Chapitre 9

Espaces probabilisés

Dans de nombreuses situations, une expérience, reproduite plusieurs fois dans des conditions apparemment identiques, peut fournir des résultats différents et imprévisibles. Lorsque l'on lance une pièce en l'air, si l'on avait une parfaite connaissance de toutes les données (sur la pièce, la façon dont elle est lancée, la constitution et le mouvement de l'air ambiant, les équations des différents mouvements, le moment où la personne rattrape la pièce,...) on serait peut-être en mesure de prévoir si le résultat obtenu est « pile » ou « face ». En pratique, une telle connaissance est sans doute impossible, et la moindre variation dans les conditions de l'expérience peut avoir sur le résultat une influence qui le rend impossible à prévoir.

On considère que de tels phénomènes relèvent de l'aléatoire, du hasard (parmi ces phénomènes, on peut aussi citer le comportement de particules physiques, l'évolution du cours de la bourse, la démographie, les jeux de hasard). Pour les étudier, on ne cherche pas à prévoir leur résultat mais on s'attache à mesurer les « chances » ou le « risque » qu'un événement se réalise. La théorie des probabilités donne un cadre mathématique à ce que l'on entend par « expérience aléatoire » et développe des outils permettant l'étude des phénomènes associés.

Dans tout le chapitre, Ω est un ensemble ; $\mathscr{P}(\Omega)$ désigne la collection de toutes les parties de Ω.

I. Ensembles dénombrables

En première année ont été étudiées des expériences aléatoires ayant un nombre fini de résultats possibles. 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 – Ensemble dénombrable

Soit E un ensemble. On dit que E est **dénombrable** si E est en bijection avec $\mathbb{N}$, c'est-à-dire s'il existe une bijection φ de $\mathbb{N}$ sur E.

Dans ce cas, on peut noter, pour tout $n \in \mathbb{N}$, $x_n = \varphi(n)$, et on a donc $E = \{x_n ;\ n \in \mathbb{N}\}$.

C'est ce que l'on appelle décrire E **en extension**.

Exemples

• L'ensemble $\mathbb{N}$ est bien sûr dénombrable (choisir $\varphi = \mathrm{Id}$), c'est en quelque sorte le modèle d'ensemble dénombrable.

• L'ensemble $2\mathbb{N}$ des entiers naturels pairs est dénombrable : $\varphi : n \mapsto 2n$ est une bijection de $\mathbb{N}$ sur $2\mathbb{N}$.

Remarques

• Quitte à faire un changement d'indice, on peut toujours se ramener à une bijection de $\mathbb{N}^*$ sur E dans la définition précédente.

• On montre facilement que les ensembles finis ou dénombrables sont les ensembles qui sont en bijection avec une partie I de $\mathbb{N}$. Dans le cas où E est fini, on peut choisir $I = [\![1,m]\!]$ avec $m = \mathrm{card}(E)$; on peut aussi décrire E en extension sous la forme $E = \{x_1, \ldots, x_m\}$.

Propriété

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

Démonstration – Soit φ l'application ainsi définie : pour tout $k \in \mathbb{N}$,

$$\varphi(2k) = -k, \quad \varphi(2k+1) = k+1.$$

Il s'agit d'une bijection de $\mathbb{N}$ sur $\mathbb{Z}$, ce qui prouve le résultat. $\square$

Propriété

Un produit cartésien d'ensembles dénombrables est dénombrable.

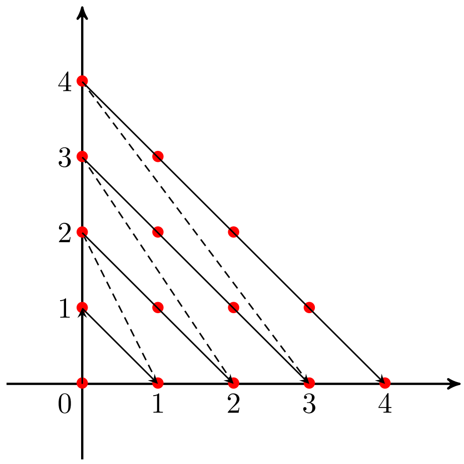
Démonstration – Soient E_1 et E_2 deux ensembles dénombrables, φ une bijection de $\mathbb{N}$ sur E_1, ψ une bijection de $\mathbb{N}$ sur E_2. L'idée est la suivante : si E_1 et E_2 sont décrits en extension sous la forme

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

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. $\square$

Exemples

• Pour tout $n \in \mathbb{N}^*$, $\mathbb{N}^n$, $\mathbb{Z}^n$ sont dénombrables.

