

I - Introduction

- objectif : quantification du hasard par la modélisation probabiliste.
- Cette modélisation permet de passer de la notion de probabilité, concrète dans notre vie quotidienne, à la notion de modèle probabiliste, abstraite.
- Les proba. reposent sur deux théories : la théorie de la mesure et l'intégration.

① - Phénomène aléatoire

- L'objectif de la théorie des probabilités est d'analyser mathématiquement les phénomènes dans lesquels le hasard intervient : ce sont des phénomènes aléatoires.
- Un phénomène est aléatoire si, reproduit plusieurs fois dans les mêmes conditions, il se déroule chaque fois différemment avec un résultat imprévisible.

② - Idées sous-jacentes à la théorie des probabilités

a) La loi des grands nombres :

L'aléa est souvent lié à la méconnaissance des paramètres intervenant dans une expérience (ou leur trop grand nombre). Bien que ces comportements aléatoires soient a priori sujets à des variations imprévisibles, on pourra donner des renseignements sur ce type de phénomènes en répétant l'expérience. Cette procédure permet de déceler généralement des lois régissant les résultats.

b) Indépendance et conditionnement.

La construction d'un modèle probabiliste repose sur l'information connue a priori sur l'expérience aléatoire - Par exemple choisir au hasard un homme de + de 80kg parmi 1000 hommes dans la pop. générale est plus probable si le groupe est composé d'hommes de plus d'1m85 que d'hommes de \leq 1m75.

\Rightarrow C'est la notion de conditionnement!

\Rightarrow Si l'information a priori sur un phénomène aléatoire n'a aucune incidence sur la réalisation d'un autre phénomène, on dit que ces phénomènes sont II.

③ - Les variables aléatoires

Ce sont des fonctions qui dépendent du résultat de l'expérience aléatoire sous-jacente. Au lieu de chercher les antécédents de chaque valeur de la fonction, on s'intéresse à la chance de réalisation de l'ensemble des antécédents qui conduisent à cette valeur : c'est la loi de la variable aléatoire.

④ - Espace de probabilité

a) Expérience aléatoire :

\rightarrow Une expérience aléatoire E peut conduire à plusieurs résultats possibles, imprévisibles, même en étant menée dans des conditions identiques.

\rightarrow L'espace d'états associé à l'expérience, noté Ω , représente l'ensemble des résultats possibles.

\rightarrow Un résultat de l'expérience est souvent noté w , avec $w \in \Omega$.

- Exemples: • Lancer de 2 pièces: $\Omega = \{PP, FF, PF, FP\}$ (2)
- Prix d'un actif financier sur l'intervalle de temps $[t_1, t_2]$:
- $\Omega =$ ensemble des fonctions continues sur $[t_1, t_2]$ à valeurs réelles ≥ 0 .

b) Evenement aléatoire:

→ Un événement aléatoire est un sous-ensemble de Ω dont on peut dire au vu de l'expérience s'il est réalisé ou non. C'est une partie de Ω , donc c'est un ensemble.

- Exemples: • On considère l'expérience aléatoire $E =$ lancer de 2 dés à 6 faces.
- $A = \{ \text{somme des 2 dés} < 4 \}$ est un événement aléatoire.
- $B = \{ \text{Le résultat du 1er dé est} < 4 \}$ n'est pas un événement! (si Ω ne contient que les résultats non-ordonnés des tirages).

(5) - Rappel sur les ensembles

On énumère ici certaines propriétés d'opérations sur des ensembles, et leur connexion avec les événements aléatoires. Soient A, B, C des ensembles.

→ $A \cup B = B \cup A$

→ $A \cup (B \cap C) = (A \cup B) \cap C = A \cup B \cap C$

→ $A \cap (B \cap C) = A \cap B \cap C$

→ $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

→ $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

→ Une partition de Ω est une famille $(A_i)_{i \in I}$ telle que

$$\begin{cases} \bigcup_{i \in I} A_i = \Omega \\ A_i \cap A_j = \emptyset \text{ pour } i \neq j, i, j \in I \end{cases}$$

→ Pour une famille $(A_i)_{i \in I}$ d'ensembles, indexée par un ensemble I quelconque, on a:

- $\bigcup_{i \in I} A_i$ est la réunion de cette famille: c'est l'ensemble des points qui appartiennent à au moins l'un des A_i .
- $\bigcap_{i \in I} A_i$ est l'intersection de la famille: c'est l'ensemble des points qui appartiennent à tous les A_i en même temps.

