



Chapitre 1 - le modèle probabiliste

Nous considérons un espace aléatoire. Nous notons Ω l'ensemble des résultats possibles : c'est un ensemble quelconque a priori

1) Evénements :

Un événement associé à l'expérience aléatoire est un ensemble de résultats dont on peut calculer la probabilité.

Notons $\mathcal{F}$ la collection de tous les événements.

Ainsi, $\mathcal{F} \subset \mathcal{P}(\Omega)$, l'ensemble des parties de Ω . Nous demandons en fait que $\mathcal{F}$ soit une tribu ou σ -algèbre sur Ω , i.e. $\mathcal{F}$ est une collection non vide de parties de Ω qui vérifie :

- $\forall A \in \mathcal{F}, A^c \in \mathcal{F}$ (stabilité par passage complémentaire)
- Si $(A_n)_{n \in \mathbb{N}}$ est une suite d'éléments de $\mathcal{F}$, alors $\bigcup_{n \in \mathbb{N}} A_n$ est encore dans $\mathcal{F}$ (stabilité par $\cup$ dénombrable)

Rq: une tribu contient forcément $\emptyset$ et Ω .

($\mathcal{F}$ non vide : $\exists A \in \mathcal{F}, A^c \in \mathcal{F}, A \cup A^c \in \mathcal{F}, \Omega^c \in \mathcal{F}$)

Rappel: 2 événements A et B sont dits **disjoints** ou **incompatibles** s'ils ne peuvent arriver simultanément, i.e. si $A \cap B$ est l'événement impossible $\emptyset$.

2) Probabilité

Soit Ω un ensemble équipé d'une tribu $\mathcal{F}$.

def: Une probabilité sur $(\Omega, \mathcal{F})$ est une application $P: \mathcal{F} \rightarrow [0, 1]$ et qui vérifie : i) $P(\Omega) = 1$

ii) Si $(A_n)_{n \in \mathbb{N}}$ est une suite d'éléments 2 à 2 disjoints,

$$\text{alors } P\left(\bigcup_{n \in \mathbb{N}} A_n\right) = \sum_{n \in \mathbb{N}} P(A_n)$$

Regardons dans le cas où Ω est fini. Dans ce cas, on prend toujours $\mathcal{F} = \mathcal{P}(\Omega)$

Soient A, B 2 événements disjoints. Appliquons ii) à la suite $A_0 = A, A_1 = B, A_2 = \dots = A_n = \dots = \emptyset$

iii) $\forall A, B$ disjoints, $P(A \cup B) = P(A) + P(B)$

La théorie moderne des probabilités est construite avec l'axiome ii), plus fort que iii), ne heurte pas l'intuition. Mais surtout, c'est une condition fondamentale qui permet de montrer de nombreux théorèmes limites.

Conclusion: Un espace de probabilité est un triplet $(\Omega, \mathcal{F}, P)$ où :

• $\mathcal{F}$ est une tribu sur Ω

[Axiomatique de Kolmogorov 1930]

• P est une probabilité sur $(\Omega, \mathcal{F})$

3) Tribu borelienne

Soit Ω un ensemble. Si $\mathcal{F}, \mathcal{G}$ sont 2 tribus sur Ω , on définit leur intersection $\mathcal{F} \cap \mathcal{G} = \{A \subset \Omega; A \in \mathcal{F} \text{ et } A \in \mathcal{G}\}$

C'est encore une tribu. (ew)

Plus généralement, si $(\mathcal{F}_i)_{i \in I}$ est une famille de tribus sur Ω indexée par l'ensemble I quelconque, alors $\bigcap_{i \in I} \mathcal{F}_i = \{A \subset \Omega; \forall i \in I, A \in \mathcal{F}_i\}$ est encore une tribu.

Prop-def: Soit $\mathcal{A}$ une collection de parties de Ω . Il existe une unique tribu $\sigma(\mathcal{A})$ qui vérifie:

i) $\mathcal{A} \subset \sigma(\mathcal{A})$

ii) $\forall \mathcal{F}$ tribu, $\mathcal{A} \subset \mathcal{F} \Rightarrow \sigma(\mathcal{A}) \subset \mathcal{F}$

La tribu $\sigma(\mathcal{A})$ s'appelle la tribu engendrée par $\mathcal{A}$, c'est la plus petite tribu qui contient $\mathcal{A}$.

preuve: Considérons la tribu suivante: $\bigcap_{\mathcal{F} \text{ tribu}, \mathcal{A} \subset \mathcal{F}} \mathcal{F}$ est une tribu, elle contient $\mathcal{A}$. Elle est contenue dans toute tribu qui contient $\mathcal{A}$.
C'est donc la plus petite tribu qui contient $\mathcal{A}$.

Rem: C'est le procédé utilisé pour définir le ss groupe engendré par une famille d'éléments.

(ss engendré...)

Cas de $\mathbb{R}$

La tribu borelienne sur $\mathbb{R}$, notée $\mathcal{B}(\mathbb{R})$, est la tribu engendrée par une des familles suivantes:

- les intervalles $[a, b]$, $a < b \in \mathbb{R}$
- $]a, b[$, $a, b \in \mathbb{R}$
- les semi intervalles $] -\infty, \lambda]$, $\lambda \in \mathbb{R}$
- $[\lambda, +\infty[$, $\lambda \in \mathbb{R}$

4) Les variables aléatoires

Def: Les variables aléatoires sont les nymphes des espaces de probas.

Une v.a. est une fonction de l'expérience aléatoire, c'est donc à priori une application $X: \Omega \rightarrow \mathbb{R}$.

Nous voulons pouvoir calculer la probabilité qu'une variable aléatoire tombe dans un intervalle arbitraire de $\mathbb{R}$.

Une CNS pour cela est de demander que X soit mesurable.

def: Soit $(\Omega, \mathcal{F}, P)$ un espace de probas. Une variable aléatoire X sur $(\Omega, \mathcal{F}, P)$ est une application $X: \Omega \rightarrow \mathbb{R}$ qui est mesurable de $(\Omega, \mathcal{F})$ dans $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$,

i.e., $\forall B \in \mathcal{B}(\mathbb{R}), X^{-1}(B) = \{\omega \in \Omega: X(\omega) \in B\} \in \mathcal{F}$.

Ainsi, si $B \in \mathcal{B}(\mathbb{R})$, alors $X^{-1}(B)$ est un événement, on le note simplement $\{X \in B\}$ ou même $X \in B$.

De manière générale, on s'efforce de simplifier les formules en enlevant la variable ω chaque fois qu'elle n'est pas indispensable.

Convention: On note les v.a. par des lettres majuscules X, Y, Z, U, V, W , et les valeurs correspondantes par des lettres minuscules x, y, z, u, v, w .

Plus généralement, on peut considérer des v.a. à valeurs $\mathbb{C}$, dans $\mathbb{R}^n$, - et en fait, dans n'importe quel ensemble équipé d'une tribu.

Une fonction $f: \mathbb{R} \rightarrow \mathbb{R}$ est dite borelienne si elle est mesurable de $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, dans $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, i.e., si $\forall B \in \mathcal{B}(\mathbb{R}), f^{-1}(B) = \{x \in \mathbb{R}: f(x) \in B\} \in \mathcal{B}(\mathbb{R})$.

ex: Si X est une v.a., et f fctn borelienne, alors $f(X)$ est encore une v.a.

Opérations sur les v.a.:

X v.a. αX v.a.

X, Y v.a. $X+Y$ v.a.
 XY

Si (X_n) suite de v.a., $\liminf X_n$, $\limsup X_n$ et $\lim X_n$ sont v.a.

En fait, on utilise le langage et les résultats de la théorie de la mesure.

Un espace E équipé d'une tribu $\mathcal{E}$ est appelé espace mesurable.

Si $(E, \mathcal{E})$ et $(F, \mathcal{F})$ sont 2 espaces mesurables, une fctn $f: E \rightarrow F$ est mesurable si $\forall F \in \mathcal{F}, f^{-1}(F) \in \mathcal{E}$.

Une proba P est une mesure (positive) de masse totale 1.

5) La loi d'une v.a.

Soit $(\Omega, \mathcal{F}, P)$ un espace de probabilité et soit X une v.a. définie sur Ω à valeurs ds $\mathbb{R}$.

Def: La loi de la v.a. X , notée P_X , est la mesure de probabilité sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ définie par:

$$\forall B \in \mathcal{B}(\mathbb{R}) \quad P_X(B) = P(X \in B) (= P(X^{-1}(B))) = P(\{\omega \in \Omega: X(\omega) \in B\})$$

Insistons sur le fait que la loi d'une v.a. est une probabilité définie sur l'espace des valeurs de la v.a.; dans notre cas, $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

On dit que 2 v.a. X et Y ont même loi si $P_X = P_Y$.

L'objet d'intérêt primordial, quand on étudie une variable aléatoire, c'est sa loi.

En effet, nous voulons évaluer les probabilités d'événements associés à la v.a.

"La loi, c'est fondamental!"

6) Espérance $(\Omega, \mathcal{F}, P)$

Une variable aléatoire $S: \Omega \rightarrow \mathbb{R}$ est dite 'étagée' si elle prend un nombre fini de valeurs. Elle s'écrit donc:

$$S = \sum_{i=1}^n a_i \cdot 1_{A_i} \quad \text{où } a_1, \dots, a_n \in \mathbb{R} \\ A_1, \dots, A_n \in \mathcal{F}$$

Son espérance est son intégrale par rapport à P , donnée par $E(S) = \int S dP = \sum_{i=1}^n a_i \cdot P(A_i)$

Si $X: \Omega \rightarrow [0, +\infty[$ est une v.a. positive ou nulle, éventuellement elle prend la valeur $+\infty$, on définit son espérance par:

$E(X) = \sup \{ E(S); S \text{ étagée}, 0 \leq S \leq X \}$. C'est aussi l'intégrale de X par rapport à P . On note $E(X) = \int X dP$. Attention, on peut avoir $E(X) = +\infty$

Si $X: \Omega \rightarrow \mathbb{R}$ est de signe quelconque, on écrit $X = X^+ - X^-$ où: $X^+ = \max(X, 0)$ partie positive

X^+ et X^- sont 2 v.a. positives. $X^- = -\min(X, 0)$ partie négative

Lorsque $E(X^+) < +\infty$ et $E(X^-) < \infty$, on pose: $E(X) = \int X dP = E(X^+) - E(X^-)$

Comme de plus, $|X| = X^+ + X^-$, on a équivalence: $E(X^+) < \infty, E(X^-) < \infty \Leftrightarrow E(|X|) < \infty$

Dans ce cas, son espérance est $E(X) = \int X dP$. On note $\mathcal{L}^1(\Omega, \mathcal{F}, P)$ l'espace des v.a. intégrables

7) Variance

Si X est une v.a. intégrable, on définit sa variance:

$$\text{Var}(X) = [E(X^2) - E(X)^2] \\ = E[(X - E(X))^2]$$

Amphi 2 - 12/09/25

1) La mesure de Lebesgue

Une mesure sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ est une application de $\mathcal{B}(\mathbb{R})$ dans $[0, +\infty[$ telle que $\mu(\emptyset) = 0$ et qui est σ additive:

$$\forall (B_n)_{n \in \mathbb{N}} \in \mathcal{B}(\mathbb{R})^{\mathbb{N}}, [\forall n \neq m, B_n \cap B_m = \emptyset \Rightarrow \mu(\bigcup_n B_n) = \sum_n \mu(B_n)]$$

Théorème: Il existe une unique mesure μ sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ qui vérifie:

$$\forall a, b \in \mathbb{R}, a < b, \mu([a, b]) = b - a$$

Cette mesure est appelée la mesure de Lebesgue et est notée λ

Il s'agit d'un résultat difficile! $B \in \mathcal{B}(\mathbb{R}) \Rightarrow \lambda(B)$

Idee de la preuve: Nous définissons pour A une partie quelconque de $\mathbb{R}$

$$\lambda^*(A) = \inf \left\{ \sum_{i \in I} b_i - a_i; (a_i)_{i \in I}, (b_i)_{i \in I} \text{ familles de réels tq: } \forall i \in I, a_i < b_i, A \subset \bigcup_{i \in I}]a_i, b_i[\right\}$$

λ^* n'est pas une mesure, il est sous additif mais pas σ additif

On montre que la restriction de λ^* à $\mathcal{B}(\mathbb{R})$ est σ additive et on prend $\lambda = \lambda^*|_{\mathcal{B}(\mathbb{R})}$

Premières propriétés

$$\bullet \forall x \in \mathbb{R}, \lambda(\{x\}) = 0$$

$$\bullet \text{ Si } A \subset \mathbb{R} \text{ est dénombrable, alors } \lambda(A) = 0 = \sum_{x \in A} \lambda(\{x\}) \text{ par } \sigma \text{ additivité}$$

Il existe des boreliens C de $\mathbb{R}$ non dénombrables tq $\lambda(C) = 0$

2) Loi uniforme sur $[0, 1]$. Nous souhaitons construire un modèle mathématique de l'expérience qui consiste à tirer un nb au hasard dans $[0, 1]$, uniformément.

Nous prenons $\Omega = [0, 1]$ muni de sa tribu borelienne $\mathcal{B}([0, 1])$

$\mathcal{B}([0, 1])$ est, au choix, la tribu σ (I intervalle, $I \subset [0, 1]$) engendrée par les sous-intervalles de $[0, 1]$ ou bien la tribu trace de $\mathcal{B}(\mathbb{R})$ sur $[0, 1]: \{B \in \mathcal{B}(\mathbb{R}); B \subset [0, 1]\} = \{B \cap [0, 1]: B \in \mathcal{B}(\mathbb{R})\}$ et muni de la probabilité $P = \lambda|_{\mathcal{B}([0, 1])}$

Nous définissons alors une v.a. X en posant:

$$X: \Omega \rightarrow [0, 1]$$

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

Je dis que X est un nb choisi au hasard unif^{ts} ds $[0, 1]$

En fait la loi P_X de X est $\lambda|_{[0, 1]}$: En effet, $\forall B \in \mathcal{B}([0, 1]), P_X(B) = P(X \in B) = P(\{\omega: X(\omega) \in B\}) = \lambda(B)$ donc $P_X = \lambda$.

3) Lois à densité: Une loi de probabilité μ sur $\mathbb{R}$ est dite à densité s'il existe une fonction f borelienne positive tq:

$$\forall A \in \mathcal{B}(\mathbb{R}), \mu(A) = \int_A f d\lambda = \int_A f(x) dx (\neq)$$

La fonction f est la densité, ou plutôt, une densité de μ . Une densité est une fonction borelienne positive qui

$$\text{vérifie } \int_{\mathbb{R}} f d\lambda = \int_{\mathbb{R}} f(x) dx = 1 \quad (0)$$

Inversement, étant donnée une fonction f borélienne ≥ 0 qui vérifie (0), on peut construire une loi de probabilité sur $\mathbb{R}$ associée à f via la formule ν .

Exemples fondamentaux:

• loi uniforme sur un intervalle $[a, b]$, c'est la loi à densité, de densité $\frac{1}{b-a} \mathbb{1}_{[a, b]}(x)$ (par rapport à la mesure de Lebesgue)

On la note $\mathcal{U}([a, b])$

• loi exponentielle de paramètre $\alpha > 0$

loi de densité $\alpha e^{-\alpha x} \mathbb{1}_{\mathbb{R}^+}(x)$ notée $\text{Exp}(\alpha)$

• loi normale ou Gaussienne de paramètres m, σ^2

loi de densité $\frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2} \frac{(x-m)^2}{\sigma^2}\right)$ notée $\mathcal{N}(m, \sigma^2)$

4) Classification des lois sur $\mathbb{R}$

Un point x est un atome de la loi μ si $\mu(\{x\}) > 0$

On dit que la loi μ est diffuse si elle n'a pas d'atomes.

On distingue 3 types "purs" de lois.

• loi discrètes: On dit que μ est discrète s'il existe un ensemble D fini ou dénombrable tq $\mu(D) = 1$.

On a alors $\mu = \sum_{x \in D} p_x \delta_x$

• lois à densité: On dit que μ est à densité s'il existe f borélienne ≥ 0 tq $\forall A \in \mathcal{B}(\mathbb{R}), \mu(A) = \int_A f d\lambda$

• lois singulières: On dit que μ est singulière si elle est diffuse et s'il existe un ensemble C tq $\mu(C) = 1$ et $\lambda(C) = 0$

Toute combinaison convexe de ces 3 types est une loi de proba sur $\mathbb{R}$. En fait toute mesure de probas sur $\mathbb{R}$ se décompose comme

$\mu = \alpha_d \mu_d + \alpha_a \mu_a + \alpha_s \mu_s$ où $\alpha_d, \alpha_a, \alpha_s \in [0, 1]$ $\alpha_d + \alpha_a + \alpha_s = 1$ et cette écriture est unique.
 $\uparrow$ $\uparrow$ $\uparrow$
 discrète densité singulière

5) Sur le choix de l'intégrale

Supposons que nous disposions d'une théorie de l'intégration qui permet d'intégrer les fctns continues positives.

Étant donnée une fonction $f \geq 0$ telle que $\int_{\mathbb{R}} f(x) dx = 1$. Nous pouvons définir $\mu(A) = \int_A f(x) dx$ pour toutes les parties A autorisées par la théorie.

Exemple: $\int$ Riemann, fonctions continues.

Dans ce cas, la probabilité est définie sur tous les intervalles, les $\cup$ finies d'intervalles.

Mais le choix de prédilection pour la théorie des probas c'est $\int$ de Lebesgue: $\int_A f d\lambda$
 $\uparrow$ $\uparrow$ $\uparrow$
 borélienne mesure de Lebesgue $\mathcal{B}(\mathbb{R})$

Resume de la construction:

Une fonction borélienne s est dite étagée si elle prend un nombre fini de valeurs, elle est de la forme

$s = \sum_{i=1}^n a_i \mathbb{1}_{B_i}$ où $a_1, \dots, a_n \in \mathbb{R}$ et $B_1, \dots, B_n \in \mathcal{B}(\mathbb{R})$

Ensuite, si f est une fonction borélienne ≥ 0 , on définit $\int f d\lambda = \sup \left\{ \int s d\lambda, s \text{ étagée}; 0 \leq s \leq f \right\}$

Attention, le sup peut être ∞

Soit maintenant f borélienne de signe quelconque. On dit que f est λ intégrale si $\int |f| d\lambda < +\infty$.

