Chapitre 9 : Espace probabilisé

I Ensembles dénombrables

En sup, le cours était limité au cas où l'univers Ω était un ensemble fini. De nombreuses expériences aléatoires ont un nombre infini de résultats possibles. Mais il convient de distinguer plusieurs types d'infinis, ce qui mène à définir la notion d'ensemble dénombrable.
Intuitivement, un ensemble est dénombrable si l'on peut « étiqueter » ses éléments, c'est-à-dire en dresser une liste exhaustive où chaque élément est repéré par un nombre, l'ensemble de ces nombres parcourant $\mathbb{N}$. Mathématiquement, cela s'écrit ainsi :

Définition 1

On dit qu'un ensemble E est **dénombrable** s'il est en bijection avec $\mathbb{N}$, i.e. s'il existe une application bijective $\phi : \mathbb{N} \to E$.
On peut alors noter $x_n = \phi(n)$, $\forall n \in \mathbb{N}$, et on peut écrire $E = \{x_n, n \in \mathbb{N}\}$, c'est ce qu'on appelle écrire E en **extension**

Remarque 2

Pour un ensemble fini, son écriture en extension est $E = \{x_1, \ldots, x_n\}$.

Exemple 3

$\mathbb{N}$ est dénombrable
L'ensemble des entiers naturels pairs $2\mathbb{N}$ est dénombrable

Exemple 4

On lance un dé jusqu'à obtenir 2 fois de suite un 6. On veut compter le nombre de jets avant d'avoir réussi à obtenir ces deux 6 consécutifs. Quel univers prendre pour modéliser cette expérience ?

Propriété 5

L'ensemble $\mathbb{Z}$ est dénombrable.

Démonstration :

$\blacksquare$

Remarque 6

Pendant le parcours des éléments de $\mathbb{Z}$, on ne peut pas avancer sur $\mathbb{N}$ sinon on pourra pas revenir à $-\mathbb{N}$.
Ce genre de parcours se généralise :

Propriété 7

Un produit cartésien de deux ensembles dénombrables est dénombrable.

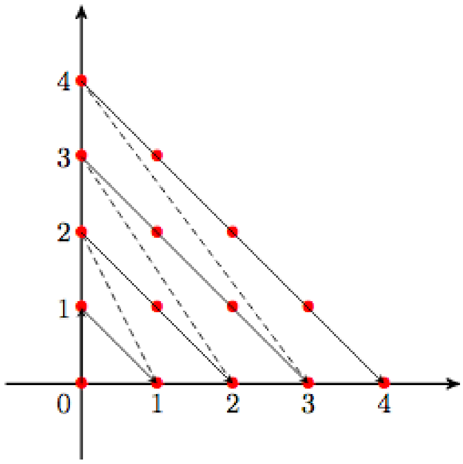
Démonstration – Soient E_1 et E_2 deux ensembles dénombrables, φ une bijection de $\mathbb{N}$ sur E_1, ψ une bijection de $\mathbb{N}$ sur E_2. L'idée est la suivante : si E_1 et E_2 sont décrits en extension sous la forme

$$E_1 = \{x_n;\ n \in \mathbb{N}\}, \quad E_2 = \{y_n;\ n \in \mathbb{N}\},$$

on peut décrire $E_1 \times E_2$ en extension sous la forme

$$E_1 \times E_2 = \{(x_0,y_0),(x_0,y_1),(x_1,y_0),(x_0,y_2),(x_1,y_1),(x_2,y_0),(x_3,y_0),\ldots\}.$$

Ce principe est illustré sur le graphique suivant dans le cas de $\mathbb{N}^2$:



Pour construire explicitement une bijection Φ de $\mathbb{N}$ sur $E_1 \times E_2$ qui correspond à la description précédente, on peut procéder ainsi : pour tout $n \in \mathbb{N}$, soit k l'unique entier naturel tel que

$$1 + 2 + \cdots + k \leqslant n < 1 + 2 + \cdots + k + (k+1)$$

($k = 0$ si $n = 0$, $k = 1$ si $n \in [\![1,2]\!]$), et soient $i = n - (1 + 2 + \cdots + k)$, $j = k - i$. On pose alors $\Phi(n) = (x_i,y_j)$. On vérifie facilement que Φ convient. $\square$

Exemple 8

On déduit de nombreux ensembles dénombrables par ces constructions :
$\mathbb{N}^n$, $\mathbb{Z}^n$, $\mathscr{M}_n(\mathbb{Z})$, $\mathbb{Q}$,...
HP : En revanche $\mathbb{R}$ n'est pas dénombrable.

Jusqu'à la fin du chapitre, on note Ω un ensemble fini ou dénombrable.