→ On rappelle de plus que les événements élémentaires sont des ensembles, ce qui permet de connecter facilement toutes ces notions: Soient A et B 2 événements.

- Implication: la réalisation de A entraîne la réalisation de $B \Rightarrow A \subset B$:
en effet, si $\omega \in A$ alors $\omega \in B$. $\left(\begin{smallmatrix} \textcircled{C} \\ \textcircled{A} \end{smallmatrix} \right)^{\omega} B$

- On note en général \mathcal{A} l'ensemble de tous les événements. Souvent, on prend $\mathcal{A} = \mathcal{P}(\Omega)$ (ensemble des parties de Ω), sauf si Ω n'est pas fini ou dénombrable.

→ Pour que la modélisation probabiliste soit cohérente, \mathcal{A} doit être stable pour les opérations suivantes:

- $A \in \mathcal{A} \left\{ \begin{array}{l} B \in \mathcal{A} \end{array} \right\} \Rightarrow A \cap B \in \mathcal{A}$
- $A \in \mathcal{A} \left\{ \begin{array}{l} B \in \mathcal{A} \end{array} \right\} \Rightarrow A \cup B \in \mathcal{A}$
- $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$
- $\Omega \in \mathcal{A}$
- $\emptyset \in \mathcal{A}$

→ En probabilité, on cherche à associer à chaque événement A un nombre $P(A) \in [0, 1]$, représentant la chance que cet événement se réalise avec l'expérience. Intuitivement, en répétant cette expérience, on pourrait approximer ce nombre par la fréquence empirique de réalisation.

II - Probabilité

(3)

Soit A un événement aléatoire, et l'ensemble des événements, \mathcal{A} l'espace d'états de l'expérience aléatoire.

① - Cas où \mathcal{A} est fini (ou dénombrable)

Dans ce cas, on choisira toujours $\mathcal{A} = \mathcal{P}(\mathcal{A})$.

Définition : Une probabilité sur \mathcal{A} fini est une application $P: \mathcal{P}(\mathcal{A}) \rightarrow [0,1]$ qui est caractérisée par :

- $0 \leq P(A) \leq 1$
- $P(\mathcal{A}) = 1$
- $P(A \cup B) = P(A) + P(B)$ si $A \cap B = \emptyset$

De plus, on a : les propriétés suivantes :

- $P(\emptyset) = 0$
- $P(A) + P(A^c) = 1$
- $P(A) \leq P(B)$ si $A \subset B$
- $P(\bigcup_{i \in I} A_i) = \sum_{i \in I} P(A_i)$ si les A_i sont disjoints 2 à 2
- $P(A \cup B) + P(A \cap B) = P(A) + P(B)$

→ Un modèle probabiliste est un triplet $(\mathcal{A}, \mathcal{A}, P)$

→ Cas discret : $\mathcal{A} = \{w_1, \dots, w_n\}$.

- Une probabilité P sur \mathcal{A} est entièrement caractérisée par $\{P(\{w_i\}) = p_{w_i}, w_i \in \mathcal{A}\}$
- Étant donné une famille $(p_i)_{i \in I}$ de nombres réels, il lui correspond une unique probabilité P telle que $\forall w_i \in \mathcal{A}, p_i = P(\{w_i\}) \Leftrightarrow \begin{cases} 0 \leq p_i \leq 1 \\ \sum_i p_i = 1 \end{cases}$

On a alors : $\forall A \in \mathcal{P}(\mathcal{A}), P(A) = \sum_{w \in A} P(\{w\}) = \sum_{w \in A} p_w$.

- Exemple de la probabilité uniforme = chaque singleton de Ω a la même chance de réalisation. Ainsi, $P_w = IP(\{w\})$ ne dépend pas de w .

$$\left\{ \forall w, P_w = \frac{1}{\text{Card}(\Omega)} \text{ avec Card}(\Omega) \text{ désigne le nb d'éléments de } \Omega. \right.$$

② - Cas général

On fait ici référence au cas où Ω est infini.

→ Prenons un exemple concret: on joue à Pile ou Face, n fois - $\Omega = \{P, F\}^n$.

Ω est fini, de cardinal 2^n . Sans trichage, on obtient $\forall A \in \mathcal{A}$ ou $\forall A \subset \Omega$,

$$IP_n(A) = \text{Card}(A) / 2^n \quad \text{et indénombrable!}$$

→ On joue maintenant une infinité de fois - Ainsi $\Omega = \{P, F\}^{\mathbb{N}^+}$ est infini. ↘

On s'intéresse à la proba. de l'événement A : "On ne tire jamais Pile.",

et on introduit l'événement A_n = "On ne tire pas Pile lors des n premières tirages."