Dans ce cas, on définit $\int f d\lambda = \int f^+ d\lambda - \int f^- d\lambda$ avec $f^+ = \max(f, 0)$ la partie ≥ 0
 et $f^- = -\min(f, 0)$

et on note $\mathcal{L}^1(\mathbb{R}, \lambda)$ l'ensemble des fonctions λ intégrables.

6) Formule de transfert.

Lorsqu'on travaille avec des v.a., on effectue tout d'abord des manipulations dans l'espace $(\Omega, \mathcal{F}, P)$. Lorsqu'il s'agit de faire intervenir les lois spécifiques à des v.a., on transfère le calcul dans $\mathbb{R}$.

PROP. [formule de transfert $(\Omega, \mathcal{F}, P)$]

Soit X une v.a. $\Omega \rightarrow \mathbb{R}$ et $f: \mathbb{R} \rightarrow \mathbb{R}$ borélienne tq $f(X)$ soit dans $\mathcal{L}^1(\Omega, \mathcal{F}, P)$. Alors

$$\int_{\Omega} f(X) dP = \int_{\mathbb{R}} f dP_X \text{ ou encore } \int_{\omega \in \Omega} f(X(\omega)) dP(\omega) = \int_{x \in \mathbb{R}} f(x) dP_X(x)$$

$\uparrow$ loi de X

preuve: Commençons par $f = 1_B$ où $B \in \mathcal{B}(\mathbb{R})$

$$\int_{\Omega} 1_B(x) dP = \int_{\Omega} 1_B(X(\omega)) dP(\omega) = P(\{\omega; X(\omega) \in B\}) = P(X \in B) = P_X(B) = \int_B dP_X = \int_{\mathbb{R}} 1_B dP_X$$

Si s est étagée, s vérifie la formule de transfert.

lemme: Soit $\mathcal{L}$ une collection de fctns boréliennes tq:

i) $\forall B \in \mathcal{B}(\mathbb{R}), 1_B \in \mathcal{L}$

ii) $\forall f, g \in \mathcal{L} \quad \forall \alpha, \beta > 0 \quad \alpha f + \beta g \in \mathcal{L}$

iii) $\mathcal{L}$ est stable par limite croissante: si (f_n) est une suite de $\mathcal{L}$ $\nearrow, f_n \leq f_{n+1}$ et $f_n \xrightarrow{n \rightarrow +\infty} f$, alors $f \in \mathcal{L}$
Alors $\mathcal{L}$ contient ttes les fonctions boréliennes ≥ 0 .

Application: $(\Omega, \mathcal{F}, P)$ espace probas, X une v.a. intégrable de $\Omega \rightarrow \mathbb{R}$

Alors $E(X) = \int_{\Omega} X dP = \int_{\Omega} X(\omega) dP(\omega)$

Transfert dans $\mathbb{R} = \int_{\mathbb{R}} \text{Id}_{\mathbb{R}} dP_X = \int_{x \in \mathbb{R}} x dP_X(x)$

Supposons que la loi de X est μ , proba sur $\mathbb{R}$. On note cela $X \sim \mu$.

Alors $P_X = \mu$ et on a: $E(X) = \int_{\mathbb{R}} x d\mu(x)$

Si de plus μ a une densité f , il vient $E(X) = \int_{x \in \mathbb{R}} x f(x) d\lambda(x)$

De même pour la variance:

Si X est intégrable, de loi μ , qui a une densité f .

$\text{Var}(X) = \int_{\mathbb{R}} (x-m)^2 d\mu(x) = \int_{\mathbb{R}} (x-m)^2 f(x) d\lambda(x)$ où $m = E(X)$

7) Méthode de la fctn muette

Supposons que l'on ait construit une v.a. X par un certain procédé.

Nous voulons calculer la loi de X . C'est une mesure de probas sur $\mathbb{R}$.

En tant que telle, elle est déterminée par son action sur les fctns continues bornées.

prop: Soient μ_1, μ_2 2 mesures de probas sur $\mathbb{R}$ tq pour tte fonction $f: \mathbb{R} \rightarrow \mathbb{R}$ continue bornée $\int f d\mu_1 = \int f d\mu_2$

Alors $\mu_1 = \mu_2$.

Pour trouver la loi de X , nous prenons une fctn $f: \mathbb{R} \rightarrow \mathbb{R}$ continue bornée et nous évaluons $E(f(X)) = \int f(x) dP$

Par la formule de transfert, $= \int_{\mathbb{R}} f(x) dP_X(x)$. On s'efforce de mettre le résultat sous cette forme, et on peut lire la loi de X

Exemple: $X \sim U([0,1])$, $K > 1$, loi de X^K ?

Soit $f: \mathbb{R} \rightarrow \mathbb{R}$ continue bornée. Calculons $E(f(X^K))$.

$$\begin{aligned} E(f(X^K)) &= \int_{\Omega} f(X^K) dP = \int_{\mathbb{R}} f(x^K) dP_X(x) \quad , P_X \text{ densité } 1_{[0,1]} \\ &= \int_{[0,1]} f(x^K) dx \quad (y = x^K, x = y^{1/K}) \\ &\quad dx = \frac{1}{K} y^{1/K - 1} dy \\ &= \int_{[0,1]} f(y) \frac{1}{K} y^{1/K - 1} dy \\ &\quad \text{lit le résultat} \end{aligned}$$

Conclusion: X^K a la loi de densité $\frac{1}{K} y^{1/K - 1} 1_{[0,1]}(y)$

Chapitre 3 - Indépendance :

1) Tribu engendrée par une variable aléatoire.

Soit $(\Omega, \mathcal{F}, P)$ un espace de probabilité

Definition: Une sous-tribu $\mathcal{G}$ de $\mathcal{F}$ est une tribu qui est incluse dans $\mathcal{F}$, i.e. $\mathcal{G} \subset \mathcal{F} : \forall A \in \mathcal{G}, A \in \mathcal{F}$

Def: Soit X une v.a. de $(\Omega, \mathcal{F}, P)$ dans $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. La tribu engendrée par la v.a. X , notée $\sigma(X)$, est la tribu $\sigma(X) = \sigma(X^{-1}(B), \mathcal{B}(\mathbb{R}))$

Comme X est une v.a., $\forall B \in \mathcal{B}(\mathbb{R}), X^{-1}(B) \in \mathcal{F}$ et donc $\sigma(X)$ est une sous-tribu de $\mathcal{F}$

2) Définitions fondamentales

1) n événements $A_1, \dots, A_n$ de $\mathcal{F}$ sont indépendants ssi pour toute partie finie I de $\{1, \dots, n\}$, on a

$$P\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} P(A_i) \quad (2^n \text{ conditions})$$

Attention, être indépendant $\neq$ être indépendants 2 à 2

2) n v.a. $X_1, \dots, X_n : \Omega \rightarrow \mathbb{R}$ sont dites indépendantes ssi

$$\forall B_1 \in \mathcal{B}(\mathbb{R}), \dots, \forall B_n \in \mathcal{B}(\mathbb{R}), P(X_1 \in B_1, \dots, X_n \in B_n) = P(X_1 \in B_1) \times \dots \times P(X_n \in B_n)$$

3) n sous-tribus sont indépendantes ssi $\mathcal{F}_1, \dots, \mathcal{F}_n$ de $\mathcal{F}$

$$\forall A_1 \in \mathcal{F}_1, \dots, \forall A_n \in \mathcal{F}_n, P(A_1 \cap \dots \cap A_n) = P(A_1) \times \dots \times P(A_n)$$

En fait, on peut récrire la 1^{ère} définition à l'aide de la 2^{ème}. En effet, les événements $A_1, \dots, A_n$ sont indépendants $\Leftrightarrow$ les v.a. $1_{A_1}, \dots, 1_{A_n}$ sont indep et la 2^{ème} à l'aide de la 3^{ème}

Les v.a. $X_1, \dots, X_n$ sont indep ssi les tribus associées $\sigma(X_1), \dots, \sigma(X_n)$ sont indépendantes

C'est donc la définition avec les tribus qui est la plus générale.

Elle permet de traiter des situations plus complexes, par exemple le cas de v.a. à valeurs dans des espaces quelconques.

3) Variables aléatoires générales

Soit E un espace quelconque équipé d'une tribu $\mathcal{E}$. Une v.a. à valeurs dans E est une application X qui est mesurable de $(\Omega, \mathcal{F})$ dans $(E, \mathcal{E})$, i.e., $\forall B \in \mathcal{E}, X^{-1}(B) \in \mathcal{F}$.

La loi de X est la mesure de probas sur $(E, \mathcal{E})$ définie par

$$\forall B \in \mathcal{E}, P_X(B) = P(X \in B) = P(X^{-1}(B))$$

On peut ainsi construire des vecteurs aléatoires et même des fonctions aléatoires

Exemple: X, Y 2 v.a. à valeurs dans $\mathbb{R}$. On pose $Z = (X, Y)$. Alors Z est un vecteur aléatoire de $\mathbb{R}^2$. On dit que la loi de Z admet une densité

$$\text{si il existe } f : \mathbb{R}^2 \rightarrow [0, +\infty[\text{ tq } \forall B \in \mathcal{B}(\mathbb{R}^2), P_Z(B) = \iint_B f(x, y) dx dy$$

4) Lois Produits

Soient $(E_1, \mathcal{E}_1, \mu_1), \dots, (E_n, \mathcal{E}_n, \mu_n)$ n espaces de probabilité.

Q: existe-t-il un espace de probas $(\Omega, \mathcal{F}, P)$ sur lequel sont définies n v.a. $X_1, \dots, X_n$ telles que

$$\bullet X_1, \dots, X_n \text{ sont indep. } \bullet \forall i \in \{1, \dots, n\} X_i \text{ est à valeurs dans } E_i \text{ et a pour loi } \mu_i$$

Considérons le produit cartésien $E = E_1 \times \dots \times E_n$. Nous le munissons de la tribu produit des tribus $\mathcal{E}_1, \dots, \mathcal{E}_n$ qui est la tribu engendrée par les rectangles mesurables $B_1 \times \dots \times B_n$ $B_1 \in \mathcal{E}_1, \dots, B_n \in \mathcal{E}_n$

Cette tribu est notée $\mathcal{E}_1 \otimes \mathcal{E}_2 \otimes \dots \otimes \mathcal{E}_n$

Thm: Il existe une unique mesure de probabilité μ sur $(E_1 \times \dots \times E_n, \mathcal{E}_1 \otimes \dots \otimes \mathcal{E}_n)$ qui vérifie: $\forall B_1 \in \mathcal{E}_1, \dots, \forall B_n \in \mathcal{E}_n, \mu(B_1 \times \dots \times B_n) = \mu_1(B_1) \times \dots \times \mu_n(B_n)$

Cette proba μ est appelée la mesure produit de $\mu_1, \dots, \mu_n$ et est notée $\mu = \mu_1 \otimes \dots \otimes \mu_n$

Prenons alors $\Omega = E_1 \times \dots \times E_n, \mathcal{F} = \mathcal{E}_1 \otimes \dots \otimes \mathcal{E}_n$, et $P = \mu_1 \otimes \dots \otimes \mu_n$ et définissons X_i pour $1 \leq i \leq n$ par $X_i : \Omega \rightarrow E_i$

projection sur la i -ème composante

$$\omega = (\omega_1, \dots, \omega_n)$$

Je dis que $X_1, \dots, X_n$ sont indep et que X_i a pour loi μ_i pour $1 \leq i \leq n$

$$\rightarrow X_i(\omega) = \omega_i$$

preuve: Soit $i \in \{1, \dots, n\}$ et regardons la loi de X_i

$$\forall B_i \in \mathcal{E}_i, P_{X_i}(B_i) = P(X_i \in B_i) = \mu(\{\omega = (\omega_1, \dots, \omega_n) \in \Omega : \omega_i \in B_i\}) = \mu(E_i \times \dots \times E_{i-1} \times B_i \times E_{i+1} \times \dots \times E_n)$$

Vérifions que $X_1, \dots, X_n$ sont indep. soient $B_1 \in \mathcal{E}_1, \dots, B_n \in \mathcal{E}_n$. $= \mu(E_1) \dots \mu_{i-1}(E_{i-1}) \mu_i(B_i) \mu_{i+1}(E_{i+1}) \dots \mu_n(E_n) = \mu_i(B_i)$ et donc $P_{X_i} = \mu_i$

$$\text{On a } P(X_1 \in B_1, \dots, X_n \in B_n) = \mu(\{\omega = (\omega_1, \dots, \omega_n) : \omega_1 \in B_1, \dots, \omega_n \in B_n\}) = \mu(B_1 \times \dots \times B_n)$$

Application au cas où $(E_i, \mathcal{E}_i) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ $1 \leq i \leq n$. Soient $\mu_1, \dots, \mu_n$ n lois de probas sur $\mathbb{R}$

Alors il existe $(\Omega, \mathcal{F}, P)$ sur lequel sont définies n v.a. indépendantes $X_1, \dots, X_n$ à valeurs dans $\mathbb{R}$, de lois où $\mu_i = \dots = \mu_n = \mu$. Étant donnée une loi μ sur $\mathbb{R}$, il existe $(\Omega, \mathcal{F}, P)$ sur lequel $\mu_1, \dots, \mu_n$ sont définies n.v.a. $X_1, \dots, X_n$ indépendantes identiquement distribuées (i.i.d en abrégé) de loi μ .

5) Regroupement par blocs

Théorème: Soient $\mathcal{F}_1, \dots, \mathcal{F}_n$ n tribus indépendantes. Soient I, J deux parties de $\{1, \dots, n\}$ disjointes. Alors les tribus $\mathcal{J} = \sigma(\mathcal{F}_i, i \in I)$ et $\mathcal{J}' = \sigma(\mathcal{F}_j, j \in J)$ sont indépendantes.

pu: il s'agit de montrer que $\forall D \in \mathcal{J}, \forall E \in \mathcal{J}', P(D \cap E) = P(D)P(E)$

Soit D un élément de la forme $D = \bigcap_{i \in I} D_i$, où $D_i \in \mathcal{F}_i$ pour $i \in I$. Si $P(D) = 0$, rien à montrer

Supposons $P(D) > 0$ et définissons, pour $E \in \mathcal{J}'$ $P_1(E) = P(E|D), P_2(E) = P(E)$

Si E est de la forme $E = \bigcap_{j \in J} E_j$ où $E_j \in \mathcal{F}_j$

$$\begin{aligned} \text{Alors } P(D \cap E) &= P\left(\bigcap_{i \in I} D_i \cap \bigcap_{j \in J} E_j\right) = \prod_{i \in I} P(D_i) \times \prod_{j \in J} P(E_j) \\ &= P\left(\bigcap_{i \in I} D_i\right) P\left(\bigcap_{j \in J} E_j\right) = P(D)P(E) \end{aligned}$$

Lemme: (unicité de l'extension)

Soient P_1, P_2 2 probas sur Ω . Supposons qu'il existe une collection C d'événements tq:

$$i) \sigma(C) = \mathcal{F}$$

$$ii) \forall A, B \in C, A \cap B \in C$$

$$iii) \forall C \in C, P_1(C) = P_2(C)$$

$$\text{Alors } P_1 = P_2$$

preuve du lemme: on introduit la collection $\mathcal{G} = \{A \in \mathcal{F} : P_1(A) = P_2(A)\}$. $C \subset \mathcal{G}$, $\mathcal{G}$ stable par différence

union décroissante, on mq $\mathcal{G}$ est tribu. D'où $\mathcal{F} \subset \mathcal{G}$.

On a donc $P_1(E) = P_2(E)$ pour E de la forme $E = \bigcap_{j \in J} E_j$. On prend pour C la collection de ces ensembles.

On a $\sigma(C) = \mathcal{J}$, C stable par $\cap$. Par le lemme, on a $P_1(E) = P_2(E) \forall E \in \mathcal{J}$

Donc on a $P(D \cap E) = P(D)P(E)$ pour $\forall E \in \mathcal{J}$ et tout D de la forme $D = \bigcap_{i \in I} D_i$

Pour cela, on fixe E dans $\mathcal{J}$ et on itère l'argument précédent sur D .

Proposition: $\mathcal{F}_1, \dots, \mathcal{F}_n$ n tribus indépendantes. $I_1, \dots, I_n$ n parties de $\{1, \dots, n\}$ à 2 disjointes. Alors les tribus $\sigma(\mathcal{F}_i, i \in I_k), 1 \leq k \leq r$ sont indépendantes

preuve: récurrence sur n .

Corollaire: $X_1, \dots, X_n$ n.v.a. indépendants. $I_1, \dots, I_n$ n parties disjointes de $\{1, \dots, n\}$

Les vecteurs $(X_i, i \in I_k), 1 \leq k \leq r$ sont indépendants

6) Espérance et indépendance

Proposition: Soient $X, Y: \Omega \rightarrow \mathbb{R}$ deux variables aléatoires. On a équivalence entre: i) X, Y sont indépendantes

ii) Si f, g sont 2 fonctions bornées, $\mathbb{R} \rightarrow \mathbb{R}$ tq $E(|f(X)|) < \infty$ et $E(|g(Y)|) < \infty$

$$\text{ors } E(|f(X)g(Y)|) < \infty \text{ et } E(f(X)g(Y)) = E(f(X))E(g(Y))$$

pu: ii) $\Rightarrow$ i) Soient $A, B \in \mathcal{B}(\mathbb{R})$

$$\text{Appliquons ii) à } f = 1_A, g = 1_B. E(f(X)g(Y)) = E(1_A(X)1_B(Y)) = P(X \in A, Y \in B) = E(f(X))E(g(Y)) = P(X \in A)P(Y \in B)$$

i) $\Rightarrow$ ii) si $f = 1_A$ et $g = 1_B$, c'est vrai par def de l'indép. C'est en core vrai pour les fonctions étagées.

Fixons $g = 1_B$ et considérons $\mathcal{L} = \{f \text{ les fctrs qui vérifient ii}\}$.