II Espace probabilisé

1 Tribu

On réalise une expérience, qu'on modélise par un univers Ω fini ou dénombrable, et un événement est un sous ensemble $A \subset \Omega$.
On note $\mathscr{A}$ l'ensemble de tous les événements qui peuvent se réaliser, alors on sait que $\mathscr{A} \subset \mathscr{P}(\Omega)$, et dans le cas dénombrable, on peut trouver un Ω tel qu'on a l'égalité (limite du programme de prépa), mais ce n'est pas toujours le plus simple à modéliser.
On réalise un lancer de pile ou face infini (= jusqu'à obtenir un résultat voulu), qu'on peut

modéliser par l'univers $\Omega = \{0,1\}^{\mathbb{N}^*}$ qui n'est pas dénombrable. (0 représente Pile et 1 représente Face).
Une suite d'élements dans $\{0,1\}$ est un événement. Il sera difficile de définir une probabilité sur $\mathscr{P}(\Omega)$, et on se donne le principe de tribu, qui est un sous-ensemble de $\mathscr{P}(\Omega)$ dont les éléments sont les événements observables.
On peut prendre $\mathscr{A} = \mathscr{P}\left(\left\{(x_n) \in \{0,1\}^{\mathbb{N}^*}, \exists n_0 \in \mathbb{N}, \forall n \geq n_0, x_n = 0\right\}\right)$ les parties de l'ensembles des suites presque nulles (= nulle àpcr).

Définition 9

On appelle **tribu** sur un univers Ω une partie $\mathscr{A}$ de $\mathscr{P}(\Omega)$ telle que

1. $\Omega \in \mathscr{A}$
2. pour tout $A \in \mathscr{A}$, le complémentaire de A est dans la tribu : $\overline{A} = \Omega \setminus A \in \mathscr{A}$
3. Pour tout suite $(A_n)_n$ d'élements de $\mathscr{A}$, la réunion $\bigcup_{n=0}^{+\infty} A_n$ est dans $\mathscr{A}$.

   On appelle **événements** les élements de la tribu $\mathscr{A}$ (qui sont des sous-ensembles de Ω).

Exemple 10

• Les parties de Ω est une tribu.
• L'ensemble $\{\emptyset, \Omega\}$ est une tribu de Ω (appelée tribu grossière)
• Sur $\Omega = [\![1,6]\!]$, on peut choisir $\mathscr{A} = \mathscr{P}(\Omega)$ qui est une tribu, ou encore $\mathscr{A} = \{\emptyset, \{2,4,6\}, \{1,3,5\}, \Omega\}$ si on ne s'intéresse qu'à la parité des résultats.
Le choix de la tribu dépend donc de l'expérience à modéliser. Intuitivement, le choix de la tribu va donc dépendre de l'information dont on dispose sur l'expérience aléatoire qu'on observe.

Remarque 11

$\emptyset = \overline{\Omega}$ est toujours un événement de la tribu

Jusqu'à la fin du chapitre on note $\mathscr{A}$ une tribu sur Ω.

Propriété 12 (*Stabilité par intersection dénombrable*)

La tribu $\mathscr{A}$ est stable par intersection dénombrable.
i.e. si $(A_n)_n$ est une suite d'événements, alors $\bigcap_{n=0}^{+\infty} A_n \in \mathscr{A}$.

Démonstration – Notons $B = \bigcap_{n=0}^{+\infty} A_n$. Alors

$$\overline{B} = \bigcup_{n=0}^{+\infty} \overline{A_n} \in \mathscr{A}$$

car pour tout $n \in \mathbb{N}$, $\overline{A_n} \in \mathscr{A}$, et $\mathscr{A}$ est stable par réunion dénombrable. Alors, par passage au complémentaire, $B \in \mathscr{A}$. □

Remarque 13

Si $A_0, \ldots, A_n$ sont des événements d'une tribu, en posant $\forall k > n,\ A_k = A_n$, alors
$\bigcup_{n=0}^{n} A_n = \bigcup_{n=0}^{+\infty} A_n$ et $\bigcap_{n=0}^{n} A_n = \bigcap_{n=0}^{+\infty} A_n$
donc une tribu est aussi stable par union et intersection finie d'événements

Comme pour le cas fini en sup, on peut faire un parallèle entre le vocabulaire ensembliste et le vocabulaire probabiliste :

Vocabulaire ensembliste	Vocabulaire des probabilités
Ensemble Ω	Univers, événement certain
Élément ω de Ω	Issue (ou résultat possible, ou réalisation)
$A \in \mathscr{A}$ ($A \in \mathcal{P}(\Omega)$ si Ω est fini)	Événement A
$\omega \in A$	L'issue ω réalise l'événement A
Si Ω est fini par exemple : singleton $\{\omega\}$	Événement élémentaire
Ensemble vide $\varnothing$	Événement impossible (jamais réalisé)
Réunion $A \cup B$	Événement « A ou B »
Réunion $\bigcup_{n=0}^{+\infty} A_n$	Événement « l'un au moins des A_n est réalisé »
Intersection $A \cap B$	Événement « A et B »
Intersection $\bigcap_{n=0}^{+\infty} A_n$	Événement « tous les A_n sont réalisés »
Complémentaire $\overline{A} = \Omega \setminus A$	Événement contraire
Parties disjointes : $A \cap B = \varnothing$	Événements incompatibles