On remarque que $IP(A_n) = 1/2^n$.

→ Naturellement, A est la limite des A_n : en effet si je réalise A_{n+1} alors je réalise A_n - Ainsi, $A_{n+1} \subset A_n$; c'est donc la limite d'une suite décroissante.

Donc $A = \bigcap_n A_n$ - On a envie d'écrire ensuite: $IP(A) = \lim_{n \rightarrow \infty} IP(A_n) = 0$.

Mais l'intervention entre limite et proba IP nécessite d'ajouter un axiome pour les propriétés d'une probabilité - Pour le moment, on ne peut pas écrire:

$$A = \lim_{n \rightarrow \infty} \downarrow A_n \Rightarrow IP(A) = IP(\lim_{n \rightarrow \infty} \downarrow A_n) = \lim_{n \rightarrow \infty} IP(A_n) !$$

⇒ Ici, Ω est infini donc \mathcal{A} ne peut pas être $\mathcal{P}(\Omega)$.

⇒ Cela renvoie aux notions d'ensembles dénombrables, de tribus (et de tribu borélienne en particulier sur \mathbb{R}). En effet, il faut pouvoir mesurer la chance de n'importe quelle partie de Ω pour définir une modélisation probabiliste (IP)

Rp: On rappelle que si $\Omega = \mathbb{R}$, la tribu borélienne est la tribu engendrée par la classe des intervalles ouverts de \mathbb{R} de la forme $] -\infty, a[$ avec $a \in \mathbb{Q}$. En effet, la tribu est stable par passage au complémentaire, et par union ou intersection dénombrable (en bijection avec \mathbb{N}) car:

- $]a, +\infty[=]-\infty, a]^c$ est un ouvert de \mathbb{R} .
- l'union dénombrable d'intervalles ouverts est un intervalle ouvert
- idem pour l'intersection.
- soit (x_n) suite \downarrow de rationnels vers x , et (y_n) une suite \uparrow de rationnels vers y :

$$]x, y[= \bigcup_n (]-\infty, y_n[\cap]-\infty, x_n]^c)$$
 en rappelant que \mathbb{Q} est dense dans \mathbb{R} .
 \Rightarrow tout intervalle ouvert \in à la tribu.

Définition: Une probabilité sur l'espace (Ω, \mathcal{F}) est une application de \mathcal{F} dans $[0, 1]$, notée P , telle que:

- $P(\Omega) = 1$
- Pour toute suite dénombrable $(A_n)_{n \in \mathbb{N}}$ d'éléments de \mathcal{F} disjoints 2 à 2,

$$P\left(\bigcup_n A_n\right) = \sum_n P(A_n)$$

\rightarrow Ce dernier axiome, appelé " σ -additivité", permet le passage à la limite précédemment évoqué car il assure que la série de terme général $P(A_n)$ est convergente.

\rightarrow Ainsi,

- en considérant $(A_n)_n$ une suite \downarrow de parties de Ω ($A_{n+1} \subset A_n$) avec $A = \bigcap_n A_n$, on écrit $A_n \downarrow A$, on a $P(\bigcap_n A_n) = \lim_{n \rightarrow \infty} P(A_n)$.
- inversement, pour une suite \uparrow ($A_n \subset A_{n+1}$) avec $A = \bigcup_n A_n$, on a $P(A) = P(\bigcup_n A_n) = \lim_{n \rightarrow \infty} P(A_n)$.

→ Ce résultat entraîne que si $(A_n)_n$ est une suite \uparrow ou \downarrow d'événements, la suite $(\mathbb{P}(A_n))_n$ admet une limite quand n tend vers l'infini.

Définition : Une propriété est vraie \mathbb{P} -presque sûrement (\mathbb{P} p.s.) si l'ensemble des $\omega \in \Omega$ pour lesquels elle est vraie est de probabilité 1. Dans ce cas, l'ensemble des ω pour lesquels elle est fausse est négligeable.

III - Autres notions fondamentales

① - Loi d'une variable aléatoire (v.a.)

En pratique, on utilise les variables aléatoires plutôt que les événements. La variable aléatoire dépend du résultat de l'expérience.

→ On considère un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. Une v.a. X est une application de (Ω, \mathcal{A}) dans un ensemble F : $\omega \in \Omega \mapsto X(\omega) \in F$.