• L'ensemble des matrices de $\mathscr{M}_n(\mathbb{R})$ dont les coefficients appartiennent à $\mathbb{Z}$ est en bijection avec $\mathbb{Z}^{n^2}$ qui est dénombrable, il est donc également en bijection avec $\mathbb{N}$, et ainsi dénombrable.

• L'idée mise en œuvre dans la démonstration précédente peut être utilisée pour montrer que $\mathbb{Q}$ est dénombrable. En revanche, $\mathbb{R}$ n'est pas dénombrable.

II. Espaces probabilisés

1. Tribu, probabilité

Modéliser une expérience aléatoire (afin de pouvoir l'étudier), c'est choisir :

- un **ensemble** Ω qui permet de représenter toutes les issues de l'expérience, c'est-à-dire tous les résultats possibles de l'expérience. L'ensemble Ω est appelé **univers**.
- une **probabilité** sur Ω, qui est une fonction ayant certaines propriétés qui font que cette fonction peut être choisie pour mesurer les chances ou le risque qu'un résultat ou ensemble de résultats possibles de l'expérience (ce que l'on appelle sous certaines conditions un **événement**), se réalise.

Exemples

• Une expérience aléatoire ayant deux issues, l'une (interprétée comme succès) de probabilité p, et l'autre (échec) de probabilité $q = 1-p$, est appelée épreuve de Bernoulli de paramètre p. C'est le cas de l'expérience consistant à lancer une fois une pièce non nécessairement équilibrée (avec par exemple, p la probabilité d'obtenir « pile », q celle d'obtenir « face »).

• L'expérience aléatoire consistant à lancer une fois un dé équilibré et à noter le résultat obtenu peut être modélisée de la façon suivante : l'ensemble des issues est $\Omega = \{1,2,3,4,5,6\}$, le fait que le dé soit équilibré se traduit par le choix de la probabilité uniforme sur Ω, c'est-à-dire que chacun des résultats possibles a la probabilité $1/6$ de se réaliser. Le sous-ensemble $\{2,4,6\}$ de Ω est l'événement que l'on peut décrire en français comme « le résultat est un nombre pair ».

• Une personne se lève de façon aléatoire à 7h00 ou 7h05 ou 7h10 ou 7h15. Pour son petit déjeuner, elle choisit au hasard soit des tartines, soit des céréales. En numérotant 1, 2, 3 et 4 les horaires possibles de lever, et en notant T et C les deux petits déjeuners possibles, on peut modéliser l'expérience aléatoire consistant à observer, un jour, l'heure de réveil et le choix de petit déjeuner de cette personne, par le choix de

$$\Omega = \{(1,C),(1,T),(2,C),(2,T),(3,C),(3,T),(4,C),(4,T)\},$$

chaque élément ayant par exemple une probabilité $1/8$ de se produire. Selon la connaissance que l'on a de la situation, on peut bien sûr être amené à choisir des valeurs de probabilités différentes.

On peut bien sûr imaginer des expériences aléatoires plus complexes, par exemple des lancers successifs de pièces jusqu'à obtenir « pile » trois fois de suite, l'observation du déplacement d'un insecte sur une surface plane, la trajectoire d'une balle de tennis. Dans ce cas, déterminer l'ensemble des issues peut être très complexe, cet ensemble peut notamment être *infini*. Pour cette raison, on est amené à préciser ce que l'on entend par événement :

Définition – Tribu

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

- $\Omega \in \mathscr{A}$,
- Pour tout $A \in \mathscr{A}$, le complémentaire de A, *i.e.* $\overline{A} = \Omega \setminus A$, appartient à $\mathscr{A}$.
- Pour toute suite $(A_n)_{n\in\mathbb{N}}$ d'éléments de $\mathscr{A}$, la réunion $\displaystyle\bigcup_{n=0}^{+\infty} A_n$ appartient à $\mathscr{A}$.

Lorsque $\mathscr{A}$ est une tribu sur Ω, l'ensemble Ω est appelé **univers**, et les éléments de $\mathscr{A}$ (qui sont des parties de Ω) sont appelés les **événements**.

Remarques

• Les opérations ensemblistes correspondent bien sûr à des opérations logiques : le passage au complémentaire traduit la négation, la réunion correspond à « ou ». Une tribu rassemble tous les événements **observables** lors de l'expérience aléatoire considérée, et la définition précédente fixe les règles fondamentales de logique permettant de combiner ces événements.

• Il est notamment important de savoir passer de la description d'un événement par une phrase en français à sa description par opérations ensemblistes à partir d'autres événements, et inversement.

