

Introduction aux Probabilités

Shen LIN

2021



Shen Lin

Email : shen.lin@sorbonne-universite.fr

Organisation

polycopié + feuilles de TD : pdf sur Moodle
exemplaires imprimés à chercher au secrétariat

Cours d'amphi : 2h → 1h30
vidéos courtes sur Moodle

TD : corrections sur Moodle

Évaluation :

CC (quizzes sur Moodle)	20%
Partiel (fin mars)	30%
Examen (fin mai)	50%

Introduction

Théorie des probabilités : l'étude des phénomènes caractérisés par **le hasard** et **l'incertitude**

Exemples : lancer une pièce Pile ou Face

lancer un dé

tirage au loto

marche aléatoire (promenade d'un ivrogne)

trajectoire d'une poussière de pollen sur la surface de l'eau

Formalisation mathématique : 20^e siècle avec
l'axiomatique de Kolmogorov

Programme du cours

Chap. 1 Ensembles, dénombrement, dénombrabilité

Chap. 2 Espaces probabilisés pour décrire une expérience aléatoire

Chap. 3 Variables aléatoires discrètes

Chap. 4 Variables aléatoires réelles générales

Chap. 5 Variables aléatoires à densité

Chap. 6 Loi des grands nombres et applications

Chap. 7. Simulation des variables aléatoires.

Chap1 Ensembles, Dénombrement et Dénombrabilité

§ 1.1 Opérations sur les ensembles

Un ensemble = collection d'objets

Soit Ω un ensemble.

$x \in \Omega$ "x appartient à Ω "

Ensemble vide $\emptyset$

$A \subset \Omega$ sera appelé sous-ensemble ou partie de Ω

$A, B \subset \Omega$

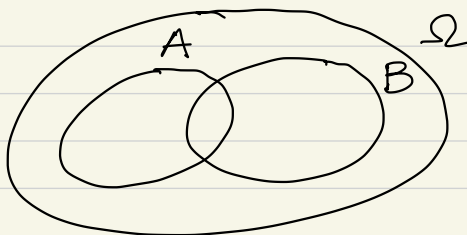
$A \cup B := \{\omega \in \Omega : \omega \in A \text{ ou } \omega \in B\}$ union

$A \cap B := \{\omega \in \Omega : \omega \in A \text{ et } \omega \in B\}$ intersection

$A^c := \{\omega \in \Omega : \omega \notin A\}$ complémentaire

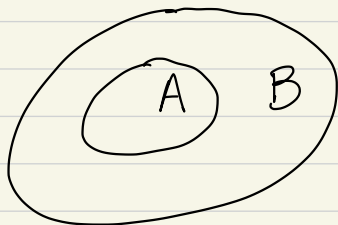
$A \setminus B := A \cap B^c$ différence de A avec B

$A \Delta B := (A \setminus B) \cup (B \setminus A)$ différence symétrique



$A \subset B$ $A \subseteq B$ inclusion

pour tout $\omega \in A$ on a $\omega \in B$

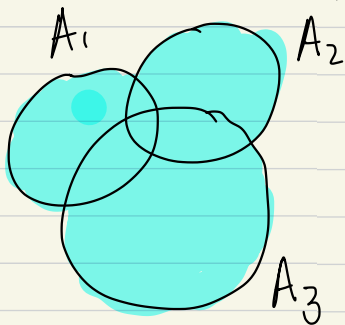


I ensemble d'indices arbitraire

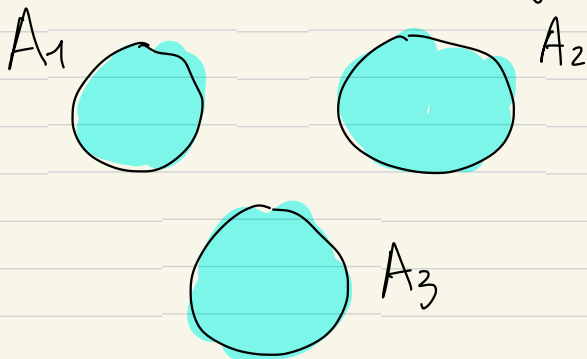
$(A_i)_{i \in I}$ famille de sous-ensembles de Ω

$\bigcup_{i \in I} A_i := \{ \omega \in \Omega : \exists i \in I \text{ tel que } \omega \in A_i \}$ union
"il existe"

$\bigcap_{i \in I} A_i := \{ \omega \in \Omega : \forall i \in I \text{ on a } \omega \in A_i \}$ intersection.
"pour tout"



$\bigcup_{i \in I} A_i$ est disjointe si les $(A_i)_{i \in I}$ sont 2-à-2 disjoints
c-à-d. $A_i \cap A_j = \emptyset$ pour tous $i \neq j$
 $i, j \in I$



Dans ce cas, on écrit aussi

" $\bigsqcup_{i \in I} A_i$ "

Distributivité

$$\left(\bigcup_{i \in I} A_i \right) \cap B = \bigcup_{i \in I} (A_i \cap B)$$

$$\left(\bigcap_{i \in I} A_i \right) \cup B = \bigcap_{i \in I} (A_i \cup B)$$

$$\left(\bigcup_{i \in I} A_i \right)^c = \bigcap_{i \in I} A_i^c$$

$$\left(\bigcap_{i \in I} A_i \right)^c = \bigcup_{i \in I} A_i^c$$

Produit cartésien

$$A \times B = \{ (a, b) : a \in A \text{ et } b \in B \} \neq B \times A$$

↑ couple ordonné

$n \geq 2$ entier

$A^n =$ le produit $A \times \dots \times A$ n fois

$\mathcal{P}(\Omega) := \{ A : A \subseteq \Omega \}$ l'ensemble de tous les sous-ensembles de Ω

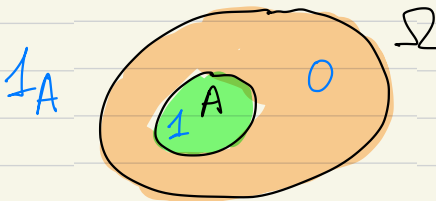
$$A \in \mathcal{P}(\Omega) \Leftrightarrow A \subseteq \Omega \Leftrightarrow A \subset \Omega$$

$$\emptyset \in \mathcal{P}(\Omega) \text{ et } \Omega \in \mathcal{P}(\Omega)$$

$\forall A \in \mathcal{P}(\Omega)$ on définit la fonction indicatrice de A

$$1_A : \Omega \rightarrow \mathbb{R}$$

$$1_A(x) = \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{si } x \notin A \\ & (x \in \Omega \setminus A) \end{cases}$$



§ 1.2 Dénombrement

A ensemble

Le nombre d'éléments que A contient s'appelle la **cardinalité** de A, notée **|A|**

$$|A| \in \mathbb{N} := \{0, 1, 2, \dots\} \quad \text{si } A \text{ fini}$$

$$|A| = \infty \quad \text{si } A \text{ infini}$$

- Si $A_1, \dots, A_n$ finis et **2-à-2 disjoints**, alors

$$|A_1 \cup \dots \cup A_n| = \sum_{i=1}^n |A_i|$$

- Si $A_1, \dots, A_n$ finis $|A_1 \times \dots \times A_n| = \prod_{i=1}^n |A_i|$

Déf: $f: A \rightarrow B$ **injective** si tout élément de B est l'image d'**au plus** un élément de A

surjective si tout élément de B est l'image d'**au moins** un élément de A

bijjective si à la fois injective et surjective.

[Dans ce cas, l'application inverse f^{-1} est définie]

Si $\exists A \xrightarrow{\text{injection}} B$ alors $|A| \leq |B|$

$\exists A \xrightarrow{\text{surjection}} B$ alors $|A| \geq |B|$

$\exists A \xrightarrow{\text{bijection}} B$ alors $|A| = |B|$

Exercice 1.1 $\{0,1\}^\Omega := \{ \text{fonctions } f: \Omega \rightarrow \{0,1\} \}$

$\varphi: \mathcal{P}(\Omega) \rightarrow \{0,1\}^\Omega$ est une bijection

$$A \mapsto 1_A$$

[Voir vidéo sur Moodle]

Si $|\Omega| < \infty$, on a $|\mathcal{P}(\Omega)| = 2^{|\Omega|}$

Remarque: l'ensemble E est dit infini s'il existe $x_0 \in E$ et une injection de E dans $E \setminus \{x_0\}$

E est dit fini sinon

$$\mathbb{N}^* = \mathbb{N} \setminus \{0\}.$$

$\mathbb{N}$ infini car $\mathbb{N} \rightarrow \mathbb{N}^*$ est injective.
 $n \mapsto n+1$

Soit $|A| = n \geq 1$.

$$(a_1, \dots, a_k) \neq (a_k, \dots, a_1)$$

Arrangement (sans répétition) de k éléments de A : $(a_1, \dots, a_k)$
un k -uplet ordonné d'éléments distincts de A .

Nb. d'arrangements de k éléments de A :

$$= \begin{cases} n \times (n-1) \times (n-2) \times \dots \times (n-k+1) & \text{si } 0 \leq k \leq n \\ 0 & \text{si } k > n \end{cases}$$

Permutation de A : une bijection de A dans A .

Nb. de permutations de $A = n! = n \times (n-1) \times \dots \times 1$

Notation: $\mathcal{S}_n = \{ \text{permutations de } \{1, 2, \dots, n\} \}$

Convention: $0! = 1$.

$$\{a_1, a_2, \dots, a_k\} = \{a_2, a_k, \dots, a_1\}$$

Combinaisons de k éléments de A : un sous-ensemble

$\{a_1, \dots, a_k\}$ d'exactly k éléments de A

$a_1, \dots, a_k$ sont distincts, mais non ordonnés.

Nb. de combinaisons de k éléments de $A = \binom{n}{k} = C_n^k$

Coefficient binomial

$$\binom{n}{k} = \frac{n(n-1)\dots(n-k+1)}{k!} = \frac{n!}{k!(n-k)!} \quad \text{si } 0 \leq k \leq n.$$

Convention $\binom{n}{0} = 1$

$$\binom{n}{k} = 0 \text{ si } k > n.$$

Binôme de Newton : $\forall x, y \in \mathbb{R}$

$$(x+y)^n = \sum_{k=0}^n \binom{n}{k} x^k y^{n-k}$$

§ 1.3 Dénombrabilité

Déf : Un ensemble E est dit **dénombrable** s'il est fini ou **infini dénombrable** c'est-à-dire en bijection avec $\mathbb{N}$.

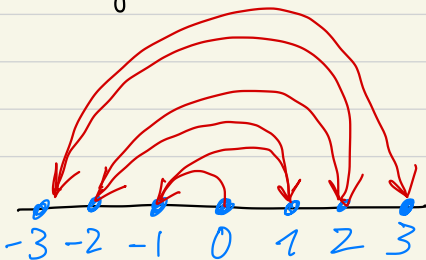
Si E fini, énumérer ses éléments comme suite finie $x_1, \dots, x_n$

Si E infini dénombrable, énumérer comme une suite infinie $(x_i)_{i \geq 0}$

Exemple $\mathbb{Z}$ infini dénombrable

$$\mathbb{N} \rightarrow \mathbb{Z}$$

$$n \mapsto \begin{cases} \frac{n}{2} & \text{si } n \text{ pair} \\ -\frac{n+1}{2} & \text{si } n \text{ impair.} \end{cases}$$



Lemme:

Un sous-ensemble (infini) d'un ensemble dénombrable est dénombrable.
Tout $A \subset \mathbb{N}$ est dénombrable.

Lemme: (1) $A \xrightarrow{\text{injection}} B$ avec B dénombrable $\Rightarrow A$ dénombrable

(2) $A \xrightarrow{\text{surjection}} B$ avec A dénombrable $\Rightarrow B$ dénombrable

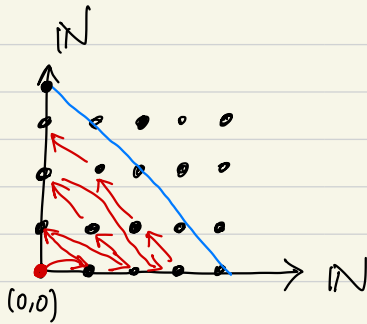
Théorème 1.3.

(1) Si $A_1, \dots, A_n$ sont dénombrables, le produit $A_1 \times \dots \times A_n$ l'est aussi

(2) Si $(A_i)_{i \in I}$ est une famille dénombrable (c-à-d. I est dénom.) d'ensembles dénombrables (c-à-d. A_i est dénom. pour tout $i \in I$), alors l'union $\bigcup_{i \in I} A_i$ est dénombrable.

[Voir vidéo pour la preuve].

Applications (1) $\mathbb{N} \times \mathbb{N}$ dénombrable
preuve directe :



(2) $\mathbb{Q} = \{\text{nb. rationnel}\}$ dénombrable

$$\begin{array}{ccc} \mathbb{Z} \times \mathbb{N}^* & \longrightarrow & \mathbb{Q} \\ \text{dénombrable.} \quad (p, q) & \longmapsto & \frac{p}{q} \end{array} \quad \text{surjection}$$

(3) $\mathbb{Z}[X] := \{ \text{polynômes à coeff. entiers} \}$

$\mathbb{Z}_n[X] := \{ \text{polynômes à coeff. dans } \mathbb{Z} \text{ de degré } n \}$

$\mathbb{Z}^* \times \mathbb{Z}^n \longrightarrow \mathbb{Z}_n[X]$ bijection.
 $(a_n, a_{n-1}, \dots, a_0) \longmapsto \sum_{k=0}^n a_k X^k$
 (car $\mathbb{Z}^*$, $\mathbb{Z}^n$ dénombrables)

Donc $\mathbb{Z}_n[X]$ dénombrable

$\mathbb{Z}[X] = \bigcup_{n \in \mathbb{N}} \mathbb{Z}_n[X] = \bigcup_{n=0}^{+\infty} \mathbb{Z}_n[X]$ union disjointe
 $= \bigsqcup_{n \in \mathbb{N}} \mathbb{Z}_n[X]$ est une union dénombrable d'ensembles dénom.

Donc $\mathbb{Z}[X]$ dénombrable.

Théorème 1.7 : $\mathbb{R}$ n'est pas dénombrable

[voir vidéo pour la preuve]



Corollaire : $[0, 1[$ pas dénombrable

$[0, 1[\longrightarrow \{0, 1\}^{\mathbb{N}}$ injection.

$x \longmapsto$ son développement en base 2

$$x = \sum_{k=1}^{+\infty} a_k 2^{-k} \quad a_k \in \{0, 1\}, \forall k \geq 1$$

$\Rightarrow \{0, 1\}^{\mathbb{N}}$ pas dénombrable

Comme $\mathcal{P}(\mathbb{N})$ en bijection avec $\{0, 1\}^{\mathbb{N}}$, $\mathcal{P}(\mathbb{N})$ pas dénombrable

————— FIN SÉANCE 1 —————

Chap 2. Espaces probabilisés $(\Omega, \mathcal{F}, \mathbb{P})$

§2.1 Axiomes des probabilités

§2.1.1 Espaces probabilisés

Pour modéliser une expérience aléatoire :

- L'**espace d'état** Ω est un ensemble contenant tous les résultats possibles
- L'**ensemble des événements** $\mathcal{F}$ contient des parties de Ω
un événement = $\{\omega \in \Omega \text{ tel que " } \omega \text{ vérifie la propriété "}\}$
- La **probabilité** $\mathbb{P}$ est une application de $\mathcal{F}$ à $[0, 1]$.
 $\forall A \in \mathcal{F} \quad \mathbb{P}(A) \in [0, 1]$

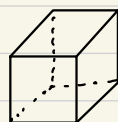
Triplet $(\Omega, \mathcal{F}, \mathbb{P})$ appelé **espace probabilisé**

Exemples (1) Lancer un dé **régulier** à 6 faces

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

Le dé donne un nb. pair $A = \{2, 4, 6\} \subset \Omega \quad A \in \mathcal{F}$

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|} = \frac{3}{6} = \frac{1}{2} \quad \text{probabilité uniforme}$$



(2) Si le dé **irrégulier**, $\mathbb{P}$ non uniforme

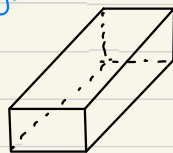
Suite **infinie** de Pile ou Face **équilibré**

$$\Omega = \{(u_i)_{i \geq 1} : u_i \in \{\text{Pile}, \text{Face}\}, \forall i \geq 1\}$$

Les 2 premiers sont Pile : $B = \{(u_i)_{i \geq 1} \in \Omega : u_1 = u_2 = \text{Pile}\}$

Il y a au moins un Face : $C = \bigcup_{k=1}^{+\infty} \{(u_i)_{i \geq 1} \in \Omega : u_k = \text{Face}\}$

$B, C \in \mathcal{F}$.



ici Ω en bijection avec $\{0,1\}^{\mathbb{N}}$, donc non dénombrable

Remarque Si Ω dénombrable, on prend toujours $\mathcal{F} = \mathcal{P}(\Omega)$

Mais si Ω non dénombrable, plus compliqué (voir plus tard)

ici $\mathcal{F} \neq \mathcal{P}(\Omega)$ $A_k, A_k^c \in \mathcal{F}$

$\mathcal{F}$ = tribu engendrée par les événements de base $(A_k, A_k^c)_{k \geq 1}$.

$k \geq 1$ A_k = "le k^e lancer donne Pile"
 A_k^c = "le k^e lancer donne face" $\mathbb{P}(A_k) = \mathbb{P}(A_k^c) = \frac{1}{2}$

$$\mathbb{P}(A_1 \cap A_2 \cap \dots \cap A_n) = \left(\frac{1}{2}\right)^n, \quad \forall n \geq 1$$

= "n premiers lancers donnent Pile"

$$\mathbb{P}(A_1^{c \text{ ou pas}} \cap A_2^{c \text{ ou pas}} \cap \dots \cap A_n^{c \text{ ou pas}}) = \left(\frac{1}{2}\right)^n$$

• Propriétés pour $\mathcal{F}$:

l'événement certain = $\Omega \in \mathcal{F}$

l'événement impossible = $\emptyset \in \mathcal{F}$

$A, B \in \mathcal{F}$

"A et B sont tous les 2 vérifiés" $\rightarrow A \cap B \in \mathcal{F}$

"au moins un parmi A et B est vérifié" $\rightarrow A \cup B \in \mathcal{F}$

"A n'est pas vérifié" $\rightarrow A^c \in \mathcal{F}$

Def: On dit que $\mathcal{F}$ est une **tribu** (ou une **σ -algèbre**) sur Ω si $\mathcal{F}$ est une partie non vide de $\mathcal{P}(\Omega)$ vérifiant:

(1) $\Omega \in \mathcal{F}$

(2) Si $A \in \mathcal{F}$, alors $A^c \in \mathcal{F}$ (**stabilité par passage au complémentaire**)

(3) Si $(A_i)_{i \in \mathbb{N}}$ est une **suite** d'éléments de $\mathcal{F}$, alors
(= "famille dénombrable")

$\bigcup_{i \in \mathbb{N}} A_i \in \mathcal{F}$ (**stabilité par union dénombrable**)

Remarque: $\emptyset = \Omega^c \in \mathcal{F}$

$$\bigcap_{i \in \mathbb{N}} A_i = \left(\bigcup_{i \in \mathbb{N}} A_i^c \right)^c \in \mathcal{F}$$

Ω dénombrable $\mathcal{F} = \mathcal{P}(\Omega)$

Si Ω non dénombrable, on prendra $\mathcal{F}$ une tribu construite à partir des événements de base en utilisant

{ passage au complémentaire
union / intersection dénombrable

Potentiellement $\mathcal{F} \neq \mathcal{P}(\Omega)$ c-à-d. $\mathcal{F} \subsetneq \mathcal{P}(\Omega)$

• Propriétés de $\mathbb{P}$:

Interprétation fréquentiste: Répéter N fois, N grand.

$N_A \in \{0, 1, \dots, N\}$ nb. de fois l'événement A vérifié

$$" \mathbb{P}(A) \approx \frac{N_A}{N} "$$

$$\text{Si } A = \Omega, N_\Omega = N \rightarrow \mathbb{P}(\Omega) = 1$$

$$\text{Si } A, B \in \mathcal{F}, A \cap B = \emptyset \Rightarrow N_{A \cup B} = N_A + N_B$$

$$\text{" } \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) \text{"}$$

Déf Axiomes des probabilités

Ω ensemble non vide muni d'un ensemble d'événements (tribu) $\mathcal{F}$
 Une fonction $\mathbb{P}: \mathcal{F} \rightarrow [0,1]$ est appelée (mesure de) probabilité

Si $\boxed{A1}: \mathbb{P}(\Omega) = 1$

$\boxed{A2}$ (σ -additivité) Pour tt suite $(A_n)_{n \geq 1}$ d'événements de $\mathcal{F}$
 σ signifie "denombrable"
 2-à-2 disjoints (c-à-d. $A_n \cap A_m = \emptyset$ pour $n \neq m$)

on a

$$\mathbb{P}\left(\bigcup_{n=1}^{+\infty} A_n\right) = \mathbb{P}\left(\bigcup_{n=1}^{+\infty} A_n\right) = \sum_{n=1}^{+\infty} \mathbb{P}(A_n)$$

Remarque 1: $\mathbb{P}(\emptyset) = 0$ car $\mathbb{P}(\emptyset) = \sum_{n=1}^{+\infty} \mathbb{P}(\emptyset)$ par A2

Remarque 2: $\mathbb{P}$ vérifie l'additive finie :

Si $A_1, \dots, A_k$ sont k événements disjoints ("2-à-2 disjoints")
 alors $\mathbb{P}(A_1 \cup \dots \cup A_k) = \sum_{j=1}^k \mathbb{P}(A_j)$

Preuve: Appliquer A2 à $(B_n)_{n \geq 1}$ où $B_n = A_n$ si $n \leq k$
 $= \emptyset$ si $n > k$.

□

Remarque 3 $\sum_{n=1}^{+\infty} \underbrace{P(A_n)}_{\geq 0}$ dans $\mathcal{A}$ ne dépend pas de l'ordre de sommation.

§ 2.1.2 Propriétés fondamentales de $\mathbb{P}$

Prop 2.7 (1) $A, B \in \mathcal{F}$ $A \subset B$

$$0 \leq P(B \setminus A) = P(B) - P(A)$$

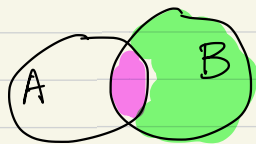
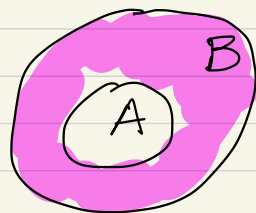
Par conséquent, $P(A) \leq P(B)$

$$\forall A \in \mathcal{F}, P(A^c) = 1 - P(A)$$

(2) $\forall A, B \in \mathcal{F}$

$$P(A \cup B) \stackrel{(*)}{=} P(A) + P(B) - \underbrace{P(A \cap B)}_{\geq 0}$$

Par conséquent, $P(A \cup B) \leq P(A) + P(B)$



Preuve (1) $A \subset B$ $B = A \cup (B \setminus A)$ union disjointe

Par additivité de $\mathbb{P}$,

$$P(B) = P(A) + P(B \setminus A)$$

En prenant $B = \Omega$, on a $P(A) + P(A^c) = P(\Omega) = 1$

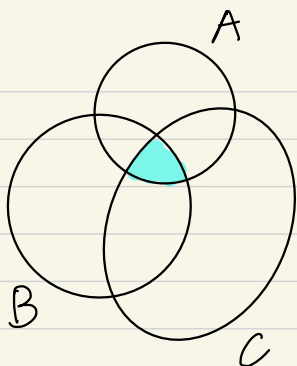
(2) $A \cup B = A \cup (B \setminus A)$ union disjointe

$$P(A \cup B) = P(A) + P(B \setminus A)$$

Or $B = (B \setminus A) \cup (A \cap B)$ union disjointe

$$P(B) = P(B \setminus A) + P(A \cap B)$$

□



$$(A \cup B) \cup C$$

$$\begin{aligned}
 \mathbb{P}(A \cup B \cup C) &\stackrel{||}{=} \mathbb{P}(A \cup B) + \mathbb{P}(C) - \mathbb{P}((A \cup B) \cap C) \\
 &\stackrel{(*)}{=} \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B) + \mathbb{P}(C) - \mathbb{P}((A \cap C) \cup (B \cap C)) \\
 &\stackrel{(*)}{=} \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B) + \mathbb{P}(C) - \mathbb{P}(A \cap C) - \mathbb{P}(B \cap C) \\
 &\quad + \mathbb{P}(\underbrace{(A \cap C) \cap (B \cap C)}_{= A \cap B \cap C}) \\
 &= \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) - \mathbb{P}(A \cap B) - \mathbb{P}(A \cap C) \\
 &\quad - \mathbb{P}(B \cap C) + \mathbb{P}(A \cap B \cap C)
 \end{aligned}$$

Cas général

Prop 2.9 Formule d'inclusion-exclusion

$A_1, \dots, A_n \in \mathcal{F}$

$$\mathbb{P}(A_1 \cup A_2 \cup \dots \cup A_n) = \sum_{k=1}^n \sum_{\substack{J \subseteq \{1, \dots, n\} \\ \text{tel que } |J|=k}} (-1)^{k+1} \mathbb{P}\left(\bigcap_{i \in J} A_i\right)$$

[Voir vidéo dans Moodle]

Application: Exemple 2.19 Père Noël secret

[Voir vidéo dans Moodle] .