On va retrouver la pluspart des notions de premières années.
Dans les exercices théoriques la notion de tribu est importante, mais dans les exercices pratiques, elle peut ne pas être nommée, parce qu'on choisi souvent l'univers Ω tel que $\mathscr{A} = \mathscr{P}(\Omega)$.

Définition 14

On appelle **système complet (dénombrable) d'événements** toute suite $(A_n)_n$ d'événements vérifiant

1. Les événements A_n sont deux à deux incompatibles
2. $\displaystyle\bigcup_{n=0}^{+\infty} A_n = \Omega$

Remarque 15

C'est la même definition qu'en sup, sauf que l'union est infinie.
Un SC(D)E est une partition de l'univers Ω.

Exemple 16

• Si A est un événement, $(A, \overline{A})$ est un SCE.
• L'ensemble des singletons est un SCDE
• Reprenons l'expérience de lancer une infinité de fois une pièce de monnaie.
On note les événements $P_\infty =$ « on obtient une infinité de piles », et pour $n \in \mathbb{N}, P_n =$ « on on obtient exactement n piles ».
Alors la famille $(P_\infty, P_0, P_1, P_2, ..., P_n, ...)$ est un SCDE.

Propriété 17

Tout événement se décompose sur un SC(D)E :
Soit $B \in \mathscr{A}$ un événement, et (A_n) un SC(D)E, alors

$$B = \bigcup_{n\in\mathbb{N}} (B \cap A_n) \text{ et}$$

les événements $(B \cap A_n)$ sont deux à deux incompatibles.
i.e. la famille $(B \cap A_n)_n$ est une partition de B.

2 Probabilité

Définition 18

Si Ω est un ensemble et $\mathscr{A}$ une tribu sur Ω, on appelle **probabilité** sur $(\Omega, \mathscr{A})$ toute application $P : \mathscr{A} \to [0,1]$ vérifiant :

— $P(\Omega) = 1$

— σ-additivité : pour toute suite $(A_n)_{n\in\mathbb{N}}$ d'événements incompatibles deux à deux, la série $\sum P(A_n)$ est convergente et

$$P\left(\bigcup_{n=0}^{+\infty} A_n\right) = \sum_{n=0}^{+\infty} P(A_n).$$

On appelle **espace probabilisé** le triplet $(\Omega, \mathscr{A}, P)$.

Remarque 19

1. la propriété de σ-additivité contient la **convergence** de la série $\sum P(A_n)$, pour toute suite $(A_n)_{n\in\mathbb{N}}$ d'événements incompatibles 2 à 2.
   Elle est bien entendu vraie aussi pour toute famille **finie** d'événements incompatibles deux à deux ; on parle alors d'additivité :
   Si $A_0, \ldots, A_m$ sont des événements incompatibles 2 à 2, $P\left(\bigcup_{n=0}^{m} A_n\right) = \sum_{n=0}^{m} P(A_n)$.
2. $P(\varnothing) = 0$ et $\forall A \in$ $P(\overline{A}) = 1 - P(A)$.
3. Pour A, B dans $\mathscr{A}$ tels que $A \subset B$, $P(B \backslash A) = P(B) - P(A)$.
4. $\forall (A,B) \in \mathscr{A}^2$ $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Exemple 20

• **Jeu de pile ou face infini.** L'expérience consistant à lancer indéfiniment une pièce peut-être modélisée par le choix de $\Omega = \{0,1\}^{\mathbb{N}^*}$ des suites à termes dans $\{0,1\}$ indexées à partir de 1 (0 représente « face », 1 représente « pile »). Cet ensemble n'est pas dénombrable, il n'est alors pas évident de définir une tribu $\mathscr{A}$ sur Ω et une probabilité sur $(\Omega, \mathscr{A})$. On peut montrer qu'il existe une tribu $\mathscr{A}$ sur Ω qui contient toutes les parties de Ω constituées des éléments dont les premiers termes sont imposés, c'est-à-dire les parties

$$C_{u_1,\ldots,u_k} = \{\omega = (\omega_n)_{n\geqslant 1};\ \omega_1 = u_1,\ \ldots,\ \omega_k = u_k\}$$

où $k \in \mathbb{N}^*$ et $(u_1, \ldots, u_k) \in \{0,1\}^k$ représente les k premiers termes imposés. Ce sont des événements naturels. Il existe alors une probabilité P sur $(\Omega, \mathscr{A})$ telle que, avec les notations précédentes,

$$P(C_{u_1,\ldots,u_k}) = \frac{1}{2^k}.$$

Par exemple :