• D'après les deux premiers points, $\varnothing = \overline{\Omega}$ est un événement.

• Si $(A_n)_{n\in\mathbb{N}}$ est une suite d'éléments de $\mathscr{A}$,

$$\bigcup_{n=0}^{+\infty} A_n = \{\omega \in \Omega;\ \exists\, n \in \mathbb{N}, \omega \in A_n\}.$$

C'est l'ensemble des éléments de Ω qui appartiennent à au moins l'un des A_n, *i.e.*, l'événement « l'un au moins des événements A_n est réalisé ».

• La collection $\mathscr{P}(\Omega)$ des parties de Ω est une tribu sur Ω (tribu triviale). De même, $\{\varnothing,\Omega\}$ est une tribu sur Ω (tribu grossière).

• Pour modéliser une expérience aléatoire ayant un nombre fini de résultats possibles, on choisit Ω de sorte que $\mathscr{A} = \mathscr{P}(\Omega)$. C'est aussi souvent le cas lorsque l'ensemble des résultats est dénombrable. Par exemple, considérons un dé à six faces sur lequel la face portant le numéro 1 est gravée de façon habituelle, et les autres faces de sorte que le numéro ne soit lisible qu'au microscope. Si l'expérience consistant à lancer une fois ce dé est réalisée avec microscope, on choisira pour univers $\Omega = \{1,2,3,4,5,6\}$, mais si elle est réalisée sans microscope, on choisira par exemple $\Omega = \{1, A\}$, où A représente l'ensemble des autres faces que celle numérotée 1. Dans ce cas, par exemple, 2 ne doit pas être considéré comme une issue, sinon $\{2\}$ serait un événement (puisque $\mathscr{A} = \mathscr{P}(\Omega)$), alors que le résultat 2 n'est pas observable dans les conditions de l'expérience.

• En revanche, dans le cas général, choisir $\mathscr{A} = \mathscr{P}(\Omega)$ est possible mais pas toujours adapté.

Propriété – Stabilité par intersection dénombrable

Soit $\mathscr{A}$ une tribu sur Ω et $(A_n)_{n\in\mathbb{N}}$ une suite d'événements. Alors

$$\bigcap_{n=0}^{+\infty} A_n \in \mathscr{A}.$$

Ainsi, $\mathscr{A}$ est stable par intersection dénombrable : une intersection dénombrable d'événements est un événement.

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, $B \in \mathscr{A}$. □

Remarques

• Avec les notations précédentes,

$$\bigcap_{n=0}^{+\infty} A_n = \{\omega \in \Omega;\ \forall\, n \in \mathbb{N}, \omega \in A_n\}.$$

C'est l'ensemble des éléments de Ω qui appartiennent à tous les A_n, *i.e.*, l'événement « tous les événements A_n sont réalisés ».

• Si $A_0, \dots, A_n$ sont des événements, en posant $A_k = A_n$ pour tout $k \geqslant n+1$, on a

$$\bigcup_{k=0}^{+\infty} A_k = \bigcup_{k=0}^{n} A_k \quad \text{et} \quad \bigcap_{k=0}^{+\infty} A_k = \bigcap_{k=0}^{n} A_k.$$

On voit donc qu'une tribu est également stable par réunion et intersection finie.

Le tableau suivant définit un certain nombre de termes du vocabulaire des probabilités, en parallèle avec le vocabulaire ensembliste :

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

Définition – Système complet d'événements

On appelle **système complet (dénombrable) d'événements** toute suite $(A_n)_{n\in\mathbb{N}}$ d'événements telle que :

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

Remarques

- On définit comme en première année les systèmes complets (finis) d'événements, les A_n étant en nombre fini.
- Un système complet d'événement permet de partitionner l'univers en plusieurs événements, ce qui permet de faire des disjonctions de cas dans les raisonnements.

Exemples

- Si A est un événement, $(A,\overline{A})$ est un système complet d'événements.
- On lance un dé à six faces. Pour $n \in [\![1,6]\!]$, on note A_n l'événement « le numéro obtenu est n ». La famille $(A_i)_{1\leqslant i\leqslant 6}$ est un système complet d'événements.

