, de second moment fini ( $\text{Var}(X_i) < +\infty$ ).

Alors:  $\frac{1}{\sqrt{n}} (X_1 + \dots + X_n - n E(X_1)) \xrightarrow{n \rightarrow \infty} \mathcal{N}(0, \text{Var}(X_1))$ .