$\bullet$ $\mathcal{L}$ contient $1_A, A \in \mathcal{B}(\mathbb{R})$. $\bullet$ $\mathcal{L}$ est stable par limite $\nearrow$

$\mathcal{L}$ stable par CL à ∞ et $f \geq 0$. Donc $\mathcal{L}$ contient toutes les fonctions bornées ≥ 0

On fixe ensuite f et on réitère l'argument sur g . On décompose ensuite $f = f^+ + f^-$, $g = g^+ - g^-$
 $\geq 0 \geq 0$

Cor: Soient X, Y 2 v.a. indépendantes tq $E(|X|) < \infty$ et $E(|Y|) < \infty$. Alors $E(|XY|) < \infty$ et $E(XY) = E(X)E(Y)$.

7) Covariance

X, Y 2 v.a. $(\Omega, \mathcal{F}, P)$ intégrables XY intégrable

Def: la covariance de X, Y est $\text{cov}(X, Y) = E((X - E(X))(Y - E(Y)))$

Ainsi si X, Y sont indep alors $\text{cov}(X, Y) = 0 \Rightarrow E(XY) = E(X)E(Y)$

Mais la réciproque est fautive.

Chapitre 4

Soit $(\Omega, \mathcal{F}, P)$ un espace de probabilité

1) Modes de convergence

Soit $(X_n)_{n \in \mathbb{N}}$ une suite de v.a. définies sur $(\Omega, \mathcal{F}, P)$

Soit également X une v.a. supplémentaire définie sur $(\Omega, \mathcal{F}, P)$

Définition: La suite $(X_n)_{n \in \mathbb{N}}$ converge presque sûrement vers X , ce que l'on note $X_n \xrightarrow{p.s.} X$ ssi

$$P(\{\omega \in \Omega; \lim_{n \rightarrow +\infty} X_n(\omega) = X(\omega)\}) = 1$$

Soit $r \geq 1$. La suite $(X_n)_{n \in \mathbb{N}}$ converge vers X dans $\mathcal{L}^r(\Omega, P)$, ce que l'on note $X_n \xrightarrow{L^r} X$

$$\text{ssi } \lim_{n \rightarrow +\infty} E(|X_n - X|^r) = 0$$

La suite $(X_n)_{n \in \mathbb{N}}$ converge vers X en probabilité, ce que l'on note $X_n \xrightarrow{P} X$, ssi

$$\forall \varepsilon > 0 \lim_{n \rightarrow +\infty} P(|X_n - X| > \varepsilon) = 0$$

2) Liens entre convergences

La convergence la plus difficile est la convergence p.s.

Prop: Si $X_n \xrightarrow{p.s.} X$, alors $X_n \xrightarrow{P} X$

preuve: Supposons $X_n \xrightarrow{p.s.} X$. Alors $P(\{\omega \in \Omega; \lim_{n \rightarrow +\infty} X_n(\omega) = X(\omega)\}) = 1$

$$\Leftrightarrow P(\lim_{n \rightarrow +\infty} X_n = X) = 1 \Leftrightarrow P(\forall \varepsilon > 0 \exists N \in \mathbb{N} \forall n \geq N |X_n - X| < \varepsilon) = 1$$

$$\Leftrightarrow P(\bigcap_{\varepsilon > 0} \bigcup_{n \in \mathbb{N}} \bigcap_{n \geq n} \{|X_n - X| < \varepsilon\}) = 1 \quad (\mathbb{Q}: \text{dénombrable})$$

$$\Leftrightarrow \forall \varepsilon > 0 \quad P(\bigcup_{n \in \mathbb{N}} \bigcap_{n \geq n} \{|X_n - X| < \varepsilon\}) = 1 \quad [\text{une intersection dénombrable d'événements a proba. 1 ssi ils ont tous proba. 1}]$$

$$\Leftrightarrow \forall \varepsilon > 0, P(\left(\bigcup_{n \in \mathbb{N}} \bigcap_{n \geq n} \{|X_n - X| < \varepsilon\}\right)^c) = 0 \quad [P(A) = 1 \Leftrightarrow P(A^c) = 0]$$

$$\Leftrightarrow \forall \varepsilon > 0 \quad P(\underbrace{\bigcap_{n \in \mathbb{N}} \bigcup_{n \in \mathbb{N}} \{|X_n - X| \geq \varepsilon\}}_{\text{suite décroissante d'éléments}}) = 0$$

$$\Leftrightarrow \forall \varepsilon > 0 \quad \lim_{N \rightarrow +\infty} P(\bigcup_{n \geq N} \{|X_n - X| \geq \varepsilon\}) = 0 \quad [\text{car si } A_n \searrow \quad P(\bigcap_n A_n) = \lim_n P(A_n), \quad A_n = \bigcup_{n \geq N} \{|X_n - X| \geq \varepsilon\}]$$

$$\Rightarrow \forall \varepsilon > 0 \quad \lim_{N \rightarrow +\infty} P(|X_N - X| > \varepsilon) = 0 \quad \text{i.e. } X_n \xrightarrow{P} X$$

Ensuite: Prop: Soit $r \geq 1$. Si $X_n \xrightarrow{L^r} X$ alors $X_n \xrightarrow{P} X$

preuve: Soit $\varepsilon > 0$. Écrivons

$$E(|X_n - X|^r) = \int_{\Omega} |X_n - X|^r dP \geq \int_{\{|X_n - X| > \varepsilon\}} |X_n - X|^r dP \geq \varepsilon^r P(|X_n - X| > \varepsilon)$$

$$\geq \int \mathbb{1}_{|X_n - X| > \varepsilon} \mathbb{E}^r dP = \mathbb{E}^r(\mathbb{1}_{|X_n - X| > \varepsilon})$$

$$\text{d'où } P(|X_n - X| > \varepsilon) \leq \frac{1}{\varepsilon^r} \mathbb{E}(|X_n - X|^r) \xrightarrow{n \rightarrow +\infty} 0$$

Prop: [Compatibilité avec l'addition:]

Si $X_n \xrightarrow{*} X$, $Y_n \xrightarrow{*} Y$, alors $X_n + Y_n \xrightarrow{*} X + Y$
 où $*$ = p.s, L^r ou P (avec $*$ le même choix partout).

indic: $P(|X_n + Y_n - (X + Y)| > \varepsilon) \leq P(|X_n - X| > \frac{\varepsilon}{2}) + P(|Y_n - Y| > \frac{\varepsilon}{2})$

Nous pouvons également traduire dans le langage des v.a. les théorèmes classiques de théorie de la mesure.

ex: Thm convergence dominée:

Soit (X_n) une suite de v.a. tq:

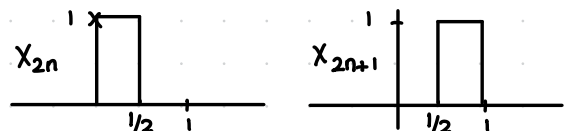
i) $\exists$ une v.a. X telle que $X_n \xrightarrow{p.s.} X$

ii) $\exists$ une v.a. $Y > 0$ intégrable tq $|X_n| \leq Y$ p.s.

Alors $X_n \xrightarrow{L^1} X$, et en particulier: $\lim_{n \rightarrow \infty} \mathbb{E}(X_n) = \mathbb{E}(X)$

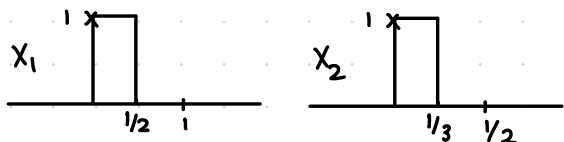
3) Exemples:

On se place sur $(\Omega, \mathcal{F}, P) = ([0, 1], \mathcal{B}([0, 1]), \lambda|_{[0, 1]})$, λ mesure de Lebesgue

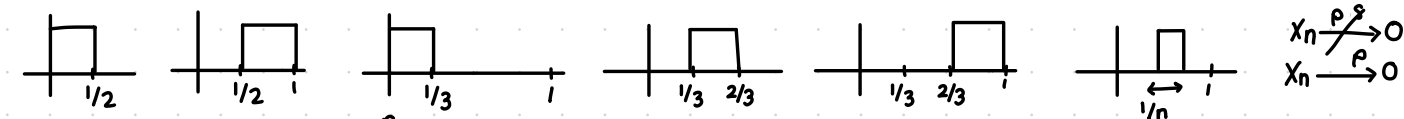


$\leadsto$ ne converge en aucun sens (ex)

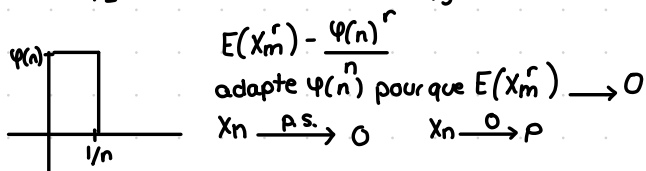
$$X_n \xrightarrow[P]{p.s.} 0$$



$$\times L^r$$



$$\begin{matrix} X_n \xrightarrow{P} 0 \\ X_n \not\xrightarrow{p.s.} 0 \end{matrix}$$



$$\mathbb{E}(X_n^r) = \frac{\psi(n)^r}{n}$$

adapte $\psi(n)$ pour que $\mathbb{E}(X_n^r) \rightarrow 0$

$$X_n \xrightarrow{p.s.} 0 \quad X_n \xrightarrow{0} P$$

4) Indépendance pour une suite infinie

Def: Une suite infinie de tribus $(\mathcal{F}_n)_{n \in \mathbb{N}}$ est dite indépendante (resp. d'événements, de v.a.) ssi toute famille extraite de la suite est indépendante.

Rem: analogie avec la liberté en algèbre linéaire.

Théorème: regroupement par blocs infinis

Soit $(X_n)_{n \in \mathbb{N}}$ une suite de v.a. indépendantes.

Soit $(I_k, k \in \mathbb{N})$ des parties de $\mathbb{N}$ deux à deux disjointes. Alors les tribus $\mathcal{F}_k = \sigma(X_n, n \in I_k)$, $k \in \mathbb{N}$ sont indépendantes.

5) Le modèle analytique pour le jeu de pile ou face

Nous voudrions construire un espace de probas $(\Omega, \mathcal{F}, P)$ sur lequel est définie une suite de v.a. $(X_n)_{n \in \mathbb{N}}$ de loi $\text{Ber}(1/2)$ et qui sont toutes indépendantes.

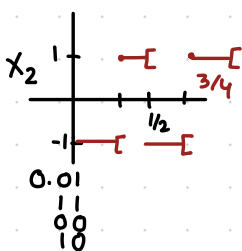
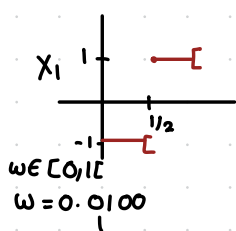
Nous prenons $(\Omega, \mathcal{F}, P) = ([0, 1], \mathcal{B}([0, 1]), \lambda)$, λ mesure de Lebesgue

A un membre réel ω de $[0, 1[$, nous associons son développement dyadique ou binaire:

$$\omega = \sum_{n \geq 1} \frac{\omega_n}{2^n} \text{ où } \omega_n \in \{0, 1\}$$

L'écriture est unique si on exclut les développements égaux à 1 à partir d'un certain rang.

Nous définissons alors, pour $n \geq 1$, $\forall \omega \in [0,1[$ $X_n(\omega) = 2\omega_{n-1} \in \{-1,1\}$. Je dis que la suite (X_n) répond à la question.



De manière générale, $X_n^{-1}(\{1\})$ est l'union de 2^{n-1} intervalles dyadiques de longueur $\frac{1}{2^n}$

$$\text{d'où } P(X_n = +1) = \lambda(\{\omega : X_n(\omega) = +1\}) = \frac{2^{n-1}}{2^n} = \frac{1}{2}$$

Mq les $(X_n)_{n \geq 1}$ sont indépendantes. Soient $\varepsilon_1, \dots, \varepsilon_n$ dans $\{-1,1\}$ et regardons

$$P(X_1 = \varepsilon_1, \dots, X_n = \varepsilon_n)$$

$$= \lambda(\{\omega \in [0,1[: \omega_1 = \frac{1+\varepsilon_1}{2}, \dots, \omega_n = \frac{1+\varepsilon_n}{2}\}) = \lambda(\{\omega : \frac{1+\varepsilon_1}{2} + \dots + \frac{1+\varepsilon_n}{2^{2^n}} \leq \omega < \frac{1+\varepsilon_1}{2} + \dots + \frac{1+\varepsilon_n}{2^n} + \frac{1}{2^n}\})$$

$$= \frac{1}{2^n} = P(X_1 = \varepsilon_1) \times \dots \times P(X_n = \varepsilon_n) \cdot \text{Nous possédons une suite } \infty \text{ de pièces indépendantes}$$

6) Marche aléatoire sur les entiers

• démarre de 0

• lance une pièce

• pile face: 1 pas à droite / gauche

• il répète l'opération à partir de la nouvelle position.

Notons S_n la position du marcheur après n pas.

Formellement, $S_0 = 0$ $S_n = X_1 + \dots + X_n$, où $X_1, \dots, X_n$ suite iid $\text{Ber}(1/2)$ sur $\{-1,1\}$



Quelle est la probabilité pour que l'allumeur repasse par le reverber original?

$$\text{Regardons, c'est } P(\exists n \in \mathbb{N}^* S_n = 0) = P(\bigcup_{n \in \mathbb{N}^*} \{S_n = 0\})$$

$$= P(\{S_2 = 0\} \cup (\{S_4 = 0\} \setminus \{S_2 = 0\}) \cup \dots)$$

$$= P(\bigcup_{n \in \mathbb{N}} (\{S_{2n} = 0\} \setminus (\bigcup_{1 \leq k \leq n} \{S_{2k} = 0\}))) = P(\bigcup_{n \geq 1} \{S_2 \neq 0, \dots, S_{2n-2} \neq 0, S_{2n} = 0\})$$

$$\text{Par } \sigma \text{ additivité: } = \sum_{n \geq 1} P(S_2 \neq 0, \dots, S_{2n-2} \neq 0, S_{2n} = 0)$$

Calculons d'abord $P(S_{2n} = 0)$

$$\text{Compte nb de chemins } \binom{2n}{0} \rightarrow \binom{2n}{0} \begin{cases} \alpha + \beta = \text{nb pas} = 2n \\ \alpha - \beta = \text{pas finale} = 0 \end{cases}$$

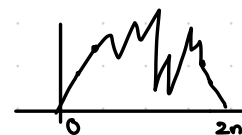
$\alpha = n$ nb de chemins

$\beta = n$ nb de choix pour monter $\binom{2n}{n}$

$$\text{d'où } P(S_{2n} = 0) = \frac{\binom{2n}{n}}{2^{2n}} = \frac{1}{2^{2n}} \frac{(2n)!}{(n!)^2} \xrightarrow[n \rightarrow \infty]{\text{Stirling}} \frac{1}{\sqrt{\pi n}} \rightarrow 0$$

$$\text{Lemme: } P(S_2 \neq 0, \dots, S_{2n-2} \neq 0, S_{2n} = 0) = \frac{1}{2n-1} P(S_{2n} = 0)$$

$$P(\exists n \in \mathbb{N}^* S_n = 0) = \sum_{n \geq 1} \frac{1}{2n-1} \frac{(2n)!}{(n!)^2} \frac{1}{2^{2n}}$$



Cours 6

Prop: Si la marche revient en 0 avec probabilité 1, alors elle repasse par 0 une infinité de fois avec proba 1

preuve: Soit A l'instant de la dernière visite de 0, i.e. $R = \sup \{n \geq 0; S_n = 0\}$

Supposons que $P(\exists n \geq 1, S_n = 0) = 1$ et regardons, pour $K \geq 1$

$$P(R = K) = P(S_K = 0, \forall n > K, S_n - S_K \neq 0) = P(X_1 + \dots + X_K = 0, \forall n > K, X_{K+1} + \dots + X_n \neq 0)$$

or $\{ \forall n > K, X_{K+1} + \dots + X_n \neq 0 \} \in \sigma(X_i, i \geq K+1)$

Par le thm de regroupement par blocs, cet événement est indépendant de $\{X_1 + \dots + X_K\} = 0$

$$\text{d'où } P(R = K) = P(S_K = 0) P(\forall n > K, S_n - S_K \neq 0)$$

de plus $(X_{K+1}, X_{K+2}, \dots) \stackrel{\text{loi}}{=} (X_1, X_2, \dots)$

donc $P(\forall n > K, S_n - S_K \neq 0) = P(\forall n \geq 1, S_n \neq 0)$ d'où: (sachant $P(S_K = 0) \leq 1$)

$$P(R = K) \leq P(\forall n \geq 1, S_n \neq 0) \leq 1 - P(\exists n \geq 1, S_n = 0) = 0$$

$$\text{or, } P(\text{la marche visite 0 un nb fini de fois}) = P(R < +\infty) = P(\bigcup_{K \in \mathbb{N}} \{R = K\}) \leq \sum_K P(R = K) = 0$$

corollaire: Equivalence entre

$$\textcircled{1} P(\exists n \geq 1, S_n = 0) = 1 \quad \textcircled{2} P(S_n \text{ visite 0 une } \infty \text{ de fois}) = 1$$

Souvent, il est plus facile de mq un événement arrive une infinité de fois avec proba plutôt que de mq' il arrive une fois avec proba 1

2) Le lemme de Borel - Cantelli

Soit $(\Omega, \mathcal{F}, P)$ un espace de probabilité. Soit $(A_n)_{n \in \mathbb{N}}$ une suite d'événements de $\mathcal{F}$.

Nous définissons:

$$\liminf_n A_n = \bigcup_{n \geq 0} \bigcap_{m \geq n} A_m = \{ \omega \in \Omega : \omega \text{ appartient à tous les } A_n \text{ sauf un nombre fini } \}$$

$$\limsup_n A_n = \bigcap_{n \geq 0} \bigcup_{m \geq n} A_m = \{ \omega \in \Omega : \omega \text{ appartient à une infinité d}'A_n \}$$

Par ailleurs, $\limsup_n A_n$ se note également:

$$\limsup_n A_n = \{ \hat{A}_n \text{ arrive infiniment souvent} \} = \{ A_n \text{ infiniment souvent} \} = \{ A_n \text{ i. so} \} \text{ en abrégé}$$

ex: (X_n) suite de v.a. réelles $B \in \mathcal{B}(\mathbb{R})$

$$\limsup_n \{ X_n \in B \} = \{ X_n \in B \text{ i. so} \}$$

ex de la marche: $\limsup \{ S_n = 0 \} = \{ S_n = 0 \text{ i. so} \}$

Lemme de Borel - Cantelli: Soit $(A_n)_{n \in \mathbb{N}}$ une suite d'événements.