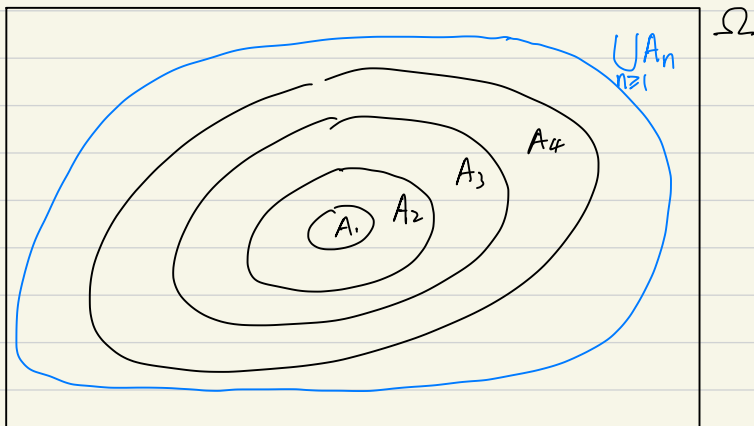
Prop 2.10 Continuité par le bas et par le haut des probabilités

(1) $(A_n)_{n \geq 1}$ suite croissante d'événements

(c-à-d $A_n \subset A_{n+1}$, $\forall n \geq 1$)

alors
$$\mathbb{P}\left(\underbrace{\bigcup_{n=1}^{\infty} A_n}_{\in \mathcal{F}}\right) = \lim_{n \rightarrow \infty} \mathbb{P}(A_n)$$

$$\mathbb{P}(A_n) \leq \mathbb{P}(A_{n+1})$$



$$\forall n \geq 1 \quad A_n \subset \bigcup_{n \geq 1} A_n$$

(2) $(A_n)_{n \geq 1}$ suite **décroissante** d'événements
(c-à-d. $A_{n+1} \subset A_n \quad \forall n \geq 1$)

Alors $\mathbb{P}\left(\underbrace{\bigcap_{n \geq 1} A_n}_{\in \mathcal{F}}\right) = \lim_{n \rightarrow \infty} \downarrow \mathbb{P}(A_n)$ $\mathbb{P}(A_n) \geq \mathbb{P}(A_{n+1})$



$$\forall n \geq 1, \quad \bigcap_{i \geq 1} A_i \subset A_n$$

[Voir vidéo sur Moodle]

Corollaire 2.11

$(A_n)_{n \geq 1}$ suite d'événements *non nécessairement croissante ou décroissante*

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} A_n\right) = \lim_{n \rightarrow \infty} \uparrow \mathbb{P}\left(\bigcup_{k=1}^n A_k\right)$$

$$\mathbb{P}\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n \rightarrow \infty} \downarrow \mathbb{P}\left(\bigcap_{k=1}^n A_k\right)$$

Preuve: Poser $B_n = \bigcup_{k=1}^n A_k$

$(B_n)_{n \geq 1}$ suite croissante d'événements.

$$B_n \subset B_{n+1} \text{ car } \bigcup_{k=1}^n A_k \subset \bigcup_{k=1}^{n+1} A_k = A_{n+1} \cup \left(\bigcup_{k=1}^n A_k\right)$$

$$\text{Par Prop 2.10 } \mathbb{P}\left(\bigcup_{n=1}^{\infty} B_n\right) = \lim_{n \rightarrow +\infty} \uparrow \mathbb{P}(B_n) = \lim_{n \rightarrow +\infty} \uparrow \mathbb{P}\left(\bigcup_{k=1}^n A_k\right)$$

$$\text{Or } \bigcup_{n=1}^{+\infty} A_n = \bigcup_{n=1}^{+\infty} B_n$$

car $B_n \subset \bigcup_{n=1}^{+\infty} A_n$ et $A_n \subset B_n$ pour tout $n \geq 1$

$$\text{Donc } \mathbb{P}\left(\bigcup_{n=1}^{+\infty} A_n\right) = \lim_{n \rightarrow +\infty} \uparrow \mathbb{P}\left(\bigcup_{k=1}^n A_k\right)$$

$$2^{\text{e}} \text{ formule : } C_n = \bigcap_{k=1}^n A_k$$

exercice : compléter la preuve.

□

Exemple 2.13 Suite infinie de Pile ou Face équilibré

$A_k = "$ k^e lancer est Pile"

$A = \bigcap_{k=1}^{+\infty} A_k = "$ on n'obtient que des Piles dans la suite infinie"

$$\mathbb{P}(A) = \lim_{n \rightarrow +\infty} \downarrow \mathbb{P}\left(\bigcap_{k=1}^n A_k\right) = \lim_{n \rightarrow +\infty} \left(\frac{1}{2}\right)^n = 0$$

Donc la probabilité d'avoir au moins un Face vaut 1
($\mathbb{P}(A^c) = 1$)

Un événement $B \in \mathcal{F}$ de probabilité $\mathbb{P}(B) = 1$ est appelé
presque sûr



Si $\Omega \cong \{0,1\}^{\mathbb{N}}$ "Pile = 0"

$$A = \{(0,0,0, \dots)\} \neq \emptyset$$

$$\mathbb{P}(A) = 0$$

$$A^c \neq \Omega \quad \mathbb{P}(A^c) = 1.$$

Pour un événement $B \in \mathcal{F}$,

Si $\mathbb{P}(B) = 1$, on n'a pas nécessairement $B = \Omega$

Si $\mathbb{P}(B) = 0$, $B = \emptyset$



Rappel: $(\Omega, \mathcal{F}, \mathbb{P})$

proba $\mathbb{P}: \mathcal{F} \rightarrow [0, 1]$

$A \mapsto \mathbb{P}(A) \in [0, 1]$

(A1): $\mathbb{P}(\Omega) = 1$

(A2): Soit $(A_n)_{n \geq 1}$ une suite d'événements 2-à-2 disjoints

$$\mathbb{P}\left(\bigsqcup_{n \geq 1} A_n\right) = \sum_{n \geq 1} \mathbb{P}(A_n) \quad \text{\textcolor{blue}{}\sigma\text{-additivit }}$$

Rappel: $A, B \in \mathcal{F}$

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B) \leq \mathbb{P}(A) + \mathbb{P}(B) \quad (*)$$

Corollaire 2.12 (Sous-additivit  de $\mathbb{P}$)

Soit $(A_n)_{n \geq 1}$ une suite d' v nements. Alors

$$\mathbb{P}\left(\bigcup_{n \geq 1} A_n\right) \leq \sum_{n \geq 1} \mathbb{P}(A_n)$$

Preuve: Par r currence, on  tend l'in galit  (*) et on obtient

$$\mathbb{P}\left(\bigcup_{k=1}^n A_k\right) \leq \sum_{k=1}^n \mathbb{P}(A_k) \quad \text{pour tout } n \geq 1$$

Alors

$$\begin{aligned} \mathbb{P}\left(\bigcup_{n \geq 1} A_n\right) &= \lim_{n \rightarrow +\infty} \mathbb{P}\left(\bigcup_{k=1}^n A_k\right) \\ &\leq \lim_{n \rightarrow +\infty} \sum_{k=1}^n \mathbb{P}(A_k) = \sum_{k=1}^{\infty} \mathbb{P}(A_k) \quad \square \end{aligned}$$

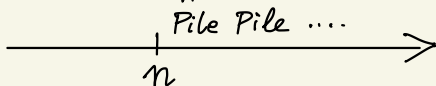
Application: Suite infinie de Pile ou Face  quilibr 

$A_k = \text{"le } k^{\text{ }} \text{ lancer est Pile"}$

$$(\text{Exemple 2.13}) \quad \mathbb{P}\left(\bigcup_{k=1}^{\infty} A_k^c\right) = 1 \quad \text{et} \quad \mathbb{P}\left(\bigcap_{k=1}^{\infty} A_k\right) = 0$$

par m me argument

$$\mathbb{P}\left(\bigcup_{k=n}^{\infty} A_k^c\right) = 1 \quad \text{et} \quad \mathbb{P}\left(\bigcap_{k=n}^{\infty} A_k\right) = 0, \quad \forall n \geq 1$$



Par sous-additivité de $\mathbb{P}$

$$\mathbb{P}\left(\bigcup_{n=1}^{\infty} \left(\bigcap_{k=n}^{\infty} A_k\right)\right) \leq \sum_{n=1}^{\infty} \mathbb{P}\left(\bigcap_{k=n}^{\infty} A_k\right) = \sum_{n=1}^{\infty} 0 = 0$$

En passant au complémentaire

$$\mathbb{P}\left(\bigcap_{n=1}^{\infty} \left(\bigcup_{k=n}^{\infty} A_k^c\right)\right) = 1$$

= " $\forall n \geq 1, \exists k \geq n$ t.q. le k^e est Face"

= "il y a une infinité de Face"

ceci est un événement presque sûr

§ 2.2 Le cas d'espaces d'états dénombrable

Dans cette partie, Ω dénombrable non vide

$$\mathcal{F} = \mathcal{P}(\Omega)$$

caractériser $\mathbb{P}$ par une densité discrète

Déf (Densité discrète)

On appelle densité discrète sur Ω toute fonction $p: \Omega \rightarrow \mathbb{R}$ qui satisfait

$$\begin{cases} p(\omega) \geq 0 & \forall \omega \in \Omega \\ \sum_{\omega \in \Omega} p(\omega) = 1 \end{cases}$$

↑ cette somme ne dépend pas de l'ordre de sommation

Prop Soit p une densité discrète sur Ω

Alors $\mathbb{P}: \mathcal{P}(\Omega) \rightarrow \mathbb{R}$

définit une proba sur $\mathcal{P}(\Omega)$

$$A \mapsto \mathbb{P}(A) := \sum_{\omega \in A} p(\omega)$$

Preuve: $\mathbb{P}(\Omega) = \sum_{\omega \in \Omega} p(\omega) = 1$

Soit $(A_k)_{k \in \mathbb{N}}$ une suite de sous-ensembles de Ω 2-à-2 disjoints

Pose $A = \bigcup_{k \in \mathbb{N}} A_k$

$$\mathbb{P}\left(\bigcup_{k \in \mathbb{N}} A_k\right) = \sum_{\omega \in A} p(\omega) = \sum_{k \in \mathbb{N}} \left(\sum_{\omega \in A_k} p(\omega) \right) = \sum_{k \in \mathbb{N}} \mathbb{P}(A_k)$$

$p(\omega) \geq 0, \forall \omega \in \Omega$
 $(A_k)_{k \in \mathbb{N}}$ disjoints $\Rightarrow$ On peut sommer par paquets

□

La réciproque est aussi vraie: Soit $\mathbb{P}$ une proba. sur Ω ,
 $p: \Omega \rightarrow \mathbb{R}$
 $\omega \mapsto p(\omega) := \mathbb{P}(\{\omega\})$ est une densité discrète sur Ω

$\uparrow$ singleton $\{\omega\}$
 (événement élémentaire)

Déf (Équiprobabilité) Soit Ω un ensemble fini
 On définit $\mathbb{P}(A) := \frac{|A|}{|\Omega|} \quad \forall A \subset \Omega$

$\mathbb{P}$ est une probabilité, appelée probabilité uniforme sur Ω

Remarque 1: densité discrète associée
 $p(\omega) = \mathbb{P}(\{\omega\}) = \frac{1}{|\Omega|}$ ne dépend pas de ω

Remarque 2: Soit Ω fini et $\mathbb{P}$ une proba. sur $\mathcal{P}(\Omega)$ tq.
 $p(\omega) = \mathbb{P}(\{\omega\}) = c$ ne dépend pas de $\omega \in \Omega$
 Alors $\mathbb{P}$ est la probabilité uniforme sur Ω
 (nécessairement $c = \frac{1}{|\Omega|}$)

Remarque 3: Si Ω est infini (dénombrable), on ne peut pas définir de proba. uniforme

Si $p(\omega) = a, \forall \omega \in \Omega$

alors $\sum_{\omega \in \Omega} p(\omega) = \begin{cases} +\infty & \text{si } a > 0 \\ 0 & \text{si } a = 0 \end{cases}$ ne vaut pas 1

Exemple: choisir un entier naturel au hasard n'est jamais de façon uniforme (Il n'existe pas de probabilité uniforme sur $\mathbb{N}$)

Par exemple, $P(\{n\}) = 2^{-(n+1)}$ définit une proba sur $\mathbb{N}$, mais elle n'est pas uniforme.

Exemple 2.18 Paradoxe des anniversaires [vidéo courte]

Exemple 2.19 Père Noël secret [vidéo courte]

§ 2.3 Probabilité conditionnelle et indépendance

§ 2.3.1 Probabilité conditionnelle

$(\Omega, \mathcal{F}, P)$, Ω n'est plus supposé dénombrable.

Déf. Soient $A, B \in \mathcal{F}$ avec $P(B) > 0$

On appelle **probab. conditionnelle de A sachant B**

$$P(A|B) := \frac{P(A \cap B)}{P(B)}$$

Remarque: Interprétation fréquentiste $P(A) \approx \frac{N_A}{N}$

$$P(A|B) \approx \frac{N_{A \cap B}}{N_B} = \frac{\frac{N_{A \cap B}}{N}}{\frac{N_B}{N}} \approx \frac{P(A \cap B)}{P(B)}$$

Exercice 2.21: La fonction $\mathbb{P}(\cdot | B): A \in \mathcal{F} \mapsto \mathbb{P}(A|B)$
est une proba. sur $(\Omega, \mathcal{F})$

$$\mathbb{P}(A^c | B) = 1 - \mathbb{P}(A | B)$$

$$\mathbb{P}(A | B^c) \neq 1 - \mathbb{P}(A | B) \quad \triangle !$$

$\mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_1) \times \mathbb{P}(A_2 | A_1)$ par définition
Prop 2.22 (proba conditionnelle en chaîne)

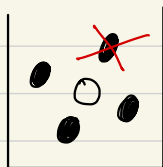
Soient $n \geq 2$ et $A_1, \dots, A_n$ des événements tq.

$$\mathbb{P}(A_1 \cap A_2 \cap \dots \cap A_{n-1}) > 0.$$

$$\begin{aligned} \text{Alors } \mathbb{P}\left(\bigcap_{k=1}^n A_k\right) &= \mathbb{P}(A_1) \times \mathbb{P}(A_2 | A_1) \times \dots \times \mathbb{P}(A_n | A_1 \cap A_2 \cap \dots \cap A_{n-1}) \\ &= \mathbb{P}(A_1) \times \prod_{k=2}^n \mathbb{P}(A_k | A_1 \cap \dots \cap A_{k-1}) \end{aligned}$$

Preuve: vidéo courte

Application:



tirer deux boules sans remise
 $i = 1, 2$

$A_i = \text{"je tire une boule noire au } i^{\text{e}} \text{ tour"}$

$A = A_1 \cap A_2 = \text{"je ne tire pas la boule blanche"}$

$$\mathbb{P}(A) = \mathbb{P}(A_1) \mathbb{P}(A_2 | A_1)$$

$$= \frac{4}{5} \times \frac{3}{4} = \frac{3}{5}$$

Prop Soit $(B_i)_{i \in I}$ une partition dénombrable de Ω

(c'est-à-dire I est dénombrable

$B_i \cap B_j = \emptyset$ pour tous $i \neq j$, et $\bigsqcup_{i \in I} B_i = \Omega$)

Pour un événement $A \in \mathcal{F}$, on a la formule de décomposition

$$\mathbb{P}(A) = \sum_{i \in I} \mathbb{P}(A \cap B_i)$$

Si de plus $\forall i \in I, \mathbb{P}(B_i) > 0$, on a la formule des proba. totales

$$\mathbb{P}(A) = \sum_{i \in I} \mathbb{P}(A|B_i) \mathbb{P}(B_i)$$

Preuve $A = A \cap \Omega = A \cap \left(\bigsqcup_{i \in I} B_i \right) = \bigsqcup_{i \in I} (A \cap B_i)$

$\swarrow \quad \searrow$
 disjointes

Par σ -additivité de $\mathbb{P}$, $\mathbb{P}(A) = \sum_{i \in I} \mathbb{P}(A \cap B_i)$

Si $\mathbb{P}(B_i) > 0$, $\mathbb{P}(A \cap B_i) = \mathbb{P}(A|B_i) \mathbb{P}(B_i)$ □

Théorème (Formule de Bayes)

Si $A, B \in \mathcal{F}$ avec $\mathbb{P}(A) > 0$ et $\mathbb{P}(B) > 0$

Alors $\mathbb{P}(B|A) = \frac{\mathbb{P}(A|B) \mathbb{P}(B)}{\mathbb{P}(A)}$

Preuve: $\mathbb{P}(B|A) \mathbb{P}(A) = \mathbb{P}(B \cap A) = \mathbb{P}(A \cap B) = \mathbb{P}(A|B) \mathbb{P}(B)$ □

Exemple: Test d'un virus

Si le virus est présent, résultat positif dans 100% des cas.

Si le virus est absent, _____ 2% des cas (Faux positif)

On sait que 5 personnes sur 10000 ont le virus.

On fait ce test sur un individu au hasard, avec un résultat positif

Question: avec quelle confiance peut-on affirmer qu'il soit malade?

$A = \text{"résultat positif"}$

$B = \text{"l'individu choisi est malade"}$

$$P(B) = \frac{5}{10000} \quad P(B^c) = 1 - P(B) = \frac{9995}{10000}$$

$$P(A) = P(A|B)P(B) + P(A|B^c)P(B^c)$$

$$P(A|B) = 1 \quad \text{et} \quad P(A|B^c) = 0,02$$

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}$$

$$= \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|B^c)P(B^c)}$$

$$= \frac{1 \times 0,0005}{1 \times 0,0005 + 0,02 \times 0,9995} \approx 0,025$$

Ce test donne en réalité beaucoup de faux positifs.

————— FIN SÉANCE 3 —————

Continuation sur l'exemple test du virus

$A = \text{"test positif"}$

$B = \text{"a le virus"}$

$$P(A|B) = 1 \quad P(A|B^c) = \alpha \quad \text{taux de faux positif}$$

$$P(B) = \beta$$

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|B^c)P(B^c)}$$

$$= \frac{\beta}{\beta + \alpha(1-\beta)} = \frac{1}{1 + (1-\beta)\frac{\alpha}{\beta}}$$

$$\approx \frac{1}{1 + \frac{\alpha}{\beta}} \quad \text{comme } 1-\beta \approx 1$$

Si $\alpha \ll \beta$ $\frac{\alpha}{\beta} \approx 0$ alors $P(B|A) \approx 1$
 Si $\alpha \gg \beta$ $\frac{\alpha}{\beta}$ très grand alors $P(B|A) \approx \frac{1}{\frac{\alpha}{\beta}} \approx 0$

Ici $\alpha = 0,02$ $\frac{\alpha}{\beta} = 40$
 $\beta = 0,0005$

$P(B|A) \approx 0,025$

§ 2.3.2 Indépendance d'événements

Déf: $(\Omega, \mathcal{F}, P)$ espace probabilisé
 $A, B \in \mathcal{F}$ sont dit **indépendants** si $P(A \cap B) = P(A)P(B)$

Remarque 1: Si $P(A) > 0$, cela signifie $P(B|A) = P(B)$
 Si $P(B) > 0$, cela signifie $P(A|B) = P(A)$
 "l'un soit vérifié ne modifie pas la probab. de l'autre"

Remarque 2: indép entre A et B $\nRightarrow$ A et B disjoints
 Si $P(A) > 0$, $P(B) > 0$, alors l'indép. implique que
 $P(A \cap B) > 0$, donc $A \cap B \neq \emptyset$.

Remarque 3: Si A, B indép, alors A^c et B le sont aussi
 car $P(A^c \cap B) = P(B \setminus (A \cap B))$
 $= P(B) - P(A \cap B)$
 $= P(B) - P(A)P(B)$
 $= P(B)(1 - P(A)) = P(B)P(A^c)$

De même, A et B^c sont indép, ainsi que A^c et B^c

Def. Indépendance d'événements $\neq$ Indépendance 2-à-2

Soit $(A_i)_{i \in I}$ une famille d'événements d'un espace proba. $(\Omega, \mathcal{F}, \mathbb{P})$ où l'ensemble d'indice I est arbitraire. Cette famille d'événements est dite indépendante si pour tout sous-ensemble fini $J \subseteq I$ ($|J| \geq 2$), on a

$$\mathbb{P}\left(\bigcap_{j \in J} A_j\right) = \prod_{j \in J} \mathbb{P}(A_j)$$



Pour l'indép. de n événements $A_1, \dots, A_n$, il ne suffit pas de vérifier que $\mathbb{P}(A_1 \cap A_2 \cap \dots \cap A_n) = \mathbb{P}(A_1) \mathbb{P}(A_2) \dots \mathbb{P}(A_n)$

Il faut montrer que cette propriété de factorisation est vraie pour toute sous-famille.

Exemple : Deux lancers de Pile/Face

$A = \{\text{pile au 1}^{\text{er}} \text{ lancer}\}$

$B = \{\text{pile au 2nd lancer}\}$

$C = \{\text{même résultat aux 2 lancers}\}$

$$\mathbb{P}(A) = \mathbb{P}(B) = \frac{1}{2}$$

$$\begin{aligned} \mathbb{P}(C) &= \mathbb{P}(\text{"pile pile"}) + \mathbb{P}(\text{"face face"}) \\ &= \frac{1}{4} + \frac{1}{4} = \frac{1}{2} \end{aligned}$$

$$\begin{aligned} \mathbb{P}(A \cap B \cap C) &= \mathbb{P}(\text{"pile pile"}) = \frac{1}{4} \neq \mathbb{P}(A) \mathbb{P}(B) \mathbb{P}(C) = \frac{1}{8} \\ &(\mathbb{P}(C | A \cap B) = 1 \neq \mathbb{P}(C)) \end{aligned}$$

$$\mathbb{P}(A \cap B) = \mathbb{P}(\text{"pile pile"}) = \frac{1}{4} = \mathbb{P}(A) \mathbb{P}(B)$$

$$\mathbb{P}(A \cap C) = \mathbb{P}(A) \mathbb{P}(C) \text{ et } \mathbb{P}(B \cap C) = \mathbb{P}(B) \mathbb{P}(C)$$

Conclusion : A, B, C sont indép 2-à-2 mais non indép.

Prop. Si $(A_1, \dots, A_n)$ est une famille d'événements indép, il en est de même pour les 2^n familles $(A'_1, \dots, A'_n)$ avec $A'_i = A_i$ ou A_i^c , $1 \leq i \leq n$

Preuve Il suffit de montrer que si $A_1, \dots, A_n$ sont indép, alors $A_1^c, A_2, \dots, A_n$ le sont aussi

$$\forall \{i_1, \dots, i_p\} \subset \{2, \dots, n\}$$

$$\begin{aligned} \mathbb{P}(A_1^c \cap A_{i_1} \cap \dots \cap A_{i_p}) &= \mathbb{P}(A_{i_1} \cap \dots \cap A_{i_p}) - \mathbb{P}(A_1 \cap A_{i_1} \cap \dots \cap A_{i_p}) \\ &= \mathbb{P}(A_{i_1}) \dots \mathbb{P}(A_{i_p}) - \mathbb{P}(A_1) \mathbb{P}(A_{i_1}) \dots \mathbb{P}(A_{i_p}) \\ &= \mathbb{P}(A_1^c) \mathbb{P}(A_{i_1}) \dots \mathbb{P}(A_{i_p}) \end{aligned}$$

Ce calcul est suffisant pour dire que $A_1^c, A_2, \dots, A_n$ sont indép. $\square$

Exemple 2.32 [vidéo sur Moodle]

Lors de n tentatives indép. de proba. de succès $p \in [0, 1]$

$$\mathbb{P}(\text{il y a exactement } k \text{ succès}) = \binom{n}{k} p^k (1-p)^{n-k}, \quad 0 \leq k \leq n$$

$\rightsquigarrow$ loi binomiale (voir chap. 3)

Chap III Variables aléatoires discrètes

"variable aléatoire" (v.a.) est une quantité associée au résultat d'une expérience aléatoire

Exemples : • lancer 2 dés $\rightarrow$ la somme des résultats des 2 dés

• Suite ∞ de Pile/Face $\rightarrow$ Nb. de Pile avant d'obtenir Face pour la 1^{re} fois.

- Tirer un nombre ω au hasard dans $]0,1[$
 $\rightarrow$ le plus petit entier $\geq \frac{1}{\omega}$

Événement $\neq$ variable aléatoire

On peut construire des événements à partir de v.a.

exemple : $\{\text{la somme des résultats des 2 dés} \geq 5\} \in \mathcal{F}$

§ 3.1 V.a. discrètes

$(\Omega, \mathcal{F}, \mathbb{P})$ espace probabilisé E ensemble quelconque

Déf. On appelle **v.a. discrète** toute fonction $X: \Omega \rightarrow E$ tq.

$X(\Omega) := \{X(\omega) : \omega \in \Omega\}$ est dénombrable

et tq. pour tout $x \in X(\Omega)$, les ensembles
 $\{\omega \in \Omega : X(\omega) = x\}$ sont des événements de $\mathcal{F}$.