→ Intérêt : Comme F est un espace bien ^{mieux} connu ($\mathbb{R}, \mathbb{N}, \mathbb{R}^d, \dots$), il est plus facile de travailler en s'intéressant aux réalisations ^{préférées} des valeurs de X dans F plutôt qu'aux chances de réalisation des résultats de l'expérience.

Ainsi, en posant $B \subset F$, on note souvent $\mathbb{P}_X(B) = \mathbb{P}(X^{-1}(B))$

→ En notant \mathcal{F} la famille des parties B de F telles que $X^{-1}(B) \in \mathcal{A}$, \mathcal{F} est une tribu de F et \mathbb{P}_X définie sur (F, \mathcal{F}) est la loi de la v.a. X . C'est la mesure image de \mathbb{P} par l'application mesurable X .
 $\underbrace{\{\omega, X(\omega) \in B\}}_{\text{désigne donc un ensemble.}}$

② - Conditionnement

→ C'est la notion qui marque la principale différence entre la théorie des probabilités et la théorie de la mesure.

L'idée est qu'une information supplémentaire sur l'expérience modifie la vraisemblance accordée à l'événement étudié.

→ Exemple : on lance 2 dés et on cherche par un lancer la proba. de l'événement : "La somme des dés est ≥ 10 ".

Cette proba. vaut $\frac{1}{6}$ à la base, $\frac{1}{2}$ si l'on sait que l'un des 2 dés a donné 6...

→ On rappelle les résultats suivants : soient A et B deux événements, t.p. $P(B) > 0$:

- $P(A|B) = P(A \cap B) / P(B)$
- Si A_1, \dots, A_n sont des événements de Ω tels que $P(A_1 \cap A_2 \cap \dots \cap A_{n-1}) > 0$, alors $P(A_1 \cap A_2 \cap \dots \cap A_n) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \dots P(A_n|A_1 \cap \dots \cap A_{n-1})$.
- Si $(B_i)_{i \in I}$ est une partition finie ou dénombrable d'événements de Ω , telle que $P(B_i) > 0$ pour chaque i , Alors $\forall A \in \mathcal{A}$, $P(A) = \sum_{i \in I} P(A \cap B_i) = \sum_{i \in I} P(A|B_i)P(B_i)$

③ - Indépendance

Deux événements A et B sont \perp si le fait de savoir que A est réalisé ne donne aucune information sur la réalisation de B (et réciproquement).

→ L'indépendance est une notion liée au choix de la probabilité P , mais n'est pas une notion sur les ensembles. Par exemple, rien ne dit que A et B sont disjoints!

On peut donner l'exemple suivant pour illustrer cela : on tire au hasard une carte dans un jeu classique de 52 cartes.

Soient les événements: $A = \{ \text{la carte est une dame} \}$
On a $P(A) = \frac{4}{52}$ et $P(B) = \frac{13}{52}$ $B = \{ \text{la carte est un cœur} \}$

Or $P(A \cap B) = \frac{1}{52}$, puisque $P(A \cap B) = P(A) \times P(B) \Rightarrow A \perp B$ pour la probabilité uniforme.

Considérons maintenant un jeu truqué qui engendre une nouvelle probabilité de tirage telle $Q(\text{valet de pique}) = \frac{1}{2}$ et $Q(\text{autre carte}) = \frac{1}{2} \times \frac{1}{51} = \frac{1}{102}$.
n'importe laquelle parmi les 51 cartes restantes

Alors $Q(A \cap B) = \frac{1}{102} \neq Q(A) Q(B) = \frac{2}{51} \times \frac{13}{102}$, donc $A \not\perp B$ sous Q .
dame de cœur dame cœur

→ On rappelle un quelques propriétés importantes sur l'indépendance:

• $P(A \cap B \cap C) = P(A) P(B) P(C) \nRightarrow A, B \text{ et } C \perp$. Il faut avoir

$$\begin{cases} P(A \cap B \cap C) = P(A) P(B) P(C) \\ P(A \cap B) = P(A) P(B) \\ P(A \cap C) = P(A) P(C) \\ P(B \cap C) = P(B) P(C) \end{cases}$$

$\Rightarrow A, B \text{ et } C \text{ sont } \perp$.

IV - Le lemme de Borel-Cantelli

Le lemme de Borel-Cantelli est un résultat très important en théorie des probabilités. Il est notamment utilisé pour démontrer des théorèmes de convergence.