• partie I: si la série $\sum_n P(A_n)$ converge, alors $P(A_n \text{ i. so}) = 0$

• partie II: si la série $\sum_n P(A_n)$ diverge et si les (A_n) sont indépendants, alors $P(A_n \text{ i. so}) = 1$

pv: I) Supposons que la série $\sum_n P(A_n)$ cv. Alors

$$P(A_n \text{ arrive infiniment souvent}) = P(\bigcap_{n \geq 0} \bigcup_{m \geq n} A_m). \text{ La séquence } \bigcup_{m \geq n} A_m \text{ est décroissante}$$

Par le thm de convergence monotone:

$$P(\bigcap_{n \geq 0} \bigcup_{m \geq n} A_m) = \lim_{n \rightarrow +\infty} P(\bigcup_{m \geq n} A_m)$$

$$\text{Par la } \sigma\text{-additivité: } P(\bigcup_{m \geq n} A_m) \leq \sum_{m \geq n} P(A_m)$$

$$\text{Or la série } \sum_n P(A_n) \text{ converge, donc } \lim_{n \rightarrow \infty} \sum_{m \geq n} P(A_m) = 0, \text{ donc } P(A_n \text{ i. so}) = 0$$

II. Supposons que $\sum P(A_n) = +\infty$.

$P(A_n \text{ i. so}) = P(\bigcap_{m \geq n} A_m)$. Regardons le complémentaire

$$P(\{A_n \text{ i. so}\}^c) = \bigcup_{n \geq 1} P(A_n \text{ f. so}) = P\left(\bigcup_{n \geq 1} \bigcap_{m \geq n} A_m^c\right)$$

Soit n fixé, regardons

$$P\left(\bigcap_{m \geq n} A_m^c\right) = \lim_{N \rightarrow +\infty} P\left(\bigcap_{n \leq m \leq N} A_m^c\right).$$

$$\text{Ensuite, } P\left(\bigcap_{n \leq m \leq N} A_m^c\right) = \prod_{n \leq m \leq N} P(A_m^c) \text{ par indépendance} = \prod_{n \leq m \leq N} (1 - P(A_m))$$

$$\text{d'où, } \ln P\left(\bigcap_{n \leq m \leq N} A_m^c\right) = \sum_{n \leq m \leq N} \ln(1 - P(A_m)) \quad (\text{rappel: } \ln(1+x) \leq x)$$

$$\leq - \sum_{n \leq m \leq N} P(A_m) \xrightarrow{N \rightarrow +\infty} -\infty$$

$$\text{d'où } P(A_n \text{ f. so}) = 0$$

3) Le singe et Shakespeare

Mettons un singe à taper à la machine à écrire. Quelle est la proba qu'il tape les œuvres complètes de Shakespeare

Soit $(X_n)_{n \geq 1}$ une suite i.i.d. Bernoulli(p), $0 < p < 1$

$$\forall n \geq 1 \quad P(X_n = 1) = p \quad P(X_n = 0) = 1 - p$$

Soit $K \geq 1$ et $\varepsilon_1, \dots, \varepsilon_K$ une suite de 0 et de 1 fixée

Regardons l'événement $A_n = \{X_n = \varepsilon_1, X_{n+1} = \varepsilon_2, \dots, X_{n+K-1} = \varepsilon_K\}$

$$P(A_n) = \prod_{i=0}^{K-1} p^{\varepsilon_i} (1-p)^{1-\varepsilon_i} \text{ ne dépend } \phi \text{ de } n! \text{ donc } \sum_n P(A_n) = +\infty$$

Cependant, les (A_n) ne sont pas indépendants on ne peut pas appliquer BC II

Considérons les événements $B_n = A_{nK}$, $n \geq 1$. Par le théorème de regroupement par blocs, B_n sont indépendants

$$\text{et } \sum_n P(B_n) = +\infty \text{ Par BC II, } P(B_n \text{ i. so}) = 1$$

4) Un critère de récurrence

Revenons à notre marche aléatoire $S = 0, \dots, S_n = X_1 + \dots + X_n$

Soit T l'instant du premier retour en 0 : i.e. $T = \min\{n \geq 1 : S_n = 0\}$, convention $\inf \emptyset$

Posons, pour $n \geq 1$ $u_n = P(S_n = 0)$, $f_n = P(T = n)$

$$u_n = P(S_n = 0) = P(\{S_n = 0\} \cap \bigcup_{K=1}^n \{T = K\}) = \sum_{K=1}^n P(S_n = 0, T = K)$$

$$= \sum_{K=1}^n P(S_1 \neq 0, \dots, S_{K-1} \neq 0, S_K = 0, S_{n-K} = 0)$$

$$\text{d'où } u_n = \sum_{K=1}^n P(S_1 \neq 0, \dots, S_{K-1} \neq 0, S_K = 0) P(S_n - S_K = 0) \text{ par indép}$$

$$= \sum_{K=1}^n P(T = K) P(S_{n-K} = 0)$$

$$u_n = \sum_{K=0}^n f_K u_{n-K} \quad \forall n \geq 1$$

Au suites (u_n) et (f_n) nous associons leurs séries génératrices définie par:

$$U(s) = \sum_{n \geq 0} u_n s^n \quad F(s) = \sum_{n \geq 1} f_n s^n = E(s^T)$$

$$P(f=n)$$

Ce sont des séries entières dont le rayon de convergence est ≥ 1

$$\text{On a } U(s) - 1 = \sum_{n \geq 1} u_n s^n = \sum_{n \geq 0} \sum_{K=0}^n u_K f_{n-K} s^n \quad [\text{on peut rajouter le terme } n=0 \quad u_0 f_0 = 0 \quad f_n = 0]$$

$$= \sum_{n \geq 0} \sum_{K=0}^n (u_K s^K) (f_{n-K} s^{n-K}) \leftarrow [\text{produit Cauchy de 2 séries}]$$

$$= U(s) F(s)$$

$$d'où U(s) \cdot 1 = U(s) F(s) \Leftrightarrow U(s)(1 - F(s)) = 1$$

$$\Leftrightarrow \left(\sum_{n \geq 0} U_n s^n \right) \left(1 - \sum_{n \geq 0} f_n s^n \right) = 1 \quad 0 < s < 1$$

$$s \rightarrow 1$$

$$\left(\sum_{n \geq 0} U_n \right) \left(1 - \sum_{n \geq 0} f_n \right) = 1$$

$$\left(\sum_{n \geq 0} P(S_n = 0) \right) \left(1 - \sum_{n \geq 0} P(T = n) \right) = 1 \quad \text{mais } \sum_{n \geq 0} P(T = n) = P(T < +\infty) = P(\exists n \geq 1, S_n = 0)$$

$$\left(\sum_{n \geq 0} P(S_n = 0) \right) \left(1 - P(\exists n \geq 1, S_n = 0) \right) = 1$$

$$\Leftrightarrow P(\exists n \geq 1, S_n = 0) = 1 \Leftrightarrow \sum_{n \geq 0} P(S_n = 0) = +\infty$$

Th Equivalence:

$$① P(\exists n \geq 1, S_n = 0) = 1$$

$$② P(S_n = 0 \text{ i.s.o.}) = 1$$

$$③ \sum_{n \geq 0} P(S_n = 0) = +\infty$$

5) Application à la marche dans $\mathbb{Z}$:

$0 < p < 1$ (X_n) i.i.d. Ber(p), $P(X_n = -1) = 1-p$, $P(X_n = 1) = p$, $S_n = X_1 + \dots + X_n$

$$P(S_{2n} = 0) = \binom{2n}{n} p^n (1-p)^n \sim \frac{(4p(1-p))^n}{\sqrt{\pi n}} \quad (\text{faire calcul en utilisant Stirling})$$

$$p = 1/2 \rightarrow P(S_{2n} = 0) \sim \frac{1}{\sqrt{\pi n}} \text{ diverge} \Rightarrow \text{récurrente}$$

$$p \neq 1/2 \text{ converge} \Rightarrow \text{transiente}$$

6) Cas de $\mathbb{Z}^2$

$$(X_n) \text{ i.i.d. } P(X_n = \begin{pmatrix} 1 \\ 0 \end{pmatrix}) = 1/4, P(X_n = \begin{pmatrix} 0 \\ 1 \end{pmatrix}) = 1/4, P(X_n = \begin{pmatrix} 0 \\ -1 \end{pmatrix}) = 1/4, P(X_n = \begin{pmatrix} -1 \\ 0 \end{pmatrix}) = 1/4$$

$$\text{calcul } P(S_{2n} = 0) = \sum_{u+d+e+f=2n} \frac{(2n)!}{u!d!e!f!} = \dots = (P(S_{2n} = 0) \text{ marche ds } \mathbb{Z}) \sim 1/\pi n$$

$$\text{dans } \mathbb{Z}^3: P(S_{2n} = 0) \sim \frac{c}{n^{3/2}} \Rightarrow \text{converge, transiente}$$

Cours 6

Chapitre 6: le problème des signes aléatoires

1) Le problème:

Soit $(u_n)_{n \in \mathbb{N}}$ une suite (déterministe) de nombres réels.

Soit $(\varepsilon_n)_{n \in \mathbb{N}}$ une suite infinie i.i.d. de Bernoulli (1/2) tq:

$$\forall n \in \mathbb{N}, P(\varepsilon_n = -1) = P(\varepsilon_n = 1) = 1/2$$

$\rightarrow$ La série $\sum_{n \geq 0} \varepsilon_n$ converge-t-elle?

• si $u_n \not\rightarrow 0$, la série ne converge pas.

• si la série $\sum u_n$ est absolument convergente, alors $\sum \varepsilon_n u_n$ est p.s. absolument convergente.

Regardons la série harmonique: $\sum_{n \geq 1} 1/n = +\infty$

$$\sum_{n \geq 1} \frac{(-1)^{n-1}}{n} = 1 - 1/2 + 1/3 - \dots = \ln(2)$$

2) Convergence dans L^2

Proposition:

La série de v.a. $\sum_n u_n \varepsilon_n$ converge dans L^2 ssi la série $\sum_n u_n^2$ converge.

$$\text{PREUVE: Posons } S_n = \sum_{k=0}^n \varepsilon_k u_k$$

Alors, la série $\sum_n \varepsilon_n u_n$ converge dans L^2 ssi la suite $(S_n)_{n \in \mathbb{N}}$ converge dans L^2 .

L'espace $\mathcal{L}^2(\Omega, \mathcal{F}, \mathbb{P})$ est complet pour $\|\cdot\|_2$.

Ainsi, la suite $(S_n)_{n \in \mathbb{N}}$ converge ssi elle vérifie le critère de Cauchy.

Soit $m < n$, regardons:

$$\begin{aligned} \|S_n - S_m\|_{L^2}^2 &= \int_{\Omega} |S_n - S_m|^2 d\mathbb{P} \\ &= \mathbb{E} \left((S_n - S_m)^2 \right) \\ &= \mathbb{E} \left[\left(\sum_{k=m+1}^n \varepsilon_k u_k \right)^2 \right] \\ &= \mathbb{E} \left[\sum_{m < k, \ell \leq n} \varepsilon_k \varepsilon_\ell u_k u_\ell \right] = \sum_{\substack{m < k \\ \ell \leq n}} u_k u_\ell \mathbb{E}(\varepsilon_k \varepsilon_\ell) \\ &\quad k \neq \ell \text{ et } \mathbb{E}(\varepsilon_k \varepsilon_\ell) = \mathbb{E}(\varepsilon_k) \mathbb{E}(\varepsilon_\ell) = 0 \end{aligned}$$

Donc

$$= \sum_{m < k \leq n} u_k^2 = s_n - s_m \text{ où } s_n = \sum_{k=0}^n u_k^2$$

Ainsi: la suite (S_n) est de Cauchy dans L^2 ssi la suite (s_n) est de Cauchy dans $\mathbb{R}$.

3) La loi du 0-1 de Kolmogorov

Def: Soit $(X_n)_{n \in \mathbb{N}}$ une suite de v.a. définies sur l'espace $(\Omega, \mathcal{F}, \mathbb{P})$.

La tribu de queue, ou tribu asymptotique, associée à la suite (X_n) est la tribu:

$$\mathcal{L} = \bigcap_{n \geq 0} \sigma(X_n, X_{n+1}, \dots)$$

Un événement A de $\mathcal{F}$ est dit asymptotique pour la suite (X_n) s'il est dans la tribu asymptotique.

EXEMPLE: $\{\limsup X_n \geq 100\}$, $\{\liminf X_n \in [0, 1]\}$, $\{\omega: \lim_{n \rightarrow \infty} X_n(\omega) \text{ existe}\}$, $\{\omega: \sum_n \varepsilon_n(\omega) u_n \text{ converge}\}$

$\mathcal{L}$ est asymptotique pour la suite $(\varepsilon_n)_{n \in \mathbb{N}}$

Théorème. Soit $(X_n)_{n \in \mathbb{N}}$ une suite de v.a. indépendantes.

Soit A un événement asymptotique pour la suite (X_n) . Alors $\mathbb{P}(A)$ vaut 0 ou 1.

Application au pb des signes:

Soit $\sum_n \varepsilon_n u_n$ converge p.s, soit elle ne converge pas p.s.

4) Un résultat d'approximation:

La différence symétrique de 2 ensembles A et B est: $A \Delta B = (A \cup B) \setminus (A \cap B)$

$$\bullet A \Delta \emptyset = \emptyset \Delta A = A$$

$$\bullet A \Delta A = \emptyset$$

$$\bullet \forall A, B, C, A \Delta (B \Delta C) = (A \Delta B) \Delta C$$

$$\bullet (\mathcal{P}(\Omega), \Delta) \text{ groupe abélien}$$

PROP: Soit $(X_n)_{n \in \mathbb{N}}$ une suite de v.a. et soit A un événement dans la tribu $\sigma(X_0, X_1, \dots)$,

Alors il existe une suite d'événements $(A_n)_{n \in \mathbb{N}}$ tq:

$$\bullet \forall n \in \mathbb{N}, \exists k(n), A_n \in \sigma(X_0, \dots, X_{k(n)})$$

$$\bullet \lim_{n \rightarrow \infty} \mathbb{P}(A_n \Delta A) = 0$$

$$\text{RQ: } \forall A, B \quad \mathbb{P}(A) - \mathbb{P}(B) \leq \mathbb{P}(A \Delta B),$$

donc on aura aussi $\lim_n \mathbb{P}(A_n) = \mathbb{P}(A)$.

PREUVE: Considérons $\mathcal{G}$ la collection des événements pour lesquels le résultat est vrai.

je dis que :

$\forall m \sigma(X_0, \dots, X_m) \subset G$ si $A \in \sigma(X_0, \dots, X_m)$, il suffit de prendre $A_m = A$

On montre ensuite que G est tribu, donc contient $\sigma(\bigcup_m \sigma(X_0, \dots, X_m))$

Ainsi, $\sigma(X_0, X_1, \dots) \subset G$

En effet, $\emptyset, \Omega \in G$

Si $A \in G$ et (A_n) suite approximée A , alors la suite (A_n^c) approxime A^c

$$\bullet \quad \mathbb{P}(A_n^c \Delta A^c) = \mathbb{P}(A_n \Delta A) \longrightarrow 0$$

si $(A_m)_{m \geq 1}$ est une suite de G , alors $\bigcup_m A_m \in G$

Posons $A = \bigcup_m A_m$

Comme $A_m \in G$, il existe $(A_{m,n})$ approximant pour A_m :

$\forall n, \forall m \geq 1, \exists K(m,n); A_{m,n} \in \sigma(X_0, \dots, X_{K(m,n)})$

$$\lim_n \mathbb{P}(A_m \Delta A_{m,n}) = 0$$

$$\bullet \quad \text{Comme } \lim_{n \rightarrow +\infty} \mathbb{P}(A) \setminus \bigcup_{m=1}^n A_m = 0, \forall M \geq 1 \exists \psi(M) \text{ tq } \mathbb{P}(A \setminus \bigcup_{m=1}^{\psi(M)} A_m) \leq 1/M$$

Ensuite: $\forall m \in \{1, \dots, \psi(M)\};$

$$\exists N(M,m) \text{ tq } \forall n \geq N(M,m), \mathbb{P}(A_m \Delta A_{m,n}) \leq \frac{1}{\psi(M)M}$$

Prenons alors $N(M) = \max \{N(M,m), 1 \leq m \leq \psi(M)\}$ et $A_M = \bigcup_{m=1}^{\psi(M)} A_{m,N(M)}$

$$\text{Alors, } \mathbb{P}(A \Delta A_M) \leq \mathbb{P}(A \setminus \bigcup_{m=1}^{\psi(M)} A_m) + \mathbb{P}(\bigcup_{m=1}^{\psi(M)} A_m \Delta A_M)$$

$$\leq \frac{1}{M} + \sum_{m=1}^{\psi(M)} \mathbb{P}(A_m \Delta A_{m,N(M)})$$

$$\leq \frac{1}{M} + \sum_{m=1}^{\psi(M)} \frac{1}{\psi(M)M} = 2/M \quad \square$$

(X_n) indep, $A \in G = \sigma(X_0, X_1, \dots)$

On applique le résultat d'approx à A :

$\exists (A_n) \forall n, A_n \in \sigma(X_0, X_1, \dots, X_{K(n)})$ et $\lim \mathbb{P}(A \Delta A_n) = 0$

D'abord, $|\mathbb{P}(A) - \mathbb{P}(A_n)| \leq \mathbb{P}(A \Delta A_n)$

donc $\mathbb{P}(A_n) \rightarrow \mathbb{P}(A)$

De plus, $A_n \in \sigma(X_0, \dots, X_{K(n)})$

$A \in G \subset \sigma(X_{K(n)+1}, X_{K(n)+2}, \dots)$

donc $\mathbb{P}(A \cap A_n) \rightarrow \mathbb{P}(A)$

Comme (X_n) indep, les 2 tribus sont indep. Donc A et A_n aussi

$$\mathbb{P}(A \cap A_n) = \mathbb{P}(A) \mathbb{P}(A_n) \rightarrow \mathbb{P}(A)^2$$

$$|\mathbb{P}(A) - \mathbb{P}(A \cap A_n)| = \mathbb{P}(A \setminus A_n) \leq \mathbb{P}(A \Delta A_n) \rightarrow 0, \text{ d'où } \mathbb{P}(A)^2 = \mathbb{P}(A)$$

5) Inégalité de Kolmogorov :

Soient $X_1, \dots, X_n$ n v.a. indep, centrées et de variances finies.