Si $E = \mathbb{R}$, X v.a. réelle

Si $E = \mathbb{R}^n$, X vecteur aléatoire de dimension n

Remarque: $\forall A \subseteq E$, $\{\omega \in \Omega : X(\omega) \in A\} \in \mathcal{F}$

car $\{\omega \in \Omega : X(\omega) \in A\}$

$= \{\omega \in \Omega : X(\omega) \in A \cap X(\Omega)\}$

$= \bigcup_{x \in A \cap X(\Omega)} \{\omega \in \Omega : X(\omega) = x\} \in \mathcal{F}$

$\uparrow$ dénombrable

Exemples: 1) Fonction constante $c \in \mathbb{R}$ fixé

$$X(\omega) = c \text{ pour tout } \omega \in \Omega$$

2) Soit $A \in \mathcal{F}$, L'indicatrice de l'événement A

$$X = 1_A: \Omega \rightarrow \mathbb{R} \text{ est une v.a. discrète}$$

Toute v.a. réelle Y qui ne prend que les valeurs 0 et 1 est nécessairement une fonction indicatrice

$$Y = 1_B \text{ où } B = \{\omega \in \Omega: Y(\omega) = 1\}$$

On l'appelle une v.a. de Bernoulli.

Notation: $\{X \in A\} = X^{-1}(A) = \{\omega \in \Omega: X(\omega) \in A\}$

↑

préimage (image réciproque) de A par X

ce type d'événements sont appelés événements générés par X

$$\{X = x\} = \{X \in \{x\}\}$$

$$\{X > a\} = \{X \in]a, \infty[\}$$

Si X et Y sont des v.a. réelles définies sur le même $(\Omega, \mathcal{F}, \mathbb{P})$

$$\{X = Y\} = \{X - Y = 0\}$$

$$\mathbb{P}(X \in A) = \mathbb{P}(\{X \in A\})$$

$$\mathbb{P}(X = x) = \mathbb{P}(\{X = x\})$$

Prop (Admise, vraie pour des v.a. générales)

(1) Si X et Y sont des v.a. réelles définies sur $(\Omega, \mathcal{F}, \mathbb{P})$, les fonctions suivantes sont des v.a. :

$$f(X): \omega \mapsto f(X(\omega)) \text{ où } f \text{ une fonction arbitraire}$$

$$aX + bY: \omega \mapsto aX(\omega) + bY(\omega) \text{ où } a, b \in \mathbb{R}.$$

$$\max(X, Y) : \omega \mapsto \max\{X(\omega), Y(\omega)\}$$

$$\min(X, Y) : \omega \mapsto \min\{X(\omega), Y(\omega)\}$$

(2) Si $(X_n)_{n \in \mathbb{N}}$ est une suite de v.a. à valeurs dans $\overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$ définies sur le même espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$,

$$\sup_{n \in \mathbb{N}} X_n : \omega \mapsto \sup\{X_n(\omega), n \in \mathbb{N}\}$$

$$\inf_{n \in \mathbb{N}} X_n : \omega \mapsto \inf\{X_n(\omega), n \in \mathbb{N}\}$$

$$\lim_{n \rightarrow \infty} X_n : \omega \mapsto \lim_{n \rightarrow \infty} X_n(\omega) \text{ si la limite existe pour tout } \omega \in \Omega$$

(par exemple si $(X_n(\omega))$ est monotone)

$$\sum_{n \in \mathbb{N}} X_n : \omega \mapsto \sum_{n \in \mathbb{N}} X_n(\omega) \text{ si la somme est bien définie pour tout } \omega \in \Omega$$

(par exemple si $X_n(\omega) \geq 0, \forall n \in \mathbb{N}$)

Def. Soit $X: \Omega \rightarrow E$ une v.a. discrète, on appelle **loi de X** l'application $\mu_X: \mathcal{P}(E) \rightarrow [0, 1]$ définie par

$$\mu_X(A) = \mathbb{P}(X \in A), \quad \forall A \subseteq E$$

On appelle **densité discrète de X** l'application

$$p_X : E \rightarrow [0, 1]$$

$$p_X(x) = \mathbb{P}(X=x) = \mu_X(\{x\}) \quad \forall x \in E$$

Loi μ_X décrit quelles valeurs prend X avec quelle probabilité.

Exemple Si $X = 1_A$, $X(\Omega) = \{0, 1\}$

$$p_X(1) = \mathbb{P}(X=1) = \mathbb{P}(A)$$

$$p_X(0) = \mathbb{P}(X=0) = \mathbb{P}(A^c) = 1 - \mathbb{P}(A)$$

$$p_X(x) = 0 \quad \forall x \notin \{0, 1\}.$$

Si $p_X(1) = p_X(0) = \frac{1}{2} = \mathbb{P}(A)$, alors

$X = 1_A$ et $1-X = 1_{A^c}$ ont la même loi

c-à-d. $\mu_X = \mu_{1-X}$

Mais X et X' ne sont jamais égales

Prop 3.7 [vidéo sur Moodle]

p_X est une densité discrète sur $X(\Omega)$ et est nulle en dehors de $X(\Omega)$

μ_X est la probabilité sur $(X(\Omega), \mathcal{P}(X(\Omega)))$ associée à p_X
étendue à une probabilité sur $(E, \mathcal{P}(E))$ par

$$\mu_X(A) = \sum_{x \in A} p_X(x) = \sum_{x \in A \cap X(\Omega)} p_X(x), \quad \forall A \subseteq E$$

Exemple 3.11 [vidéo sur Moodle]

Rappel: Un événement $A \in \mathcal{F}$ de proba $\mathbb{P}(A)=1$ est appelé *presque sûr*

Déf: Deux v.a. X et X' définies sur le même espace $(\Omega, \mathcal{F}, \mathbb{P})$
et à valeurs dans le même ensemble E sont dites

presque sûrement égales si

$$\mathbb{P}(X=X') = \mathbb{P}(\{\omega \in \Omega : X(\omega) = X'(\omega)\}) = 1$$

Autrement dit, X et X' diffèrent sur un ensemble de proba nulle

$$\mathbb{P}(X \neq X') = 0$$

Exercice 3.9: Montrer que si X et X' sont presque sûrement égales
alors X et X' ont la même loi (et la même densité discrète)

$$\mu_X = \mu_{X'} \quad \text{et} \quad p_X = p_{X'}$$

[vidéo sur Moodle]



X et X' ont même loi $\not\Rightarrow X$ et X' sont presque sûrement égales



Deux v.a. réelles définies sur les espaces probabilisés *différents* peuvent avoir la même loi.

Voir Exemple 3.8 du polycopié.

FIN SÉANCE 4

Exemple: Pile / Face

↓ ↓
0 1

$$\text{v.a. } X: \Omega = \{0, 1\} \longrightarrow \{0, 1\}$$

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

Dé à 6 faces $A = \text{"le résultat est pair"}$

$$\text{v.a. } Y = 1_A: \Omega' = \{1, 2, 3, 4, 5, 6\} \longrightarrow \{0, 1\}$$

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

X et Y définies sur deux espaces d'états différents
Pourtant, X et Y ont la même loi.

$$P_X(0) = P_X(1) = \frac{1}{2}$$

$$P_Y(0) = P(Y=0) = P(A^c) = \frac{|A^c|}{|\Omega'|} = \frac{3}{6} = \frac{1}{2}$$

$$P_Y(1) = P(Y=1) = P(A) = \frac{|A|}{|\Omega'|} = \frac{1}{2}$$

$$P_X = P_Y,$$

$$\mu_X = \mu_Y = \text{proba uniforme sur } \{0, 1\}$$

§3.2 Indépendance des v.a.

§3.2.1 Loi jointe et lois marginales

$$\begin{aligned} X: \Omega &\rightarrow E \\ Y: \Omega &\rightarrow F \end{aligned} \quad \text{v.a. discrètes}$$

Le couple $Z = (X, Y): \Omega \rightarrow E \times F$ est une v.a. (discrète)

La loi de Z , notée $\mu_{X,Y}$, est appelée **loi jointe** de X et Y
($\mu_{X,Y}$ est une proba. sur $E \times F$)

Les lois μ_X et μ_Y sont appelées **lois marginales**

Si X et Y sont des v.a. discrètes, (X, Y) l'est aussi

La densité discrète

$$p_{X,Y}(x, y) = \mathbb{P}((X, Y) = (x, y)) = \mathbb{P}(X = x, Y = y)$$

$$\{X = x, Y = y\} = \{X = x\} \cap \{Y = y\}$$

s'appelle **densité jointe**

$p_X(x) = \mathbb{P}(X = x)$ et $p_Y(y) = \mathbb{P}(Y = y)$ sont **densités marginales**

Prop $\forall x \in E \quad p_X(x) = \sum_{y \in F} p_{X,Y}(x, y)$

$$\sum_{x \in E} = \sum_{x \in E \cap X(\Omega)}$$

$$\forall y \in F \quad p_Y(y) = \sum_{x \in E} p_{X,Y}(x, y)$$

Preuve: $\{Y = y\}, y \in Y(\Omega)$ est une partition dénombrable de Ω

$$\{X = x\} = \bigsqcup_{y \in Y(\Omega)} \{X = x, Y = y\} \quad \text{où l'union est disjointe et dénombrable.}$$

$$p_X(x) = \mathbb{P}(X = x) = \sum_{y \in Y(\Omega)} \mathbb{P}(X = x, Y = y) \quad (\text{additivité de } \mathbb{P})$$

$$= \sum_{y \in Y(\Omega)} \mathbb{P}((X, Y) = (x, y))$$

$$= \sum_{y \in Y(\Omega)} p_{X,Y}(x, y)$$

$$= \sum_{y \in F} p_{X,Y}(x, y) \quad \text{car si } y \notin Y(\Omega), \\ \{(X, Y) = (x, y)\} \subset \{Y = y\} = \emptyset \\ \text{alors } p_{X,Y}(x, y) = 0. \quad \square$$

Conclusion: Lois marginales sont déterminées par la loi jointe
mais pas réciproquement (voir l'exemple 3.14)

Généralisation: $X_1, \dots, X_n$ v.a. discrètes définies sur Ω
à valeurs dans $E_1, \dots, E_n$

$(X_1, \dots, X_n): \Omega \rightarrow E_1 \times \dots \times E_n$ est une v.a. discrète

La densité jointe $p_{X_1, \dots, X_n}(x_1, \dots, x_n) = \mathbb{P}(X_1 = x_1, \dots, X_n = x_n)$

$$\forall x_i \in E_i \quad p_{X_i}(x_i) = \sum_{j \neq i} \sum_{x_j \in E_j} p_{X_1, \dots, X_n}(x_1, \dots, x_n)$$

↑
densité marginale

§ 3.2.2. Indépendance de v.a.

Déf Soient $X_1, \dots, X_n$ des v.a. définies sur Ω , à valeurs dans $E_1, \dots, E_n$. Elles sont dites indép. si pour tout choix de sous-ensembles $A_1 \subset E_1, \dots, A_n \subset E_n$, on a

$$\mathbb{P}(X_1 \in A_1, \dots, X_n \in A_n) = \prod_{i=1}^n \mathbb{P}(X_i \in A_i) \quad (*)$$

Remarque $\forall \{i_1, \dots, i_k\} \subset \{1, 2, \dots, n\}$
 on pose $A_j = E_j \quad j \notin \{i_1, \dots, i_k\}$

$(\{X_j \in A_j\} = \Omega \text{ pour ces } j)$

$$\mathbb{P}(X_{i_1} \in A_{i_1}, \dots, X_{i_k} \in A_{i_k})$$

$$= \mathbb{P}(X_1 \in A_1, \dots, X_n \in A_n)$$

$$\stackrel{(*)}{=} \prod_{i=1}^n \mathbb{P}(X_i \in A_i) = \prod_{\ell=1}^k \mathbb{P}(X_{i_\ell} \in A_{i_\ell})$$

Si $X_1, \dots, X_n$ sont indép., toute sous-famille $X_{i_1}, \dots, X_{i_k}$ le sont aussi

Déf. Soient $(X_i)_{i \in I}$ des v.a. définies sur le même $(\Omega, \mathcal{F}, \mathbb{P})$
 où l'ensemble des indices I est *arbitraire*.

Ces v.a. sont dites indép si pour tout sous-ensemble $J \subset I$
fini ($|J| < \infty$), les v.a. $(X_j)_{j \in J}$ sont indép.

Proposition 3.17 [Vidéo sur Moodle]

Soient $X: \Omega \rightarrow E$ et $Y: \Omega \rightarrow F$ des v.a. discrètes

X et Y sont indép si et seulement si

$$p_{X,Y}(x,y) = p_X(x) p_Y(y), \quad \forall x \in E, y \in F$$

Généralisation: $X_1: \Omega \rightarrow E_1, \dots, X_n: \Omega \rightarrow E_n$ indép si et seulement si

$$p_{X_1, \dots, X_n}(x_1, \dots, x_n) = \prod_{i=1}^n p_{X_i}(x_i) \quad \forall (x_1, \dots, x_n) \in E_1 \times \dots \times E_n$$

Exemple 3.14 [vidéo sur Moodle]

n boules, de 1 à n , dans une urne

1) On tire 2 boules avec remise.

X_1, X_2 les numéros des boules tirées

$$P_{X_1, X_2}(x_1, x_2) = \frac{1}{n^2}, \forall (x_1, x_2) \in \{1, 2, \dots, n\}^2$$

$$P_{X_1}(x) = P_{X_2}(x) = \frac{1}{n} \quad \forall 1 \leq x \leq n$$

X_1 et X_2 sont indép comme $P_{X_1, X_2}(x_1, x_2) = P_{X_1}(x_1) P_{X_2}(x_2)$

2) On tire 2 boules sans remise

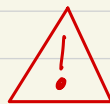
X'_1, X'_2 les numéros des boules tirées $X'_1 \neq X'_2$

$$P_{X'_1, X'_2}(x_1, x_2) = \frac{1}{n(n-1)} \quad \forall (x_1, x_2) \in \{1, 2, \dots, n\}^2$$

t.q. $x_1 \neq x_2$.

$$P_{X'_1}(x) = P_{X'_2}(x) = \frac{1}{n} \quad \forall 1 \leq x \leq n$$

X'_1 et X'_2 ne sont pas indép.



Conclusion: $X_1 \stackrel{(\text{loi})}{=} X'_1$

$X_2 \stackrel{(\text{loi})}{=} X'_2$

MAIS (X_1, X_2) n'a pas la même loi
que (X'_1, X'_2)

Une famille de v.a. indép. divisée en sous-familles disjointes
reste des v.a. indép.

Pour simplifier les notations, on regarde seulement le cas
de 2 paquets :

Proposition 3.19 [Vidéo sur Moodle] (Indép par paquets)

Soient $X_1, \dots, X_n$ des v.a. (discrètes) indép, et soient $I = \{i_1, \dots, i_\ell\}$ et $J = \{j_1, \dots, j_k\}$ des parties non vides et **disjointes** de $\{1, 2, \dots, n\}$.

Alors les v.a.

$X_I := (X_{i_1}, \dots, X_{i_\ell})$ et $X_J := (X_{j_1}, \dots, X_{j_k})$ sont indép.

Proposition 3.20 (Indép par transformation)

$X: \Omega \rightarrow E$, $Y: \Omega \rightarrow F$ deux v.a. indép

$f: E \rightarrow H$ $g: F \rightarrow K$ des fonctions (arbitraires)

Alors les v.a. $f(X)$ et $g(Y)$ sont aussi indép.

Preuve: $\forall A \subset H, B \subset K$,

$$\begin{aligned} \mathbb{P}(f(X) \in A, g(Y) \in B) &= \mathbb{P}(X \in f^{-1}(A), Y \in g^{-1}(B)) \\ &= \mathbb{P}(X \in f^{-1}(A)) \mathbb{P}(Y \in g^{-1}(B)) \end{aligned}$$

$$"f^{-1}(A)" := \{x \in E: f(x) \in A\}$$

par l'indép entre X et Y

" f^{-1} " n'est pas la fonction inverse!

$$= \mathbb{P}(f(X) \in A) \mathbb{P}(g(Y) \in B)$$

Donc $f(X)$ et $g(Y)$ sont indép.

□

Notation: $X \perp\!\!\!\perp Y$ signifie que X et Y sont indép.

Corollaire: Soient $X_1, \dots, X_n, X_{n+1}, \dots, X_{n+m}$ des v.a. réelles indép, avec $n, m \geq 1$. Alors

$$(X_1 + \dots + X_n) \perp\!\!\!\perp (X_{n+1} + \dots + X_{n+m})$$

Preuve: $X_I := (X_1, \dots, X_n)$

$$X_J := (X_{n+1}, \dots, X_{n+m})$$

Par l'indép par paquets, $X_I \perp\!\!\!\perp X_J$

$$\forall k \geq 1, f_k: \mathbb{R}^k \rightarrow \mathbb{R}$$

fonction somme

$$f_k(x_1, \dots, x_k) = x_1 + \dots + x_k$$

$$\text{Alors } X_1 + \dots + X_n = f_n(X_I)$$

$$X_{n+1} + \dots + X_{n+m} = f_m(X_J)$$

Par l'indép par transformation, $f_n(X_I) \perp\!\!\!\perp f_m(X_J)$ $\square$

§ 3.3 Exemples de v.a. discrètes.

1) Loi uniforme discrète

$$X: \Omega \rightarrow E, |E| < \infty.$$

"est de loi"

X est une v.a. uniforme (discrète) dans E (notée $X \sim \text{Unif}(E)$)

si et seulement la loi μ_X est la proba. uniforme sur E

$$\mu_X(A) = \frac{|A|}{|E|} \quad \forall A \subseteq E$$

$$p_X(x) = \frac{1}{|E|} \quad \forall x \in E$$

2) Loi de Bernoulli

X est une v.a. de Bernoulli si $X(\Omega) = \{0, 1\}$

$X = \mathbb{1}_{\{X=1\}}$ sa loi μ_X déterminée par $p = p_X(1) \in [0, 1]$

$$p_X(0) = 1 - p$$

$$p_X(x) = 0 \quad \text{si } x \notin \{0, 1\}.$$

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

3) Loi binomiale

$$n \geq 1, \quad p \in [0, 1].$$

X suit une loi binomiale de paramètre n et p ($X \sim \text{Bin}(n, p)$)
Si $X(\Omega) = \{0, 1, \dots, n\}$ et

$$p_X(k) = \binom{n}{k} p^k (1-p)^{n-k} \quad \forall k \in \{0, 1, \dots, n\}$$

Prop (voir Exemple 2.32)

Soient $X_1, \dots, X_n$ des v.a. $\text{Bern}(p)$ indép. Alors.

$X_1 + \dots + X_n$ suit la loi $\text{Bin}(n, p)$

compte le nb. de succès parmi n épreuves indép. de proba de succès p .

$$\text{Bin}(1, p) = \text{Bern}(p)$$

Exercice 3.22 Soient $X \sim \text{Bin}(n, p)$ et $Y \sim \text{Bin}(m, p)$ des v.a. indép.
Montrer que $X + Y \sim \text{Bin}(n+m, p)$

Lemme : Si $X: \Omega \rightarrow E$ sont deux v.a. de même loi,
 $Y: \Omega' \rightarrow E$

alors $\forall$ fonction $f: E \rightarrow F$,

$f(X): \Omega \rightarrow F$ et $f(Y): \Omega' \rightarrow F$ ont la même loi.

Solution de l'exercice 3.22 :

Soient $(X_i)_{1 \leq i \leq n+m}$ des v.a. $\text{Bern}(p)$ indép

Considérons $X' = \sum_{i=1}^n X_i$ et $Y' = \sum_{i=n+1}^{n+m} X_i$

Alors (X', Y') a la même loi que (X, Y) (pourquoi?)

Donc $X+Y$ a même loi que $X'+Y'$
 $X'+Y' = \sum_{i=1}^{n+m} X_i \sim \text{Bin}(n+m, p)$

□

4) Loi de Poisson

$$\lambda \in]0, \infty[$$

X suit une loi de Poisson de paramètre λ ($X \sim \text{Poi}(\lambda)$)
 si X est à valeurs dans $\mathbb{N}$ et

$$p_X(k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!} \quad \forall k \geq 0$$

Une binomiale avec beaucoup de tentatives mais proba de succès faible $\approx$ Poisson.

Prop Soit $(p_n)_{n \geq 1}$ une suite de réels dans $]0, 1[$

On suppose que $p_n \sim \frac{\lambda}{n}$ quand $n \rightarrow \infty$ où $\lambda > 0$

Alors si $X_n \sim \text{Bin}(n, p_n)$, on a pour tout $k \geq 0$,

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{P}(X_n = k) &= \mathbb{P}(X = k) \quad \text{où } X \sim \text{Poi}(\lambda). \\ &= e^{-\lambda} \cdot \frac{\lambda^k}{k!} \end{aligned}$$

Prop $X \sim \text{Poi}(\lambda)$, $Y \sim \text{Poi}(\mu)$

Si $X \perp Y$ alors $X+Y \sim \text{Poi}(\lambda+\mu)$

5) Loi géométrique

$$p \in]0, 1[$$

X suit la loi géométrique de paramètre p ($X \sim \text{Géom}(p)$)

si $X(\Omega) = \mathbb{N}^*$ et $p_X(k) = p \underbrace{(1-p)^{k-1}}_{\substack{\text{succès} \\ k-1 \text{ échecs}}} \quad \forall k \geq 1.$

X représente l'instant du premier succès lors de tentatives indép de proba de succès p .

Variant de la loi géométrique : $Y \sim \text{Géom}_0(p)$

$$p_Y(k) = p(1-p)^k \quad \forall k \geq 0$$

Prop Absence de mémoire

$X \sim \text{Géom}(p)$ Alors $\forall n, m \geq 0$

$$\mathbb{P}(X > n+m | X > n) = \mathbb{P}(X > m)$$

Exercice 3.26 Soit X une v.a. à valeurs dans $\mathbb{N}^*$ tel que

$$\mathbb{P}(X > n+m | X > n) = \mathbb{P}(X > m) \quad \forall n, m \geq 0$$

Montrer que $X \sim \text{Géom}(p)$ avec $p = \mathbb{P}(X=1)$

FIN SÉANCE 5

Rappel : $X \sim \text{Poi}(\lambda) \quad \lambda \in]0, +\infty[$ $X(\Omega) = \mathbb{N}$
 $\forall k \geq 0 \quad p_X(k) = \mathbb{P}(X=k) = e^{-\lambda} \cdot \frac{\lambda^k}{k!}$

Prop Soient $X \sim \text{Poi}(\lambda)$ et $Y \sim \text{Poi}(\mu)$

Si $X \perp Y$ alors $X+Y \sim \text{Poi}(\lambda+\mu)$

Preuve: $\forall n \geq 0 \quad \{X+Y=n\} = \bigsqcup_{k=0}^n \{X=k, Y=n-k\}$
union disjointe

$$\begin{aligned}
\mathbb{P}(X+Y=n) &= \sum_{k=0}^n \mathbb{P}(X=k, Y=n-k) && \text{additivité de } \mathbb{P} \\
&= \sum_{k=0}^n \mathbb{P}(X=k) \mathbb{P}(Y=n-k) && X \perp\!\!\!\perp Y \\
&= \sum_{k=0}^n e^{-\lambda} \frac{\lambda^k}{k!} e^{-\mu} \frac{\mu^{n-k}}{(n-k)!} && \begin{array}{l} X \sim \text{Poi}(\lambda) \\ Y \sim \text{Poi}(\mu) \end{array} \\
&= e^{-(\lambda+\mu)} \frac{1}{n!} \sum_{k=0}^n \frac{n!}{k!(n-k)!} \lambda^k \mu^{n-k} \\
&= e^{-(\lambda+\mu)} \frac{1}{n!} (\lambda+\mu)^n
\end{aligned}$$

Ainsi $X+Y \sim \text{Poi}(\lambda+\mu)$. □

"Binomiale avec **beaucoup** de tentative mais proba de succès **faible**"
 $\approx$ Poisson

Prop Soit $(X_n)_{n \geq 1}$ une suite de v.a. t.q. $X_n \sim \text{Bin}(n, p_n)$
 où $p_n \sim \frac{\lambda}{n}$ quand $n \rightarrow \infty$, $\lambda \in]0, \infty[$
 Alors $\forall k \geq 0$ $\lim_{n \rightarrow \infty} \mathbb{P}(X_n = k) = e^{-\lambda} \frac{\lambda^k}{k!}$

Preuve (Cas $p_n = \frac{\lambda}{n}$)

$k \geq 0$ fixé, $n \geq k$

($\mathbb{P}(X_n = k) = 0$ si $k > n$)

$$\begin{aligned}
\mathbb{P}(X_n = k) &= \binom{n}{k} p_n^k (1-p_n)^{n-k} \\
&= \frac{n!}{k!(n-k)!} \frac{\lambda^k}{n^k} \left(1 - \frac{\lambda}{n}\right)^{n-k} \\
&= \frac{\lambda^k}{k!} \frac{n(n-1) \cdots (n-k+1)}{n^k} \frac{1}{\left(1 - \frac{\lambda}{n}\right)^k} \left(1 - \frac{\lambda}{n}\right)^n
\end{aligned}$$

$$= 1 \cdot \left(1 - \frac{1}{n}\right) \cdots \left(1 - \frac{k-1}{n}\right)$$

$$k \text{ étant fixé: } \lim_{n \rightarrow \infty} \frac{\overbrace{n(n-1) \cdots (n-k+1)}^{n(n-1) \cdots (n-k+1)}}{n^k} = 1$$

$$\lim_{n \rightarrow \infty} \left(1 - \frac{\lambda}{n}\right)^k = 1$$

$$\lim_{n \rightarrow \infty} \left(1 - \frac{\lambda}{n}\right)^n = e^{-\lambda}$$

Analyse: $\forall x \in \mathbb{R}, \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n = \lim_{n \rightarrow \infty} e^{n \ln(1 + \frac{x}{n})} = e^x$

Donc $\lim_{n \rightarrow \infty} \mathbb{P}(X_n = k) = \frac{\lambda^k}{k!} \cdot \underset{1}{\underset{1}{\cdot}} \cdot e^{-\lambda} = e^{-\lambda} \frac{\lambda^k}{k!} \quad \square$

Exercice traiter le cas général $np_n \xrightarrow{n \rightarrow \infty} \lambda$.