Définition – Probabilité

Soient Ω un ensemble et $\mathscr{A}$ une tribu sur Ω. On appelle **probabilité** sur $(\Omega, \mathscr{A})$ une application $P : \mathscr{A} \to [0,1]$ telle que :

- $P(\Omega) = 1$,
- Pour toute suite $(A_n)_{n\in\mathbb{N}}$ d'éléments de $\mathscr{A}$ deux à deux incompatibles, la série $\sum_{n\geqslant 0} P(A_n)$ converge et

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

Lorsque P est une probabilité sur $(\Omega, \mathscr{A})$, on dit que le triplet $(\Omega, \mathscr{A}, P)$ est un **espace probabilisé**.
Deux événements A et B tels que $P(A) = P(B)$ sont dits **équiprobables**.

Remarques

• La probabilité d'un événement A s'interprète comme la « mesure » de l'ensemble des issues constituant A relativement à l'ensemble des issues. C'est, de façon imagée, le « poids relatif », la proportion de A dans l'univers Ω.

• Si $(A_n)_{n\in\mathbb{N}}$ est un système complet d'événements, $\sum_{n\geqslant 0} P(A_n)$ converge et a pour somme 1.

▶ 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 :

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

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.

• **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

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

où $k \in \mathbb{N}^*$ et $(u_1, \dots, u_k) \in \{0,1\}^k$ 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,

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

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 $A_0 = \bigcap_{k=1}^{+\infty} C_{u_1,\dots,u_k}$ 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 $A_n = C_{0,\dots,0,1}$ (0 apparaissant $n-1$ fois).

La famille $(A_n)_{n\in\mathbb{N}}$ est un système complet dénombrable d'événements.

• Il existe une tribu $\mathscr{A}$ sur [0,1] qui contient les segments inclus dans [0,1], et une probabilité P sur $(\Omega,\mathscr{A})$ telle que pour tout segment $[a,b]$ inclus dans [0,1], on ait $P([a,b]) = b-a$. L'espace probabilisé $([0,1],\mathscr{A},P)$ peut modéliser par exemple l'expérience consistant à noter le moment où une particule se désintègre, l'intervalle de temps étant ramené à [0,1] si l'on est sûr que la désintégration a lieu avant un temps connu.

Remarques

• Un événement peut tout à fait avoir une probabilité nulle sans être impossible. C'est le cas de tous les singletons dans l'exemple précédent. En particulier, la définition $P(A) = \sum_{\omega\in A} P(\{\omega\})$ est impossible à généraliser dans ce cadre.

• Lorsque Ω n'est pas dénombrable, P n'est presque jamais définie en donnant la probabilité de tous les événements ; on peut par exemple donner (en analysant les conditions de l'expérience) la probabilité d'événements fondamentaux à partir desquels on peut retrouver toutes les probabilités souhaitées, en utilisant les règles de calculs imposées. Dans l'exemple du jeu de pile ou face infini, l'événement A : « le résultat du second lancer est pile » est la réunion des deux événements incompatibles $C_{0,1}$ et $C_{1,1}$, chacun de probabilité $1/4$; on a donc (voir la propriété suivante) $P(A) = 1/2$.

Ce raisonnement se généralise et montre que pour tout $k\in\mathbb{N}^*$, la probabilité d'obtenir k résultats fixés est $1/2^k$ (et en particulier, à chaque lancer, la probabilité d'obtenir « pile » est $1/2$) : en fait, cette modélisation porte en elle le fait que la pièce est équilibrée et que chaque lancer est indépendant de tous les autres (cette notion sera précisée dans la suite).

2. Propriétés élémentaires

Propriété

Soit $(\Omega,\mathscr{A},P)$ un espace probabilisé. Alors :

• $P(\varnothing) = 0$.

• Pour tout événement A, $\quad P(\overline{A}) = 1 - P(A)$.

• Si $n\in\mathbb{N}$ et $A_0,\dots,A_n$ sont des événements deux à deux incompatibles, l'événement $\bigcup_{k=0}^{n} A_k$ vérifie

$$P\left(\bigcup_{k=0}^{n} A_k\right) = \sum_{k=0}^{n} P(A_k).$$

• Si A et B sont des événements avec $A\subset B$, alors $P(A)\leqslant P(B)$.

• Si A et B sont des événements, l'événement $A\cup B$ vérifie

$$P(A\cup B) = P(A) + P(B) - P(A\cap B).$$

• Si $n\in\mathbb{N}$ et $A_0,\dots,A_n$ sont des événements,

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

Démonstration

• Posons $B_n = \varnothing$ pour tout $n\in\mathbb{N}$. Les événements B_n sont deux à deux incompatibles donc $\sum_{n\geqslant 0} P(B_n) = \sum_{n\geqslant 0} P(\varnothing)$ converge. Cette série étant à termes constant, on a $P(\varnothing) = 0$.