– « le résultat du second lancer est pile » est un événement : il s'agit de $C_{0,1} \cup C_{1,1}$;
– « on n'obtient jamais pile » est un événement : il s'agit de $A_0 = \bigcap_{k=1}^{+\infty} C_{u_1,\ldots,u_k}$ où tous les u_n sont nuls ;
– pour tout $n \in \mathbb{N}^*$, « on obtient pile pour la première fois au n-ième lancer » est un événement : il s'agit de $A_n = C_{0,\ldots,0,1}$ (0 apparaissant $n-1$ fois).

La famille des (A_n) est un SCDE

On note $(\Omega, \mathscr{A}, P)$ un espace probabilisé jusqu'à la fin du chapitre

Définition 21

Un événement de probabilité 1 autre que Ω est dit **presque sûr** ou **presque certain**.
Un événement de probabilité 0 autre que $\varnothing$ est dit **négligeable** ou **presque impossible**.

Propriété 22 **(*des probabilités*)**

1. Croissance : pour A, B dans $\mathscr{A}$ tels que $A \subset B$, $P(A) \leq P(B)$.
2. Continuité croissante :
   Si $(A_n)_{n\in\mathbb{N}}$ est une suite croissante d'événements (i.e. $\forall n,\ A_n \subset A_{n+1}$), alors
   $$\lim_{n\to+\infty} P(A_n) = P\left(\bigcup_{n=0}^{+\infty} A_n\right).$$
3. Continuité décroissante :
   Si $(A_n)_{n\in\mathbb{N}}$ est une suite décroissante d'événements (i.e. $\forall n,\ A_{n+1} \subset A_n$),
   $$\lim_{n\to+\infty} P(A_n) = P\left(\bigcap_{n=0}^{+\infty} A_n\right).$$
4. Sous-additivité :
   Si $(A_n)_{n\in\mathbb{N}}$ est une suite d'événements, alors
   $$P\left(\bigcup_{n=0}^{+\infty} A_n\right) \leq \sum_{n=0}^{+\infty} P(A_n)$$
   (où cette dernière somme vaut par convention $+\infty$ si la série diverge...).
   Pour une famille finie $A_0, \dots, A_m$ d'événements, on a $P\left(\bigcup_{n=0}^{m} A_n\right) \leq \sum_{n=0}^{m} P(A_n)$.

- On écrit $B = A \cup (B \cap \overline{A})$. Les événements A et $B \cap \overline{A} = B \setminus A$ sont incompatibles, donc
$$P(B) = P(A) + P(B \cap \overline{A}) \geqslant P(A).$$
- Posons $B_0 = A_0$ et pour tout $k \in \mathbb{N}^*$, $B_k = A_k \cap \overline{A_{k-1}} = A_k \setminus A_{k-1}$. Alors
$$\bigcup_{k=0}^{+\infty} A_k = \bigcup_{k=0}^{+\infty} B_k,$$
les événements B_k étant deux à deux incompatibles : s'il existait $(n,m) \in \mathbb{N}^2$ tel que $n < m$ et $B_n \cap B_m \neq \varnothing$, on pourrait trouver un élément ω de A_n n'appartenant pas à A_{m-1}, ce qui est absurde car $A_n \subset A_{m-1}$.

Mais, d'après la démonstration de la propriété précédente, pour tout $k \in \mathbb{N}^*$,

$$P(A_k \cap \overline{A_{k-1}}) = P(A_k) - P(A_{k-1}).$$

Finalement,

$$P\left(\bigcup_{k=0}^{+\infty} A_k\right) = \sum_{k=0}^{+\infty} P(B_k) = P(B_0) + \sum_{k=1}^{+\infty} P(A_k \cap \overline{A_{k-1}}) = P(A_0) + \sum_{k=1}^{+\infty}(P(A_k) - P(A_{k-1})).$$

On reconnaît une somme de série télescopique, et on conclut en rappelant que pour tout $n \in \mathbb{N}^*$,

$$\sum_{k=1}^{n}(P(A_k) - P(A_{k-1})) = P(A_n) - P(A_0).$$

• Posons, pour tout $k \in \mathbb{N}$, $B_k = \overline{A_k}$. Alors, pour tout $k \in \mathbb{N}$, B_k est un événement et $B_k \subset B_{k+1}$. D'après le point précédent,

$$P(B_n) \underset{n\to+\infty}{\longrightarrow} P\left(\bigcup_{k=0}^{+\infty} B_k\right),$$

c'est-à-dire

$$1 - P(A_n) \underset{n\to+\infty}{\longrightarrow} 1 - P\left(\bigcap_{k=0}^{+\infty} \overline{B_k}\right) = 1 - P\left(\bigcap_{k=0}^{+\infty} A_k\right),$$

d'où le résultat. On remarquera que la suite $(P(A_n))_{n\in\mathbb{N}}$ est décroissante.