Posons $\forall k \in \{1, \dots, n\}; S_k = X_1 + \dots + X_k$

$$\text{Alors } \forall \alpha > 0, \mathbb{P}\left(\max_{1 \leq k \leq n} |S_k| \geq \alpha\right) \leq \frac{1}{\alpha^2} \text{Var}(S_n)$$

preuve:

Pour $K \in \{1, \dots, n\}$, on introduit l'événement $A_K = \{|S_1| < \alpha, \dots, |S_{K-1}| < \alpha, |S_K| \geq \alpha\}$ disjoints.

$$\begin{aligned} \text{Var}(S_n) &= \mathbb{E}(S_n^2) = \int_{\Omega} S_n^2 d\mathbb{P} \geq \int_{\bigcup_{K=1}^n A_K} S_n^2 d\mathbb{P} \\ &= \sum_{K=1}^n \int_{A_K} S_n^2 d\mathbb{P} \quad \text{car } A_K \text{ à 2 disjoints} \end{aligned}$$

$$\begin{aligned} \text{De plus, } \int_{A_K} S_n^2 d\mathbb{P} &= \int_{A_K} (S_K + S_n - S_K)^2 d\mathbb{P} \\ &= \dots \end{aligned}$$

6) Convergence de séries de v.a. indépendantes:

Thm: Soit $(X_n)_{n \in \mathbb{N}}$ une suite de v.a. indep. centrées et de variances finies, $\forall n, \mathbb{E}(X_n) = 0, \mathbb{E}(X_n^2) \leq +\infty$

Cours 7

1) Inégalité de Bienaymé - Chebyshev

Supposons que X est positive. Alors $\forall \lambda \geq 0, \mathbb{P}(X \geq \lambda) \leq \frac{1}{\lambda} \mathbb{E}[X]$

rq: Comme $X \geq 0$, $\mathbb{E}[X]$ est toujours défini, éventuellement c'est $+\infty$.

$$\begin{aligned} \text{PREUVE: } \mathbb{E}[X] &= \int_{\Omega} X d\mathbb{P} \geq \int_{\{X \geq \lambda\}} X d\mathbb{P} \geq \lambda \mathbb{P}(X \geq \lambda) \\ \int_{\{X \geq \lambda\}} X d\mathbb{P} &= \int \mathbb{1}_{\{X \geq \lambda\}} X d\mathbb{P} \geq \int \mathbb{1}_{\{X \geq \lambda\}} \lambda d\mathbb{P} = \lambda \mathbb{P}(X \geq \lambda) \end{aligned}$$

Soit X une v.a. qui admet un moment d'ordre 2, i.e., $\mathbb{E}[X^2] < +\infty$

Alors $\forall \epsilon > 0, \mathbb{P}(|X - \mathbb{E}[X]| \geq \epsilon) \leq \frac{1}{\epsilon^2} \text{Var}(X)$

$$\begin{aligned} \text{PREUVE: } \mathbb{P}(|X - \mathbb{E}[X]| \geq \epsilon) &= \mathbb{P}((X - \mathbb{E}[X])^2 \geq \epsilon^2) \\ &\stackrel{\text{Markov}}{\leq} \frac{1}{\epsilon^2} \mathbb{E}[(X - \mathbb{E}[X])^2] = \frac{1}{\epsilon^2} \text{Var } X \end{aligned}$$

L'intérêt de cette inégalité réside dans sa simplicité et le fait qu'elle est très bien adaptée aux sommes de v.a. i.

2) La loi faible des grands nombres

THM: Soit μ une mesure de probas sur $\mathbb{R}$ qui est intégrable, i.e. $\int_{\mathbb{R}} |x| d\mu(x) < +\infty$.

$$\text{Alors } m = \int_{\mathbb{R}} x d\mu(x)$$

Pour $n \geq 1$, Soient $X_1, \dots, X_n$ n v.a. i.i.d. de la loi μ .

$$\text{Alors } \forall \epsilon > 0, \lim_{n \rightarrow +\infty} \mathbb{P}\left(\left|\frac{X_1 + \dots + X_n}{n} - m\right| \geq \epsilon\right) = 0$$

Rem: Si (X_n) reste i.i.d loi μ , cela s'écrit: $\frac{X_1 + \dots + X_n}{n} \xrightarrow{\mathbb{P}} m$

En fait, par la formule de transfert appliquée au vecteur $(X_1, \dots, X_n)$, on a:

$$\mathbb{P}\left(\left|\frac{X_1 + \dots + X_n}{n} - m\right| \geq \epsilon\right) = \mathbb{P}_{(\mu \otimes \dots \otimes \mu)}\left(\left\{\left(x_1, \dots, x_n\right) \in \mathbb{R}^n; \left|\frac{x_1 + \dots + x_n}{n} - m\right| \geq \epsilon\right\}\right)$$

Ainsi, la loi faible est en fait un résultat asymptotique sur la suite de lois $(\mu \otimes n, n \geq 1)$. La loi ne nécessite pas d'avoir une suite i.i.d

PREUVE: Supposons que la loi μ admet un moment d'ordre 2

Posons $\sigma^2 = \text{Var}(\mu) = \int (x-m)^2 d\mu(x)$ et appliquons BT à $\frac{1}{n}(X_1 + \dots + X_n)$

Comme $E(\frac{1}{n}(X_1 + \dots + X_n)) = \frac{1}{n}(E(X_1) + \dots + E(X_n)) = m$

Soit $\varepsilon > 0$

$$P\left(\left|\frac{X_1 + \dots + X_n}{n} - m\right| \geq \varepsilon\right) \leq \frac{1}{\varepsilon^2} \text{Var}\left(\frac{X_1 + \dots + X_n}{n}\right)$$

$$\text{Calculons } \text{Var}\left(\frac{X_1 + \dots + X_n}{n}\right) = \frac{1}{n^2} \text{Var}(X_1 + \dots + X_n) \stackrel{\text{indép}}{=} \frac{1}{n^2} (\text{Var}(X_1) + \dots + \text{Var}(X_n))$$

$$\stackrel{\text{m loi}}{=} \frac{1}{n} \text{Var}(\mu) = \frac{\sigma^2}{n}$$

$$\text{d'où } P\left(\left|\frac{X_1 + \dots + X_n}{n} - m\right| \geq \varepsilon\right) \leq \frac{\sigma^2}{n\varepsilon^2} \xrightarrow{n \rightarrow +\infty} 0$$

3) Loi forte des grands nombres

Thm: Soit $(X_n)_{n \geq 1}$ une suite de v.a. intégrables i.e. $E(|X_1|) < +\infty$

$$\text{Alors } \frac{X_1 + \dots + X_n}{n} \xrightarrow{\text{p.s.}} E(X_1)$$

preuve: 1^{ère} étape centrage

Posons $m = E(X_1)$ quitte à remplacer X_n par $X_n - m$, pour $n \geq 1$, nous pouvons supposer la suite centrée, ou de manière équivalente, que $E(X_1) = 0$

2^e étape: troncature: Posons $\forall n \geq 1$, $Y_n = \begin{cases} X_n & \text{si } |X_n| \leq n, \text{ i.e. } Y_n = X_n \mathbb{1}_{|X_n| \leq n} \\ 0 & \text{sinon} \end{cases}$

$$\text{et écrivons } \frac{X_1 + \dots + X_n}{n} = \underbrace{\frac{(X_1 - Y_1) + \dots + (X_n - Y_n)}{n}}_{\substack{\uparrow \text{ je dis que} \\ \text{le terme tend vers 0 p.s.}}} + \frac{Y_1 + \dots + Y_n}{n}$$

Lemme: Soit $(X_n)_{n \geq 1}$ une suite de v.a. iid

On a équivalence entre: i) $E(|X_1|) < +\infty$

$$\text{ii) } P(|X_n| \geq n \text{ } \infty^{\text{t}} \text{ souvent}) = 0$$

PREUVE: Comme les événements $\{|X_n| \geq n\}$ sont indépendants, par BC on a équivalence:

$$P(|X_n| \geq n \text{ } \infty^{\text{t}} \text{ souvent}) = 0 \Leftrightarrow \sum_{n \geq 1} P(|X_n| \geq n) < +\infty \text{ car } X_1 \text{ et } X_n \text{ ont même loi}$$

$$\Leftrightarrow \sum_{n \geq 1} P(|X_1| \geq n) < +\infty$$

$$\Leftrightarrow E(|X_1|) < +\infty$$

$$E(|X_1|) = \int_{\mathbb{R}} |x| dP = \int_{\mathbb{R}} \left(\int_0^{+\infty} \mathbb{1}_{\{t \leq |x|\}} dt \right) dP$$

$$\stackrel{\text{Fubini}}{=} \int_0^{+\infty} \left(\int_{\mathbb{R}} \mathbb{1}_{\{t \leq |x|\}} dP \right) dt = \int_0^{+\infty} P(|X_1| \geq t) dt$$

$$\text{En suite } \int_0^{+\infty} P(|X_1| \geq t) dt = \sum_{n=0}^{+\infty} \int_n^{n+1} P(|X_1| \geq t) dt$$

$$P(|X_1| \geq n+1) \leq \int_n^{n+1} P(|X_1| \geq t) dt \leq P(|X_1| \geq n)$$

$$\sum_{n \geq 0} P(|X_1| \geq n+1) \leq E(|X_1|) \leq \sum_{n \geq 0} P(|X_1| \geq n)$$

$$Y_n = X_n \mathbb{1}_{|X_n| \leq n}$$

$$Y_n \neq X_n \Leftrightarrow |X_n| > n$$

$$\text{Par le lemme } P(Y_n \neq X_n \text{ i.s.}) = P(|X_n| > n \text{ i.s.}) = 0 \text{ (car } E(|X_1|) < +\infty)$$

$$\text{Ainsi } P(X_n \neq Y_n \text{ i.s.}) = 0 \text{ et donc } \frac{X_1 - Y_1 + \dots + X_n - Y_n}{n} \xrightarrow[n \rightarrow +\infty]{\text{p.s.}} 0$$

4) Critère de Kolmogorov

PROP: Soit (Z_n) une suite de v.a. indépendantes, centrées.

$$\text{Si } \sum_{n \geq 1} \frac{\mathbb{E}(Z_n^2)}{n^2} < +\infty \text{ alors } \frac{Z_1 + \dots + Z_n}{n} \xrightarrow{\text{p.s.}} 0$$

PREUVE: (...)

Cours 8 7/11/2025

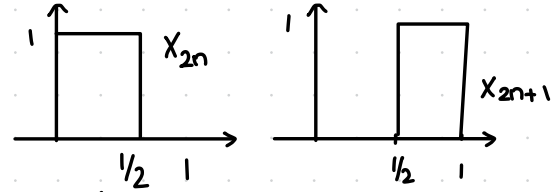
La convergence en loi:

Definition: Il existe des suites de v.a. qui ne convergent ni p.s., ni en probas, ni dans L^1 , mais dont la loi converge, voir même constante.

$$(\Omega, \mathcal{F}, \mathbb{P}) = ([0, 1], \mathcal{B}([0, 1]), \lambda)$$

Pour tout n , la loi de X_n est la loi Bernoulli $(1/2)$

$\forall n \quad \mathbb{P}(X_n=0) = \mathbb{P}(X_n=1) = 1/2$ et donc $\mathbb{P}_{X_n} = \text{Ber}(1/2)$ ne dépend pas de n !



Definition: une suite de v.a. $(X_n)_{n \in \mathbb{N}}$ converge en loi vers une v.a. X , ce que l'on note $X_n \xrightarrow{\text{loi}} X$ ssi pour toute fonction f continue bornée $\mathbb{R} \rightarrow \mathbb{R}$, $\lim_{n \rightarrow +\infty} \mathbb{E}(f(X_n)) = \mathbb{E}(f(X))$

Remarquons que $\mathbb{E}(f(X_n))$ ne dépend que de la loi de X_n , i.e.

$$\mathbb{E}(f(X_n)) = \int_{\Omega} f(X_n) d\mathbb{P} \underset{\text{transfert}}{=} \int_{\mathbb{R}} f(x) d\underbrace{\mathbb{P}_{X_n}(x)}_{\text{loi de } X_n}$$

Ainsi, la convergence en loi s'applique même dans des situations où les v.a. X_n ne sont pas définies sur le même espace de probas $(\Omega, \mathcal{F}, \mathbb{P})$, i.e., si $X_n: (\Omega_n, \mathcal{F}_n, \mathbb{P}_n) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ où $(\Omega_n, \mathcal{F}_n, \mathbb{P}_n)$ n'ont aucun lien a priori. Ceci distingue fondamentalement la convergence en loi des autres modes de convergence.

2) La fonction de répartition

Soit X une v.a. $(\Omega, \mathcal{F}, \mathbb{P}) \rightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

Def: La fonction de répartition de X , notée F_X , est la fonction définie par: $\forall x \in \mathbb{R}, F_X(x) = \mathbb{P}(X \leq x)$

$$\text{On a } F_X(x) = \mathbb{P}_X([-\infty, x])$$

Comme les intervalles $[-\infty, x], x \in \mathbb{R}$, engendrent les boréliens et forment une famille stable par $\cap$ finie, par unicité de l'extension, la loi $\mathbb{P}_X$ est déterminée par F_X .

Conclusion: la fonction de répartition F_X détermine la loi $\mathbb{P}_X$ de X .

Une fonction de répartition F vérifie:

- i) F est croissante de $\mathbb{R}$ dans $[0, 1]$
- ii) F est continue à droite
- iii) $\lim_{x \rightarrow -\infty} F(x) = 0, \lim_{x \rightarrow +\infty} F(x) = 1$

preuve: i)

$$\text{ii) Soit } x_0 \in \mathbb{R}. \text{ Regardons } \lim_{x \rightarrow x_0} F_X(x) = \lim_{\substack{x \rightarrow x_0 \\ x > x_0}} \mathbb{P}_X([-\infty, x])$$

$$\underset{\text{continuité monotone}}{=} \mathbb{P}_X\left(\bigcap_{\substack{x > x_0 \\ x \in \mathbb{Q}}} [-\infty, x]\right) = \mathbb{P}_X([-\infty, x_0]) = F_X(x_0)$$

$$\text{iii) } \lim_{x \rightarrow -\infty} F_X(x) = \lim_{\substack{x \rightarrow -\infty \\ x \in \mathbb{Z}}} \mathbb{P}_X([-\infty, x])$$

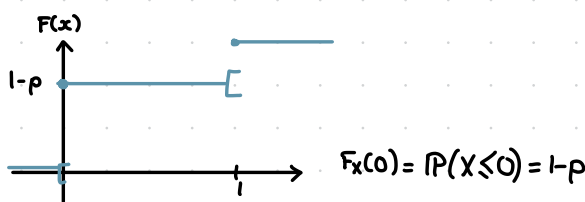
$$= \mathbb{P}_X\left(\bigcap_{x \in \mathbb{Z}} [-\infty, x]\right) = \mathbb{P}_X(\emptyset) = 0$$

$$\lim_{x \rightarrow +\infty} F_X(x) = \lim_{x \rightarrow +\infty} \mathbb{P}_X([-\infty, x]) = \mathbb{P}_X\left(\bigcup_{x \in \mathbb{N}} [-\infty, x]\right) = \mathbb{P}_X(\mathbb{R}) = 1$$

3) Exemples.

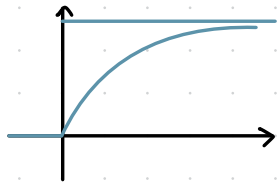
$$X \sim \text{Ber}(p)$$

$$\mathbb{P}(X=0) = 1-p$$



$$F_X(0) = \mathbb{P}(X \leq 0) = 1-p$$

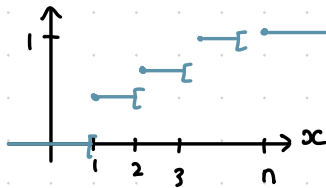
$$F_X \sim \text{Exp}(\alpha) \quad F_X(x) = \int_0^x \alpha e^{-\alpha t} dt \quad x \geq 0 = 1 - e^{-\alpha x}$$



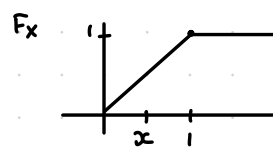
Plus généralement, si la loi de X a une densité f , alors

$$F_X(x) = \int_{-\infty}^x f(t) dt$$

$$X \sim U(\{1, \dots, n\})$$



$$X \sim U([0, 1])$$



4) La bijection.

si μ est une mesure de probabilité sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, sa fonction de répartition F_μ est l'application de $\mathbb{R}$ dans $[0, 1]$:

$$\forall x \in \mathbb{R}, F_\mu(x) = \mu([-\infty, x]).$$

$$\text{Coïncidence: si } X \text{ est une v.a., } F_X(x) = \mathbb{P}_X([-\infty, x]) = F_{\mathbb{P}_X}(x)$$

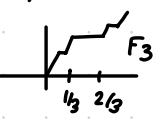
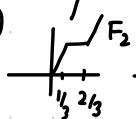
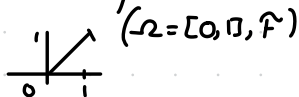
Notons $\mathcal{M}_1(\mathbb{R})$ l'ensemble des mesures de probabilité sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$

$\mathcal{R}$ l'ensemble des fonctions de répartition sur $\mathbb{R}$ (les fonctions qui vérifient i) ii) iii))

THM: L'application $\mu \in \mathcal{M}_1(\mathbb{R}) \mapsto F_\mu \in \mathcal{R}$ est une bijection

Construction d'une loi singulière

i.e. loi μ sans atomes: $\forall x, \mu(\{x\}) = 0$ et tq $\exists c, \mu(c) = 1, \lambda(c) = 0$



(F_n)

$(F_n) \rightarrow F_\infty$ sur $[0, 1]$ uniformément

5) Traduction sur les fonctions de répartition

Comme la convergence en loi ne porte que sur les lois des v.a., elle doit pouvoir s'exprimer via les fonctions de répartition.