• Posons $B_0 = A$, $B_1 = \overline{A}$ et $B_n = \varnothing$ si $n\geqslant 2$. Les B_n sont des événements deux à deux incompatibles donc la série $\sum_{n\geqslant 0} P(B_n)$ converge et

$$P\left(\bigcup_{n=0}^{+\infty} B_n\right) = \sum_{n=0}^{+\infty} P(B_n), \quad \textit{i.e.} \quad P(\Omega) = P(A) + P(\overline{A}) \quad \text{d'après le point 1.}$$

Sachant que $P(\Omega)=1$, on obtient le résultat.

• Posons $A_k=\varnothing$ pour tout $k\geqslant n+1$. Les A_k sont deux à deux incompatibles, donc

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

• 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 $B_0=A\cap\overline{B}$, $B_1=\overline{A}\cap B$, $B_2=A\cap B$. Alors B_0, B_1 et B_2 sont des événements deux à deux incompatibles et $A\cup B=B_0\cup B_1\cup B_2$, donc d'après le point précédent,

$$P(A\cup B)=P(A\cap\overline{B})+P(\overline{A}\cap B)+P(A\cap B).$$

Mais on a également

$$P(A)=P(A\cap\overline{B})+P(A\cap B)\quad\text{et}\quad P(B)=P(\overline{A}\cap B)+P(A\cap B).$$

Ainsi

$$\begin{aligned}P(A\cup B)&=P(A)-P(A\cap B)+P(B)-P(A\cap B)+P(A\cap B)\\&=P(A)+P(B)-P(A\cap B).\end{aligned}$$

On remarquera en particulier que $P(A\cup B)\leqslant P(A)+P(B)$.

• On prouve cette dernière propriété par récurrence sur n, à partir de l'inégalité ci-dessus. □

Propriété – Suites monotones d'événements, sous-additivité

Soient $(\Omega,\mathscr{A},P)$ un espace probabilisé et $(A_n)_{n\in\mathbb{N}}$ une suite d'événements.

• **Continuité croissante :** si pour tout $n\in\mathbb{N}$, $A_n\subset A_{n+1}$, alors

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

• **Continuité décroissante :** si pour tout $n\in\mathbb{N}$, $A_{n+1}\subset A_n$, alors

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

• **Sous-additivité :** si $\sum_{n\geqslant 0}P(A_n)$ converge, alors

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

Démonstration

• 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. □

Exemples

• 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. Probabilités conditionnelles

1. Conditionnement

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

Définition – Probabilité conditionnelle

Soit A un événement tel que $P(A) > 0$. Pour tout événement B, on appelle **probabilité conditionnelle** de B sachant A le réel, noté $P_A(B)$ ou $P(B \mid A)$, défini par

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

On a donc $P(A \cap B) = P(B \mid A)P(A)$.

Remarque – Sachant que $A \cap B \subset A$, on a $P(A \cap B) \leqslant P(A)$, et donc avec le fait que $P(A) > 0$, on a $P_A(B) \in [0,1]$.

Exemple – Reprenons l'exemple du petit déjeuner exposé au début de ce chapitre, avec la probabilité définie par le tableau suivant :

ω	$(1,C)$	$(1,T)$	$(2,C)$	$(2,T)$	$(3,C)$	$(3,T)$	$(4,C)$	$(4,T)$
$P(\{\omega\})$	0,2	0,05	0,1	0,15	0,05	0,3	0,05	0,1

Notons A l'événement « la personne se lève à 7h00 » (*i.e.*, l'ensemble des issues ω dont la première composante est 1) et B l'événement « la personne choisit des céréales » (*i.e.*, l'ensemble des issues ω dont la deuxième composante est C). Alors on a

$$P(A) = 0{,}2 + 0{,}05 = 0{,}25, \quad P(B) = 0{,}2 + 0{,}1 + 0{,}05 + 0{,}05 = 0{,}4$$

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{0{,}2}{0{,}25} = \frac{4}{5}, \quad P(A \mid B) = \frac{P(A \cap B)}{P(B)} = \frac{0{,}2}{0{,}4} = \frac{1}{2}.$$

On notera que le calcul d'une probabilité conditionnelle n'est pas à confondre avec un lien de cause à effet, on peut calculer $P(A \mid B)$ même si la personne se lève avant de déjeuner ! Simplement, quelqu'un arrivant chez cette personne après son déjeuner, voyant un bol vide sur la table (et disposant du tableau précédent), peut affirmer qu'il y a une chance sur deux que la personne se soit levée à 7h00. Sans cette information, on pouvait donner une probabilité $P(A) = 0{,}25$, deux fois moindre !