• Posons, pour tout $n \in \mathbb{N}$, $B_n = \bigcup_{k=0}^{n} A_k$. Alors pour tout $n \in \mathbb{N}$, B_n est un événement et $B_n \subset B_{n+1}$, donc d'après la propriété de continuité croissante,

$$P(B_n) \underset{n\to+\infty}{\longrightarrow} P\left(\bigcup_{k=0}^{+\infty} B_k\right) = P\left(\bigcup_{k=0}^{+\infty} A_k\right).$$

Mais d'après le dernier point de la propriété précédente, on obtient, pour tout $n \in \mathbb{N}$,

$$P(B_n) \leqslant \sum_{k=0}^{n} P(A_k).$$

En passant à la limite lorsque $n \to +\infty$, on obtient l'inégalité souhaitée. □

Exemple 23

• Dans le jeu de pile ou face infini, soit A l'événement « on obtient pile au moins deux fois », et pour tout $n \geqslant 2$, A_n l'événement « on obtient pile au moins deux fois au cours des n premiers lancers ». Réaliser $\overline{A_n}$ revient à obtenir pile aucune fois ou une fois exactement au cours des n premiers lancers : $\overline{A_n}$ est la réunion des événements deux à deux incompatibles $C_{u_1,\dots,u_n}$ où les u_i sont tous nuls, ou bien tous nuls sauf un. Ces événements sont au nombre de $n+1$ et ont tous pour probabilité $1/2^n$, donc

$$P(A_n) = 1 - \frac{n+1}{2^n}.$$

De plus, pour tout $n \geqslant 2$, $A_n \subset A_{n+1}$; enfin, $A = \bigcup_{n=2}^{+\infty} A_n$. Ainsi,

$$P(A) = \lim_{n\to+\infty} P(A_n) = 1$$

par croissances comparées.

• Soit $(A_k)_{k\in\mathbb{N}}$ une suite d'événements. Pour tout $n \in \mathbb{N}$, notons B_n l'événement $\bigcup_{k=n}^{+\infty} A_k$; notons également B l'événement $\bigcap_{n=0}^{+\infty} B_n$. Ainsi,

$$B = \bigcap_{n=0}^{+\infty}\left(\bigcup_{k=n}^{+\infty} A_k\right).$$

Il s'agit de l'événement « une infinité des A_k sont réalisés ». En effet, $\omega \in B$ si et seulement si pour tout $n \in \mathbb{N}$, il existe $k \geqslant n$ tel que $\omega \in A_k$, ce qui équivaut au fait que ω appartient à une infinité de A_k.

Supposons que la série $\sum_{k\geqslant 0} P(A_k)$ converge. Pour tout $n \in \mathbb{N}$, $B \subset B_n$, donc $P(B) \leqslant P(B_n)$. Or, d'après la propriété de sous-additivité et le fait que $\sum_{k\geqslant 0} P(A_k)$ converge, on a pour tout $n \in \mathbb{N}$,

$$P(B_n) \leqslant \sum_{k=n}^{+\infty} P(A_k),$$

le majorant tendant vers 0 en tant que reste d'une série convergente. Une probabilité étant positive, on en déduit que $P(B_n) \to 0$, et donc $P(B) = 0$. Cette propriété s'écrit ainsi : presque sûrement, le nombre des événements A_n qui sont réalisés est fini.

III Généralisation des théorèmes de sup

1 Probabilité conditionnelle

Lors d'une expérience aléatoire, le fait de savoir (ou d'imaginer) qu'un événement est réalisé revient à ajouter de l'information sur l'expérience, et peut modifier notre façon de calculer laprobabilité de certains événements. C'est ce que l'on appelle les probabilités conditionnelles. Soit $(\Omega, \mathscr{A}, P)$ un espace probabilisé. Supposons que B soit un événement tel que $P(B) > 0$. Calculer la probabilité qu'un événement A soit réalisé en sachant que l'événement B est réalisé revient à considérer, parmi les issues qui réalisent B, celles qui réalisent également A. Tout se passe comme si, pour ce calcul, on considérait l'expérience aléatoire à travers le « filtre » de l'événement B, comme si l'on considérait B comme univers.

Définition 24 (*/propriété*)

Soit B un événement non négligeable,
Pour tout événement $A \in \mathscr{A}$, on appelle **probabilité conditonnelle** de A sachant B le réel noté $P_B(A)$ ou $P(A|B)$ défini par

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

L'application $P_B : \begin{cases} \mathscr{A} \to [0,1] \\ A \mapsto P_B(A) \end{cases}$ est une probabilité sur $(\Omega, \mathscr{A})$, appelée **probabilité conditionnelle** sachant B.