Rappel $X \sim \text{Géom}(p), p \in]0, 1[$

$$X(\Omega) = \mathbb{N}^*$$

$$\forall k \geq 1, \mathbb{P}(X=k) = p_X(k) = (1-p)^{k-1} \cdot p$$

l'instant du premier succès ↗ d'abord $k-1$ échecs ↗ k^{e} tentative = succès ↗

Exemple: Lancers successifs de dés (à 6 faces)

X = numéro du lancer qui donne 3 pour la 1^{ère} fois

alors $X \sim \text{Géom}\left(\frac{1}{6}\right)$

Y = numéro du lancer qui donne 3 ou 4 pour la 1^{ère} fois

alors $Y \sim \text{Géom}\left(\frac{2}{6}\right)$ (voir l'Exercice 4.11 en TD)

Variation: $Z \sim \text{Géom}_0(p) \Leftrightarrow \forall k \geq 0, \mathbb{P}(Z=k) = (1-p)^k \cdot p$
 $(Z(\Omega) = \mathbb{N})$

Absence de mémoire

$$\forall n, m \geq 0 \quad \mathbb{P}(X > n+m \mid X > n) = \mathbb{P}(X > m)$$

Preuve
$$\mathbb{P}(X > n+m \mid X > n) = \frac{\mathbb{P}(X > n+m, X > n)}{\mathbb{P}(X > n)} = \frac{\mathbb{P}(X > n+m)}{\mathbb{P}(X > n)}$$

$$\begin{aligned} \forall n \geq 0, \quad \mathbb{P}(X > n) &= \sum_{k=n+1}^{\infty} \mathbb{P}(X=k) \\ &= \sum_{k=n+1}^{\infty} p(1-p)^{k-1} \\ &= (1-p)^n \sum_{\ell=0}^{\infty} p(1-p)^{\ell} = (1-p)^n \end{aligned}$$

(voir l'Exercice 4.4 en TD)

$$\begin{aligned} \text{Donc } \mathbb{P}(X > n+m \mid X > n) &= \frac{(1-p)^{n+m}}{(1-p)^n} = (1-p)^m \\ &= \mathbb{P}(X > m) \end{aligned}$$

□

Exercice : Soit Y une v.a. à valeurs dans $\mathbb{N}^*$ t.q

$$\forall n, m \geq 0 \quad \mathbb{P}(Y > n+m \mid Y > n) = \mathbb{P}(Y > m)$$

Montrer que $Y \sim \text{Géom}(p)$ avec $p = \mathbb{P}(Y=1)$

Preuve La propriété d'absence de mémoire implique que

$$\forall n, m \geq 0 \quad \mathbb{P}(Y > n+m) = \mathbb{P}(Y > n) \mathbb{P}(Y > m)$$

En prenant $m=1$: $\mathbb{P}(Y > n+1) = \mathbb{P}(Y > n) \mathbb{P}(Y > 1)$

Alors $\forall n \geq 1, \quad \mathbb{P}(Y > n) = \mathbb{P}(Y > 1)^n = p^n$

et donc $Y \sim \text{Géom}(p)$.

□

$$\begin{aligned} \uparrow \quad \mathbb{P}(Y=n) &= \mathbb{P}(Y > n-1) - \mathbb{P}(Y > n) \\ \text{car } \{Y=n\} &= \{Y > n-1\} \setminus \{Y > n\} \end{aligned}$$

§ 3.4 Espérance, variance et moments

§ 3.4.1 Définition de l'espérance

L'espérance d'une v.a. **discrète réelle** est la moyenne **pondérée** des valeurs x_i , que peut prendre X , pondérées par les poids

$$p_i = \mathbb{P}(X = x_i)$$

Interprétation fréquentiste :

on répète $N \gg 1$ fois une expérience aléatoire

$X_1, \dots, X_N$ les valeurs de la v.a. X

N_x = nb. de fois où X_i vaut $x \in \mathbb{R}$

$$\frac{1}{N} (X_1 + \dots + X_N) = \frac{1}{N} \sum_{x \in X(\Omega)} x N_x = \sum_{x \in X(\Omega)} x \frac{N_x}{N}$$

$$\approx \sum_{x \in X(\Omega)} x \mathbb{P}(X=x)$$

Notation: $x \in \mathbb{R}$ $x^+ = \max\{x, 0\}$ partie positive de x

$0 \leq x^- = \max\{-x, 0\}$ partie négative de x

$$x^+ - x^- = x \quad \text{et} \quad x^+ + x^- = |x|$$

Définition de l'espérance

X v.a. **discrète réelle**, $X(\Omega)$ dénombrable, densité discrète p_X

On dit que X admet une espérance si $\sum_{x \in X(\Omega)} x p_X(x)$ est bien définie

c-à-d. si $\sum_{x \in X(\Omega)} x^+ p_X(x) < +\infty$ ou $\sum_{x \in X(\Omega)} x^- p_X(x) < +\infty$

Si X admet une espérance, on définit l'espérance de X

$${}^{\text{''}}\mathbb{E}(X){}^{\text{''}} = {}^{\text{''}}\mathbb{E}[X]{}^{\text{''}} := \sum_{x \in X(\Omega)} x p_X(x) = \sum_{x \in \mathbb{R}} x p_X(x) \in [-\infty, +\infty]$$



$\mathbb{E}(X)$ pas toujours définie à cause du problème " $\infty - \infty$ "

Remarque 1: Si X ne prend qu'un nb. **fini** de valeurs, alors X admet une espérance finie.

Remarque 2: Toute v.a. **positive** X admet une espérance $\mathbb{E}(X) \in [0, +\infty]$ qui peut éventuellement valoir $+\infty$

Remarque 3: Pour toute v.a. réelle X , on considère les v.a. positives

$$X^+ = \max\{X, 0\} \geq 0$$

$$X^- = \max\{-X, 0\} \geq 0$$

$$|X| = X^+ + X^- \geq 0$$

$\mathbb{E}(X^+)$, $\mathbb{E}(X^-)$ et $\mathbb{E}(|X|)$ toujours bien définies

$$\mathbb{E}(X^+) = \sum_{x \in X^+(\Omega)} x \mathbb{P}(X^+ = x) = \sum_{x \in \mathbb{R}} x \mathbb{P}(X^+ = x)$$

$$= \sum_x x^+ \mathbb{P}(X = x) = \sum_x x^+ p_X(x)$$

de même $\mathbb{E}(X^-) = \sum_x x^- p_X(x)$

$$\mathbb{E}(|X|) = \sum_x |x| p_X(x)$$

- X admet une espérance $\Leftrightarrow$ au moins l'une des deux espérances $\mathbb{E}(X^+)$, $\mathbb{E}(X^-)$ est finie

Alors $\mathbb{E}(X) = \mathbb{E}(X^+) - \mathbb{E}(X^-)$

car $\sum_x x p_X(x) = \sum_x x^+ p_X(x) - \sum_x x^- p_X(x)$

- X admet une espérance **finie** $\Leftrightarrow \mathbb{E}(X^+)$ et $\mathbb{E}(X^-)$ sont finies

$$\Leftrightarrow \mathbb{E}(|X|) = \mathbb{E}(X^+) + \mathbb{E}(X^-) < +\infty$$

$$\Leftrightarrow \sum_x x p_X(x) \text{ est absolument convergente} \\ \text{c-à-d. } \sum_x |x| p_X(x) < +\infty.$$

Exemple 1 Soit X v.a. réelle presque sûrement constante
 $\exists c \in \mathbb{R} \quad p_X(c) = 1$ et $p_X(x) = 0$ si $x \neq c$
 $\mathbb{E}(X) = c p_X(c) = c$

Exemple 2 Soit X v.a. de Bernoulli de paramètre p
 $\mathbb{E}(X) = 0 \times \mathbb{P}(X=0) + 1 \times \mathbb{P}(X=1) = p$

En particulier, si $X = 1_A$, $A \in \mathcal{F}$
 $p_X(1) = \mathbb{P}(A)$ et $p_X(0) = 1 - \mathbb{P}(A)$
 $\mathbb{E}(X) = \mathbb{P}(A)$

Exemple 3 $X =$ nb. obtenu en lançant un dé à 6 faces
 $p_X(x) = \frac{1}{6} \quad \forall x \in \{1, 2, \dots, 6\} = X(\Omega)$
 $\mathbb{E}(X) = 1 \times \frac{1}{6} + 2 \times \frac{1}{6} + \dots + 6 \times \frac{1}{6} = \frac{7}{2}$



$\mathbb{E}(X) \notin X(\Omega)$

En général, $\mathbb{E}(X)$ n'est pas la valeur la plus probable de X
ni une valeur que l'on s'attend à observer.

§ 3.4.2 Propriétés de l'espérance

Prop Soit X une v.a. admettant une espérance.

On a $|\mathbb{E}(X)| \leq \mathbb{E}(|X|)$. En particulier,

Si $\mathbb{E}(|X|) < +\infty$, alors $\mathbb{E}(X)$ est finie

Preuve: $|\mathbb{E}(X)| = |\sum x p_X(x)| \leq \sum |x| p_X(x) = \mathbb{E}(|X|) \quad \square$

Proposition 3.35 [Vidéo sur Moodle]

Soit X une v.a. discrète à valeurs dans $X(\Omega) \subset \mathbb{E}$

Soit $g: \mathbb{E} \rightarrow \mathbb{R}$ une fonction

La v.a. $g(X)$ admet une espérance si et seulement si

$\sum_{x \in X(\Omega)} g(x) p_X(x)$ est bien définie. Dans ce cas,

$$\mathbb{E}(g(X)) = \sum_{x \in X(\Omega)} g(x) p_X(x) \in [-\infty, +\infty]$$

Formule de transfert

Remarque: Cette formule s'applique toujours à $g(X)^+$, $g(X)^-$, $|g(X)|$
 $g(X)$ admet une espérance finie $\Leftrightarrow \sum_{x \in X(\Omega)} |g(x)| p_X(x) < +\infty$

Proposition 3.36. Soit $X: (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow \mathbb{R}$ v.a. discrète avec Ω dénombrable

Alors X admet une espérance si et seulement si la somme

$\sum_{\omega \in \Omega} X(\omega) \mathbb{P}(\{\omega\})$ est définie. Si c'est le cas, on a

$$\mathbb{E}(X) = \sum_{\omega \in \Omega} X(\omega) \mathbb{P}(\{\omega\}) \quad (\text{formule alternative de l'espérance})$$

Preuve v.a. discrètes $\gamma: \Omega \rightarrow \Omega$ identité
 $\omega \mapsto \omega$

$X: \Omega \rightarrow \mathbb{R}$ (vue comme une fonction)
 $\omega \mapsto X(\omega)$

$$X(\gamma)(\omega) = (X \circ \gamma)(\omega) = X(\gamma(\omega)) = X(\omega), \quad \forall \omega \in \Omega$$

$$\Rightarrow \mathbb{E}(X(\gamma)) = \mathbb{E}(X)$$

$$\begin{aligned} \text{Or } \mathbb{E}(X(Y)) &= \sum_{y \in Y(\Omega)} X(y) P_Y(y) \quad (\text{formule de transfert}) \\ &= \sum_{y \in Y(\Omega)} X(y) P(\{y\}) \quad (\forall \omega, Y(\omega) = y) \end{aligned}$$

$$\text{Donc } \mathbb{E}(X) = \sum_{y \in Y(\Omega)} X(y) P(\{y\}) = \sum_{\omega \in \Omega} X(\omega) P(\{\omega\}). \quad \square$$

Prop (Linéarité de l'espérance)

- Si $\mathbb{E}(X), \mathbb{E}(Y)$ sont finies, $\forall a, b \in \mathbb{R}$, $aX + bY$ admet une espérance finie et $\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y)$.
- Si $X, Y \geq 0$, $a, b \in \mathbb{R}^+$, $aX + bY$ admet une espérance et $\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y)$.

"Preuve": D'abord, si $c \in \mathbb{R}$ est une constante, $\mathbb{E}(cX) = c\mathbb{E}(X)$

$$\text{car } \mathbb{E}(cX) = \sum_{x \in X(\Omega)} cx P_X(x) = c \sum_{x \in X(\Omega)} x P_X(x) = c\mathbb{E}(X)$$

↑
formule de transfert

Ensuite, il suffit de montrer que $\mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y)$

On note $X(\Omega) = \{x_j, j \in \mathbb{N}\}$, $Y(\Omega) = \{y_k, k \in \mathbb{N}\}$

$$\mathbb{E}(X+Y) \stackrel{\text{formule de transfert}}{=} \sum_{j,k \in \mathbb{N}} (x_j + y_k) P(X=x_j \text{ et } Y=y_k)$$

$$\begin{aligned} &= \sum_j \left(\sum_k x_j P(X=x_j \text{ et } Y=y_k) \right) + \sum_k \left(\sum_j y_k P(X=x_j \text{ et } Y=y_k) \right) \\ &\stackrel{\text{Fubini (admis)}}{=} \sum_j x_j P(X=x_j) + \sum_k y_k P(Y=y_k) = \mathbb{E}(X) + \mathbb{E}(Y) \quad \square \end{aligned}$$

Prop (Monotonie de l'espérance)

$X: (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow \mathbb{R}$, $Y: (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow \mathbb{R}$ v.a.

Si $\forall \omega \in \Omega$, $X(\omega) \leq Y(\omega)$, alors $\mathbb{E}(X) \leq \mathbb{E}(Y)$

"Preuve" $Y - X \geq 0$ v.a. positive

$$\mathbb{E}(Y - X) + \mathbb{E}(X) = \mathbb{E}(Y) \quad \text{par linéarité}$$

Comme $\mathbb{E}(Y - X) \geq 0$, on a $\mathbb{E}(X) \leq \mathbb{E}(Y)$. $\square$

Application de monotonie:

Si X **bornée** c-à-d. $\exists M < +\infty$ t.q. $|X(\omega)| \leq M$ pour tout $\omega \in \Omega$
alors $|\mathbb{E}(X)| \leq \mathbb{E}(|X|) \leq M$.

En particulier, X admet une espérance finie

Applications de linéarité:

1) $X \sim \text{Bin}(n, p)$

$X = X_1 + \dots + X_n$ où $X_i \sim \text{Bern}(p)$ $1 \leq i \leq n$.
sont indépendantes

$$\mathbb{E}(X) = \mathbb{E}(X_1) + \dots + \mathbb{E}(X_n) = np \quad (\text{sans utiliser l'indép})$$

2) $A_1, \dots, A_n$ événements (pas nécessairement indép)

$$X = 1_{A_1} + \dots + 1_{A_n} = \text{nb. des } A_i \text{ vérifiés}$$

loi de X peut être très compliquée

$$\mathbb{E}(X) = \mathbb{E}(1_{A_1}) + \dots + \mathbb{E}(1_{A_n}) = \mathbb{P}(A_1) + \dots + \mathbb{P}(A_n)$$

application au paradoxe des anniversaires

(Exemple 3.39 [vidéo en Moodle])

3) une nouvelle preuve de la formule d'inclusion-exclusion
voir Exemple 3.38, page 37 du polycopié

Prop: Soit $X: (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow \mathbb{R}^+$ v.a. **positive**
Si $\mathbb{E}(X) = 0$, alors X est presque sûrement égale à 0
c-à-d. $\mathbb{P}(X=0) = 1$

Preuve: $\mathbb{E}(X) = \sum_{x \in X(\Omega)} x \mathbb{P}_X(x)$, $X(\Omega) \subset \mathbb{R}^+$

↑
une somme de termes positifs vaut 0
 $\Leftrightarrow$ tous les termes sont nuls

Donc $\forall x \in X(\Omega)$, $x > 0$ on a $\mathbb{P}_X(x) = 0$
 $\mathbb{P}(X > 0) = \sum_{\substack{x \in X(\Omega) \\ x > 0}} x \mathbb{P}_X(x) = 0$

Ainsi $\mathbb{P}(X=0) = \mathbb{P}(X \geq 0) - \mathbb{P}(X > 0) = 1 - 0 = 1 \quad \square$

Linéarité "infinie"

Prop (Admise) Soit $(X_i)_{i \geq 1}$ une suite de v.a. **positives** définies sur le même espace $(\Omega, \mathcal{F}, \mathbb{P})$.

$\sum_{i=1}^{\infty} X_i$ est une v.a. bien définie (qui peut valoir $+\infty$)

$$\text{Alors } \mathbb{E}\left(\sum_{i=1}^{\infty} X_i\right) = \sum_{i=1}^{\infty} \mathbb{E}(X_i)$$

(à voir en L3, conséquence du théorème de convergence monotone)

§ 3.4.3 Moments, variance et covariance

Définition (Moments)

Soit $k \geq 1$. On dit que X admet un moment d'ordre k fini si $\mathbb{E}(|X|^k) < +\infty$. Si c'est le cas, $\mathbb{E}(X^k)$ est finie, et appelée **moment d'ordre k** de X

Définition (Variance et covariance)

$X: (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow \mathbb{R}$, $Y: (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow \mathbb{R}$ v.a.

- Si X admet un moment d'ordre 2 fini, on appelle **variance** de X

$$\text{Var}(X) := \mathbb{E}((X - \mathbb{E}(X))^2) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 \geq 0$$

La racine carrée $\sigma_X := \sqrt{\text{Var}(X)}$ s'appelle **l'écart type** de X

(σ_X quantifie
" l'écart de X à son espérance $\mathbb{E}(X)$ "
ou " la dispersion des valeurs de X autour de $\mathbb{E}(X)$ ")

- Si X , Y et XY admettent une espérance finie, on appelle **covariance** de X et Y

$$\begin{aligned} \text{Cov}(X, Y) &:= \mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))) \\ &= \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) \end{aligned}$$

Remarque $\text{Var}(X) = \text{Cov}(X, X)$

— FIN SÉANCE 6 —

Rappel: X v.a. réelle discrète

Espérance $\mathbb{E}(X) := \sum_{x \in X(\Omega)} x p_X(x)$ si bien définie
(ne dépend que de la loi de X)
$$= \sum_{x \in X(\Omega)} x^+ p_X(x) - \sum_{x \in X(\Omega)} x^- p_X(x) \neq "(+\infty) - (+\infty)"$$

$\mathbb{E}(X) \in [-\infty, +\infty]$ est un chiffre déterministe, n'est pas une v.a.

Formule de transfert g fonction à valeurs réelles

$$\mathbb{E}(g(X)) = \sum_{x \in X(\Omega)} g(x) p_X(x)$$

- Monotonie de l'espérance
- Linéarité de l'espérance
- $|\mathbb{E}(X)| \leq \mathbb{E}(|X|)$

Pour une v.a. positive X , $\mathbb{E}(X) = 0 \Rightarrow X$ est presque sûrement égale à 0

§3.4.3 Moments, Variance et covariance

X admet un moment d'ordre $k \geq 1$ fini si $\mathbb{E}(|X|^k) < +\infty$

Dans ce cas, $\mathbb{E}(X^k) \in \mathbb{R}$ s'appelle moment d'ordre k de X

Prop (1) Si X et Y sont v.a. admettant un moment d'ordre k fini.

alors $X+Y$ admet aussi un moment d'ordre k fini

(2) Si X admet un moment d'ordre k fini, alors X admet un moment d'ordre j fini pour tout $1 \leq j \leq k$.

Preuve (1) $\forall x, y \in \mathbb{R}$, $|x+y| \leq 2 \max\{|x|, |y|\}$

$$|x+y|^k \leq 2^k \max\{|x|^k, |y|^k\} \leq 2^k |x|^k + 2^k |y|^k$$

$$\Rightarrow \forall \omega \in \Omega, |X(\omega) + Y(\omega)|^k \leq 2^k |X(\omega)|^k + 2^k |Y(\omega)|^k$$

$$\text{notée } |X+Y|^k \leq 2^k |X|^k + 2^k |Y|^k$$

Monotonie + Linéarité de l'espérance

$$\Rightarrow \mathbb{E}(|X+Y|^k) \leq 2^k \mathbb{E}(|X|^k) + 2^k \mathbb{E}(|Y|^k) < +\infty$$

$$(2) \forall 1 \leq j \leq k, x \in \mathbb{R} \quad |x|^j \leq \begin{cases} 1 & \text{si } |x| \leq 1 \\ |x|^k & \text{si } |x| > 1 \end{cases}$$

$$\Rightarrow |x|^j \leq 1 + |x|^k \text{ pour tout } x \in \mathbb{R}$$

$$X \text{ v.a.} \quad |X|^j \leq 1 + |X|^k$$

$$\Rightarrow \mathbb{E}(|X|^j) \leq 1 + \mathbb{E}(|X|^k) < +\infty$$

□

Remarque Il est possible que $\mathbb{E}(X^{k+1}) = +\infty$ mais $\mathbb{E}(X^k) < +\infty$

exemple: X à valeurs dans $\mathbb{N}^*$

$$\forall n \geq 1, \mathbb{P}(X=n) = p_X(n) = c \cdot n^{-(k+2)}$$

où c est la constante telle que $\sum_{n \geq 1} p_X(n) = 1$

$$\mathbb{E}(X^k) = \sum_{n \geq 1} n^k p_X(n) = \sum_{n \geq 1} \frac{c}{n^2} < +\infty$$

$$\mathbb{E}(X^{k+1}) = \sum_{n \geq 1} n^{k+1} p_X(n) = \sum_{n \geq 1} \frac{c}{n} = +\infty$$

Définition $X: (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow \mathbb{R}$, $Y: (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow \mathbb{R}$ v.a.