Propriété/Définition – Probabilité conditionnelle

Soit A un événement tel que $P(A) > 0$. L'application

$$P_A : \begin{cases} \mathscr{A} & \to & [0{,}1] \\ B & \mapsto & P_A(B) \end{cases}$$

est une probabilité sur $(\Omega, \mathscr{A})$, appelée **probabilité conditionnelle** sachant A.

Démonstration – On a remarqué plus haut que P_A est à valeurs dans $[0{,}1]$. On a

$$P_A(\Omega) = \frac{P(\Omega \cap A)}{P(A)} = \frac{P(A)}{P(A)} = 1.$$

Enfin, si $(B_n)_{n\in\mathbb{N}}$ est une suite d'événements deux à deux incompatibles, on a

$$P_A\left(\bigcup_{n=0}^{+\infty} B_n\right) = \frac{1}{P(A)} P\left(A \cap \bigcup_{n=0}^{+\infty} B_n\right) = \frac{1}{P(A)} P\left(\bigcup_{n=0}^{+\infty} (A \cap B_n)\right).$$

Les événements B_n sont deux à deux incompatibles, donc les événements $A \cap B_n$ également ; P étant une probabilité, on a alors

$$P_A\left(\bigcup_{n=0}^{+\infty} B_n\right) = \frac{1}{P(A)} \sum_{n=0}^{+\infty} P(A \cap B_n) = \sum_{n=0}^{+\infty} \frac{P(A \cap B_n)}{P(A)} = \sum_{n=0}^{+\infty} P_A(B_n).$$

On a vérifié les différentes propriétés qui font de P_A une probabilité sur $(\Omega, \mathscr{A})$. $\square$

Remarque – Si $P(A) = 0$, afin que l'égalité $P(A \cap B) = P(B \mid A)P(A)$ reste valable, on pose *par convention* $P(B \mid A)P(A) = 0$ (mais le terme $P(B \mid A)$ seul *n'est pas défini* dans ce cas).

2. Propriétés et utilisation des probabilités conditionnelles

Propriété – Formule des probabilités composées

Soient $A_1, \dots, A_p$ des événements ($p \geqslant 2$) tels que $P(A_1 \cap \dots \cap A_{p-1}) > 0$. Alors

$$P(A_1 \cap \dots \cap A_p) = P(A_1)\, P(A_2 \mid A_1)\, P(A_3 \mid A_1 \cap A_2) \cdots P(A_p \mid A_1 \cap \dots \cap A_{p-1}).$$

Démonstration – On procède par récurrence sur le nombre $p \geqslant 2$ d'événements :

Initialisation ($p = 2$) : cela résulte de la définition de $P(A_2 \mid A_1)$ (on a $P(A_1) > 0$ par hypothèse).

Hérédité : supposons le résultat vrai pour un nombre $p \geqslant 2$ d'événements, et considérons $A_1, \dots, A_{p+1}$ des événements tels que $P(A_1 \cap \dots \cap A_p) > 0$. Alors, par définition

$$P(A_1 \cap \dots \cap A_{p+1}) = P(A_{p+1} \mid A_1 \cap \dots \cap A_p)\, P(A_1 \cap \dots \cap A_p).$$

Or on a également $P(A_1 \cap \dots \cap A_{p-1}) > 0$, et donc par hypothèse de récurrence,

$$P(A_1 \cap \dots \cap A_p) = P(A_1)\, P(A_2 \mid A_1)\, P(A_3 \mid A_1 \cap A_2) \cdots P(A_p \mid A_1 \cap \dots \cap A_{p-1}).$$

Des deux égalités précédentes, on déduit le résultat au rang $p + 1$ et finalement pour tout $p \geqslant 2$ par principe de récurrence. $\square$

Remarque – On réalise parfois des arbres pour représenter une expérience aléatoire. La formule des probabilités composées traduit le fait que la probabilité d'un chemin est le produit des probabilités des arêtes qui le composent.