Démonstration :

■

Remarque 25

Si $P(B) = 0$, afin que l'égalité $P(A \cap B) = P(A|B)P(B)$ reste valable, on pose **par convention** la valeur du produit $P(A|B)P(B) = 0$. (Mais le terme $P(A|B)$ n'est pas défini, s'il est seul, dans ce cas).

2 Propriétés et théorèmes

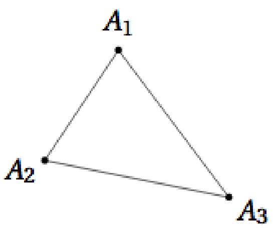
Théorème 26 (*Formule des probabilités composées*)

Soit des événements $A_1, \dots, A_n \in \mathscr{A}$ telles que $P(A_1 \cap \dots \cap A_n) > 0$. alors

$$P(A_1 \cap A_2 \cap \dots \cap A_n) = P(A_1) \times P_{A_1}(A_2) \times P_{A_1 \cap A_2}(A_3) \times \cdots \times P_{A_1 \cap \cdots \cap A_{n-1}}(A_n)$$

Exemple 27

Chaîne de Markov. On considère un point qui se déplace sur les sommets d'un triangle $A_1A_2A_3$:



On suppose qu'initialement le point se trouve en A_1. Ensuite les déplacements s'effectuent de la manière suivante :

- si le point est en A_i alors il passe en A_j ($j \neq i$) avec probabilité $\frac{2}{5}$ dans les deux cas ;
- le point reste en A_i avec probabilité $\frac{1}{5}$.

On peut résumer ceci grâce à un diagramme de transition :

Le point commence en A_1, quelle est la probabilité que le premier mouvement du point soit exactement A_1 puis A_2 puis A_3 ?

Théorème 28 (*Formule des probabilités totales*)

Soit $(A_n)_{n\in\mathbb{N}}$ un **système complet (D) d'événements**, alors pour tout $B \in \mathscr{A}$, la série $\sum P_{A_n}(B)P(A_n) = P(B \cap A_n)$ converge, et

$$P(B) = \sum_{n=0}^{+\infty} P_{A_n}(B)P(A_n) = \sum_{n=0}^{+\infty} P(B \cap A_n)$$

Tout d'abord, la série $\sum_{n\geqslant 0} P(B\cap A_n)$ converge, car les événements $B \cap A_n$ sont deux à deux incompatibles. De plus, notons N l'événement $\Omega \setminus \bigcup_{n=0}^{+\infty} A_n$. Les A_n étant deux à deux incompatibles,

$$P\left(\bigcup_{n=0}^{+\infty} A_n\right) = \sum_{n=0}^{+\infty} P(A_n) = 1$$

et donc $P(N) = 0$. En particulier, $P(B \cap N) = 0$. On a alors

$$B = B \cap \Omega = B \cap \left(N \cup \bigcup_{n=0}^{+\infty} A_n\right) = (B \cap N) \cup \bigcup_{n=0}^{+\infty} (B \cap A_n).$$

Les A_n et N forment une famille d'événements deux à deux incompatibles, donc c'est aussi le cas des $B \cap A_n$ et de $B \cap N$, et on a finalement

$$P(B) = P(B\cap N) + \sum_{n=0}^{+\infty} P(B\cap A_n) = \sum_{n=0}^{+\infty} P(B\cap A_n) = \sum_{n=0}^{+\infty} P(B \,|\, A_n)P(A_n). \qquad \square$$

Remarque 29

L'événement N dans cette preuve n'est pas obligatoire, (noté N pour négligeable), il permet de retirer le besoin d'avoir l'hypothèse que les événements recouvre l'univers Ω, on remarque qu'il suffit que leur union soit certaine.
On a une réécriture plus simple de ce théorème.

Définition 30

On dit que la suite d'événements (A_n) forme un **système quasi-complet d'événements** si Les (A_n) sont deux à deux incompatibles et

$$P(\bigcup_{n\in\mathbb{N}} A_n) = \sum_{n=0}^{+\infty} P(A_n) = 1$$

Théorème 31 (*Formule des probabilités totales*)

Soit $(A_n)_{n\in\mathbb{N}}$ un **système quasi complet d'événements**, alors pour tout $B \in \mathscr{A}$,

$$P(B) = \sum_{n=0}^{+\infty} P_{A_n}(B)P(A_n) = \sum_{n=0}^{+\infty} P(B \cap A_n)$$

Remarque 32

Pour appliquer la formule des proba totales, il suffit de vérifier que les événements sont deux à deux incompatibles et la somme des probas vaut 1.