Thm: Soit $(X_n)_{n \in \mathbb{N}}$, X des v.a. réelles et notons $(F_{X_n})_{n \in \mathbb{N}}$, F_X les fonctions de répartition.

Equivalence: i) La suite (X_n) converge en loi vers X

ii) en tout point x de $\mathbb{R}$ où F_X est continue, la suite $(F_n(x))_{n \in \mathbb{N}}$ converge $F_X(x)$ i.e. $\lim_{n \rightarrow +\infty} F_{X_n}(x) = F_X(x)$

preuve: i $\Rightarrow$ ii) je prends $f(x) = 1_{[-\infty, x]}$ alors $\mathbb{E}[f(X_n)] = \mathbb{P}(X_n \leq x) = F_{X_n}(x)$

$$\downarrow$$

$$\mathbb{E}[f(X)]$$

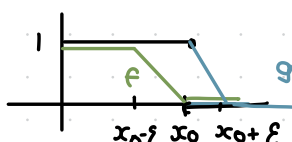
$$\downarrow$$

$$= F_X(x)$$

(pb: f pas continue en x)

vraie preuve: Soit x_0 un point où F_X est continue. Il s'agit de mq $\lim_{n \rightarrow \infty} F_{X_n}(x_0) = F_X(x_0)$

Soit $\varepsilon > 0$ et considérons les fonctions f, g suivantes:



Alors $\forall x \in \mathbb{R} \quad 1_{[-\infty, x_0 - \varepsilon]}(x) \leq f(x) \leq 1_{[-\infty, x_0 + \varepsilon]}(x) \leq g(x) \leq 1_{[-\infty, x_0 + \varepsilon]}(x)$

et f, g sont continues, donc $\mathbb{E}[f(X_n)] \rightarrow \mathbb{E}[f(X)]$

$$\mathbb{E}[g(X_n)] \rightarrow \mathbb{E}[g(X)]$$

$$\mathbb{E}[f(X_n)] \leq F_{X_n}(x_0) \leq \mathbb{E}[g(X_n)] \quad \forall n$$

$$F_X(x_0 - \varepsilon) \leq \mathbb{E} f(X) \leq \liminf_{n \rightarrow +\infty} F_{X_n}(x_0) \leq \limsup_{n \rightarrow +\infty} F_{X_n}(x_0) \leq \mathbb{E}[g(X)] \leq F_X(x_0 + \varepsilon) \quad \forall \varepsilon > 0$$

$\downarrow$ car x_0 p.t. de continuité $\quad \quad \quad \downarrow$ cont à droite $\varepsilon \rightarrow 0$
 $F_X(x_0) \quad \quad \quad F_X(x_0)$

ii) $\Rightarrow$ i) $F_{X_n}(x) \rightarrow F_X(x)$ en tout point x où F_X est continue

Soit $f: \mathbb{R} \rightarrow \mathbb{R}$ continue bornée. Il faut mq $\mathbb{E}[f(X_n)] \rightarrow \mathbb{E}[f(X)]$

Soit $\varepsilon > 0$. Comme $\lim_{-\infty} F_X = 0, \lim_{+\infty} F_X = 1$, il existe $a < b$ tq $F_X(a) < \varepsilon, F_X(b) > 1 - \varepsilon$, F_X continue en a et b

$$\text{Ensuite on écrit } \mathbb{E} f(X_n) = \int f(X_n) dP = \int_{X_n \leq a} f(X_n) dP + \int_{a \leq X_n \leq b} f(X_n) dP + \int_{X_n > b} f(X_n) dP$$

$$\left| \int_{X_n \leq a} f(X_n) dP \right| \leq \|f\|_{\infty} P(X_n \leq a) = \|f\|_{\infty} F_{X_n}(a)$$

$$\left| \int_{X_n > b} f(X_n) dP \right| \leq \|f\|_{\infty} P(X_n > b) = \|f\|_{\infty} (1 - F_{X_n}(b))$$

Soit $x_0 = a < x_1 < \dots < x_k = b$ une subdivision de $[a, b]$ tq F_X est continue en $x_0, x_1, \dots, x_k$.

$$\text{Définissons } f_k^+ \text{ par: } f_k^+(x) = \sum_{\ell=1}^k 1_{]x_{\ell-1}, x_{\ell}]}(x) \sup_{]x_{\ell-1}, x_{\ell}]} f = \sum_{\ell=1}^k \sup_{]x_{\ell-1}, x_{\ell}]} f \int_{x_{\ell-1} < X_n \leq x_{\ell}} dP$$

$$\text{Alors } f \leq f_k^+ \text{ sur } [a, b] \text{ et } \int_{a < X_n \leq b} f(X_n) dP \leq \int_{a < X_n \leq b} f_k^+(X_n) dP = \sum_{\ell=1}^k \sup_{]x_{\ell-1}, x_{\ell}]} f \int_{x_{\ell-1} < X_n \leq x_{\ell}} dP$$

(... il a effacé)

Cours 9

Les fonctions caractéristiques

Si f est une fonction de $\mathbb{R}$ dans $\mathbb{C}$ et si μ est une mesure de probabilité sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, on définit

$$\int f d\mu = \int \operatorname{Re} f d\mu + i \int \operatorname{Im} f d\mu \text{ sous réserve que } \operatorname{Re} f \text{ et } \operatorname{Im} f \text{ sont intégrables.}$$

1) Définition

si μ est une mesure de probabilité sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, sa fonction caractéristique, notée Ψ_{μ} , est la fonction $\mathbb{R} \rightarrow \mathbb{C}$ définie par: $\forall u \in \mathbb{R}, \Psi_{\mu}(u) = \int_{\mathbb{R}} e^{iux} d\mu(x)$ (bien définie car $|e^{iux}| = 1$ et donc $\operatorname{Re} e^{iux} = \cos ux$ est μ -intégrable)

$\rightarrow$ Si X est une v.a.r. définie sur un espace de probas $(\Omega, \mathcal{F}, P)$, sa fonction caractéristique, notée Ψ_X , est la fonction de $\mathbb{R}$ dans $\mathbb{C}$ définie par $\forall u \in \mathbb{R} \Psi_X(u) = \mathbb{E}[e^{iux}]$
 $= \int_{\Omega} e^{iux} dP = \int_{\omega \in \Omega} e^{iux(\omega)} dP(\omega)$ (bien définie car $|e^{iux}| = 1$ donc e^{iux} P -intégrable)

Lien entre les 2 def: Par la formule de transfert

$$\mathbb{E}[e^{iux}] = \int_{\mathbb{R}} e^{iux} dP_X(x) \text{ et donc } \Psi_X = \Psi_{P_X}$$

La fonction caract d'une v.a. X est, à un signe près, la transformée de Fourier de sa loi P_X .

2) Exemples :

• Si $X \sim \text{Ber}(p)$ $P(X=1)=p, P(X=0)=1-p$
 $\Psi_X(u) = \mathbb{E}[e^{iux}] = 1-p + pe^{iu}$

• si $X \sim P(\theta)$: $\forall k \in \mathbb{N} P(X=k) = e^{-\theta} \frac{\theta^k}{k!}$

$$\Psi_X(u) = \mathbb{E}[e^{iux}] = \sum_{k=0}^{+\infty} e^{iuk} e^{-\theta} \frac{\theta^k}{k!} = \sum_{k \geq 0} e^{-\theta} \frac{(e^{iu}\theta)^k}{k!} = e^{-\theta} e^{\theta e^{iu}} = e^{\theta(e^{iu}-1)}$$

$$X \sim \text{Exp}(\alpha), \alpha > 0$$

$$\mathbb{E}[e^{i\mu x}] = \int_{\mathbb{R}^+} e^{i\mu x} \alpha e^{-\alpha x} d\lambda(x) = \int_{\mathbb{R}^+} \alpha e^{(i\mu - \alpha)x} d\lambda(x) = \frac{\alpha}{i\mu - \alpha} e^{(i\mu - \alpha)x} \Big|_0^{+\infty} = \frac{\alpha}{i\mu - \alpha}$$

3) Propriétés:

Si X est une v.a. sa fonction caractéristique φ_X vérifie:

$$i) \varphi_X(0) = 1$$

$$ii) \forall \mu \in \mathbb{R}, |\varphi_X(\mu)| \leq 1$$

$$iii) \forall \mu \in \mathbb{R}, \varphi_X(\mu) = \overline{\varphi_X(-\mu)}$$

$$iv) \varphi_X \text{ est continue sur } \mathbb{R}$$

preuve: ii) $|\mathbb{E}(e^{i\mu x})| \leq \mathbb{E}(|e^{i\mu x}|) = 1$

iv) Soit $\mu_0 \in \mathbb{R}$ et montrons que φ_X est continue en μ_0 .

$$\varphi_X(\mu) = \int e^{i\mu x} dP_X(x)$$

Il s'agit d'une $\int$ dépendant d'un paramètre.

Ici, $\forall \mu \in \mathbb{R} \quad \forall x \in \mathbb{R} \quad |e^{i\mu x}| \leq 1$ [majoration uniforme en μ par une fonction P -intégrable]

Par le TCD : $\lim_{\mu \rightarrow \mu_0} \varphi_X(\mu) = \varphi_X(\mu_0)$

Effet d'une transformation affine:

Soit $\alpha, \beta \in \mathbb{R}$

Alors $\forall \mu \in \mathbb{R}, \varphi_{\alpha X + \beta}(\mu) = e^{i\mu\beta} \varphi_X(\alpha\mu)$

$$(\mathbb{E}[e^{i\mu(\alpha X + \beta)}]) = \mathbb{E}(e^{i\mu\beta} e^{i\mu\alpha X}) = e^{i\mu\beta} \mathbb{E}(e^{i\alpha\mu X})$$

4) La loi gaussienne

Soit $m \in \mathbb{R}, \sigma > 0$ et X une v.a. de loi $N(m, \sigma^2)$

* Posons $Z = \frac{X-m}{\sigma}$ ssi $X = m + \sigma Z$. alors $Z \sim N(0, 1)$. Calculons φ_Z

$$\varphi_Z(\mu) = \mathbb{E}(e^{i\mu Z}) = \int_{\mathbb{R}} e^{i\mu x} \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx$$

* on met sous forme canonique: $i\mu x - \frac{x^2}{2} = -\frac{1}{2}(x - i\mu)^2 - \frac{\mu^2}{2}$

$$= \int_{\mathbb{R}} e^{-1/2(x-i\mu)^2} e^{-\mu^2/2} \frac{dx}{\sqrt{2\pi}} = e^{-\mu^2/2} \int_{\mathbb{R}} e^{-1/2(x-i\mu)^2} \frac{dx}{\sqrt{2\pi}}$$

* Calcul de $I(\mu) = \int_{\mathbb{R}} e^{-1/2(x-i\mu)^2} dx$

$\rightarrow I(0) = \int_{\mathbb{R}} e^{-1/2 x^2} dx = \sqrt{2\pi}$

Calculons $I'(\mu)$

$$I(\mu) = \int f(x, \mu) dx \text{ avec } f(x, \mu) = e^{-1/2(x-i\mu)^2}$$

(on espère: $I'(\mu) = \int \frac{\partial f}{\partial \mu}(x, \mu) dx$)

$f(x, \mu)$ est C^1 de $\mathbb{R} \times \mathbb{R} \rightarrow \mathbb{C}$

$$\frac{\partial f}{\partial \mu}(x, \mu) = i(x-i\mu)e^{-1/2(x-i\mu)^2}$$

[on veut que la majoration soit uniforme en μ]

$$\left| \frac{\partial f}{\partial \mu}(x, \mu) \right| \leq (|x| + |\mu|) e^{-1/2(x^2 + \mu^2)}$$

La fonction $u \in \mathbb{R} \mapsto |u| e^{-u^2/2}$ est bornée sur $\mathbb{R}$, et donc
 $\exists M > 0 \quad \forall u \in \mathbb{R} \quad \forall x \in \mathbb{R} \quad \left| \frac{\partial f}{\partial u}(x, u) \right| \leq (|x| + M) e^{-x^2/2}$

Cette fonction est λ intégrable et ne dépend pas de u . Alors I est dérivable et sa dérivée vaut :

$$I'(u) = \int_{\mathbb{R}} \frac{\partial f}{\partial u}(x, u) dx = \int_{\mathbb{R}} i(x - iu) e^{-1/2 (x - iu)^2} dx$$

$$\frac{\partial f}{\partial u} = i(x - iu) e^{-1/2 (x - iu)^2}$$

$$\frac{\partial f}{\partial x} = -(x - iu) e^{-1/2 (x - iu)^2}$$

$$= -i \int_{\mathbb{R}} \frac{\partial f}{\partial x}(x, u) dx = -i [f(x, u)]_{-\infty}^{+\infty} = 0$$

d'où $I(u) = cte = I(0) = \sqrt{2\pi}$

On trouve finalement $\psi_Z(u) = e^{-u^2/2}$ si $Z \sim N(0, 1)$
 si $X \sim N(m, \sigma^2)$, $\psi_X(u) = e^{i u m - \frac{\sigma^2 u^2}{2}}$ ($X = \sigma Z + m$)

Posons $g_0(x) = \frac{e^{-x^2/2\sigma^2}}{\sqrt{2\pi}\sigma}$

$$\int e^{i u x} g_0(x) dx = \frac{\sqrt{2\pi}}{\sigma} g_{1/\sigma}(u)$$

5) Thm d'injectivité

Notons $M_1(\mathbb{R})$ l'ensemble des mesures de probas sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$

L'application qui à une mesure de probas μ associe sa f. c. ψ_μ est injective

$\mu \in M_1(\mathbb{R}) \longrightarrow \psi_\mu$. Ainsi si X est une v.a., la fonction ψ_X caractérise la loi de X .

cas particulier: si μ admet une densité f_μ , alors $\psi_\mu(u) = \int e^{i u x} f(x) d\lambda(x)$

preuve: Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probas sur lequel sont définies 2 v.a. indépendantes X et Y telles que $X \sim \mu$ et $Y \sim N(0, 1)$ et soit $\sigma > 0$

Idee: Nous allons convoluer μ avec une gaussienne pour la régulariser et nous allons mq cette loi convolée est déterminée par ψ_μ . Pour cela, nous étudions $X + \sigma Y$.

Soit $f: \mathbb{R} \rightarrow \mathbb{R}$ continue à support compact.

$$\mathbb{E}[f(X + \sigma Y)] = \int_{\Omega} f(X + \sigma Y) d\mathbb{P} = \int_{\mathbb{R}} \int_{\mathbb{R}} f(x + \sigma y) d\mathbb{P}_{(X, Y)}(x, y) \quad \mathbb{P}_{(X, Y)} = \mathbb{P}_X \otimes \mathbb{P}_Y$$

On sait que $\mathbb{P}_X = \mu$ et $\mathbb{P}_Y = N(0, 1)$

$$\mathbb{E}[f(X + \sigma Y)] = \iint f(x + \sigma y) d\mu(x) g_1(y) dy = \iint f(x + y') d\mu(x) g_\sigma(y') dy'$$

On a vu que $\int e^{i u x} g_\sigma(x) dx = \frac{\sqrt{2\pi}}{\sigma} g_{1/\sigma}(u)$

$$g_\sigma(y') = \frac{1}{\sqrt{2\pi}\sigma} \int e^{i y' t} g_{1/\sigma}(t) dt \quad \begin{matrix} x \rightarrow t \\ u \rightarrow y' \\ \sigma \rightarrow 1/\sigma \end{matrix}$$

$$\mathbb{E}[f(X + \sigma Y)] = \iiint f(x + y') \frac{1}{\sqrt{2\pi}\sigma} e^{i y' t} g_{1/\sigma}(t) dt dy' d\mu(x) \quad \begin{matrix} z = x + y' \\ \text{garder } x' \end{matrix}$$

$$= \iiint f(z) \frac{1}{\sqrt{2\pi}\sigma} e^{i(z-x)t} g_{1/\sigma}(t) dt dz d\mu(x)$$

THM Fubini: $|f(z) e^{i(z-x)t} g_{1/\sigma}(t)| \leq \|f\|_\infty g_{1/\sigma}(t)$

f support compact $L^1(\mathbb{R} \times \mathbb{R} \times \mathbb{R}, \mu \otimes \lambda \otimes \lambda)$

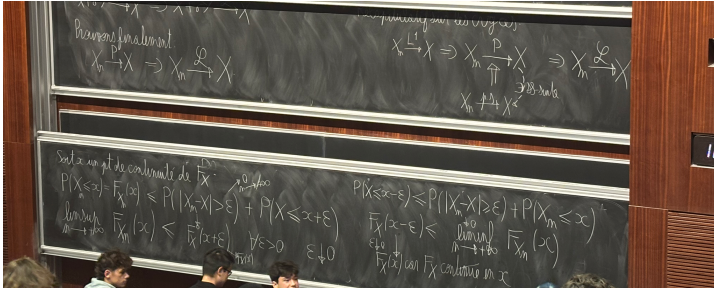
$$\begin{aligned}\mathbb{E}[f(X+\sigma Y)] &= \int \int_{\mathbb{C}} f(z) \frac{1}{\sqrt{2\pi}\sigma} e^{izt} g_{1/\sigma}(t) \left(\int_{\mathbb{R}} e^{-itx} d\mu_n(x) \right) dt dz \\ &= \int \int_{\mathbb{C}} f(z) \frac{1}{\sqrt{2\pi}\sigma} e^{izt} g_{1/\sigma}(t) \psi_n(-t) dt dz\end{aligned}$$

Ainsi la loi de $X + \sigma Y$ est déterminée par ψ_n .

Faisons $\sigma \rightarrow 0$. On dit que $X + \sigma Y \xrightarrow{\mathcal{L}} X$

et donc la loi de X est déterminée $\sigma \rightarrow 0$ par la famille de v.a. $X + \sigma Y, \sigma > 0$ dont les lois sont déterminées par ψ_n .