Exemple – Une personne qui se rend au restaurant prend uniquement un plat une fois sur trois, un menu sinon (événement M). Lorsqu'elle prend un menu, elle choisit de la viande (événement V) une fois sur deux. Dans ce cas, elle prend un café (événement C) trois fois sur quatre.
Les données du problème se traduisent de la manière suivante :

$$P(M) = 1 - \frac{1}{3} = \frac{2}{3}, \quad P(V \mid M) = \frac{1}{2}, \quad P(C \mid M \cap V) = \frac{3}{4}.$$

La probabilité pour que la personne choisisse un menu avec viande puis café est, d'après la formule des probabilités composées,

$$P(M \cap V \cap C) = P(M)P(V \mid M)P(C \mid M \cap V) = \frac{2}{3}\frac{1}{2}\frac{3}{4} = \frac{1}{4}.$$

Propriété – Formule des probabilités totales

Soient $(A_n)_{n\in\mathbb{N}}$ un système complet d'événements et B un événement. Alors la série $\sum_{n\geqslant 0} P(B \cap A_n)$ converge, et on a

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

Le résultat précédent reste valable dans le cas plus général suivant :

$(A_n)_{n\in\mathbb{N}}$ est une suite d'événements deux à deux incompatibles tels que $\displaystyle\sum_{n=0}^{+\infty} P(A_n) = 1$.

Démonstration – Il suffit de faire la démonstration sous la deuxième hypothèse, puisqu'elle est plus générale. On se place donc dans ce cadre. 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

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

Cas particulier – Lorsque A est un événement, $(A, \overline{A})$ est un système complet d'événements, donc pour tout événement B,

$$P(B) = P(B \mid A)P(A) + P(B \mid \overline{A})P(\overline{A}).$$

Propriété – Formules de Bayes

• Soient A et B deux événements tels que $P(B) > 0$. Alors

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)}.$$

• Soient B un événement tel que $P(B) > 0$ et $(A_n)_{n\in\mathbb{N}}$ une suite d'événements deux à deux incompatibles tels que $\sum_{n=0}^{+\infty} P(A_n) = 1$. Alors, pour tout $j \in \mathbb{N}$,

$$P(A_j \mid B) = \frac{P(B \,|A_j)P(A_j)}{\sum_{n=0}^{+\infty} P(B \,|A_n)P(A_n)}.$$

Démonstration

• Sachant que $P(B) > 0$, on peut écrire

$$\frac{P(B \mid A)P(A)}{P(B)} = \frac{P(A \cap B)}{P(B)} = P(A \,|B).$$

• Il suffit de reprendre la même idée en écrivant de plus que $P(B) = \sum_{n=0}^{+\infty} P(B \,|A_n)P(A_n)$ d'après la formule des probabilités totales. □

Exemple – On présente à un candidat trois coffres fermés. L'un des coffres contient un lingot d'or, les deux autres sont vides. Le candidat choisit un coffre ; l'organisateur, qui connaît l'emplacement du lingot, dévoile, parmi les coffres non choisis, un coffre vide (de façon équiprobable lorsque le candidat a choisi le bon coffre). On propose alors au candidat de maintenir son choix ou de changer de coffre, puis d'ouvrir le coffre choisi. Quelle est la meilleure stratégie ?

Numérotons 1 le coffre choisi par le candidat au début du jeu, et 2, 3 les deux autres coffres. Pour $i \in \{2,3\}$, notons C_i l'événement « l'organisateur ouvre le coffre i » et pour $i \in \{1,2,3\}$, L_i l'événement « le lingot se trouve dans le coffre i ».
Le problème revient à comparer $P(L_1 \mid C_2)$ et $P(\overline{L_1} \mid C_2) = 1 - P(L_1 \mid C_2)$. D'après la seconde formule de Bayes,

$$\begin{aligned} P(L_1 \mid C_2) &= \frac{P(C_2 \mid L_1)P(L_1)}{P(C_2 \mid L_1)P(L_1) + P(C_2 \mid L_2)P(L_2) + P(C_2 \mid L_3)P(L_3)} \\ &= \frac{\frac{1}{2}\cdot\frac{1}{3}}{\frac{1}{2}\cdot\frac{1}{3} + 0\cdot\frac{1}{3} + 1\cdot\frac{1}{3}} = \frac{1}{3}. \end{aligned}$$

On obtient le même résultat pour $P(L_1 \mid C_3)$. La meilleure stratégie est donc de changer de coffre !

On remarquera que l'on n'a pas eu besoin de préciser Ω pour répondre à la question, mais simplement de traduire les conditions de l'expérience. On peut souvent admettre l'existence de $(\Omega, \mathscr{A}, P)$.

IV. Événements indépendants

Dans de nombreuses situations, le fait de savoir qu'un événement A est réalisé n'apporte rien pour le calcul de la probabilité d'un événement B. C'est la notion d'événements indépendants :

Propriété/Définition : Événements indépendants

Soient A et B deux événements. On dit que A et B sont **indépendants** si

$$P(A \cap B) = P(A)\,P(B).$$

Si $P(A) > 0$, ceci équivaut à : $P(B \,|\, A) = P(B)$.