Exemple 33

Pour tout $n \in \mathbb{N}$, on note

- U_n = "Après n déplacements, le point se trouve en A_1" ;
- V_n = « Après n déplacements le point se trouve en A_2 » ;
- W_n = « Après n déplacements le point se trouve en A_3 ».

On pose alors, pour tout $n \geq 1$: $u_n = \mathbb{P}(U_n)$, $v_n = \mathbb{P}(V_n)$ et $w_n = \mathbb{P}(W_n)$.

Les conditions initiales sont $u_0 = 1$, $v_0 = 0$, $w_0 = 0$, et grâce à la formule des probabilités totales, on a :

$$\forall n \in \mathbb{N}^*, \quad \begin{cases} u_{n+1} = \frac{1}{5}u_n + \frac{2}{5}v_n + \frac{2}{5}w_n \\ v_{n+1} = \frac{2}{5}u_n + \frac{1}{5}v_n + \frac{2}{5}w_n \\ w_{n+1} = \frac{2}{5}u_n + \frac{2}{5}v_n + \frac{1}{5}w_n \end{cases}$$

On peut alors déterminer les expressions de u_n, v_n et w_n en fonction de n, grâce au calcul matriciel. En effet, on a :

$$\forall n \in \mathbb{N}, \quad \begin{pmatrix} u_{n+1} \\ v_{n+1} \\ w_{n+1} \end{pmatrix} = A \times \begin{pmatrix} u_n \\ v_n \\ w_n \end{pmatrix}$$

où A est la matrice donnée par $A = \frac{1}{5}\begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{pmatrix}$.

On en déduit que : $\forall n \in \mathbb{N}, \begin{pmatrix} u_n \\ v_n \\ w_n \end{pmatrix} = A^n \times \begin{pmatrix} u_0 \\ v_0 \\ w_0 \end{pmatrix}$.

Le problème est donc ramené au calcul des puissances de la matrice A.

Théorème 34 (*Formule de Bayes*)

Soit (A_n) un SCE (ou un quasi-SCE), et $B \in \mathscr{A}$ un événement tel que $P(B) > 0$, alors

1. Pour tout $A \in \mathscr{A}$ tel que $P(A) > 0$,

$$P_B(A) = \frac{P_A(B)P(A)}{P(B)}$$

2.

$$\forall j \in \mathbb{N}, P_B(A_j) = \frac{P_{A_j}(B)P(A_j)}{\sum_{i=1}^{n} P_{A_i}(B)P(A_i)}$$

Avec la convention $P_{A_i}(B) \times P(A_i) = 0$ si $P(A_i) = 0$

Exemple 35

Application classique de la formule de Bayes, si A est un événement non négligeable non certain, avec le SCE $(A, \overline{A})$, $P_B(A) = \dfrac{P_A(B)P(A)}{P(B)} = \dfrac{P_A(B)P(A)}{P_A(B)P(A) + P_{\overline{A}}(B)P(\overline{A})}$

3 Construction de probabilité

▶ **Cas des univers finis**

Si Ω est un ensemble fini de cardinal N, la définition précédente est équivalente à la définition donnée en première année, dans laquelle le deuxième point était remplacé par la propriété :

si A et B sont deux événements incompatibles, $P(A \cup B) = P(A) + P(B)$.

Dans ce cas, on choisit toujours $\mathscr{A} = \mathscr{P}(\Omega)$. On dit alors simplement que le couple (Ω, P) est un espace probabilisé fini. Avec la règle de calcul ci-dessus, la fonction P est entièrement déterminée par la donnée des probabilités des événements élémentaires : pour tout $A \in \mathscr{P}(\Omega)$,

$$P(A) = \sum_{\omega \in A} P(\{\omega\}).$$

On définit la **probabilité uniforme** sur Ω en posant, pour tout $\omega \in \Omega$, $P(\{\omega\}) = 1/N$, c'est-à-dire que tous les événements élémentaires sont équiprobables. C'est le cas dans le deuxième exemple décrit plus haut (lancer de dé). On a alors, pour tout événement A,

$$P(A) = \sum_{\omega \in A} P(\{\omega\}) = \mathrm{card}(A)\,\frac{1}{N} = \frac{\mathrm{card}(A)}{\mathrm{card}(\Omega)},$$

ce que l'on résume souvent ainsi :

$$P(A) = \frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}}.$$

Le fait de choisir la probabilité uniforme est souvent signalé par des expressions comme « la pièce est équilibrée », « le dé est équilibré », « les billes sont indiscernables au toucher et le contenu de l'urne est soigneusement mélangé », etc...

On remarque immédiatement que la situation est plus complexe lorsque l'univers est infini : il n'est pas possible de généraliser la notion précédente de probabilité uniforme.