Prouvons finalement $X_n \xrightarrow{\mathcal{L}} X \Rightarrow X_n \xrightarrow{\mathcal{L}} X$



Cours 10

1) Soit le théorème de Paul Lévy

Soit $X, X_n, n \in \mathbb{N}$ des v.a.s réelles et soit $\psi_X, \psi_{X_n}, n \in \mathbb{N}$ leur fonctions caractéristique

Nous avons équivalence entre :

i) $X_n \xrightarrow{\mathcal{L}} X$

ii) $\forall u \in \mathbb{R}, \lim_{n \rightarrow +\infty} \psi_{X_n}(u) = \psi_X(u)$

Dém:

Sens facile i) $\Rightarrow$ ii)

$$\psi_{X_n}(u) = \mathbb{E}(e^{iuX_n}) = \mathbb{E}(\cos(uX_n)) + i \mathbb{E}(\sin(uX_n))$$

les fonctions $x \mapsto \cos ux, \sin ux$ (u étant fixe) sont continues bornées et donc par définition de la convergence en loi :

$$\mathbb{E}(\cos uX_n) \xrightarrow{n \rightarrow +\infty} \mathbb{E}(\cos uX) \quad \text{et donc} \quad \psi_{X_n}(u) \xrightarrow{n \rightarrow +\infty} \psi_X(u)$$

$$\mathbb{E}(\sin uX_n) \xrightarrow{n \rightarrow +\infty} \mathbb{E}(\sin uX)$$

Sens réciproque : ii) $\Rightarrow$ i)

Nous utilisons la même stratégie que pour prouver l'injectivité de la transformation de Fourier des mesures de probas

Nous supposons que $X, X_n, n \in \mathbb{N}$ sont toutes définies sur un même espace de probas $(\Omega, \mathcal{F}, \mathbb{P})$ sur lequel est aussi définie une v.a. Y indépendante de $X, X_n, n \in \mathbb{N}$ et de loi $\mathcal{N}(0,1)$

Soit $f: \mathbb{R} \rightarrow \mathbb{R}$ une fonction continue à support compact. Nous supposons ii) vraie et nous voulons $X_n \xrightarrow{\mathcal{L}} X$

Soit $\sigma > 0$ et écrivons

$$\mathbb{E}(f(X_n)) - \mathbb{E}(f(X)) = \mathbb{E}(f(X_n) - f(X_n + \sigma Y)) + \mathbb{E}(f(X_n + \sigma Y)) - \mathbb{E}(f(X + \sigma Y)) + \mathbb{E}(f(X + \sigma Y) - f(X))$$

Comme f est continue à support compact, elle est uniformément continue.

$$\forall \varepsilon > 0 \exists \delta > 0 \forall x, y \in \mathbb{R} |x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon$$

Soit $\varepsilon > 0$ fixé. Soit δ associé à ε comme ci-dessus. Écrivons :

$$|\mathbb{E}(f(X_n) - f(X_n + \sigma Y))| \leq \mathbb{E}(|f(X_n) - f(X_n + \sigma Y)|)$$

(on peut majorer chacune des intégrales)

$$= \int_{|\sigma Y| \geq \delta} |f(X_n) - f(X_n + \sigma Y)| d\mathbb{P} + \int_{|\sigma Y| < \delta} |f(X_n) - f(X_n + \sigma Y)| d\mathbb{P}$$

$$\int_{|\sigma y| \geq \delta} |1 dP| \leq 2 \|f\|_\infty \int_{|\sigma y| \geq \delta} 1 dP = 2 \|f\|_\infty P(|\sigma y| \geq \delta)$$

je dis que $\lim_{\sigma \rightarrow 0} P(|\sigma y| \geq \delta) = \lim_{\sigma \rightarrow 0} P(|y| \geq \frac{\delta}{\sigma}) = 0$ donc $\exists \sigma_0$ $2 \|f\|_\infty P(|\sigma_0 y| \geq \delta) < \varepsilon$

$$\int_{|\sigma y| < \delta} |f(x_n) - f(x_n + \sigma y)| dP < \varepsilon$$

On a donc $|E(f(x_n) - f(x_n + \sigma_0 y))| < 2\varepsilon$ (vraie pour tout n !)

Par le même argument, on a $|E(f(x + \sigma_0 y) - f(x))| < 2\varepsilon$. Reste le terme du milieu.

On a donc:

$$\forall n \in \mathbb{N}, |E(f(x_n)) - E(f(x))| \leq 2\varepsilon + |E(f(x_n + \sigma_0 y) - f(x + \sigma_0 y))| + 2\varepsilon$$

$$\text{Nous utilisons la formule du cours précédent: } E(f(x_n + \sigma_0 y)) = \int_{\mathbb{R}} \int_{\mathbb{R}} f(z) e^{iuz} \frac{1}{\sqrt{2\pi}\sigma} g_{1/\sigma}(u) \varphi_{x_n}(-u) du dz$$

$$\begin{aligned} & \text{Posons } F_n(u, z) = f(z) e^{iuz} \frac{1}{\sqrt{2\pi}\sigma} g_{1/\sigma}(u) \varphi_{x_n}(-u) \\ & \text{C'est majoré par une fct qui dépend pas de } n \text{ (car } |\varphi_{x_n}(-u)| \leq 1 \text{)} \\ & |F_n(u, z)| \leq |f(z)| \cdot \frac{1}{\sqrt{2\pi}\sigma} g_{1/\sigma}(u) \text{ domination par une fonction qui ne dépend pas de } n \text{ et qui est dans } \mathcal{L}^1(\mathbb{R}^2, \lambda^{\otimes 2}) \end{aligned}$$

$\Rightarrow$ d'où la convergence souhaitée (en appliquant le thm de convergence dominée)

Par TCD:

$$\exists N \forall n \geq N |E(f(x_n + \sigma_0 y)) - E(f(x + \sigma_0 y))| < \varepsilon$$

Alors pour $n \geq N$

$$|E(f(x_n)) - E(f(x))| \leq 2\varepsilon + \varepsilon + 2\varepsilon < 5\varepsilon$$

3) Moments

PROP: Soit X une v.a. réelle. Soit $K \geq 1$ et supposons que $E(|X|^K) < +\infty$.

Alors φ_X est K fois dérivable, et de plus $E(X^K) = i^K \varphi_X^{(K)}(0)$

$$\text{PREUVE: } \varphi_X(u) = \int_{\mathbb{R}} e^{iux} dP_X(x)$$

Faisons le cas où $K=1$: $E(|X|) < +\infty$.

Il s'agit d'une $\int$ à paramètre, et on veut dériver / u .

On pose $F(x, u) = e^{iux}$

Or $\frac{\partial F}{\partial u}(x, u) = ix e^{iux} \sim |x|$ fonction majorante qui est P_X -intégrable et qui ne dépend pas de u .

$$\begin{aligned} \text{Thm de dérivation: } \varphi_X \text{ dérivable et } \varphi_X'(u) &= \frac{\partial F}{\partial u}(x, u) = i \int x e^{iux} dP_X(x) \\ \text{d'où } \varphi_X'(0) &= i \int x dP_X(x) = i E(X) \end{aligned}$$

Ce résultat permet de trouver facilement la suite des moments.

4) Exemple:

$X \sim U([-1, 1])$

$$\varphi_X(u) = \int_{\mathbb{R}} e^{iux} 1_{[-1, 1]}(x) \frac{dx}{2} = \frac{e^{iu} - e^{-iu}}{2iu} = \frac{\sin u}{u}$$

$$\left(\sin = \sum_{n \geq 0} \frac{(-1)^n u^{2n+1}}{(2n+1)!} \right) \leadsto \frac{\sin u}{u} = \sum_{n \geq 0} \frac{(-1)^n u^{2n}}{(2n+1)!}$$

$$\text{d'où } \varphi_X^{(2n)}(0) = \frac{(-1)^n}{(2n+1)!} (2n)! = \frac{(-1)^n}{2n+1}$$

$$\mathbb{E}(X^{2n}) = \frac{1}{2^{n+1}} \int_{-1}^1 x^{2n} \frac{dx}{2}$$

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

$$\varphi_X(u) = e^{-u^2/2} = \sum_{n \geq 0} \frac{1}{n!} \left(\frac{-u^2}{2}\right)^n = \sum_{n \geq 0} \frac{(-1)^n u^{2n}}{n! 2^n}$$

(nulle si impaire)

$$\text{d'où } \varphi_X^{(2n)}(0) = \frac{(-1)^n 2n!}{n! 2^n} = (-1)^n \mathbb{E}(X^{2n}) \text{ d'où } \mathbb{E}(X^{2n}) = \frac{(2n)!}{n! 2^n}$$

À chaque distribution symétrique par rapport à 0 (comme $\mathcal{N}(0,1)$ ou $\mathcal{U}(-1,1)$) les moments impairs = 0

5) Convolution

Lemme: Soient X, Y 2 v.a. indépendantes. Alors $\forall u \in \mathbb{R} \quad \varphi_{X+Y}(u) = \varphi_X(u) \varphi_Y(u)$

$$\begin{aligned} \text{Preuve: } \varphi_{X+Y}(u) &= \mathbb{E}(e^{iu(X+Y)}) = \mathbb{E}(e^{iuX} e^{iuY}) \\ &= \mathbb{E}(e^{iuX}) \mathbb{E}(e^{iuY}) = \varphi_X(u) \varphi_Y(u) \end{aligned}$$

indep

Corollaire: Si $X_1, \dots, X_n$ sont n v.a. i.i.d. alors $\varphi_{X_1 + \dots + X_n}(u) = (\varphi_{X_1}(u))^n$

Supposons que X a une loi de densité f , Y a une loi de densité g et X, Y indépendantes.

$$\text{Alors } \varphi_{X+Y}(u) = \iint_{\mathbb{R}^2} e^{iu(x+y)} f(x) g(y) dx dy$$

transfert x, y

Pose $x = x, z = x + y$ garde x , remplace y

$$\varphi_{X+Y}(u) = \iint_{\mathbb{R}^2} e^{iu z} f(x) g(z-x) dx dz \quad (\text{Fubini})$$

$$= \int_{\mathbb{R}} e^{iu z} \left(\int_{\mathbb{R}} f(x) g(z-x) dx \right) dz = \int_{\mathbb{R}} e^{iu z} h(z) dz \quad \text{où } h(z) = \int_{\mathbb{R}} f(x) g(z-x) dx = f * g(z) \text{ produit de convolution}$$

de f et g . Par identification des 2 formules on conclut que $X+Y$ a une loi de densité $f * g$

$$(f * g)(z) = \int_{\mathbb{R}} f(x) g(z-x) dx \quad \text{2 gaussienne} = \int_{\mathbb{R}} f(z-x) g(x) dx$$

Dès que f et g sont dans $\mathcal{L}^1$, le produit de convolution $f * g$ est bien défini et aussi dans $\mathcal{L}^1(\mathbb{R})$

6) Le thm limite central

Soit $(X_n)_{n \geq 1}$ une suite de v.a. i.i.d. qui admettent un moment d'ordre 2.

$$\text{Posons } m = \mathbb{E}(X_1) \text{ et } \sigma^2 = \text{Var}(X_1). \text{ Alors } \frac{X_1 + \dots + X_n - nm}{\sqrt{n}} \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2)$$

Attention, il faut absolument centrer les v.a. en soustrayant $n\mathbb{E}(X_1) = nm$.

$$\text{Par contre on peut normaliser par } \sqrt{n \text{Var}(X_1)}, \text{ i.e. } \frac{X_1 + \dots + X_n - nm}{\sqrt{n} \sigma} \xrightarrow{\mathcal{L}} \mathcal{N}(0, 1)$$

Il s'agit d'un résultat universel.

Dès que les v.a. (X_n) ont un moment d'ordre 2, les fluctuations de $X_1 + \dots + X_n$ autour de sa moyenne sont décrites par la loi gaussienne. Le TLC ou TCL est un raffinement de la loi des grands nombres sous l'hypothèse d'existence d'un moment d'ordre 2. $\frac{X_1 + \dots + X_n}{n} = \mathbb{E}(X_1) + \frac{R_n}{\sqrt{n}}, R_n \text{ v.a. tq } R_n \xrightarrow{\mathcal{L}} \mathcal{N}(0, \sigma^2)$

7) Preuve TCL

$$\text{Posons } Y_n = \frac{X_1 + \dots + X_n - nm}{\sqrt{n}} \text{ et aussi } \tilde{X}_K = X_K - m, 1 \leq K \leq n, \text{ et étudions la convergence en loi}$$

de Y_n à l'aide des fonctions caractéristiques.

$$\varphi_{Y_n}(u) = \mathbb{E}(e^{iu Y_n}) = \mathbb{E}(e^{iu \frac{\tilde{X}_1 + \dots + \tilde{X}_n}{\sqrt{n}}}) = \mathbb{E}(e^{iu \tilde{X}_1 / \sqrt{n}}) \prod_{i=2}^n \mathbb{E}(e^{iu \tilde{X}_i / \sqrt{n}})$$

$$= \left(\varphi_{\tilde{X}_1} \left(\frac{u}{\sqrt{n}} \right) \right)^n$$

$\varphi_{\tilde{X}_1}(u)$ est 2 fois dérivable car $\mathbb{E}(\tilde{X}_1^2) < +\infty$, donc elle admet un développement limité à l'ordre 2

$$\psi_{\tilde{X}_1}(\mu) = \psi_{\tilde{X}_1}(0) + \mu \psi'_{\tilde{X}_1}(0) + \frac{\mu^2}{2} \psi''_{\tilde{X}_1}(0) + o(\mu^2)$$

$$= 1 + i \mathbb{E}(\tilde{X}_1) \mu - \frac{\mu^2}{2} \mathbb{E}(\tilde{X}_1^2) + o(\mu^2)$$

$$= 1 + 0 - \frac{\sigma^2 \mu^2}{2} + o(\mu^2), \quad \sigma^2 = \mathbb{E}(\tilde{X}_1^2)$$

$$\text{d'où } \psi_{\tilde{X}_1}\left(\frac{\mu}{\sqrt{n}}\right) = 1 - \frac{\sigma^2 \mu^2}{2n} + o\left(\frac{1}{n}\right) \quad n \longrightarrow +\infty$$

$$\psi_{Y_n}(\mu) = \left(1 - \frac{\sigma^2 \mu^2}{2n} + o\left(\frac{1}{n}\right)\right)^n$$

$$= \exp n \ln\left(1 - \frac{\sigma^2 \mu^2}{2n} + o\left(\frac{1}{n}\right)\right)$$

$$= \exp -\frac{\sigma^2 \mu^2}{2} + o(1) \longrightarrow e^{-\frac{\sigma^2 \mu^2}{2}} = \psi_{N(0,1)}(\mu)$$

on conclut par le Thm de Levy

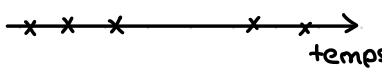
$$(\text{Ici, on utilise le ln complexe}) \quad (\ln(1+z)) = z - \frac{z^2}{2} + \frac{z^3}{3} - \dots = \sum_{n \geq 1} \frac{(-1)^{n-1}}{n} z^n \quad \text{rayon 1}$$

Cours 11

Le processus de Poisson

Nous voulons modéliser des événements qui arrivent à des dates précises mais imprévisibles.

Exemples: séismes, désintégration d'atomes, pluie, début ou fin d'appels téléphoniques.

À chaque occurrence d'un événement, nous marquons un impact sur l'axe des temps: 

Hypothèses: i) Les conditions de l'expérience ne changent pas au cours du temps

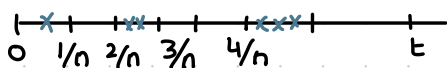
ii) les nbs d'impacts dans des intervalles de temps disjoints sont indépendants

Pour $t \geq 0$, nous notons $N(t)$ le nombre d'impacts dans $[0, t]$

Nous avons alors une famille de v.a. $(N(t))_{t \geq 0}$ indexée par un paramètre continu. c'est un processus stochastique

1) Construction d'un modèle discret

Soit $n \geq 1$ un entier. Nous subdivisons l'axe réel en intervalles de longueur $1/n$: $\left[\frac{k}{n}, \frac{k+1}{n}\right[$; $k \geq 0$



Un intervalle $\left[\frac{k}{n}, \frac{k+1}{n}\right[$ est dit occupé s'il contient au moins un impact, et vide sinon.

Notons p_n la probabilité qu'un intervalle $\left[\frac{k}{n}, \frac{k+1}{n}\right[$ soit occupé. Par l'hypothèse i), p_n ne dépend pas de l'intervalle.

Pour $t > 0$, nous notons $N_n(t)$ le nombre d'intervalles $\left[\frac{k}{n}, \frac{k+1}{n}\right[$ qui sont inclus dans $[0, t]$ ($\frac{k+1}{n} \leq t$) et qui sont occupés. Par l'hypothèse ii), les états vides ou occupés de ces intervalles sont indépendants et $N_n(t)$ suit la loi binomiale $B(\lfloor nt \rfloor, p_n)$

Étudions la limite en loi de $N_n(t)$, par exemple à l'aide des fonctions caractéristiques.

$$\psi_{N_n(t)}(\mu) = \mathbb{E}(e^{i\mu N_n(t)}) = \left((1-p_n) + p_n e^{i\mu}\right)^{\lfloor nt \rfloor} \quad \text{Hypothèse: } p_n \longrightarrow 0, \text{ sinon le modèle explose}$$

$$= e^{\lfloor nt \rfloor \ln(1 + p_n(e^{i\mu} - 1))}$$

$$= e^{\lfloor nt \rfloor (p_n(e^{i\mu} - 1) + o(p_n))}$$

$$= e^{\lfloor nt \rfloor p_n(e^{i\mu} - 1)} + o(np_n)$$

Supposons $n p_n \xrightarrow{n \rightarrow +\infty} \lambda$

$$\ln t \sim \ln n t \quad \ln t \sim \ln p_n \rightarrow \lambda t$$

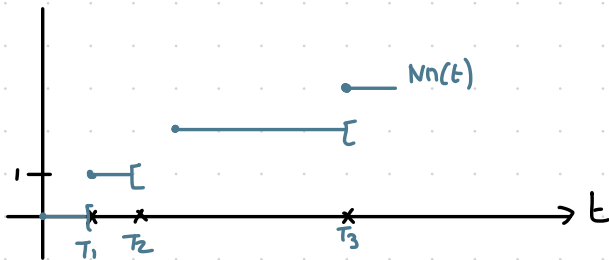
Alors $\lim_{n \rightarrow +\infty} \varphi_{N_n}(t)(\mu) = e^{\lambda(e^{\mu}-1)}$ et donc $N_n(t) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} P(\lambda)$

Rappel: $X \sim P(\lambda)$ si X est à valeurs dans $\mathbb{N}$ et $\forall k \in \mathbb{N}$, $P(X=k) = e^{-\lambda} \frac{\lambda^k}{k!}$
 $E(X) = \lambda$, $\text{Var}(X) = \lambda$, $\varphi_X(\mu) = \exp(\lambda(e^{\mu}-1))$

2) Instants de saut

Nous avons montré que, à t fixe, $N_n(t) \rightarrow P(\lambda)$

Mais en fait, la fonction $t \mapsto N_n(t)$ est une fonction aléatoire nulle en 0, croissante, d'accroissements 0 ou 1



La loi de la fonction aléatoire converge-t-elle quand $n \rightarrow \infty$?