L'équivalence des deux propriétés lorsque $P(A) > 0$ est immédiate car

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

On remarquera cependant que la deuxième formulation n'est pas symétrique en A et B, alors que la première l'est.

Définition – Événements mutuellement indépendants

Soient $A_1, \ldots, A_p$ des événements. On dit que $A_1, \ldots, A_p$ sont **mutuellement indépendants** si pour tout sous-ensemble J de $[\![1,p]\!]$, on a

$$P\left(\bigcap_{n \in J} A_n\right) = \prod_{n \in J} P(A_n).$$

Propriété – Indépendance mutuelle / indépendance deux à deux

Des événements $A_1, \ldots, A_p$ mutuellement indépendants sont deux à deux indépendants.

La réciproque est fausse en général : si $n \geqslant 3$, l'indépendance de n événements deux à deux n'entraîne pas leur indépendance mutuelle.

Démonstration – Si $A_1, \ldots, A_p$ sont mutuellement indépendants, alors pour tout $(i,j) \in [\![1,p]\!]$ tels que $i \neq j$, en choisissant $J = \{i,j\}$ dans la définition, on obtient

$$P(A_i \cap A_j) = P(A_i)P(A_j).$$

Donc $A_1, \ldots, A_p$ sont deux à deux indépendants.

En revanche, considérons l'exemple suivant : on dispose de quatre livres, un livre de mathématiques, un livre de physique, un livre de chimie, et un livre mathématiques-physique-chimie. On choisit au hasard, avec la probabilité uniforme, un livre parmi les quatre. Notons M, φ et C les événements « le livre choisi traite notamment de mathématiques » (respectivement physique, chimie). On a

$$P(M \cap \varphi) = P(M \cap C) = P(\varphi \cap C) = \frac{1}{4}$$

$$P(M)P(\varphi) = P(M)P(C) = P(\varphi)P(C) = \left(\frac{2}{4}\right)^2 = \frac{1}{4}$$

donc les événements M, φ et C sont deux à deux indépendants. Pourtant, ils ne sont pas mutuellement indépendants car

$$P(M \cap \varphi \cap C) = \frac{1}{4} \quad \text{et} \quad P(M)P(\varphi)P(C) = \left(\frac{2}{4}\right)^3 = \frac{1}{8}. \qquad \square$$

Remarque – Si A et B sont indépendants, alors A et $\overline{B}$ sont indépendants : en effet,

$$P(A) = P(A \cap B) + P(A \cap \overline{B}) = P(A)P(B) + P(A \cap \overline{B})$$

et donc

$$P(A \cap \overline{B}) = P(A)(1 - P(B)) = P(A)P(\overline{B}).$$

Plus généralement, si $A_1, \ldots, A_p$ sont mutuellement indépendants, et si pour tout $i \in [\![1,p]\!]$, $B_i = A_i$ ou $B_i = \overline{A_i}$, alors $B_1, \ldots, B_p$ sont mutuellement indépendants.

Exemples

• Lors d'un parcours à vélo, les événements « le trajet est parcouru en moins de n minutes » et l'événement « il y a un vent de face de 40 km/h » ne sont sans doute pas toujours indépendants !

• L'indépendance entre événements relève parfois de la *modélisation* : on postule que certains événements fondamentaux sont indépendants.

Par exemple, dans un jeu de pile ou face, on considère dans la plupart des cas que les lancers sont mutuellement indépendants. Ce type d'expérience sera d'ailleurs plutôt modélisé ainsi, en faisant l'hypothèse qu'à chaque lancer, « pile » et « face » ont des probabilités d'apparition respectives p et $q = 1 - p$, et l'hypothèse d'indépendance mutuelle des lancers.

Lorsque $p = q = 1/2$, le fait que pour tout $n \in \mathbb{N}^*$, les événements consistant à fixer les résultats des n premiers lancers aient pour probabilité $1/2^n$, est alors une *conséquence* de cette modélisation, ce qui est une démarche peut-être plus naturelle que de postuler ces probabilités.

Par exemple, l'événement « pile apparaît pour la première fois au n-ième lancer » a pour probabilité $1/2^n$ (car il correspond à $n-1$ premiers résultats « face » suivis d'un résultat « pile »). De plus, l'événement « tous les lancers donnent face » est de probabilité nulle : pour tout $n \in \mathbb{N}^*$, cet événement est inclus dans un événement de probabilité $1/2^n$, celui consistant à fixer n premiers résultats « face ». Il suffit alors de faire tendre n vers $+\infty$.