▶ **Cas des univers dénombrables**

Soit Ω un ensemble dénombrable, avec $\Omega = \{\omega_n;\ n \in \mathbb{N}\}$, et soit $(p_n)_{n\in\mathbb{N}}$ une suite de nombres positifs telle que la série $\sum_{n\geqslant 0} p_n$ soit convergente et de somme 1. Si $A \in \mathscr{P}(\Omega)$, on pose $P(A) = \sum_{\omega_n\in A} p_n$. Alors on pourra vérifier que $(\Omega, \mathscr{P}(\Omega), P)$ est un espace probabilisé, p_n étant pour tout $n \in \mathbb{N}$ la probabilité de l'événement élémentaire $\{\omega_n\}$.

Dans ce qui précède, la notation $\sum_{\omega_n\in A} p_n$ est intuitive, mais lorsque Ω est dénombrable, il convient de l'expliquer. Dans ce cas, A est lui-même fini ou dénombrable, et peut-être décrit en extension sous la forme $(\omega_{\varphi(1)}, \dots, \omega_{\varphi(m)})$ (où $m = \mathrm{card}(A)$) ou $\{\omega_{\varphi(k)};\ k \in \mathbb{N}\}$ (où $\varphi : \mathbb{N} \to \mathbb{N}$ est strictement croissante). Alors $\sum_{\omega_n\in A} p_n$ s'exprime comme une somme finie ou une somme de série convergente :

$$\sum_{\omega_n\in A} p_n = \sum_{k=1}^{m} p_{\varphi(k)} \quad \text{ou} \quad \sum_{\omega_n\in A} p_n = \sum_{k=0}^{+\infty} p_{\varphi(k)}.$$

Par exemple, si $\Omega = \mathbb{N}$ et $A = 2\mathbb{N} = \{2k;\ k \in \mathbb{N}\}$, alors $P(A) = \sum_{k=0}^{+\infty} P(\{2k\})$.

Exemples

• Une personne participe à un jeu dans lequel elle remporte une somme d'argent (un nombre entier naturel d'euros) déterminée de façon aléatoire. On modélise ce jeu de la façon suivante :

$\Omega = \mathbb{N}$, l'événement « la personne gagne n euros » est le singleton $\{n\}$. On pose $p_0 = 0$ et, pour tout $n \in \mathbb{N}^*$,

$$p_n = P(\{n\}) = \frac{1}{2^n}.$$

La série $\sum_{n\geqslant 1} p_n$ (série géométrique de raison $1/2$ et de premier terme $1/2$) converge et

$$\sum_{n=0}^{+\infty} p_n = \sum_{n=1}^{+\infty} \frac{1}{2^n} = \frac{1}{2}\frac{1}{1-1/2} = 1.$$

Le triplet $(\mathbb{N},\mathscr{P}(\mathbb{N}),P)$ est un espace probabilisé modélisant cette expérience.

Considérons l'événement A suivant : « la personne remporte une somme paire ». On a alors

$$A = \{\omega \in \mathbb{N};\ \exists\, k \in \mathbb{N},\, \omega = 2k\} = \bigcup_{k=0}^{+\infty} \{2k\}.$$

Les événements $\{2k\}$ sont deux à deux incompatibles, donc par définition d'une probabilité,

$$P(A) = \sum_{k=0}^{+\infty} P(\{2k\}) = p_0 + \sum_{k=1}^{+\infty} p_{2k} = \sum_{k=1}^{+\infty} \frac{1}{2^{2k}} = \frac{1}{4}\frac{1}{1-1/4} = \frac{1}{3}.$$

Fixons $p \in \mathbb{N}$ et considérons l'événement S_p suivant : « la personne remporte une somme strictement supérieure à p euros ». On a alors

$$S_p = \{\omega \in \mathbb{N};\ \exists\, n \in \mathbb{N};\ n > p,\, \omega = n\} = \bigcup_{n=p+1}^{+\infty} \{n\}.$$

Les événements $\{n\}$ pour $n > p$ sont deux à deux incompatibles, donc

$$P(S_p) = \sum_{n=p+1}^{+\infty} P(\{n\}) = \sum_{n=p+1}^{+\infty} \frac{1}{2^n} = \frac{1}{2^{p+1}}\frac{1}{1-1/2} = \frac{1}{2^p}.$$

La personne a autant de chances de remporter exactement p euros que de remporter une somme au moins égale à $p+1$ euros.

Théorème 36

Si Ω est au plus dénombrable et si $p : \Omega \to [0,1]$ vérifie $\sum_{\omega \in \Omega} p(\omega) = 1$, alors il existe une unique probabilité P sur $\big(\Omega, \mathcal{P}(\Omega)\big)$ telle que $P(\{\omega\}) = p(\omega)$. P est définie par :

$$\forall A \in \mathcal{P}(\Omega) \quad P(A) = \sum_{\omega \in A} p(\omega).$$