Dans ce théorème, la notion d'indépendance entre événements intervient fondamentalement.

① - L'ensemble \limsup

Considérons une suite $(A_n)_n$ d'événements aléatoires. On définit l'ensemble

$$\limsup_n A_n = \bigcap_p \bigcup_{n \geq p} A_n$$

↳ tous les points qui appartiennent à au moins 1 A_n .
 ↳ tous les points qui appartiennent à tous les ensembles ci-dessus en même temps.

→ Remarquons que $\limsup_n A_n$ est un ensemble !

→ De plus, $w \in \limsup_n A_n \iff \forall p, \exists n \geq p$ tel que $w \in A_n$
 $\iff w$ appartient à une infinité de A_n
 $\iff \sum_n \mathbb{1}_{A_n}(w) = +\infty$.

→ Du coup, $w \notin \limsup_n A_n \iff w$ appartient à au plus un nombre fini de A_n .

② - Énoncé du lemme

Le lemme s'énonce généralement en deux parties, suivant que les événements A_n soient \perp ou non.

→ La 1^{ère} partie du lemme (cas général : \perp ou non) est précieuse pour montrer qu'une propriété est vraie IP-presque sûrement. (cf preuve LGN).

→ La 2^e partie (cas où les A_n sont \perp) caractérise entièrement le fait que $\mathbb{P}(\limsup_n A_n)$ vaut 0 ou 1 suivant le CV ou la DV de la série de Varine général $\mathbb{P}(A_n)$.

Lemme de Borel-Cantelli: Soient $(A_n)_n$ une suite d'événements de \mathcal{A} .

i) $\sum_n^{\infty} P(A_n) < +\infty \Rightarrow P(\limsup_n A_n) = 0$.

Ainsi, P -p.s., il y a au plus un nombre fini de A_n qui sont réalisés.

ii) La suite $(A_n)_n$ est indépendante.

$\sum_n^{\infty} P(A_n) = +\infty \Rightarrow P(\limsup_n A_n) = 1$.

Ici, P -p.s., une infinité de A_n sont réalisés.

Preuve: On remarque d'abord qu'on a une suite décroissante d'événements.

i) On a $P(\limsup_n A_n) = P(\bigcap_p \bigcup_{n \geq p} A_n) = \lim_p \downarrow P(\bigcup_{n \geq p} A_n) \leq \lim_p \downarrow \sum_{n \geq p} P(A_n)$

Par hypothèse, la série $\sum_n P(A_n)$ converge, donc le reste de cette série $\rightarrow 0$.

Ainsi $\sum_{n \geq p} P(A_n) \xrightarrow[p \rightarrow \infty]{} 0$ - Par conséquent, puisque $P(\limsup_n A_n) \leq \lim_p \downarrow \sum_{n \geq p} P(A_n) \xrightarrow[p \rightarrow \infty]{} 0$,

on obtient que $P(\limsup_n A_n) = 0$.

ii) Supposons maintenant que les A_n sont \perp et que la série $\sum_n P(A_n)$ diverge.

Soit m un nombre entier, on a: $P(\underbrace{\bigcup_{i=p}^m A_i}_{\substack{w \in \mathcal{A} \text{ à au moins} \\ \text{un } A_i, \forall i \in [p, m]}}) = 1 - P(\underbrace{\bigcap_{i=p}^m A_i^c}_{\substack{w \in \mathcal{A} \text{ en m temps à ts} \\ \text{les } A_i^c = w \text{ n' } \in \text{ à aucun } A_i}}) \stackrel{\perp}{=} 1 - \prod_{i=p}^m P(A_i^c)$

Ainsi $P(\bigcup_{i=p}^m A_i) = 1 - \prod_{i=p}^m (1 - P(A_i)) \geq 1 - e^{-\sum_{i=p}^m P(A_i)}$ (on utilise que pr $x \geq 0$, $1-x \leq e^{-x}$)

Comme $\sum_n P(A_n) = +\infty$, on obtient $P(\bigcup_{i=p}^{\infty} A_i) \geq \underbrace{1 - e^{-\sum_{i=p}^{\infty} P(A_i)}}_{=1}$, ce qui permet de conclure que $P(\bigcup_{i=p}^{\infty} A_i) = 1$.

Enfin, $P(\limsup_n A_n) = \lim_p \downarrow P(\bigcup_{n \geq p} A_n) = \lim_p \downarrow 1 = 1$. (Q.F.D.)