- Si X admet un moment d'ordre 2 fini,

$$\text{Variance de } X: \text{Var}(X) := \mathbb{E}((X - \mathbb{E}(X))^2) \geq 0$$

$$\text{Écart-type de } X: \sigma_X := \sqrt{\text{Var}(X)} \geq 0$$

- Si X, Y et XY admettent une espérance finie

$$\text{Covariance de } X \text{ et } Y: \text{Cov}(X, Y) := \mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))) \in \mathbb{R}$$

$$\text{Cov}(X, X) = \text{Var}(X)$$

σ_X quantifie "l'écart de X à son espérance $\mathbb{E}(X)$ "

ou "la dispersion des valeurs de X autour de $\mathbb{E}(X)$ "

Lemme (1) $\text{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$

(2) $\text{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$

Preuve (1) On note $\mu = \mathbb{E}(X) \in \mathbb{R}$

$$(X - \mu)^2 = X^2 - 2\mu X + \mu^2$$

$$\mathbb{E}((X - \mu)^2) = \mathbb{E}(X^2) - 2\mu \mathbb{E}(X) + \mu^2 \quad \text{par linéarité}$$

$$= \mathbb{E}(X^2) - 2\mu^2 + \mu^2 = \mathbb{E}(X^2) - \mu^2$$

(2) On note $\mu_X = \mathbb{E}(X)$ et $\mu_Y = \mathbb{E}(Y)$

$$\begin{aligned}\mathbb{E}((X - \mu_X)(Y - \mu_Y)) &= \mathbb{E}(XY - \mu_X Y - \mu_Y X + \mu_X \mu_Y) \\ &= \mathbb{E}(XY) - \mu_X \mathbb{E}(Y) - \mu_Y \mathbb{E}(X) + \mu_X \mu_Y \\ &= \mathbb{E}(XY) - \mu_X \mu_Y - \mu_Y \mu_X + \mu_X \mu_Y \\ &= \mathbb{E}(XY) - \mu_X \mu_Y\end{aligned}$$

□

Remarque Si X et Y admettent un moment d'ordre 2 fini, alors XY admet une espérance finie, et ainsi $\text{Cov}(X, Y)$ est bien définie

$$\forall x, y \in \mathbb{R} \quad |xy| \leq \frac{1}{2}(x^2 + y^2)$$

$$\Rightarrow |XY| \leq \frac{1}{2}X^2 + \frac{1}{2}Y^2$$

$$\Rightarrow \mathbb{E}(|XY|) \leq \frac{1}{2}\mathbb{E}(X^2) + \frac{1}{2}\mathbb{E}(Y^2) < +\infty$$

Prop La covariance $\text{Cov}(\cdot, \cdot)$ est une forme **symétrique** et **bilinéaire**

$\forall$ v.a. X, Y, Z admettant un moment d'ordre 2 fini

$\forall a, b \in \mathbb{R}$, on a $\text{Cov}(X, Y) = \text{Cov}(Y, X)$

$$\text{Cov}(aX + bY, Z) = a\text{Cov}(X, Z) + b\text{Cov}(Y, Z)$$

De plus, si X ou Y est presque sûrement constante,

$$\text{on a } \text{Cov}(X, Y) = 0$$

Preuve • symétrique : par définition

$$\begin{aligned} \bullet \text{ bilinéaire : } \text{Cov}(aX+bY, Z) &= \mathbb{E}((aX+bY)Z) - (a\mathbb{E}(X)+b\mathbb{E}(Y))\mathbb{E}(Z) \\ &= a\mathbb{E}(XZ) + b\mathbb{E}(YZ) - a\mathbb{E}(X)\mathbb{E}(Z) - b\mathbb{E}(Y)\mathbb{E}(Z) \\ &= a(\mathbb{E}(XZ) - \mathbb{E}(X)\mathbb{E}(Z)) + b(\mathbb{E}(YZ) - \mathbb{E}(Y)\mathbb{E}(Z)) \\ &= a\text{Cov}(X, Z) + b\text{Cov}(Y, Z) \end{aligned}$$

- Si X est presque sûrement égale à $c \in \mathbb{R}$,
alors $\mathbb{E}(X) = c$ et $X - \mathbb{E}(X)$ est presque sûrement égale à 0
 $\Rightarrow (X - \mathbb{E}(X))(Y - \mathbb{E}(Y))$ est presque sûrement égale à 0
 $\Rightarrow \mathbb{E}((X - \mathbb{E}(X))(Y - \mathbb{E}(Y))) = \text{Cov}(X, Y) = 0$ □

Prop Soit X une v.a. admettant un moment d'ordre 2 fini,
alors

- (1) $\text{Var}(aX+b) = a^2 \text{Var}(X)$, $\forall a, b \in \mathbb{R}$
- (2) $\text{Var}(X) = 0 \iff X$ est presque sûrement constante
- (3) $\text{Var}(X) \leq \mathbb{E}((X-c)^2)$ pour tout $c \in \mathbb{R}$, avec égalité
si et seulement si $c = \mathbb{E}(X)$ (Ex 5.2 en TD)
- (4) Si $X_1, \dots, X_n$ sont des v.a. admettant un moment d'ordre 2 fini, alors

$$\begin{aligned} \text{Var}\left(\sum_{i=1}^n X_i\right) &= \sum_{i=1}^n \text{Var}(X_i) + \sum_{\substack{1 \leq i, j \leq n \\ i \neq j}} \text{Cov}(X_i, X_j) \\ &= \sum_{i=1}^n \text{Var}(X_i) + 2 \sum_{1 \leq i < j \leq n} \text{Cov}(X_i, X_j) \end{aligned}$$

Preuve (1) $\text{Var}(aX+b) = \text{Cov}(aX+b, aX+b) \overset{\substack{\uparrow \\ \text{proposition précédente}}}{=} a^2 \text{Cov}(X, X) = a^2 \text{Var}(X)$

(2) X est **p.s.** (presque sûrement) constante $\Rightarrow \text{Var}(X) = \text{Cov}(X, X) = 0$
 Réciproquement, si $\text{Var}(X) = \mathbb{E}(\underbrace{(X - \mathbb{E}(X))^2}_{\text{positive}}) = 0$, alors

$(X - \mathbb{E}(X))^2$ est p.s. égale à 0

$\Leftrightarrow X - \mathbb{E}(X)$ est p.s. égale à 0

$\Leftrightarrow X$ est p.s. égale à $\mathbb{E}(X)$

(3) $\forall c \in \mathbb{R}, 0 \leq \text{Var}(X) = \text{Var}(X - c) = \mathbb{E}((X - c)^2) - (\mathbb{E}(X - c))^2$

donc $\text{Var}(X) = \mathbb{E}((X - c)^2) \Leftrightarrow \mathbb{E}(X - c) = 0$

$\Leftrightarrow c = \mathbb{E}(X)$

(4) $\text{Var}\left(\sum_{i=1}^n X_i\right) = \text{Cov}\left(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i\right)$

$$= \sum_{i=1}^n \sum_{j=1}^n \text{Cov}(X_i, X_j) \quad (\text{bilinéarité})$$

$$= \sum_{1 \leq i, j \leq n} \text{Cov}(X_i, X_j)$$

$$= \sum_{\substack{1 \leq i, j \leq n \\ i=j}} \text{Cov}(X_i, X_i) + \sum_{\substack{1 \leq i, j \leq n \\ i \neq j}} \text{Cov}(X_i, X_j)$$

$$= \sum_{i=1}^n \text{Var}(X_i) + \sum_{\substack{1 \leq i, j \leq n \\ i \neq j}} \text{Cov}(X_i, X_j)$$

$$= \sum_{i=1}^n \text{Var}(X_i) + 2 \sum_{1 \leq i < j \leq n} \text{Cov}(X_i, X_j)$$

$$\text{car } \sum_{1 \leq i < j \leq n} \text{Cov}(X_i, X_j) = \sum_{1 \leq j < i \leq n} \text{Cov}(X_i, X_j)$$

(symétrie) $\square$

§ 3.4.4 Espérance et Indépendance

Proposition 3.46 [Vidéo sur Moodle]

Soient $X: \Omega \rightarrow \mathbb{R}$, $Y: \Omega \rightarrow \mathbb{R}$ v.a. **indép.**

- Si X et Y admettent toutes les deux une espérance finie, alors XY admet une espérance finie et $\mathbb{E}(XY) = \mathbb{E}(X) \mathbb{E}(Y)$
- Si $X \geq 0$ et $Y \geq 0$, alors $\mathbb{E}(XY) = \mathbb{E}(X) \mathbb{E}(Y)$ même sans l'hypothèse que $\mathbb{E}(X)$ et $\mathbb{E}(Y)$ soient finies.

Convention: $0 \times (+\infty) = 0$, $(+\infty) \times (+\infty) = (+\infty)$

"Preuve" $\mathbb{E}(XY) = \mathbb{E}(g(X, Y)) \quad g: (x, y) \mapsto xy$

$$= \sum_{\substack{x \in X(\Omega) \\ y \in Y(\Omega)}} g(x, y) p_{X,Y}(x, y) \quad \text{formule de transfert}$$
$$= \sum_{\substack{x \in X(\Omega) \\ y \in Y(\Omega)}} xy \, p_X(x) \, p_Y(y) \quad \text{indép}$$
$$= \left(\sum_{x \in X(\Omega)} x p_X(x) \right) \left(\sum_{y \in Y(\Omega)} y p_Y(y) \right)$$
$$= \mathbb{E}(X) \mathbb{E}(Y) \quad \square$$

Corollaire 1 Soient X, Y **indép.**, admettant une espérance finie

$$\text{Alors } \text{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X) \mathbb{E}(Y) = 0$$



$\boxed{\text{Cov}(X, Y) = 0 \not\Rightarrow X \perp Y}$ exemple: Ex 5.3 en TD

Corollaire 2 Si $X_1, \dots, X_n$ sont **indép.**, admettant un moment d'ordre 2 fini
alors $\text{Var}\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n \text{Var}(X_i)$ car $\text{Cov}(X_i, X_j) = 0$
pour tous $i \neq j$

§ 3.4.5 Espérance et variance des lois classiques (Ex 5.1 en TD)

Uniforme sur $\{1, \dots, n\}$	$\mathbb{E}(X) = \frac{n+1}{2}$	$\text{Var}(X) = \frac{n^2-1}{12}$
Bernoulli Bern(p)	$\mathbb{E}(X) = p$	$\text{Var}(X) = p(1-p)$
Binomiale Bin(n, p)	$\mathbb{E}(X) = np$	$\text{Var}(X) = np(1-p)$
Poisson Poi(λ)	$\mathbb{E}(X) = \lambda$	$\text{Var}(X) = \lambda$
Géométrique Géom(p)	$\mathbb{E}(X) = \frac{1}{p}$	$\text{Var}(X) = \frac{1-p}{p^2}$

§ 3.5 Inégalités probabilistes (vraies pour des v.a. réelles générales)

On a trois inégalités probabilistes à apprendre.

Théorème (Inégalité de Markov)

$$X: (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow \mathbb{R}^+ \text{ v.a. positive}$$

$$\text{Alors } \forall \varepsilon > 0, \quad \mathbb{P}(X \geq \varepsilon) \leq \frac{\mathbb{E}(X)}{\varepsilon}$$

Application: $\mathbb{P}(X \geq c \mathbb{E}(X)) \leq \frac{\mathbb{E}(X)}{c \mathbb{E}(X)} = \frac{1}{c}$ pour $c > 0$
 lorsque $\mathbb{E}(X) < +\infty$

Preuve Par hypothèse, $\forall \omega \in \Omega, X(\omega) \geq 0$

$$X \cdot 1_{\{X \geq \varepsilon\}} \geq \varepsilon \cdot 1_{\{X \geq \varepsilon\}} \quad (\text{car sur } \{X \geq \varepsilon\}, X(\omega) \geq \varepsilon)$$

$$X \cdot 1_{\{X < \varepsilon\}} = X \cdot 1_{\{\varepsilon > X \geq 0\}} \geq 0 \cdot 1_{\{\varepsilon > X \geq 0\}} = 0$$

$$\text{Donc } X = X \cdot 1_{\{X \geq \varepsilon\}} + X \cdot 1_{\{X < \varepsilon\}} \geq \varepsilon \cdot 1_{\{X \geq \varepsilon\}}$$

$$\text{Ainsi } \mathbb{E}(X) \geq \mathbb{E}(\varepsilon \cdot 1_{\{X \geq \varepsilon\}}) \underset{\text{linéarité}}{=} \varepsilon \mathbb{E}(1_{\{X \geq \varepsilon\}}) = \varepsilon \mathbb{P}(X \geq \varepsilon)$$

Remarque X v.a. positive, $\forall \varepsilon > 0, \mathbb{P}(X > \varepsilon) \leq \frac{\mathbb{E}(X)}{\varepsilon}$ □

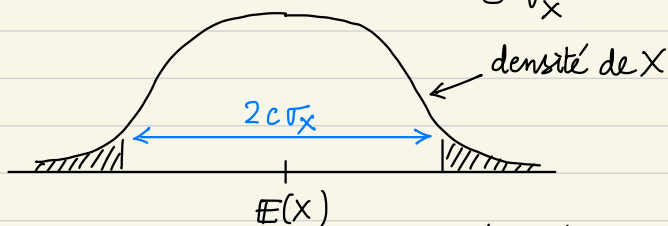
Théorème (Inégalité de Bienaymé-Tchebychev)

$X: (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow \mathbb{R}$ v.a. admettant un moment d'ordre 2 fini

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

Application $\sigma_x := \sqrt{\text{Var}(X)}$ écart-type

$$\mathbb{P}(|X - \mathbb{E}(X)| \geq c \sigma_x) \leq \frac{\text{Var}(X)}{c^2 \sigma_x^2} = \frac{1}{c^2}, \forall c > 0$$



"Typiquement, X s'écarte de son espérance $\mathbb{E}(X)$ d'au plus une constante fois σ_x "

Preuve $\{|X - \mathbb{E}(X)| \geq \varepsilon\} = \{|X - \mathbb{E}(X)|^2 \geq \varepsilon^2\}$

$$\begin{aligned} \mathbb{P}(|X - \mathbb{E}(X)| \geq \varepsilon) &= \mathbb{P}(|X - \mathbb{E}(X)|^2 \geq \varepsilon^2) \\ &\leq \frac{\mathbb{E}((X - \mathbb{E}(X))^2)}{\varepsilon^2} = \frac{\text{Var}(X)}{\varepsilon^2} \end{aligned}$$

Markov

□

Théorème 3.52 (Inégalité de Cauchy-Schwarz) [vidéo sur Moodle]

X, Y v.a. réelles ayant un moment d'ordre 2 fini

Alors XY admet une espérance finie et

$$|\mathbb{E}(XY)| \leq \mathbb{E}(X^2)^{\frac{1}{2}} \mathbb{E}(Y^2)^{\frac{1}{2}} = \sqrt{\mathbb{E}(X^2) \mathbb{E}(Y^2)}$$

Remarques (1) $\mathbb{E}(|XY|) \leq \mathbb{E}(X^2)^{\frac{1}{2}} \mathbb{E}(Y^2)^{\frac{1}{2}}$

(2) Lorsque $\mathbb{E}(Y^2) > 0$ (c-à-d. Y n'est pas p.s. égale à 0),

$$|\mathbb{E}(XY)| = \mathbb{E}(X^2)^{\frac{1}{2}} \mathbb{E}(Y^2)^{\frac{1}{2}} \iff \exists \lambda \in \mathbb{R} \text{ t.q. } X = \lambda Y \text{ p.s.}$$

$\mathbb{P}(X = \lambda Y) = 1$

(3) $(X, Y) \mapsto \mathbb{E}(XY)$ est un **produit scalaire**

bilinéaire + symétrique + définie positive

$$\mathbb{E}(X^2) = 0 \iff X = 0 \text{ p.s.}$$

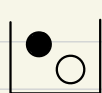
$X \mapsto \mathbb{E}(X^2)^{\frac{1}{2}}$ est une **norme** (admis)

• Méthode des moments

Pour étudier les propriétés d'une v.a., calculer son espérance et sa variance donne beaucoup d'informations.

Exemple (Exercice 5.9 en TD)

Une suite d'urnes



1^e urne

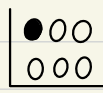


2^e urne



3^e urne

...



n^e urne

n boules blanches
et 1 boule noire

On tire successivement et indépendamment une boule uniformément au hasard de chaque urne.

Question: combien de boules noires on a tiré

$$\begin{aligned} X_n &= \text{"nombre de boules noires tirées lors de } n \text{ premiers tours"} \\ &= \sum_{k=1}^n 1_{A_k} \quad \text{où } A_k = \text{"on tire la boule noire de la } k^{\text{e}} \text{ urne"} \end{aligned}$$

$$\text{Par linéarité, } \mathbb{E}(X_n) = \sum_{k=1}^n \mathbb{E}(1_{A_k}) = \sum_{k=1}^n \mathbb{P}(A_k) = \sum_{k=1}^n \frac{1}{k+1}$$

les A_k indép $\iff$ les 1_{A_k} indép

$$\text{Var}(X_n) \overset{\substack{\uparrow \\ \text{indép}}}{=} \sum_{k=1}^n \text{Var}(1_{A_k}) = \sum_{k=1}^n \mathbb{P}(A_k)(1 - \mathbb{P}(A_k))$$

$$\begin{aligned} &= \sum_{k=1}^n \frac{1}{k+1} \left(1 - \frac{1}{k+1}\right) \\ &\leq \sum_{k=1}^n \frac{1}{k+1} = \mathbb{E}(X_n) \end{aligned}$$

Série harmonique

$$\sum_{k=1}^n \frac{1}{k+1} \leq \ln(n+1)$$

$$\sum_{k=1}^n \frac{1}{k+1} \sim \ln(n) \text{ lorsque } n \rightarrow \infty$$

$$\mathbb{E}(X_n) \sim \ln(n) \text{ et } \text{Var}(X_n) \leq \ln(n+1)$$

Par Bienaymé - Tchebychev,

$$\mathbb{P}(|X_n - \mathbb{E}(X_n)| \geq 10 \sqrt{\ln(n+1)}) \leq \frac{\text{Var}(X_n)}{100 \cdot \ln(n+1)} \leq \frac{1}{100}$$

Pour tout n (fixé), il y a au moins 99% de chances que

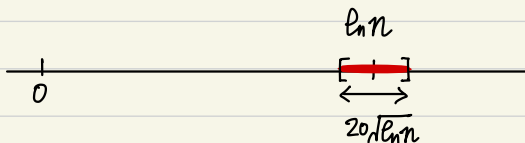
$$X_n \in [\mathbb{E}(X_n) - 10 \sqrt{\ln(n+1)}, \mathbb{E}(X_n) + 10 \sqrt{\ln(n+1)}]$$

Pour "grand" n , $\mathbb{E}(X_n) \sim \ln n$

$n \rightarrow \infty$:

$$\ln n \rightarrow +\infty$$

$$\frac{\sqrt{\ln n}}{\ln n} \rightarrow 0$$



FIN SÉANCE 7

§ 3.6 Fonction génératrice

Déf Soit X une v.a à valeurs dans $\mathbb{N}$, de densité discrète p_X .
La fonction génératrice de (la loi de) X est

$$G_X(z) := \mathbb{E}(z^X) = \sum_{n=0}^{\infty} z^n p_X(n)$$

définie $\begin{cases} \text{dans } \mathbb{R} \text{ au moins pour } z \in [-1, 1] \\ \text{dans } [0, +\infty] \text{ pour tout } z \in [0, +\infty[\end{cases}$

Convention: $0^0 = 1$ $G_X(0) = p_X(0)$

Cette définition $G_X(z) = \sum_{n=0}^{+\infty} z^n p_X(n)$ s'étend au cas $z \in \mathbb{C}$ pour tout $|z| < R$

où R est le **rayon de convergence** de la série entière $\sum_{n=0}^{+\infty} p_X(n) z^n$

Exemple $X \sim \text{Poi}(\lambda)$, $\lambda > 0$

$$p_X(n) = \frac{\lambda^n}{n!} e^{-\lambda}$$

$$\begin{aligned} G_X(z) &= \sum_{n=0}^{\infty} z^n \frac{\lambda^n}{n!} e^{-\lambda} = \left(\sum_{n=0}^{\infty} \frac{(z\lambda)^n}{n!} \right) e^{-\lambda} = e^{z\lambda} e^{-\lambda} \\ &= e^{\lambda(z-1)} \quad \text{pour tout } z \in \mathbb{C} \end{aligned}$$

Théorème Deux v.a. à valeurs dans $\mathbb{N}$ ayant la même fonction génératrice ont la même loi.

Preuve $p_X(n)$, $n \geq 0$ sont les coefficients du développement de Taylor de G_X en 0.

p_X est déterminée par G_X grâce à l'unicité de la décomposition en série entière. □

Exemple (bis)

Donc $G_X(z) = e^{\lambda(z-1)}$ implique que $X \sim \text{Poi}(\lambda)$.

Théorème $X : (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow \mathbb{N}$ v.a.

Sa fonction génératrice G_X est bien définie (au moins) sur $[-1, 1]$ et est **infiniment dérivable** (au moins) sur $] -1, 1[$.

De plus, la densité $p_X(k) = \frac{1}{k!} G_X^{(k)}(0)$ pour tout $k \in \mathbb{N}$

où $G_X^{(k)}$ est la k -ème dérivée de G_X .

(Donc la fonction génératrice **caractérise la loi**)

Preuve $\forall z \in [-1, 1], |z^n p_X(n)| \leq p_X(n)$

$$\left| \sum_{n \geq 0} z^n p_X(n) \right| \leq \sum_{n \geq 0} |z^n p_X(n)| \leq \sum_{n \geq 0} p_X(n) = 1$$

Cette série entière possède un rayon de convergence $R \geq 1$

Théorème de dérivation des séries entières implique

G_X est $\mathcal{C}^\infty$ sur $] -R, R[$.

$\forall k \geq 1, -1 < z < 1,$

$$\begin{aligned} G_X^{(k)}(z) &= \sum_{n=k}^{\infty} n(n-1) \cdots (n-k+1) z^{n-k} p_X(n) \\ &= \sum_{n=k}^{\infty} \frac{n!}{(n-k)!} z^{n-k} p_X(n) \end{aligned}$$

En prenant $z=0$, $G_X^{(k)}(0) = \frac{k!}{0!} 0^0 p_X(k) = k! p_X(k)$ □

Prop Si X et Y sont des v.a. indép à valeurs dans $\mathbb{N}$

Alors $G_{X+Y}(z) = G_X(z) G_Y(z)$

pour tout $z \geq 0$ et pour tout $z \in \mathbb{R}$ t.q. $G_X(z)$ et $G_Y(z)$ soient finies

Preuve $G_{X+Y}(z) = \mathbb{E}(z^{X+Y}) = \mathbb{E}(z^X \cdot z^Y)$

$X \perp\!\!\!\perp Y \Rightarrow z^X \perp\!\!\!\perp z^Y$ (indép par transformation)

Donc $G_{X+Y}(z) = \mathbb{E}(z^X) \mathbb{E}(z^Y) = G_X(z) G_Y(z)$ □

Application $X \sim \text{Poi}(\lambda), Y \sim \text{Poi}(\mu) \quad \lambda, \mu > 0$

$$\begin{aligned} \text{Si } X \perp\!\!\!\perp Y, G_{X+Y}(z) &= G_X(z) G_Y(z) = e^{\lambda(z-1)} e^{\mu(z-1)} \\ &= e^{(\lambda+\mu)(z-1)} \quad \forall z \in \mathbb{R} \text{ (ou } \mathbb{C}) \end{aligned}$$

fonction génératrice d'une v.a. de loi $\text{Poi}(\lambda+\mu)$

Comme la fonction génératrice caractérise la loi,

$$X + Y \sim \text{Poi}(\lambda + \mu) \quad (\text{sous l'hypothèse } X \perp\!\!\!\perp Y)$$

Rappel
 $\forall z \in]-R, R[, R \geq 1$

$$G_X^{(k)}(z) = \sum_{n=k}^{+\infty} n(n-1) \dots (n-k+1) z^{n-k} p_X(n)$$
$$= \sum_{n=0}^{+\infty} n(n-1) \dots (n-k+1) z^{n-k} p_X(n)$$

$$\mathbb{E}(X(X-1) \dots (X-k+1)) = \sum_{n=0}^{+\infty} n(n-1) \dots (n-k+1) p_X(n)$$

Prop 3.61 [Vidéo sur Moodle]

$$X: (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow \mathbb{N} \text{ v.a.}$$

- $\forall k \geq 1$, $\lim_{z \uparrow 1} G_X^{(k)}(z) = \mathbb{E}(X(X-1) \dots (X-k+1)) \in [0, +\infty]$.

- X admet un moment d'ordre k fini
 $\Leftrightarrow G_X$ est k fois dérivable à gauche en 1, c'est-à-dire

$$G_X^{(k)}(1^-) := \lim_{z \uparrow 1} G_X^{(k)}(z) < +\infty.$$

Exercice 6.1 en TD: fonctions génératrices des lois classiques
Bern(p), Bin(n,p), Géom(p), Poi(λ)

Exercice 6.2 en TD: Appliquer Prop 3.61 pour calculer les moments de Géom(p) et Poi(λ).

$$\mathbb{E}(X) = G_X'(1)$$

$$\mathbb{E}(X(X-1)) = G_X''(1)$$

$$\text{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \mathbb{E}(X(X-1)) + \mathbb{E}(X) - \mathbb{E}(X)^2$$
$$= G_X''(1) + G_X'(1) - G_X'(1)^2$$

(à faire dans la semaine suivant le partiel)

Chapitre IV Variables aléatoires réelles générales

On veut tirer un nombre au hasard dans $[0,1]$ (non dénombrable)
Le nombre obtenu X à valeurs dans $[0,1]$ n'est plus une v.a. discrète.

§4.1 Définition et Propriétés

Déf On appelle v.a. **réelle** sur un espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$ toute fonction $X: \Omega \rightarrow \mathbb{R}$ t.q. pour tout **intervalle** $I \subset \mathbb{R}$ l'ensemble $\{\omega \in \Omega: X(\omega) \in I\}$ soit un événement de $\mathcal{F}$
[$X(\Omega)$ n'est plus supposé dénombrable.]

Notation $\{X \in I\} = X^{-1}(I) = \{\omega: X(\omega) \in I\}$
 $\{X \leq t\} = \{\omega: X(\omega) \leq t\}$

Lorsque X est une v.a. réelle, on peut parler de $\mathbb{P}(X \in I)$

Remarque 1 X v.a. réelle $\Leftrightarrow \forall t \in \mathbb{R}, \{X \leq t\} \in \mathcal{F}$ (admise)

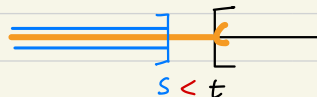
Idée: chaque intervalle I s'exprime à l'aide des demi-droite $] -\infty, t]$ avec des passages au complémentaire, union et/ou intersection dénombrable

$$]t, +\infty[=]-\infty, t]^c$$

$$]-\infty, t[= \bigcup_{\substack{s \in \mathbb{Q} \\ s < t}}]-\infty, s]$$

$$= \bigcup_{n \geq 1}]-\infty, s_n]$$

$$[s, t] = [s, +\infty[\cap]-\infty, t]$$



$$s_n < s_{n+1} < t, \forall n \geq 1$$
$$\lim_{n \rightarrow \infty} s_n = t$$

Remarque 2 Déf: $X: \Omega \rightarrow \mathbb{R}^n$ est un **vecteur aléatoire**
 $\Leftrightarrow$ Pour tout n -uplets d'**intervalles** $I_1, \dots, I_n$
 $\{X \in I_1 \times \dots \times I_n\} \in \mathcal{F}$

Remarque 3 Une v.a. discrète à valeurs dans $\mathbb{R}$ est une v.a. réelle
(l'inverse n'est pas vrai!)

Si $X(\Omega) \subset \mathbb{R}$ est dénombrable et $\forall x \in X(\Omega) \{X=x\} \in \mathcal{F}$,
alors $\{X \in I\} = \bigcup_{x \in X(\Omega) \cap I} \{X=x\} \in \mathcal{F}$

$\nwarrow$ union dénombrable

Remarque 4 Pour une v.a. réelle X , $\{X \in A\} \in \mathcal{F}$ est vraie pour une large
classe d'ensembles $A \subset \mathbb{R}$, appelé **ensemble borélien**
Un intervalle est un ensemble borélien

Déf (Tribu borélienne)

$\mathcal{B}(\mathbb{R}) = \{\text{ensembles boréliens dans } \mathbb{R}\}$

$=$ la tribu **engendrée** par les **intervalles** de $\mathbb{R}$

$=$ la **plus petite** tribu contenant les intervalles de $\mathbb{R}$

Il existe une partie de $\mathbb{R}$ qui n'est pas un ensemble borélien!