Nous définissons les instants de sauts de $N_n(t)$: $\forall k \geq 1$, $T_k^n = \inf \{ t \geq 0; N_n(t) = k \}$

La fonction $N_n(t)$ est complètement déterminée par ses instants de saut. En effet, $N_n(t) = \sum_{k \geq 0} k \mathbb{1}_{[T_k^n, T_{k+1}^n)}(t)$

Examinons la convergence des instants de saut T_k^n et commençons par T_1 .

Soit $i \geq 0$, regardons $P(T_1^n > i/n) = P([0, i/n[\text{ vide }, [i/n, i/n] \text{ vide})$

Nous introduisons une suite de v.a. X_k , $k \geq 0$ tq :

$$X_k = \begin{cases} 0 & \text{si l'intervalle } [k/n, (k+1)/n \text{ vide} \\ 1 & \text{occupé} \end{cases}$$

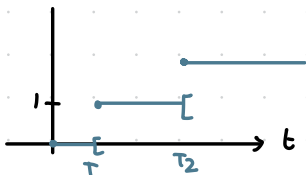
$$P(T_1^n > i/n) = P(X_0=0, \dots, X_{i-1}=0) = (1-p_n)^i$$

$$\text{d'où } \forall t > 0, P(T_1^n > t) = (1-p_n)^{\lfloor \ln t \rfloor} \xrightarrow[n \rightarrow +\infty]{} e^{-\lambda t}$$

$$\text{d'où } \forall t, F_{T_1^n}(t) \rightarrow 1 - e^{-\lambda t}$$

$$\text{d'où } T_1^n \xrightarrow{\mathcal{L}} \text{Exp}(\lambda)$$

Rappel: Une v.a. X suit la loi $\text{Exp}(\lambda)$ si sa loi est de densité $\mathbb{1}_{\mathbb{R}^+}(x) e^{-\lambda x} dx$
 et alors $E(X) = 1/\lambda$, $\text{Var}(X) = 1/\lambda^2$, $F_X(t) = 1 - e^{-\lambda t}$



je dis que $T_2^n - T_1^n$ est indépendant de T_1^n et a la même loi.

Vérifions: soit $i, j \geq 1$ et regardons

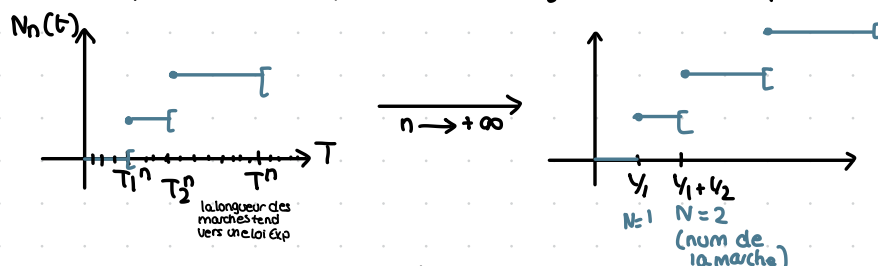
$$P(T_1^n = i/n, T_2^n - T_1^n = j/n) = P(X_0=0, X_1=0, \dots, X_{i-1}=0, X_i=1, X_{i+1}=0, \dots, X_{i+j}=0, X_{i+j+1}=1) \\ = (1-p_n)^i p_n (1-p_n)^j p_n = P(T_1^n = i/n) P(T_2^n - T_1^n = j/n)$$

d'où le résultat

ainsi, $P(T_1^n, T_2^n - T_1^n) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \text{Exp}(\lambda) \otimes \text{Exp}(\lambda)$ 2 lois exponentielles indépendantes.
 Par récurrence, on mq: $(T_1^n, T_2^n - T_1^n, \dots, T_K^n - T_{K-1}^n) \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} (\text{Exp}(\lambda))^{\otimes K}$

4) Le processus de Poisson

Les calculs précédents nous permettent d'imaginer le candidat pour la limite en loi de $N_n(t)$



Soit $(Y_n)_{n \geq 1}$ une suite i.i.d. $\text{Exp}(\alpha)$

Posons $\forall n \geq 1$ $S_n = Y_1 + Y_2 + \dots + Y_n$

Définissons $\forall t \geq 0$, $N(t) = \max \{n \geq 0; S_n \leq t\}$

Alors $(N(t))_{t \geq 0}$ est le processus de Poisson d'intensité λ

Pour chaque t , $N(t)$ est une v.a. $\sim P(\lambda t)$

Nous nous avons construit une continuum de v.a. $(N(t), t \geq 0)$ sur le même espace de probabilité.

$S_n \xrightarrow[n \rightarrow +\infty]{p.s.} N(t)$ et $N(t)$ est l'unique indice tq $S_{N(t)} \leq t < S_{N(t)+1}$

5) La loi exponentielle

La loi qui régit l'intervalle entre 2 impacts successifs est la loi $\text{exp}(\lambda)$. Elle surgit comme limite de lois géométriques renormalisées.

Si $X \sim \text{geo}(\lambda/n)$ $\frac{1}{n} X \xrightarrow[n \rightarrow +\infty]{\mathcal{L}} \text{Exp}(\lambda)$

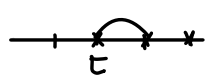
Ces 2 lois possèdent une propriété remarquable, l'absence de mémoire.

i.e. $T \sim \text{exp}(\lambda)$, $\forall s, t > 0$ $P(T > t+s | T > s) = P(T > t)$

La loi du temps résiduel à attendre sachant qu'on a déjà attendu s est la même que la loi du temps d'attente total. Elles sont utilisées pour modéliser des phénomènes qui ne présentent pas de vieillissement (ex: temps de vie d'un atome - durée d'un appel téléphonique incohérent) [$f(t+s) = f(t)f(s)e^{-\lambda t}$]

6) Applications

Autobus: les arrivées de bus à un arrêt sont modélisées par un processus de Poisson (α) . Le temps moyen entre 2 bus est $1/\alpha$. J'arrive à l'instant t . Quel sera mon temps d'attente?



1^{er} raisonnement: Vu l'absence de mémoire de la loi $\text{exp}(\alpha)$, mon temps d'attente suit une loi $\text{exp}(\alpha)$ donc en moyenne $1/\alpha$

2^e raisonnement: j'arrive au hasard entre 2 bus, par symétrie, mon temps moyen d'attente est la moitié de la longueur typique des intervalles $1/2\alpha^{-1}$

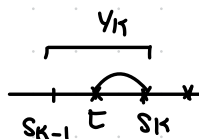
$S_n = Y_1 + \dots + Y_n$, $N(t)$. S_n instants d'arrivées des bus.

Soit K tq $S_{K-1} < t < S_K + Y_K$.

Mon temps d'attente est $S_K - t$.

tps moyen: $E(S_K - t)$

$Y_K \sim \text{Exp}(\alpha)$



Cours 12

Percolation

1) Le problème

Immergeons une pierre spongieuse dans l'eau. Elle contient des petits canaux.

molécule d'eau $\ll$ diamètre canal $\ll$ pierre

Le centre de la pierre sera-t-il mouillé? En 1957, Broadbent et Hammerly inventèrent le modèle math de la percolation, pour étudier les milieux poreux aléatoires.

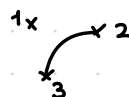
2) Le réseau carré

Définition: un graphe est un couple (V, E) où:

- V est un ensemble de points: les sommets (vertices)
- E est un ensemble de paires de points de V : les arêtes (edges)

$$V = \{1, 2, 3\}$$

$$E = \{\{2, 3\}\}$$

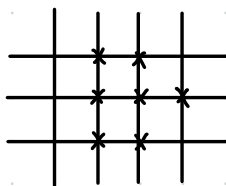


Nous munissons $\mathbb{Z}^2$ d'une structure de graphe.

L'ensemble $\mathbb{E}^2$ des arêtes est constitué des paires $\{x, y\}$ où $x, y \in \mathbb{Z}^2$ sont plus proches voisins,

i.e. $|x - y|_2 = 1$.

DEF: Le réseau carré est le graphe $(\mathbb{Z}^2, \mathbb{E}^2)$

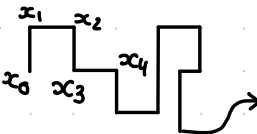


DEF: Un chemin est une suite de sommets $x_0, x_1, \dots, x_n$, $n \geq 1$ à 2 distincts tq:

pour tout $i \geq 0$, $\{x_i, x_{i+1}\}$ est une arête

Si le chemin s'arrête au sommet x_n , nous disons qu'il relie x_0 à x_n et qu'il est fini de longueur n

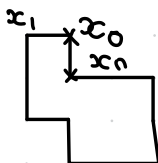
Si il est infini, nous disons qu'il relie x_0 à l'infini.



DEF: Un circuit est un chemin $x_0, \dots, x_n$

tq x_n et x_0 sont voisins

Un tel circuit a pour longueur $n+1$



3) L'espace de probabilité

Les arêtes du graphe jouent le rôle des canaux.

Nous allons les ouvrir ou fermer au hasard indépendamment, avec probas $p \in]0, 1[$

l'eau peut passer $\rightarrow$ bloqués

Nous prenons comme espace de configurations: $\Omega = \{0, 1\}^{\mathbb{E}^2}$ associe un état à chaque arête

Un élément ω de Ω est une fonction de $\mathbb{E}^2 \rightarrow \{0, 1\}$

L'arête $e \in \mathbb{E}^2$ est ouverte (resp fermée) dans la configuration ω si $\omega(e) = 1$ (resp. 0)

Nous munissons Ω de la tribu $\tilde{\mathcal{F}}$ cylindrique, engendré par les événements de la forme

$\{\omega \in \Omega; \omega(e_1) = \varepsilon_1, \dots, \omega(e_n) = \varepsilon_n\}; n \geq 1, e_1, \dots, e_n \in \mathbb{E}^2, \varepsilon_1, \dots, \varepsilon_n \in \{0, 1\}$

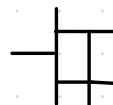
A-t-on $\mathcal{F} = \mathcal{P}(\Omega)$? Non

Nous munissons finalement l'espace mesurable $(\Omega, \tilde{\mathcal{F}})$ de la probas $P_p = \bigotimes_{e \in \mathbb{E}^2} ((1-p)\delta_{\omega(e)=0} + p\delta_{\omega(e)=1})$
(produit tensoriel des Bernouillis p)

4) Les clusters ouverts

Nous tirons au hasard l'état des arêtes. Nous enlevons les arêtes fermées.

Nous obtenons ainsi un graphe aléatoire. Précisément, soit ω une configuration.



Nous considérons le graphe de sommets les points de $\mathbb{Z}^2$, et d'arêtes les arêtes ouvertes dans ω uniquement



Le but de la percolation est de comprendre la géométrie de ce graphe aléatoire.

Si $x \in \mathbb{Z}^2$, la composante connexe de x dans ce graphe est appelé le cluster ouvert de x dans ω et est noté $C(x, \omega)$

* si $x, y \in \mathbb{Z}^2$, on note $\{x \leftrightarrow y\}$ l'événement : il existe dans ω un chemin qui relie x à y et dont toutes les arêtes sont ouvertes

Ainsi, $C(x) = C(x, \omega)$ est un ensemble aléatoire

$|C(x, \omega)|$ cardinal / diam $C(x)$...

5) La probabilité de percolation

$$[C(x, \omega) \neq C(0, \omega)]$$

La quantité fondamentale pour analyser le modèle est:

$$\theta(p) = P_p(|C(0)| = \infty) = P_p(0 \leftrightarrow \infty)$$

$$\theta: [0, 1] \rightarrow [0, 1]$$

$$\theta(0) = 0 \text{ (toutes les arêtes sont fermées)}$$

θ croissante. Pour cela, on utilise un couplage, i.e. on construit un espace

$(\Omega_1, \mathcal{F}_1, \mathbb{P}_1, \mathbb{P}_2)$ tq sous $\mathbb{P}_1$ la loi de la 1^{re} marginale ω_1 est p_1

2nd ω_2 est p_2

$$p_1 < p_2$$

$$X_1 \leq X_2$$

et tq toutes les arêtes ouvertes dans ω_1 le sont aussi dans ω_2 .

Si on a pu faire cela, alors $\theta(p_1) = P_{p_1}(0 \leftrightarrow \infty) = \mathbb{P}(0 \leftrightarrow \infty \text{ dans } \omega_1) \leq \mathbb{P}(0 \leftrightarrow \infty \text{ dans } \omega_2) = P_{p_2}(0 \leftrightarrow \infty) = \theta(p_2)$

On prend $\omega \sim \text{Unif}([0, 1])$ et $X_p = 1_{\{\omega \leq p\}}$

$$[\mathbb{P}(X_p = 1) = p]$$

$$0 \leq p \leq 1 \text{ et } X_{p_1} \leq X_{p_2}$$

6) La transition de phases

Le modèle de percolation est fascinant car c'est le modèle le plus simple qui présente une transition de phase.

THM: Il existe une valeur unique p_c strictement entre 0 et 1 tq $\theta(p) = 0$ si $p < p_c$ et $\theta(p) > 0$ si $p > p_c$

preuve:

Mq $p_c > 0$

$p_c = \sup \{p \in [0, 1]; \theta(p) = 0\}$. Soit $n \geq 1$.

$$\theta(p) = P_p(0 \leftrightarrow \infty) \leq P_p(\exists \text{ chemin de longueur } n \text{ issu de } 0 \text{ dont les arêtes sont ouvertes})$$

$$= P_p\left(\bigcup_{x_0=0, \dots, x_n \text{ chemin}} \{\text{les arêtes de } x_0, \dots, x_n \text{ sont ouvertes}\}\right)$$

$$\theta(p) \leq \sum_{x_0=0, \dots, x_n} p(\text{les arêtes de } x_0, \dots, x_n \text{ sont ouvertes})$$

$$= p^n |\{\text{chemins } x_0, \dots, x_n \text{ issus de } 0\}|$$

$$= p^n |\{\text{chemins longueur } n \text{ issus de } 0\}| \leq p^n 4 \cdot 3^{n-1} = 4 \cdot 3^{n-1} p^n$$

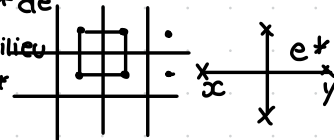
$$\leq \frac{p^n 3^{n-1}}{n} \xrightarrow{n \rightarrow +\infty} 0$$

$p_c < 1$. dualité: Le réseau dual de $\mathbb{Z}^2 + (\frac{1}{2}, \frac{1}{2}) = \mathbb{Z}^{2*}$

À une arête du réseau $\mathbb{Z}^2$, nous associons l'arête e^* de

$\mathbb{Z}^{2*}$ tq e et e^* se coupent orthogonalement en leur milieu

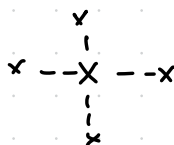
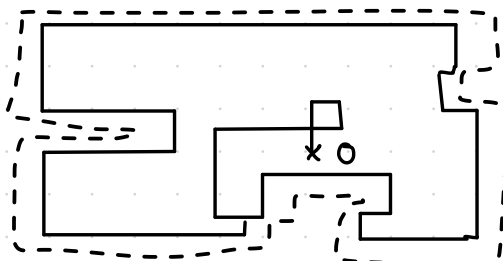
Si p est ouverte dans ω , nous déclarons e^* ouverte ds ω^*



$\theta(p) > 0$ pour p proche de 1

$$1 - \theta(p) = 1 - P_p(|C(0)| = \infty) = P_p(|C(0)| < +\infty)$$

Fait topologique: si $C(0)$ est fini, alors il existe un circuit dual d'arêtes fermées qui entoure 0

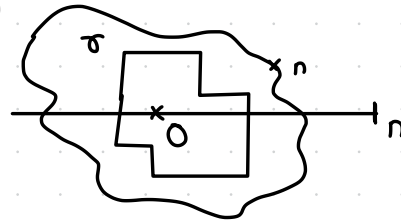


$$1 - \theta(p) \leq \mathbb{P}_p(\exists \text{ circuit dual fermé entourant } 0) \\ = \mathbb{P}_p\left(\bigcup_{\substack{\gamma = x_0^* \dots x_n^* \text{ circuit} \\ \text{dual entourant } 0}} \{\text{les arêtes de } \gamma \text{ sont fermées}\}\right) = \mathbb{P}_p\left(\bigcup_{n \geq 4} \bigcup_{\substack{\gamma \text{ dual fermé} \\ \text{longueur } n}}\right) \leq \sum_{n \geq 4} \sum_{\gamma} \mathbb{P}(\text{les arêtes de } \gamma \text{ sont fermées})$$

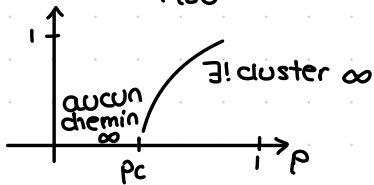
$$1 - \theta(p) \leq \sum_{n \geq 4} \sum_{\gamma} (1-p)^n = \sum_{n \geq 4} (1-p)^n; \{\text{circuits fermés de longueur } n \text{ entourant } 0\}$$

$$\leq \sum_{n \geq 4} (1-p)^n \cdot \underset{\substack{\uparrow \\ \text{choix d'un} \\ \text{pt de l'axe}}}{n} \cdot 4 \cdot 3^{n-1} \leq (1-p)^4$$

Ainsi, $\lim_{p \rightarrow 1} \theta(p) = 1$ et donc $p_c < 1$



Allure de θ Kesten 1980 $p_c = 1/2, d \geq 3$



Att-tons $\theta(p_c, \mathbb{Z}^3) = 0?$