$\mathcal{B}(\mathbb{R}) \subsetneq \mathcal{P}(\mathbb{R})$

En conséquence, pour une fonction arbitraire $g: \mathbb{R} \rightarrow \mathbb{R}$,

$g(X)$ n'est pas forcément une v.a. réelle. 

Bonne nouvelle: pour toutes les fonctions g rencontrées dans
ce cours (et dans la vie courante), $g(X)$ reste une v.a. réelle.

On ne se posera pas ce genre de questions dans la suite.

(à traiter en théorie des mesures en L3)

Déf (Loi d'une v.a. réelle)

On appelle la loi de $X: (\Omega, \mathcal{F}, \mathbb{P}) \rightarrow \mathbb{R}$ la probabilité sur $\mathbb{R}$

$$\mu_X: \mathcal{B}(\mathbb{R}) \rightarrow [0, 1]$$

$$B \mapsto \mu_X(B) := \mathbb{P}(X \in B)$$

Rappel

Si X v.a. **discrète**, sa loi μ_X est caractérisée par la densité discrète p_X :

$$\mu_X(B) = \sum_{x \in B \cap X(\Omega)} p_X(x) = \sum_{x \in B \cap X(\Omega)} \mathbb{P}(X=x)$$



Mais ce n'est pas vrai en général:

Pour X v.a. **réelle**, sa loi μ_X est caractérisée par $\mathbb{P}(X \in]-\infty, t])$, $t \in \mathbb{R}$ (Admis)

Idee: les boréliens de $\mathbb{R}$ sont "engendrés" par les intervalles du type $]-\infty, t]$ (à voir en L3)

Déf (Indépendance)

Soient $X_1, \dots, X_n$ des v.a. réelles définies sur $(\Omega, \mathcal{F}, \mathbb{P})$

Elles sont dites indépendantes si pour tous intervalles (de $\mathbb{R}$)

$I_1, \dots, I_n$, on a

$$\mathbb{P}(X_1 \in I_1, \dots, X_n \in I_n) = \prod_{i=1}^n \mathbb{P}(X_i \in I_i)$$

Pour une famille arbitraire de v.a. réelles, elles sont dites indép si toute sous-famille finie est composée de v.a. indép.

Soit $X: \Omega \rightarrow \mathbb{R}^+$ v.a. $\forall n \geq 1$, on définit la v.a. discrète

$$X_n = 2^{-n} \lfloor 2^n X \rfloor \text{ où } \lfloor x \rfloor = \text{partie entière de } x \in \mathbb{R}$$

- $0 \leq X_n \leq X \leq X_n + 2^{-n}$, $\forall n \geq 1$
- $X_n \leq X_{n+1}$ car $\lfloor 2x \rfloor \geq 2\lfloor x \rfloor$, $\forall x \geq 0$

donc $\forall \omega \in \Omega$, $\lim_{n \rightarrow \infty} \uparrow X_n(\omega) = X(\omega)$

Déf (Espérance)

1) Si X est une v.a. réelle **positive**, on définit l'espérance de X comme

$$\mathbb{E}(X) := \lim_{n \rightarrow \infty} \uparrow \mathbb{E}(X_n) \in [0, +\infty]$$

où $X_n = 2^{-n} \lfloor 2^n X \rfloor$ pour tout $n \geq 1$.

[Admis: Cette définition est cohérente avec celle quand X est discrète]

2) Si X est une v.a. réelle **générale**, on dit que X **admet une espérance** si au moins une des deux espérances $\mathbb{E}(X^+)$ et $\mathbb{E}(X^-)$ est finie.

Si c'est le cas, on définit

$$\mathbb{E}(X) := \mathbb{E}(X^+) - \mathbb{E}(X^-) \in [-\infty, +\infty]$$

Par conséquent,

X admet une espérance finie $\Leftrightarrow \mathbb{E}(X^+) < +\infty$ et $\mathbb{E}(X^-) < +\infty$

$$\Leftrightarrow \mathbb{E}[|X|] = \mathbb{E}(X^+) + \mathbb{E}(X^-) < +\infty$$

(par linéarité)



$\mathbb{E}(X) = \sum_{\omega \in \Omega} X(\omega) \mathbb{P}(\{\omega\})$ n'est plus vrai sauf Ω dénombrable

À remplacer par une intégrale $\mathbb{E}(X) = \int_{\Omega} X(\omega) \mathbb{P}(d\omega)$
(à voir en L3)

Propriétés (admis)

- Monotonie: Si $X \leq Y$ alors $\mathbb{E}(X) \leq \mathbb{E}(Y)$ pourvu que X et Y admettent une espérance
- Si X admet une espérance, alors $|\mathbb{E}(X)| \leq \mathbb{E}(|X|)$
- Linéarité: $\mathbb{E}(aX + bY) = a\mathbb{E}(X) + b\mathbb{E}(Y)$ si $X, Y \geq 0, a, b \geq 0$
ou si X et Y admettent une espérance finie, $a, b \in \mathbb{R}$
- Si $X \perp Y$, $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$ si X et Y admettent une espérance finie ou si $X, Y \geq 0$

Covariance, variance, moments : même définition

Inégalité de Markov, Bienaymé-Tchebychev, Cauchy-Schwarz : idem

§ 4.2 Fonction de répartition

Déf X v.a. réelle. On appelle *fonction de répartition* de X la fonction $F_X : \mathbb{R} \rightarrow [0, 1]$

$$t \mapsto F_X(t) := \mathbb{P}(X \leq t) = \mathbb{P}(X \in]-\infty, t]) \\ = \mu_X(]-\infty, t])$$

Exercice 4.11 [Vidéo sur Moodle]

Soient X, Y v.a. réelles *indép*, $Z = \max\{X, Y\}$ v.a. réelle

On a $F_Z(t) = F_X(t) \cdot F_Y(t)$, $\forall t \in \mathbb{R}$

Théorème (Admis)

Deux v.a. réelles X et Y ayant la même fonction de répartition

$F_X = F_Y$ ont la même loi $\mu_X = \mu_Y$.

Donc,

$$F_X = F_Y \Leftrightarrow \mu_X = \mu_Y$$

Prop 4.13 [Vidéo sur Moodle]

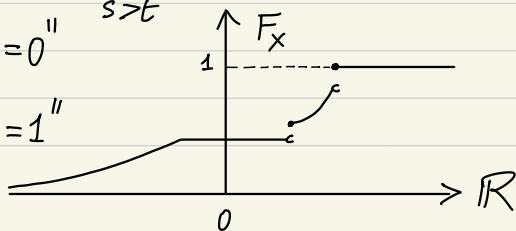
La fonction de répartition F_X a les propriétés suivantes :

- 1) F_X est *croissante*
- 2) F_X est *continue à droite* :

$$\forall t \in \mathbb{R} \quad F_X(t) = F_X(t^+) := \lim_{\substack{s \rightarrow t \\ s > t}} F_X(s) = \lim_{\substack{s \downarrow t}} F_X(s)$$

$$3) \lim_{t \rightarrow -\infty} F_X(t) = 0 \quad "F_X(-\infty) = 0"$$

$$4) \lim_{t \rightarrow +\infty} F_X(t) = 1 \quad "F_X(+\infty) = 1"$$



Remarque On verra dans Théorème 5.18 que toute fonction $\mathbb{R} \rightarrow \mathbb{R}$ qui vérifie ces propriétés est la fonction de répartition d'une v.a. réelle.

Prop On note $F_X(x^-) := \lim_{y \rightarrow x^-} F_X(y) = \lim_{y \uparrow x} F_X(y)$ la limite à gauche de F_X au point $x \in \mathbb{R}$, qui existe grâce à la croissance de F_X .

Alors $\forall x \in \mathbb{R}, \quad \mathbb{P}(X=x) = F_X(x) - F_X(x^-)$

$$F_X \text{ est discontinue en } x \Leftrightarrow \mathbb{P}(X=x) > 0$$

Preuve Si $y < x$, $\{X \in]y, x]\} = \{X \leq x\} \setminus \{X \leq y\}$

$$\mathbb{P}(X \in]y, x]) = \mathbb{P}(X \leq x) - \mathbb{P}(X \leq y) = F_X(x) - F_X(y)$$

$\{X \in]x - \frac{1}{n}, x]\}, n \geq 1$ est une suite décroissante d'événements

$$\bigcap_{n \geq 1}]x - \frac{1}{n}, x] = \{x\}$$

$$\bigcap_{n \geq 1} \{X \in]x - \frac{1}{n}, x]\} = \{X = x\}$$

$$\begin{aligned} \text{Par continuité de } \mathbb{P}, \mathbb{P}(X=x) &= \lim_{n \rightarrow \infty} \mathbb{P}(X \in]x - \frac{1}{n}, x]) \\ &= \lim_{n \rightarrow \infty} F_X(x) - F_X(x - \frac{1}{n}) \\ &= F_X(x) - \lim_{n \rightarrow \infty} F_X(x - \frac{1}{n}) \\ &= F_X(x) - F_X(x^-) \end{aligned}$$

Exercise:

$$\forall a \in \mathbb{R}, \mathbb{P}(X > a) = 1 - F_X(a), \mathbb{P}(X < a) = F_X(a^-), \mathbb{P}(X \geq a) = 1 - F_X(a^-)$$

$$\forall a \leq b, \mathbb{P}(X \in]a, b]) = F_X(b) - F_X(a), \mathbb{P}(X \in [a, b]) = F_X(b) - F_X(a^-)$$

$$\mathbb{P}(X \in]a, b[) = F_X(b^-) - F_X(a), \quad \mathbb{P}(X \in [a, b[) = F_X(b^-) - F_X(a^-)$$

FIN SÉANCE 8

Rappel: X v.a. réelle

Fonction de répartition $F_X: \mathbb{R} \rightarrow [0,1]$

$$t \mapsto F_X(t) = \mathbb{P}(X \leq t)$$

Exemple de calcul de fonction de répartition:

Expérience aléatoire: choisir un nombre ω "au hasard" dans $[0,1]$

$$\Omega = [0,1]$$

$\mathcal{F} = \mathcal{B}([0,1])$ = tribu engendrée par les intervalles $[a,b] \subset [0,1]$

$\mathbb{P}$ = "proba uniforme"

= mesure de Lebesgue sur $[0,1]$

$$(*) \quad \mathbb{P}([a,b]) = b-a, \quad \forall 0 < a < b < 1$$

$\mathbb{P}$ est la seule probabilité définie sur $\mathcal{B}([0,1])$ qui vérifie (*)

Impossible d'étendre (en restant une proba.) cette définition de $\mathbb{P}$ à $\mathcal{P}([0,1])$ ($\not\supseteq \mathcal{B}([0,1])$)

On construit un carré de côté 2ω et

on renvoie le maximum entre l'aire du carré et 1.

v.a. réelle $X: \omega \in [0,1] \mapsto X(\omega) = \max(4\omega^2, 1) \in [1,4]$

$$\text{Si } t < 1, \quad \{X \leq t\} = \emptyset \Rightarrow F_X(t) = \mathbb{P}(\emptyset) = 0$$

$$\text{Si } t > 4, \quad \{X \leq t\} = \Omega \Rightarrow F_X(t) = \mathbb{P}(\Omega) = 1$$

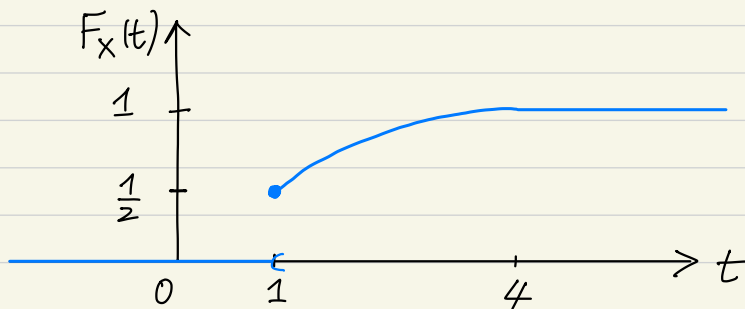
$$\text{Si } 1 \leq t \leq 4, \quad \{X \leq t\} = \{4\omega^2 \leq t\}$$

$$= \left\{ |\omega| \leq \sqrt{\frac{t}{4}} \right\}$$

$$= \left\{ \omega \leq \frac{1}{2}\sqrt{t} \right\} \quad \text{car } \omega \in [0,1]$$

$$= [0, \frac{1}{2}\sqrt{t}]$$

$$F_X(t) = \mathbb{P}\left([0, \frac{1}{2}\sqrt{t}]\right) = \frac{1}{2}\sqrt{t}$$



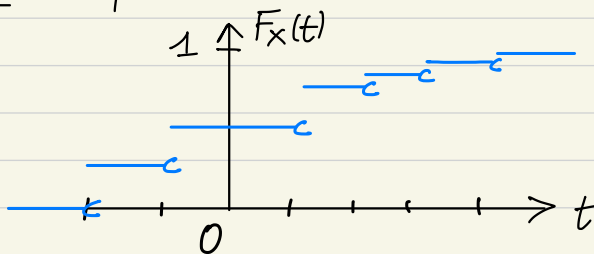
$$\mathbb{P}(X < 1) = F_X(1^-) = 0$$

$$\mathbb{P}(X = 1) = F_X(1) - F_X(1^-) = \frac{1}{2} - 0 = \frac{1}{2}$$

= "taille du saut de F_X à 1"

$$F_X(t) - F_X(t^-) = \mathbb{P}(X = t), \quad \forall t \in \mathbb{R}$$

Remarque 1: Lorsque X est une v.a. discrète:



$$\{\text{pts de discontinuité de } F_X\} = X(\Omega)$$

la taille des sauts de F_X est donnée par la densité discrète p_X

Remarque 2: Si F_X est continue, alors $\forall t \in \mathbb{R}, \mathbb{P}(X = t) = 0$

§4.3 Fonction génératrice des moments

Définition: X v.a. réelle. On appelle *fonction génératrice des moments* de X la fonction

$$M_X: \mathbb{R} \longrightarrow \mathbb{R}_+ \cup \{+\infty\}$$

$$M_X(t) = \mathbb{E}(e^{tX}) \quad \text{bien définie}$$

$\uparrow$ v.a. positive

Remarque 1: Si $X: \Omega \rightarrow \mathbb{N}$ v.a., $M_X(t) = G_X(e^t)$

où $G_X(s) = \mathbb{E}(s^X)$ est la fonction génératrice

Remarque 2: En général, M_X ne caractérise pas la loi de v.a. réelle X 

Théorème 4.16 [Vidéo sur Moodle]

Si l'existe $a > 0$ t.q. $M_X(t) < +\infty$ pour tout $t \in]-a, a[$, alors

1) X admet un moment d'ordre k fini pour tout $k \geq 1$ et pour tout $t \in]-a, a[$,

$$M_X(t) = \sum_{k=0}^{+\infty} \mathbb{E}(X^k) \frac{t^k}{k!}$$

2) M_X est infiniment dérivable sur $]-a, a[$

La k^{e} dérivée de M_X , notée $M_X^{(k)}$, vérifie que

$$M_X^{(k)}(t) = \mathbb{E}(X^k e^{tX}), \quad \forall t \in]-a, a[$$

En particulier, $M_X^{(k)}(0) = \mathbb{E}(X^k)$, $\forall k \geq 1$

"Preuve": 1) $M_X(t) = \mathbb{E}(e^{tX}) = \mathbb{E}\left(\sum_{k=0}^{+\infty} \frac{t^k X^k}{k!}\right)$

$$\stackrel{\substack{\text{!} \\ \uparrow \\ \text{à justifier}}}{=} \sum_{k=0}^{+\infty} \mathbb{E}\left(\frac{t^k X^k}{k!}\right) = \sum_{k=0}^{+\infty} \mathbb{E}(X^k) \frac{t^k}{k!}$$

1) $\Rightarrow$ 2) : par le théorème de dérivation des séries entières $\square$

Comparaison:

Fonction génératrice $X: \Omega \rightarrow \mathbb{N}$

$$G_X(z) = \mathbb{E}(z^X)$$

$$G_X^{(k)}(1) = \mathbb{E}(X(X-1) \cdots (X-k+1))$$

$$G_X(z) = \sum_{k=0}^{+\infty} \mathbb{P}(X=k) z^k$$

$$\frac{1}{k!} G_X^{(k)}(0) = \mathbb{P}(X=k)$$

$$G_{X+Y}(z) = G_X(z) G_Y(z)$$

si $X \perp\!\!\!\perp Y$

(pour $z \geq 0$ ou $z \in \mathbb{R} \text{ t.q.}$
 $G_X(z)$ et $G_Y(z)$ finies)

Fonction génératrice des moments $X: \Omega \rightarrow \mathbb{R}$

$$M_X(t) = \mathbb{E}(e^{tx})$$

$$M_X^{(k)}(0) = \mathbb{E}(X^k)$$

$$M_X(t) = \sum_{k=0}^{+\infty} \mathbb{E}(X^k) \frac{t^k}{k!}$$

(à condition que M_X soit
finie sur un voisinage de 0)

$$M_{X+Y}(t) = M_X(t) M_Y(t)$$

si $X \perp\!\!\!\perp Y$

Chapitre V Variables aléatoires à densité

§ 5.1 v.a. réelles à densité (absolument continues)

Définition: Une v.a. réelle X est dite **abs. continue** ou **à densité** s'il existe une fonction positive $f_X: \mathbb{R} \rightarrow [0, +\infty[$, Riemann intégrable dans $\mathbb{R}$, telle que $\forall t \in \mathbb{R}$,

$$F_X(t) = \mathbb{P}(X \leq t) = \int_{-\infty}^t f_X(x) dx$$

(intégrale de Riemann généralisée)
 Une telle fonction f_X est appelée **densité de X**

Remarque 1: $\lim_{t \rightarrow +\infty} F_X(t) = \mathbb{P}(X < +\infty) = 1 \Rightarrow \int_{-\infty}^{+\infty} f_X(x) dx = 1$

Remarque 2: Toute fonction positive $f: \mathbb{R} \rightarrow [0, +\infty[$ Riemann-intégrable sur $\mathbb{R}$ et t.q. $\int_{-\infty}^{+\infty} f_X(x) dx = 1$ est la densité d'une v.a. abs. continue.

(puisque $F(t) = \int_{-\infty}^t f(x) dx$ possède les propriétés d'une fonction de répartition)

Remarque 3 La densité d'une v.a. abs. continue n'est pas unique

Si f_X est une densité de X , pour toute fonction g obtenue à partir de f_X en modifiant sa valeur sur un ensemble **fini** de pts, $\int_{-\infty}^t g(x) dx = \int_{-\infty}^t f_X(x) dx$,

donc g est aussi une densité de X .

En pratique, on dira "la" densité de X , tout en gardant en tête qu'on peut changer sur quelques pts.

Proposition Si X est une v.a. abs. continue, F_X est continue et $\forall x \in \mathbb{R}, \mathbb{P}(X=x) = 0$.

Pour tous $a \leq b$,

$$\begin{aligned} \mathbb{P}(X \in [a, b]) &= \mathbb{P}(X \in [a, b[) = \mathbb{P}(X \in]a, b]) = \mathbb{P}(X \in]a, b[) \\ &= F_X(b) - F_X(a) = \int_a^b f_X(x) dx \end{aligned}$$

Preuve: $x \mapsto \int_{-\infty}^x f_X(t) dt$ est continue

car f_X bornée sur un intervalle compact, $\lim_{\varepsilon \rightarrow 0^+} \int_x^{x+\varepsilon} f_X(t) dt = 0$

$$\begin{aligned}
 \forall a \leq b, \mathbb{P}(X \in]a, b]) &= F_X(b) - F_X(a) \\
 &= \int_{-\infty}^b f_X(t) dt - \int_{-\infty}^a f_X(t) dt \\
 &= \int_a^b f_X(t) dt
 \end{aligned}$$

$$\begin{aligned}
 \mathbb{P}(X \in]a, b]) &= \mathbb{P}(X \in [a, b]) = \mathbb{P}(X \in]a, b[) = \mathbb{P}(X \in [a, b[) \\
 \text{car } \forall x \in \mathbb{R}, \mathbb{P}(X=x) &= F_X(x) - F_X(x^-) = 0. \quad \square
 \end{aligned}$$

Remarque 1: Si $f_X(x) = 0$ pour tout $x \in I$ intervalle, alors $\mathbb{P}(X \in I) = 0$

Remarque 2  $f_X(x) \neq \mathbb{P}(X=x) \leftarrow$ toujours nulle

Proposition Si F_X est continue et $\mathcal{C}^1$ par morceaux sur $\mathbb{R}$ (la restriction de F_X à tout intervalle $[a, b]$ compact est $\mathcal{C}^1$ par morceaux.), alors X est abs. continue, de densité $f_X(x) = F_X'(x)$ (f_X définie arbitrairement aux pts où F_X n'est pas dérivable).

Preuve: $\forall a < b$, F_X est continue et $\mathcal{C}^1$ par morceaux

$$\Rightarrow F_X(b) - F_X(a) = \int_a^b F_X'(x) dx$$

Comme $\lim_{a \rightarrow -\infty} F_X(a) = 0$, on trouve $F_X(b) = \int_{-\infty}^b F_X'(x) dx \quad \square$

Remarque Par le thm de différentiation de Lebesgue,

X a une densité $f_X \Leftrightarrow F_X$ est abs. continue

$\forall \varepsilon > 0, \exists \delta > 0$ t.q. pour toute suite finie $([a_n, b_n], n \leq N)$ d'intervalles d'intérieurs *disjoints*,

$$\sum_{n \leq N} (b_n - a_n) < \delta \Rightarrow \sum_{n \leq N} |F_X(b_n) - F_X(a_n)| < \varepsilon$$

Exemple (Loi uniforme continue sur $[0,1]$)

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

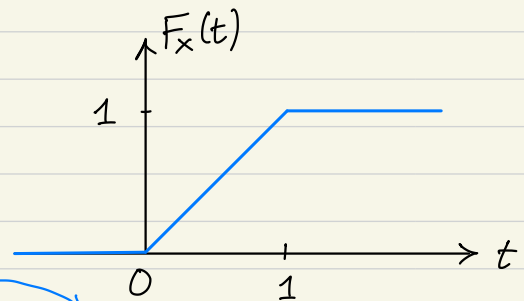
$\mathbb{P}$ probabilité t.q. $\forall 0 \leq a \leq b \leq 1$, $\mathbb{P}([a,b]) = b-a$
Une telle proba existe (à voir en L3).

Elle est *la mesure de Lebesgue* sur $[0,1]$

$$X: \omega \in \Omega \mapsto X(\omega) = \omega \quad \text{identité}$$

$$F_X(t) = \begin{cases} 0 & \text{si } t < 0 \\ \mathbb{P}(X \in [0,t]) = \mathbb{P}([0,t]) = t & \text{si } t \in [0,1] \\ 1 & \text{si } t > 1 \end{cases}$$

$$F_X(t) = \begin{cases} 0 & \text{si } t < 0 \\ 1 & \text{si } t \in]0,1[\\ 0 & \text{si } t > 1 \end{cases}$$



$$F_X(t) = \int_{-\infty}^t \underbrace{1_{[0,1]}(x)}_{\text{densité}} dx = \int_{-\infty}^t 1_{]0,1[}(x) dx$$

On dit que X est une *v.a. uniforme sur $[0,1]$* $X \sim \mathcal{U}(0,1)$

"choisir un réel au hasard dans $[0,1]$ "

Exemple 5.4 transformation linéaire de X (à densité générale)
[Vidéo sur Moodle]

Exemple 5.6 $Y = X^2$, $X \sim \mathcal{U}(0,1)$ est abs continue
[Vidéo sur Moodle] trouver sa densité

Si on pose $Z = X \cdot 1_{\{X \leq \frac{1}{2}\}} \in [0, \frac{1}{2}]$ où $X \sim \mathcal{U}(0,1)$

$$F_Z(t) = \begin{cases} \mathbb{P}(\emptyset) = 0 & \text{si } t < 0 \\ \mathbb{P}(X 1_{\{X \leq \frac{1}{2}\}} \leq 0) = \mathbb{P}(X=0) + \mathbb{P}(X > \frac{1}{2}) = 0 + \int_{\frac{1}{2}}^1 1 dx = \frac{1}{2} & \text{si } t = 0 \\ \mathbb{P}(X 1_{\{X \leq \frac{1}{2}\}} \leq t) = \mathbb{P}(X \leq t) + \mathbb{P}(X > \frac{1}{2}) = \int_0^t 1 dx + \int_{\frac{1}{2}}^1 1 dx = t + \frac{1}{2} & \text{si } 0 < t \leq \frac{1}{2} \\ \mathbb{P}(\Omega) = 1 & \text{si } t > \frac{1}{2} \end{cases}$$

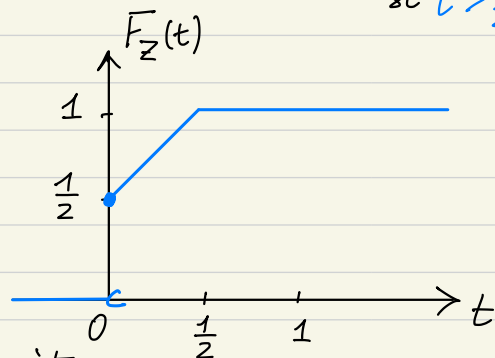
F_Z pas continue

$$\mathbb{P}(Z=0) = \frac{1}{2}$$

Donc Z n'est pas abs. continue

$Z(\Omega) = [0, \frac{1}{2}]$ non dénombrable

Z n'est pas non plus une v.a. discrète.



FAUX: une v.a. réelle est soit discrète, soit abs continue.

v.a. réelle générale ni discrète, ni abs. continue

v.a. discrète

v.a. abs. continue
(à densité)

Rappel: v.a. réelle X abs. continue ou à densité
 $\exists f_X: \mathbb{R} \rightarrow [0, +\infty[$ Riemann intégrable sur $\mathbb{R}$
 $F_X(t) = \mathbb{P}(X \leq t) = \int_{-\infty}^t f_X(x) dx$

• L'espérance de v.a. à densité

Théorème (Formule de transfert, cas positif)

X v.a. de densité f_X

$g: \mathbb{R} \rightarrow [0, +\infty[$ fonction positive et continue par morceaux

$g(x)$ v.a. positive dont l'espérance, bien définie dans $[0, +\infty]$,
 est donnée par

$$\mathbb{E}(g(X)) = \int_{-\infty}^{+\infty} g(x) f_X(x) dx$$

Idee de preuve quand X est positive:

Rappel: $\mathbb{E}(X) := \lim_{n \rightarrow \infty} \uparrow \mathbb{E}(X_n) \in [0, +\infty]$

où $X_n = 2^{-n} \lfloor 2^n X \rfloor$, $n \geq 1$ suite de v.a. discrètes
 qui converge vers X en croissant.

$$\begin{aligned} \forall k \in \mathbb{N}, \mathbb{P}(X_n = 2^{-n} k) &= \mathbb{P}(\lfloor 2^n X \rfloor = k) \\ &= \mathbb{P}(X \in [k 2^{-n}, (k+1) 2^{-n}[) \\ &= \int_{k 2^{-n}}^{(k+1) 2^{-n}} f_X(x) dx \end{aligned}$$

$$\begin{aligned} \mathbb{E}(g(X_n)) &= \sum_{k \in \mathbb{N}} g(k 2^{-n}) \mathbb{P}(X_n = k 2^{-n}) \\ &= \sum_{k \in \mathbb{N}} \int_{k 2^{-n}}^{(k+1) 2^{-n}} g(k 2^{-n}) f_X(x) dx \quad \text{somme de Riemann} \\ &\xrightarrow{n \rightarrow \infty} \int_0^{+\infty} g(x) f_X(x) dx = \int_{-\infty}^{+\infty} g(x) f_X(x) dx \quad \begin{array}{l} \text{car } X \geq 0 \\ f_X = 0 \text{ sur } \mathbb{R}^- \end{array} \end{aligned}$$

Exemple : Fonction génératrice des moments

$$M_X(t) = \mathbb{E}(e^{tx}) = \int_{-\infty}^{+\infty} e^{tx} f_X(x) dx$$

Si $g: \mathbb{R} \rightarrow \mathbb{R}$ de signe quelconque, on applique cette formule de transfert aux g^+ , g^- et $|g|$, et on obtient :

Proposition (Formule de transfert, **cas intégrable**)

Soit $g: \mathbb{R} \rightarrow \mathbb{R}$ fonction continue par morceaux

X v.a. de densité f_X

La v.a. $g(X)$ admet une espérance finie si et seulement si

$|g(x)| f_X(x)$ est intégrable sur $\mathbb{R}$

$$\mathbb{E}(|g(X)|) < +\infty \iff \int_{-\infty}^{+\infty} |g(x)| f_X(x) dx < +\infty$$

Dans ce cas, $\mathbb{E}(g(X)) = \int_{-\infty}^{+\infty} g(x) f_X(x) dx$

Exemple 5.11 Une v.a. réelle X **de Cauchy** n'admet pas d'espérance
 X de Cauchy si elle a pour densité

$$f_X(x) = \frac{1}{\pi} \frac{1}{1+x^2}$$

$$\mathbb{E}(X^+) = \int_{-\infty}^{+\infty} x^+ f_X(x) dx = \int_0^{+\infty} x f_X(x) dx = \int_0^{+\infty} \frac{x}{\pi(1+x^2)} dx = +\infty$$

car $\frac{x}{1+x^2} \sim \frac{1}{x}$ et $\forall M > 0, \int_M^{+\infty} \frac{1}{x} dx = +\infty$

De même $\mathbb{E}(X^-) = \int_{-\infty}^{+\infty} x^- f_X(x) dx = \int_{-\infty}^0 \frac{-x}{\pi(1+x^2)} dx = +\infty$

Donc $\mathbb{E}(X)$ n'existe pas, même si $\int_{-M}^M x f_X(x) dx = 0, \forall M > 0$



car $f_X(x) = f_X(-x)$

Corollaire : X admet un moment d'ordre k fini

$$\Leftrightarrow \mathbb{E}(|X|^k) = \int_{-\infty}^{+\infty} |x|^k f_X(x) dx < +\infty$$

dans ce cas, $\mathbb{E}(X^k) = \int_{-\infty}^{+\infty} x^k f_X(x) dx$

Comparaison : Cas abs. continue

$$\mathbb{E}(X) = \int_{-\infty}^{+\infty} x f_X(x) dx$$

Cas discret

$$\mathbb{E}(X) = \sum_{x \in X(\Omega)} x P_X(x)$$

$$\mathbb{E}(g(X)) = \int_{-\infty}^{+\infty} g(x) f_X(x) dx$$

$$\mathbb{E}(g(X)) = \sum_{x \in X(\Omega)} g(x) P_X(x)$$

La "réciproque" de la formule de transfert : utile pour calculer la densité
 X abs. continue et admet pour densité f_X

$\Leftrightarrow \forall h: \mathbb{R} \rightarrow \mathbb{R}$ continue par morceaux et bornée, (Remarque 5.14)

$$\mathbb{E}(h(X)) = \int_{-\infty}^{+\infty} h(x) f_X(x) dx$$

[Vidéo sur Moodle]

$\Leftrightarrow \forall h: \mathbb{R} \rightarrow \mathbb{R}$ continue et bornée,

(Remarque 5.15)

$$\mathbb{E}(h(X)) = \int_{-\infty}^{+\infty} h(x) f_X(x) dx$$

Application : Exemple 5.16 (déjà étudié dans Exemple 5.6)

$$X \sim \mathcal{U}(0,1), \quad Y = X^2$$

$$\forall h \in \mathcal{C}_b \quad \mathbb{E}(h(Y)) = \mathbb{E}(h(X^2))$$

"fonction test"

$$= \int_{-\infty}^{+\infty} h(x^2) \mathbb{1}_{]0,1[}(x) dx$$

$$= \int_0^1 h(x^2) dx = \int_0^1 h(y) \frac{1}{2\sqrt{y}} dy$$

$$= \int_{-\infty}^{+\infty} h(y) \frac{1}{2\sqrt{y}} \mathbb{1}_{]0,1[}(y) dy$$

Par la remarque, on obtient que Y est abs. continue et admet pour densité $\frac{1}{2\sqrt{y}} \mathbb{1}_{]0,1[}(y)$.

§5.2 Exemples importants de v.a. à densité

1) Loi uniforme continue

$$a, b \in \mathbb{R} \quad a < b$$

V.a. réelle abs. continue X est dite **uniforme** sur $]a, b[$ (ou $[a, b]$) et on écrira $X \sim \mathcal{U}(a, b)$, si elle admet pour densité

$$f_X(x) = \frac{1}{b-a} \mathbb{1}_{]a, b[}(x)$$

f_X constante sur $]a, b[$ $\rightsquigarrow$ "X est un réel choisi uniformément au hasard dans l'intervalle $]a, b[$ "
nulle en dehors

$$F_X(t) = \int_{-\infty}^t f_X(x) dx = \begin{cases} 0 & \text{si } t \leq a \\ \frac{t-a}{b-a} & \text{si } t \in]a, b[\\ 1 & \text{si } t \geq b \end{cases} \quad \text{continue}$$

$$\mathbb{P}(X \in]a, b[) = 1 \quad \text{p.s. } X \text{ est bornée} \quad |X| \leq \max(|a|, |b|)$$

X admet des moments finis de tout ordre

$$\mathbb{E}(X) = \frac{1}{b-a} \int_a^b x dx = \frac{b^2 - a^2}{2(b-a)} = \frac{a+b}{2}$$

$$\mathbb{E}(X^2) = \frac{1}{b-a} \int_a^b x^2 dx = \frac{b^3 - a^3}{3(b-a)} = \frac{a^2 + ab + b^2}{3}$$

$$\text{Var}(X) = \frac{(b-a)^2}{12}$$

Déf Soit $F: \mathbb{R} \rightarrow \mathbb{R}$ fonction croissante, continue à droite et t.g.

$$\lim_{x \rightarrow -\infty} F(x) = 0 \quad \text{et} \quad \lim_{x \rightarrow +\infty} F(x) = 1$$

(alors $\forall x \in \mathbb{R}, 0 \leq F(x) \leq 1$)

On définit le **pseudo-inverse** (ou **pseudo-réciproque**) de F

$$\varphi:]0,1[\rightarrow \mathbb{R}$$

$$\varphi(x) := \inf \{y \in \mathbb{R} : F(y) \geq x\}$$

Remarques

- 1) Si F est bijective de $F^{-1}(]0,1[) := \{x \in \mathbb{R} : F(x) \in]0,1[\}$ dans $]0,1[$, on a $\varphi = F^{-1}$ (la fonction réciproque de F)
- 2) φ est continue à gauche et croissante sur $]0,1[$

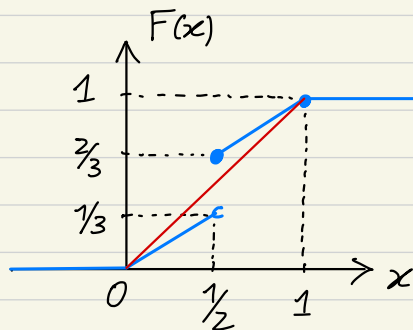
Théorème 5.18 [Vidéo sur Moodle]

Soit $X \sim \mathcal{U}(0,1)$, alors $Y := \varphi(X)$ est une v.a. réelle dont la fonction de répartition F_Y est F .

Très utile en simulation informatique, ce théorème permet de construire une v.a. de fonction de répartition donnée, à partir d'une v.a. $\sim \mathcal{U}(0,1)$.

Exemples

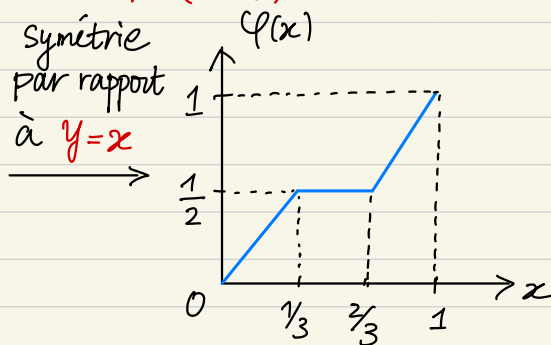
1)



les sauts de F

Restreindre d'abord le graphe de F
à l'intervalle $I = F^{-1}(]0,1[)$

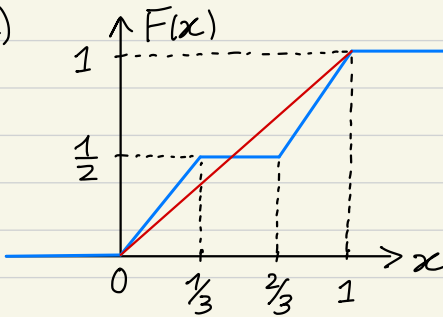
symétrie
par rapport
à $y=x$



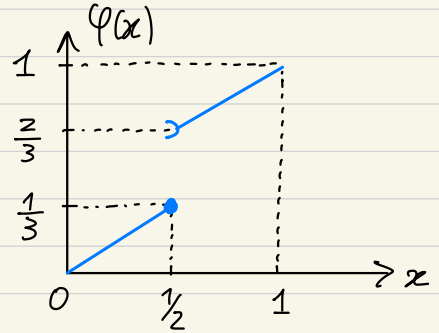
plateaux pour φ

Inversement :

2)

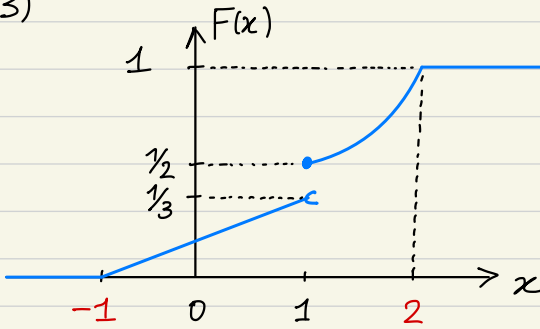


symétrie
par rapport
à $y=x$
→

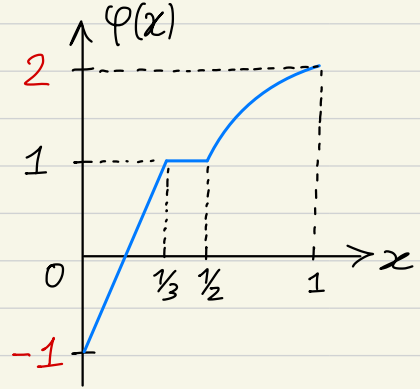


les plateaux de F → sauts pour φ

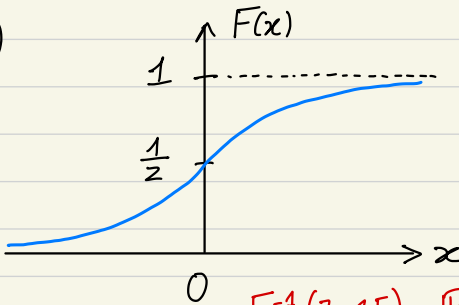
3)



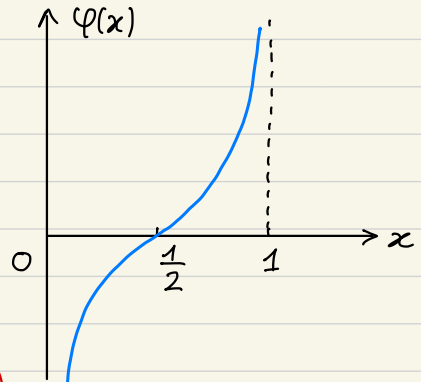
$$F^{-1}(]0,1[) =]-1,2[$$



4)



$$F^{-1}(]0,1[) = \mathbb{R}$$



$(\forall x \in \mathbb{R}, F(x) > 0)$ $(\forall x \in \mathbb{R}, F(x) < 1)$

$\lim_{x \rightarrow 0^+} \varphi(x) = -\infty$ et $\lim_{x \rightarrow 1^-} \varphi(x) = +\infty$ sont possibles

2) Loi exponentielle

X est dite **exponentielle de paramètre** $\lambda > 0$, $X \sim \mathcal{E}(\lambda)$
 $\Leftrightarrow X$ a densité $f_X(x) = \lambda e^{-\lambda x} \mathbf{1}_{]0, +\infty[}(x)$

$$\text{Si } t < 0, F_X(t) = 0$$

$$\text{Si } t \geq 0, F_X(t) = \int_{-\infty}^t f_X(x) dx = \int_0^t \lambda e^{-\lambda x} dx = [-e^{-\lambda x}]_0^t = 1 - e^{-\lambda t}$$

$$\begin{aligned} E(X) &= \int_0^{+\infty} x \lambda e^{-\lambda x} dx = [-x e^{-\lambda x}]_0^{+\infty} + \int_0^{+\infty} e^{-\lambda x} dx \quad \text{intégration par partie} \\ &= 0 + [-\frac{1}{\lambda} e^{-\lambda x}]_0^{+\infty} = \frac{1}{\lambda} \end{aligned}$$

$$\begin{aligned} E(X^2) &= \int_0^{+\infty} x^2 \lambda e^{-\lambda x} dx = [-x^2 e^{-\lambda x}]_0^{+\infty} + \frac{2}{\lambda} \int_0^{+\infty} x \lambda e^{-\lambda x} dx \\ &= 0 + \frac{2}{\lambda} E(X) = \frac{2}{\lambda^2} \end{aligned}$$

$$\text{Var}(X) = \frac{1}{\lambda^2}$$

Absence de mémoire comme la loi géométrique

Prop Si $X \sim \mathcal{E}(\lambda)$, alors $\forall s, t \geq 0$

$$\mathbb{P}(X > s+t \mid X > s) = \mathbb{P}(X > t)$$

Preuve: $\mathbb{P}(X > s) = e^{-\lambda s}$

$$\mathbb{P}(X > s+t \mid X > s) = \frac{\mathbb{P}(X > s+t)}{\mathbb{P}(X > s)} = \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} = e^{-\lambda t} = \mathbb{P}(X > t)$$

□

Exercice 5.20: $X \sim \mathcal{E}(\lambda)$, $Y \sim \mathcal{E}(\mu)$, $\lambda, \mu > 0$

Suppose que $X \perp\!\!\!\perp Y$. Soit $Z = \min\{X, Y\}$

$$\begin{aligned} \forall t \in \mathbb{R}, \mathbb{P}(Z > t) &= \mathbb{P}(X > t \text{ et } Y > t) \\ &= \mathbb{P}(X > t) \mathbb{P}(Y > t) \quad \text{par indép} \\ &= e^{-\lambda t} e^{-\mu t} \\ &= e^{-(\lambda+\mu)t} \end{aligned}$$

Donc $Z \sim \mathcal{E}(\lambda+\mu)$

3) Loi normale (gaussienne)

Z est dite **normale** (ou **gaussienne**) **standard** (ou **centrée réduite**)

si Z a pour densité $f_Z(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$, $x \in \mathbb{R}$

on écrira $Z \sim \mathcal{N}(0,1)$

Intégrale de Gauss $\int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi} \Rightarrow \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} dx = \sqrt{2\pi}$

(utilise une intégrale double qu'on exprime en coordonnées cartésiennes, puis en coordonnées polaires.)

Fonction de répartition notée

$$\Phi(t) = F_Z(t) = \int_{-\infty}^t \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx \text{ n'admet pas d'écriture plus simple}$$

$$\mathbb{E}(Z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \underline{x e^{-\frac{x^2}{2}}} dx = 0$$

fonction intégrable sur $\mathbb{R}$ et impaire

$$\begin{aligned} \text{Var}(Z) &= \mathbb{E}(Z^2) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x^2 e^{-\frac{x^2}{2}} dx \\ &= \frac{1}{\sqrt{2\pi}} \left(\left[-x e^{-\frac{x^2}{2}} \right]_{-\infty}^{+\infty} + \int_{-\infty}^{+\infty} e^{-\frac{x^2}{2}} dx \right) = 1 \end{aligned}$$

$\mu \in \mathbb{R}$, $\sigma \geq 0$ donnés

On définit $X := \sigma Z + \mu$ ($\mathbb{E}(X) = \mu$ et $\text{Var}(X) = \sigma^2$)

Si $\sigma > 0$, on vérifie (exemple 5.4) que X a pour densité

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

Une v.a. de cette densité est appelée **normale** (ou **gaussienne**) d'espérance (ou de moyenne) $\mu \in \mathbb{R}$ et de variance σ^2

On écrira $X \sim \mathcal{N}(\mu, \sigma^2)$

Par convention $X = \mu$ si $\sigma = 0$.

Si $\sigma > 0$, $X \sim \mathcal{N}(\mu, \sigma^2) \Leftrightarrow \frac{X - \mu}{\sigma} \in \mathcal{N}(0, 1)$

Prop 5.21 (exercice) Une transformation affine d'une v.a. normale reste une v.a. normale. Plus précisément,

$$\forall a, b \in \mathbb{R} \quad X \sim \mathcal{N}(\mu, \sigma^2) \Rightarrow aX + b \sim \mathcal{N}(\tilde{\mu}, \tilde{\sigma}^2)$$

$$\text{où } \tilde{\mu} = a\mu + b \text{ et } \tilde{\sigma}^2 = a^2 \sigma^2$$

Fonction génératrice des moments :

Prop Si $Z \sim \mathcal{N}(0, 1)$, $M_Z(t) = e^{\frac{t^2}{2}}$, $t \in \mathbb{R}$

Si $X \sim \mathcal{N}(\mu, \sigma^2)$, $M_X(t) = e^{\mu t + \frac{\sigma^2}{2} t^2}$, $t \in \mathbb{R}$

Preuve
$$\begin{aligned} M_Z(t) &= \int_{-\infty}^{+\infty} e^{tx} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{tx - \frac{1}{2}x^2} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{\frac{t^2}{2} - \frac{1}{2}(x-t)^2} dx \\ &= e^{\frac{t^2}{2}} \underbrace{\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-t)^2} dx}_{=1} = e^{\frac{t^2}{2}} \end{aligned}$$

$$\begin{aligned} M_X(t) &= \mathbb{E}(e^{tx}) = \mathbb{E}(e^{t(\sigma Z + \mu)}) = e^{\mu t} \mathbb{E}(e^{t\sigma Z}) \\ &= e^{\mu t} \cdot M_Z(t\sigma) = e^{\mu t + \frac{\sigma^2}{2} t^2} \end{aligned}$$

□

Corollaire Pour $Z \sim \mathcal{N}(0, 1)$, $\forall k \geq 0$ $\mathbb{E}(Z^{2k+1}) = 0$

$$\mathbb{E}(Z^{2k}) = \frac{(2k)!}{2^k \cdot k!}$$

Preuve :

$$\sum_{k \geq 0} \frac{t^{2k}}{2^k \cdot k!} = e^{\frac{t^2}{2}} = M_Z(t) = \sum_{k \geq 0} \mathbb{E}(Z^k) \frac{t^k}{k!}$$

$$\text{Donc } \mathbb{E}(Z^{2k+1}) = 0 \text{ et } \mathbb{E}(Z^{2k}) = \frac{(2k)!}{2^k \cdot k!}$$

□

Prop 5.22 (voir l'exemple 5.32 dans la section suivante)

"La somme de deux v.a. normales **indép** est normale"

Si $X \sim \mathcal{N}(\mu_1, \sigma_1^2)$, $Y \sim \mathcal{N}(\mu_2, \sigma_2^2)$ et $X \perp Y$,

alors $X+Y \sim \mathcal{N}(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$

En utilisant l'indép par paquets et un argument de récurrence, on a

Corollaire Soient $X_1, \dots, X_n$ des v.a. normales indép

Toute combinaison linéaire $X := a_1 X_1 + \dots + a_n X_n + b$

Avec $a_1, \dots, a_n, b \in \mathbb{R}$, est une v.a. normale.

FIN SÉANCE 10

§ 5.3 Vecteurs aléatoires à densité

Cas vecteur aléatoire $(X, Y) \in \mathbb{R}^2$ à étudier en détail

Cas $(X_1, \dots, X_n) \in \mathbb{R}^n$ reste analogue

Déf Un vecteur aléatoire (X, Y) à valeurs dans $\mathbb{R}^2$ est dit **absolument continu** ou **à densité** s'il existe $f_{X,Y}: \mathbb{R}^2 \rightarrow [0, +\infty[$ intégrable sur $\mathbb{R}^2$ t.q. $\forall I, J$ intervalles de $\mathbb{R}$

$$\mathbb{P}((X, Y) \in I \times J) = \mathbb{P}(X \in I, Y \in J) = \int_{I \times J} f_{X,Y}(x, y) dx dy$$

$f_{X,Y}$ appelée **densité** du vecteur aléatoire (X, Y) .

Remarque 1 En prenant $I=J=\mathbb{R}$, on a

$$\int_{\mathbb{R}^2} f_{X,Y}(x, y) dx dy = 1$$

Remarque 2 Réciproquement, toute fonction $f: \mathbb{R}^2 \rightarrow [0, +\infty[$ intégrable sur $\mathbb{R}^2$ et d'intégrale 1 est la densité d'un vecteur aléatoire dans $\mathbb{R}^2$ (admis)

Remarque 3 Densité de (X, Y) n'est pas déterminée de façon unique.

Si $g: \mathbb{R}^2 \rightarrow [0, +\infty[$ coïncide avec $f_{X,Y}$ en dehors d'un ensemble de mesure 2-dimensionnelle nulle (voir le cours d'intégrale de Lebesgue), g est aussi une densité de (X, Y)

Remarque 4 $\forall C \subset \mathbb{R}^2$ borélien

$$\begin{aligned} P((X, Y) \in C) &= \int_C f_{X,Y}(x, y) dx dy \quad (\text{intégrale au sens de Lebesgue}) \\ &= \int_{\mathbb{R}^2} f_{X,Y}(x, y) 1_C(x, y) dx dy \end{aligned}$$

La densité d'un vecteur aléatoire abs. continu caractérise sa loi.
(Admis)

Exemple (Vecteur aléatoire **uniforme continu**)

$C \subset \mathbb{R}^2$ borélien, on note $\text{mes}(C) = \int_{\mathbb{R}^2} 1_C(x, y) dx dy$

↑
mesure de Lebesgue de C

Suppose $0 < \text{mes}(C) < +\infty$.

$(X, Y) \sim \mathcal{U}(C)$ loi uniforme continue sur C

Si (X, Y) est abs. continu de densité $\frac{1}{\text{mes}(C)} 1_C(x, y)$

(X, Y) uniforme sur C représente un point uniformément choisi au hasard dans C .

Rappel **Fubini-Tonelli**

Soit $g: \mathbb{R}^2 \rightarrow \mathbb{R}^+$ une fonction (mesurable) positive

$\forall I, J$ intervalles de $\mathbb{R}$

$$\begin{aligned} \int_{I \times J} g(x, y) dx dy &= \int_I \left(\int_J g(x, y) dy \right) dx \\ &= \int_J \left(\int_I g(x, y) dx \right) dy \end{aligned}$$

Prop (Densités jointes et marginales)

Si (X, Y) est un vecteur aléatoire de densité $f_{X,Y}$,
ses composantes X et Y sont des v.a. abs. continues, de densité

$$f_X(x) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y) dy \quad f_Y(y) = \int_{-\infty}^{+\infty} f_{X,Y}(x,y) dx$$

$f_{X,Y}$ densité jointe de X, Y

f_X et f_Y densités marginales

Preuve $\forall I \subset \mathbb{R}$ intervalle

$$\mathbb{P}(X \in I) = \mathbb{P}(X \in I, Y \in \mathbb{R}) = \mathbb{P}((X, Y) \in I \times \mathbb{R})$$

$$= \int_{I \times \mathbb{R}} f_{X,Y}(x,y) dx dy$$

$$= \int_I \left(\underbrace{\int_{\mathbb{R}} f_{X,Y}(x,y) dy}_{= f_X(x)} \right) dx \quad \text{par Fubini-Tonelli}$$

analogue pour Y = $f_Y(y)$

□

Une autre application de Fubini-Tonelli : calculer la densité de $X+Y$

Proposition 5.28 [Vidéo sur Moodle]

Soit (X, Y) un vecteur aléatoire de densité $f_{X,Y}$

Alors $X+Y$ est une v.a. abs. continue de densité

$$f_{X+Y}(z) = \int_{-\infty}^{+\infty} f_{X,Y}(x, z-x) dx = \int_{-\infty}^{+\infty} f_{X,Y}(z-y, y) dy$$

à condition que $x \mapsto f_{X,Y}(x, z-x)$ soient intégrables sur $\mathbb{R}$
 $y \mapsto f_{X,Y}(z-y, y)$

Proposition 5.29 (Indépendance et densité)

1) Soit (X, Y) un vecteur aléatoire abs. continu,

Si $f_{X,Y}(x,y) = f_X(x) f_Y(y)$ pour tout $(x,y) \in \mathbb{R}^2 \setminus C$
 où C est un ensemble de mesure 2-dimensionnelle nulle (C peut être vide)
 alors X et Y sont indép.

2) Réciproquement, si X, Y sont indép et abs. continues de densité f_X, f_Y
 alors le vecteur (X, Y) est abs. continu et de densité jointe
 $f_{X,Y}(x,y) = f_X(x) f_Y(y)$

Preuve 1) $\forall I, J \subset \mathbb{R}$ intervalles

$$\begin{aligned} \mathbb{P}(X \in I, Y \in J) &= \int_{I \times J} f_{X,Y}(x,y) dx dy \\ &= \int_{I \times J} f_X(x) f_Y(y) dx dy \quad \text{par hypothèse} \\ &= \left(\int_I f_X(x) dx \right) \cdot \left(\int_J f_Y(y) dy \right) \quad \text{par Fubini-Tonelli} \\ &= \mathbb{P}(X \in I) \mathbb{P}(Y \in J) \end{aligned}$$

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

2) $\forall I, J \subset \mathbb{R}$ intervalles

$$\begin{aligned} \mathbb{P}((X,Y) \in I \times J) &= \mathbb{P}(X \in I, Y \in J) \\ &= \mathbb{P}(X \in I) \mathbb{P}(Y \in J) \quad \text{par indép} \\ &= \int_I f_X(x) dx \cdot \int_J f_Y(y) dy \\ &= \int_{I \times J} \underbrace{f_X(x) f_Y(y)}_{\text{une densité de } (X,Y)} dx dy \quad \text{par Fubini-Tonelli} \end{aligned}$$

□

Remarque 1 En pratique, pour montrer $X \perp\!\!\!\perp Y$, il suffit que $f_{X,Y}$ se factorise sous la forme $f_{X,Y}(x,y) = \alpha(x) \cdot \beta(y)$

où $\alpha: \mathbb{R} \rightarrow \mathbb{R}^+$ et $\beta: \mathbb{R} \rightarrow \mathbb{R}^+$ sont des fonctions (mesurables)
 Nécessairement, $\exists \lambda \in (0, +\infty)$ t.q. $\lambda \alpha(x)$ est la densité de X
 $x^{-1} \beta(y)$ est la densité de Y
 [Voir Remarque 3.18 du poly en cas discret]

Remarque 2

Un vecteur aléatoire
est abs. continu

$\Rightarrow$
 $\Leftarrow$

ses composantes sont des v.a.
abs. continues.

vraie si les composantes sont indép.



" $\Leftarrow$ " est fausse en général

Par exemple: soit X une v.a. abs. continue

$P((X, X) \in A) = 1$ où $A = \{(x, y) \in \mathbb{R}^2: x = y\}$ la diagonale

$$\begin{aligned} \text{Or } \int_{\mathbb{R}^2} 1_A(x, y) dx dy &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} 1_A(x, y) dx \right) dy \text{ par Fubini-Tonelli} \\ &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} 1_{\{y\}}(x) dx \right) dy = \int_{\mathbb{R}} 0 dy = 0 \end{aligned}$$

A est de mesure 2-dimensionnelle nulle.

(X, X) n'est pas un vecteur aléatoire abs. continu

Sinon, en modifiant arbitrairement les valeurs de la densité $f_{X,X}$ sur A , on obtient facilement une contradiction avec

$$1 = P((X, X) \in A) = \int_A f_{X,X}(x, y) dx dy$$

En combinant Prop 5.28 et 5.29, on a

Corollaire Si X, Y sont des v.a. indép., de densité f_X et f_Y
 alors $X+Y$ est abs. continue de densité

$$\begin{aligned} f_{X+Y}(z) &= f_X * f_Y(z) := \int_{-\infty}^{+\infty} f_X(x) f_Y(z-x) dx \\ &= f_Y * f_X(z) := \int_{-\infty}^{+\infty} f_X(z-y) f_Y(y) dy \end{aligned}$$

La fonction $f_x * f_y = f_y * f_x$ appelée *produit de convolution* de f_x et f_y

Application: Exemple 5.32 [Vidéo sur Moodle]

Si $X \sim \mathcal{N}(0,1)$ et $Y \sim \mathcal{N}(0,1)$ avec $X \perp Y$

alors $X+Y \sim \mathcal{N}(0,2)$

(voir Proposition 5.22 pour le résultat plus général)

- Formule de transfert pour calculer les espérances :

Prop 5.33 (version intégrable, admise)

Soit (X,Y) un vecteur aléatoire abs. continu de densité $f_{X,Y}$

Soit $g: \mathbb{R}^2 \rightarrow \mathbb{R}$ une fonction continue sur $\mathbb{R}^2 \setminus C$

où C est de mesure 2-dimensionnelle nulle.

Alors $g(X,Y)$ admet une espérance finie

$\Leftrightarrow$ fonction $|g(x,y)| f_{X,Y}(x,y)$ est intégrable sur $\mathbb{R}^2$

Dans ce cas, $\mathbb{E}(g(X,Y)) = \int_{\mathbb{R}^2} g(x,y) f_{X,Y}(x,y) dx dy$

Remarque : version positive

$g: \mathbb{R}^2 \rightarrow \mathbb{R}^+$ on a toujours $\mathbb{E}(g(X,Y)) = \int_{\mathbb{R}^2} g(x,y) f_{X,Y}(x,y) dx dy$.

Remarque 5.34 (Admise)

(X,Y) abs. continu de densité $f_{X,Y}$

$\Leftrightarrow \forall h: \mathbb{R}^2 \rightarrow \mathbb{R}$ fonction bornée (ou positive) et continue sauf au plus sur un ensemble de mesure 2-dim nulle, on a

$\mathbb{E}(h(X,Y)) = \int_{\mathbb{R}^2} h(x,y) f_{X,Y}(x,y) dx dy$.

Application: Exemple 5.35 [Vidéo sur Moodle]

La méthode de Box-Müller pour simuler une loi normale, où la méthode d'inversion ne s'applique pas.

U_1, U_2 indép, de loi $\sim \mathcal{U}(0,1)$

On pose $R = \sqrt{-2 \ln U_1}$ $R \perp \Theta$

$$\Theta = 2\pi U_2$$

et $X = R \cos(\Theta)$, $Y = R \sin(\Theta)$

En calculant $\mathbb{E}(h(X,Y))$ avec $h \in \mathcal{C}_b(\mathbb{R}^2, \mathbb{R})$ (ou $h \in \mathcal{C}(\mathbb{R}^2, \mathbb{R}^+)$),

montrer que $X \sim \mathcal{N}(0,1)$ et $X \perp Y$
 $Y \sim \mathcal{N}(0,1)$

Rappel: Formule de changement de variable

$U, V \subset \mathbb{R}^2$ ouverts

$\varphi: V \rightarrow U$ $\mathcal{C}^1$ -difféomorphisme

- φ bijective
- dérivées partielles de φ continues sur V
- matrice jacobienne de $\varphi = (\varphi_1, \varphi_2)$

$$J_\varphi(s,t) = \begin{pmatrix} \frac{\partial \varphi_1}{\partial s} & \frac{\partial \varphi_1}{\partial t} \\ \frac{\partial \varphi_2}{\partial s} & \frac{\partial \varphi_2}{\partial t} \end{pmatrix} \text{ est inversible } \left(\Leftrightarrow \det J_\varphi(s,t) \neq 0 \right) \\ \forall (s,t) \in V$$

$\forall g: U \rightarrow \mathbb{R}$ intégrable

$$\int_U g(x,y) dx dy = \int_V g(\varphi_1(s,t), \varphi_2(s,t)) |\det J_\varphi(s,t)| ds dt$$

Coordonnées polaires en dimension 2

$$\varphi(r,\theta) = (r \cos \theta, r \sin \theta) \quad \begin{array}{l} 0 < r < +\infty \\ 0 < \theta < 2\pi \end{array}$$

$$|\det J_\varphi(r,\theta)| = r$$

$$\int_{\mathbb{R}^2} g(x,y) dx dy = \int_{]0,+\infty[\times]0,2\pi[} g(r \cos \theta, r \sin \theta) r dr d\theta$$

Cas n -dimensionnel $(X_1, \dots, X_n) \in \mathbb{R}^n$ abs. continu

voir les résultats analogues dans le poly, pages 69-70

FIN SÉANCE 11

Chapitre VI Loi des grands nombres et Convergence en probabilité

§ 6.1 Loi des grands nombres

Soit $(X_i)_{i \geq 1}$ une suite de variables aléatoires réelles

On définit pour $n \geq 1$ la variable aléatoire réelle

$$\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i \quad \text{moyenne empirique des } n \text{ premières } X_i$$

Théorème 6.1 (Loi des grands nombres)

Soit $(X_i)_{i \geq 1}$ une suite de v.a. réelles **indép et identiquement distribuées (i.i.d.)**, qui admettent une espérance **finie**.

On note $\mu = \mathbb{E}(X_i)$. Alors $\forall \varepsilon > 0$,

$$\lim_{n \rightarrow \infty} \mathbb{P}(|\bar{X}_n - \mu| > \varepsilon) = \lim_{n \rightarrow \infty} \mathbb{P}\left(\left|\frac{\sum_{i=1}^n X_i}{n} - \mu\right| > \varepsilon\right) = 0$$

Remarque 1 $(X_i)_{i \geq 1}$ sont de même loi

$$\mathbb{E}(\bar{X}_n) = \frac{1}{n} \sum_{i=1}^n \mathbb{E}(X_i) = \mu$$

Remarque 2 Loi des grands nombres (LGN) rend rigoureuse l'idée de l'approche fréquentiste des probabilités

$X_1, \dots, X_N$ les valeurs de $X = 1_A$ pendant N répétitions indép de l'expérience.

$$\sum_{i=1}^N X_i = N_A = \text{nb. de fois qu'on a observé } A$$

$$\frac{X_1 + \dots + X_N}{N} = \frac{N_A}{N} \approx \mathbb{P}(A) = \mathbb{E}(1_A) \quad \text{quand } N \gg 1.$$

Remarque 3 La condition d'espérance finie est cruciale

1^{er} contre-exemple : $(X_i)_{i \geq 1}$ iid de loi de Cauchy
(densité $f(x) = \frac{1}{\pi(1+x^2)}$)

X_i n'admet pas d'espérance

(Admis) $\bar{X}_n = \frac{1}{n} \sum_{i=1}^n X_i$ suit aussi la loi de Cauchy !

$\forall \mu \in \mathbb{R}, \varepsilon > 0$

$$\mathbb{P}(|\bar{X}_n - \mu| > \varepsilon) = 1 - \mathbb{P}(|\bar{X}_n - \mu| \leq \varepsilon)$$

$$= 1 - \int_{\mu-\varepsilon}^{\mu+\varepsilon} \frac{dx}{\pi(1+x^2)}$$

$$= 1 - \frac{1}{\pi} (\arctan(\mu+\varepsilon) - \arctan(\mu-\varepsilon)) > 0$$

ne dépend pas de n

$(X_i)_{i \geq 1}$ ne vérifie pas la loi des grands nombres.

Une version plus faible de LGN (utilisée pour démontrer Théorème 6.1)

Proposition Soit $(X_i)_{i \geq 1}$ une suite de v.a. réelles **i.i.d.** qui admettent un **moment d'ordre 2 fini** (et donc aussi une espérance finie). On note $\mu = \mathbb{E}(X_i)$, $\sigma^2 = \text{Var}(X_i) < +\infty$

$$\forall \varepsilon > 0 \quad \lim_{n \rightarrow \infty} \mathbb{P}(|\bar{X}_n - \mu| > \varepsilon) = 0$$

Preuve $\mathbb{E}(\bar{X}_n) = \mu$

$$\begin{aligned} \text{Var}(\bar{X}_n) &= \text{Var}\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \frac{1}{n^2} \text{Var}\left(\sum_{i=1}^n X_i\right) \\ &= \frac{1}{n^2} \sum_{i=1}^n \text{Var}(X_i) \quad \text{par indép des } X_i \\ &= \frac{\sigma^2}{n} \end{aligned}$$

Par inégalité de Bienaymé-Tchebychev,

$$\mathbb{P}(|\bar{X}_n - \mu| > \varepsilon) \leq \frac{\text{Var}(\bar{X}_n)}{\varepsilon^2} = \frac{\sigma^2}{\varepsilon^2 n} \xrightarrow{n \rightarrow \infty} 0$$

□

Preuve du théorème 6.1 : voir la vidéo sur Moodle.

§ 6.2 Une application de LGN : méthode de Monte-Carlo

1) Estimer numériquement une intégrale $\int_a^b f(x) dx$

$(X_i)_{i \geq 1}$ une suite de v.a. i.i.d. de loi uniforme $\mathcal{U}(a, b)$

$$\mathbb{E}(f(X_i)) = \frac{1}{b-a} \int_a^b f(x) dx$$

Appliquer LGN à $(f(X_i), i \geq 1)$ suite i.i.d. :

$$\forall \varepsilon > 0, \lim_{n \rightarrow \infty} \mathbb{P}\left(\left|\frac{1}{n} \sum_{i=1}^n f(X_i) - \frac{1}{b-a} \int_a^b f(x) dx\right| > \varepsilon\right) = 0$$

Vitesse d'approximation :

Bienaymé-Tchebychev

$$\mathbb{P}\left(\left|\frac{1}{n} \sum_{i=1}^n f(X_i) - \frac{1}{b-a} \int_a^b f(x) dx\right| > \varepsilon\right) \leq \frac{\text{Var}\left(\frac{1}{n} \sum_{i=1}^n f(X_i)\right)}{\varepsilon^2}$$

$$= \frac{\frac{1}{n} \text{Var}(f(X_i))}{\varepsilon^2}$$

$$\leq \frac{\frac{\mathbb{E}(f(X_i)^2)}{n}}{\varepsilon^2}$$

Si f bornée ($|f| \leq M$), alors

$$\mathbb{P}\left(\left|\frac{1}{n} \sum_{i=1}^n f(X_i) - \frac{1}{b-a} \int_a^b f(x) dx\right| > \varepsilon\right) \leq \frac{M^2}{n \varepsilon^2}$$

Donc, $\forall \delta > 0$, si $n \geq \frac{\mathbb{E}(f(X_i)^2)}{\delta \varepsilon^2}$ (ou $\frac{M^2}{\delta \varepsilon^2}$), alors

avec proba $1-\delta$, $\frac{1}{b-a} \int_a^b f(x) dx$ est proche de $\frac{1}{n} \sum_{i=1}^n f(X_i)$ à ε près.

2) Estimer la probabilité d'un événement A

$$1_A \sim \text{Bern}(\mathbb{P}(A))$$

$$\mu = \mathbb{E}(1_A) = \mathbb{P}(A)$$

$$\sigma^2 = \text{Var}(1_A) = \mathbb{P}(A)(1 - \mathbb{P}(A)) \leq \frac{1}{4}$$

$(X_i)_{i \geq 1}$ iid de loi de Bernoulli de paramètre $\mathbb{P}(A)$

(X_i indicatrice que A soit réalisé lors de la i -ème expérience)

$$\forall \varepsilon > 0, \quad \mathbb{P}\left(\left|\frac{1}{n} \sum_{i=1}^n X_i - \mathbb{P}(A)\right| > \varepsilon\right) \leq \frac{\sigma^2}{n\varepsilon^2} \leq \frac{1}{4\varepsilon^2 n}$$

Exemples de mise en pratique : voir page 75 du poly
et page 81 du poly.

§ 6.3 Convergence en probabilité pour une suite de v.a.

Définition Soient $(X_n)_{n \geq 1}$ et X des v.a. réelles définies sur le même espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$. On dit que la suite $(X_n)_{n \geq 1}$ converge en ($\mathbb{P}$ -) probabilité vers X (noté $(X_n) \xrightarrow[n \rightarrow \infty]{\mathbb{P}} X$) si

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

On dit que $(X_n) \xrightarrow{\mathbb{P}} +\infty$ (resp. $-\infty$) si

$$\forall A > 0, \quad \lim_{n \rightarrow \infty} \mathbb{P}(X_n < A) = 0 \quad (\text{resp. } \lim_{n \rightarrow \infty} \mathbb{P}(X_n > -A) = 0)$$

Remarque 1 $\exists$ plusieurs probabilités $\mathbb{P}_1, \mathbb{P}_2, \dots$ sur $(\Omega, \mathcal{F})$

" $(X_n)_{n \geq 1}$ converge en $\mathbb{P}$ -probabilité" : relative à $\mathbb{P}$

Si l'il y a pas d'ambiguïté sur la probabilité de référence, on écrit " $(X_n)_{n \geq 1}$ converge en probabilité".

Remarque 2 LGN: si $(X_i)_{i \geq 1}$ i.i.d d'espérance finie $\mu = \mathbb{E}(X_i)$
alors $\left(\frac{1}{n} \sum_{i=1}^n X_i\right)_{n \geq 1} \xrightarrow{\mathbb{P}} \mu$

Exemple: $(X_i)_{i \geq 1}$ v.a. i.i.d. de loi $\mathcal{E}(1)$

$$M_n = \max\{X_1, \dots, X_n\}$$

$$\mathbb{P}(M_n < A) = \mathbb{P}(X_i < A \text{ pour tout } 1 \leq i \leq n)$$

$$= \prod_{i=1}^n \mathbb{P}(X_i < A) \quad \text{par indép}$$

$$= \mathbb{P}(X_1 < A)^n = (1 - e^{-A})^n = \mathbb{P}(M_n \leq A)$$

$$\forall A > 0, 1 - e^{-A} < 1 \Rightarrow \lim_{n \rightarrow \infty} \mathbb{P}(M_n < A) = 0$$

$$\text{Donc } (M_n) \xrightarrow{\mathbb{P}} +\infty$$

Pour faire mieux, on prend $A = (1 \pm \varepsilon) \ln n$ avec $\varepsilon > 0$

$$\mathbb{P}(M_n < (1 - \varepsilon) \ln n) = (1 - e^{-(1 - \varepsilon) \ln n})^n$$

$$= \left(1 - \frac{1}{n^{1 - \varepsilon}}\right)^n$$

$$\ln\left(\left(1 - \frac{1}{n^{1 - \varepsilon}}\right)^n\right) = n \ln\left(1 - \frac{1}{n^{1 - \varepsilon}}\right) \underset{n \rightarrow \infty}{\sim} n \left(-\frac{1}{n^{1 - \varepsilon}}\right) = -n^\varepsilon$$

$$\lim_{n \rightarrow \infty} \ln\left(\left(1 - \frac{1}{n^{1 - \varepsilon}}\right)^n\right) = -\infty$$

$$\text{Donc } \lim_{n \rightarrow \infty} \mathbb{P}(M_n < (1 - \varepsilon) \ln n) = 0$$

$$\begin{aligned} \mathbb{P}(M_n > (1 + \varepsilon) \ln n) &= 1 - \mathbb{P}(M_n \leq (1 + \varepsilon) \ln n) \\ &= 1 - (1 - e^{-(1 + \varepsilon) \ln n})^n \end{aligned}$$

$$\mathbb{P}(M_n > (1+\varepsilon) \ln n) = 1 - \left(1 - \frac{1}{n^{1+\varepsilon}}\right)^n$$

$$\ln\left(\left(1 - \frac{1}{n^{1+\varepsilon}}\right)^n\right) = n \ln\left(1 - \frac{1}{n^{1+\varepsilon}}\right) \underset{n \rightarrow \infty}{\sim} n \cdot \left(-\frac{1}{n^{1+\varepsilon}}\right)$$

$$\lim_{n \rightarrow \infty} \ln\left(\left(1 - \frac{1}{n^{1+\varepsilon}}\right)^n\right) = 0$$

Donc $\lim_{n \rightarrow \infty} \mathbb{P}(M_n > (1+\varepsilon) \ln n) = 1 - 1 = 0$

Ainsi $\forall \varepsilon > 0$, $\lim_{n \rightarrow \infty} \mathbb{P}\left(\left|\frac{M_n}{\ln n} - 1\right| > \varepsilon\right)$

$$= \lim_{n \rightarrow \infty} \left(\mathbb{P}(M_n > (1+\varepsilon) \ln n) + \mathbb{P}(M_n < (1-\varepsilon) \ln n)\right)$$

$$= 0$$

Conclusion

$$\left(\frac{M_n}{\ln n}\right)_{n \geq 2} \xrightarrow{\mathbb{P}} 1$$

Proposition 6.9 (complément à LGN) [Vidéo sur Moodle]

Soit $(X_i)_{i \geq 1}$ une suite de v.a. iid **positives, d'espérance $+\infty$** .

Alors $\left(\frac{1}{n} \sum_{i=1}^n X_i\right)_{n \geq 1} \xrightarrow{\mathbb{P}} +\infty$

Proposition 6.10 (Admise)

Si $(X_n) \xrightarrow{\mathbb{P}} X$ et $(Y_n) \xrightarrow{\mathbb{P}} Y$ (où toutes les v.a. sont définies sur le même espace probabilisé $(\Omega, \mathcal{F}, \mathbb{P})$), alors

$$\forall f: \mathbb{R}^2 \rightarrow \mathbb{R} \text{ continue} \quad (f(X_n, Y_n))_{n \geq 1} \xrightarrow{\mathbb{P}} f(X, Y)$$

En particulier, $\forall g: \mathbb{R} \rightarrow \mathbb{R}$ continue, $(g(X_n)) \xrightarrow{\mathbb{P}} g(X)$

et $(X_n + Y_n) \xrightarrow{\mathbb{P}} X + Y$ on encore $(X_n Y_n) \xrightarrow{\mathbb{P}} XY$
preuve : page 77 du poly.

Proposition (Admise)

Si $(X_n) \xrightarrow{\mathbb{P}} X$ alors $\forall f: \mathbb{R} \rightarrow \mathbb{R}$ continue **bornée**,

$$\lim_{n \rightarrow \infty} \mathbb{E}(f(X_n)) = \mathbb{E}(f(X))$$



$(X_n) \xrightarrow{\mathbb{P}} X$ n'implique pas $\mathbb{E}(X_n) \rightarrow \mathbb{E}(X)$

($f(x) = x$ n'est pas bornée)

Exemple : $(X_n)_{n \geq 1}$ v.a. t.q. $\mathbb{P}(X_n = n^2) = \frac{1}{n}$
 $\mathbb{P}(X_n = 0) = 1 - \frac{1}{n}$

$$\forall \varepsilon > 0, \quad \mathbb{P}(|X_n| > \varepsilon) = \mathbb{P}(X_n = n^2) \text{ dès que } n^2 > \varepsilon \\ = \frac{1}{n} \xrightarrow{n \rightarrow \infty} 0$$

Donc $(X_n) \xrightarrow{\mathbb{P}} 0$

$$\text{Mais, } \mathbb{E}(X_n) = 0 \cdot (1 - \frac{1}{n}) + n^2 \cdot \frac{1}{n} = n$$

$$\lim_{n \rightarrow \infty} \mathbb{E}(X_n) = +\infty \neq \mathbb{E}(0) = 0$$

FIN SÉANCE 12
