Chapitre 13. Probabilités discrètes

Plan du chapitre

1 Probabilités

1.1 Tribus

Il s'agit de généraliser la notion d'**événement** exposée en maths sup. En maths sup, l'univers des possibles Ω est fini et un événement (qui aura ensuite une probabilité une fois celle-ci définie) est une partie quelconque de Ω. La probabilité d'un événement est alors la somme finie des probabilités des événements élémentaires constituant cet événement.

En maths spé, nous passons à la situation plus générale où Ω est un ensemble quelconque et les probabilités à calculer sont des sommes (discrètes) de familles sommables. Quand on continuera à généraliser (après la maths spé), les sommes discrètes deviendront des sommes continues c'est-à-dire des intégrales et on sait qu'il existe des fonctions non intégrables même sur un segment. Il n'est donc plus possible à terme, de calculer la probabilité de n'importe quelle partie de Ω. La définition 1 ci-dessous, met en place la notion qui définit axiomatiquement les événements, dont on calculera la probabilité : c'est la notion de **tribu**. En maths spé, cette notion n'est pas encore absolument indispensable dans la pratique mais son introduction est prévue par le programme officiel.

Définition 1. Soit Ω un ensemble non vide.

Une **tribu** ou **σ-algèbre** sur Ω est un sous-ensemble $\mathscr{A}$ de $\mathscr{P}(\Omega)$ vérifiant :

1) $\mathscr{A} \neq \varnothing$,

2) $\forall A \in \mathscr{A}$, ${}^cA \in \mathscr{A}$ (stabilité par passage au complémentaire),

3) $\forall (A_n)_{n\in\mathbb{N}} \in \mathscr{A}^{\mathbb{N}}$, $\displaystyle\bigcup_{n=0}^{+\infty} A_n \in \mathscr{A}$ (stabilité par réunion dénombrable).

Un élément de $\mathscr{A}$ s'appelle alors un **événement** et le couple $(\Omega, \mathscr{A})$ est un **espace probabilisable**.

Exemples.

• $\mathscr{A} = \{\varnothing, \Omega\}$ est une tribu sur Ω appelée **tribu grossière**.
• $\mathscr{A} = \mathscr{P}(\Omega)$ est une tribu sur Ω.
• Soit $A \in \Omega$. $\mathscr{A} = \{\varnothing, A, {}^cA, \Omega\}$ est une tribu sur Ω.
• On peut montrer qu'une intersection de tribus est une tribu. En conséquence, si $\mathscr{B}$ est un sous-ensemble quelconque de $\mathscr{P}(\Omega)$, on peut montrer qu'il existe une plus petite tribu (au sens de l'inclusion) contenant $\mathscr{B}$. Elle est obtenue comme intersection de toutes les tribus contenant $\mathscr{B}$ et s'appelle la **tribu engendrée par** $\mathscr{B}$. ❏

Les axiomes de la définition 1 entrainent un certain nombre de propriétés obligatoirement vérifiées par une tribu :

Théorème 1. Soit Ω un ensemble non vide au plus dénombrable, c'est-à-dire fini ou infini et dénombrable. Soit $\mathscr{A}$ une tribu sur Ω.

1) $\Omega \in \mathscr{A}$ et $\varnothing \in \mathscr{A}$.

2) $\mathscr{A}$ est stable par réunion finie : $\forall n \in \mathbb{N}^*$, $\forall (A_k)_{1\leqslant k\leqslant n} \in \mathscr{A}^n$, $\displaystyle\bigcup_{k=1}^{n} A_k \in \mathscr{A}$.

3) $\mathscr{A}$ est stable par intersection finie ou dénombrable.

4) $\forall (A, B) \in \mathscr{A}^2$, $B \setminus A \in \mathscr{A}$.

Démonstration .

1) D'après l'axiome 1, $\mathscr{A} \neq \varnothing$. Donc, il existe $A \in \mathscr{A}$. D'après l'axiome 2, ${}^cA \in \mathscr{A}$. Donc, si on pose $A_0 = {}^cA$ et pour $n \geqslant 1$, $A_n = A$, alors d'après l'axiome 3, $\Omega = \displaystyle\bigcup_{n=0}^{+\infty} A_n \in \mathscr{A}$.

Mais alors, d'après l'axiome 2, $\varnothing = {}^c\Omega \in \mathscr{A}$.

2) Soient $n \in \mathbb{N}^*$ puis $(A_k)_{1\leqslant k\leqslant n} \in \mathscr{A}^n$. Pour $k \in \mathbb{N}^*$, on pose $B_k = A_k$ si $1 \leqslant k \leqslant n$ et $B_k = \varnothing$ si $k > n$ ou si $k = 0$. Alors, $\displaystyle\bigcup_{k=1}^{n} A_k = \bigcup_{k=0}^{+\infty} B_k \in \mathscr{A}$.

3) Soit $(A_n)_{n\in\mathbb{N}} \in \mathscr{A}^{\mathbb{N}}$. Pour $n \in \mathbb{N}$, on pose $B_n = {}^cA_n$. Alors, pour tout $n \in \mathbb{N}$, $B_n \in \mathscr{A}$ puis

$$\bigcap_{n=0}^{+\infty} A_n = \bigcap_{n=0}^{+\infty} {}^cB_n = {}^c\left(\bigcup_{n=0}^{+\infty} B_n\right) \in \mathscr{A}.$$

4) Soit $(A, B) \in \mathscr{A}^2$. Alors, $B \setminus A = B \cap {}^cA \in \mathscr{A}$.

❏

Le vocabulaire sur les événements mis en place en première année se généralise à l'identique. Par exemple, deux événements A et B sont **disjoints** ou **incompatibles** si et seulement si $A \cap B = \varnothing$, Ω est l'**événement certain** et $\varnothing$ est l'**événement impossible** ...

1.2 Espaces probabilisés

On définit maintenant la notion de probabilité sur un espace probabilisable :

DÉFINITION 2. Soit $(\Omega, \mathscr{A})$ un espace probabilisable.

Une **probabilité** sur $(\Omega, \mathscr{A})$ est une application $\mathbb{P}$ de $\mathscr{A}$ dans $[0, +\infty[$ vérifiant de plus :

1) $\mathbb{P}(\Omega) = 1$,

2) Pour toute suite $(A_n)_{n \in \mathbb{N}} \in \mathscr{A}^{\mathbb{N}}$ d'événements deux à deux disjoints,

$$\mathbb{P}\left(\bigcup_{n=0}^{+\infty} A_n\right) = \sum_{n=0}^{+\infty} P\left(A_n\right) \text{ (σ-additivité de $\mathbb{P}$).}$$

Un **espace probabilisé** est un triplet $(\Omega, \mathscr{A}, \mathbb{P})$ où $(\Omega, \mathscr{A})$ est un espace probabilisable et $\mathbb{P}$ est une probabilité sur $(\Omega, \mathscr{A})$.

⇛ **Commentaire**. *La définition ci-dessus contient implicitement le fait que pour toute suite $(A_n)_{n \in \mathbb{N}}$ d'événements deux à deux disjoints, la série de terme général $\mathbb{P}(A_n)$, $n \in \mathbb{N}$, converge. En particulier, il est obligatoire que pour toute suite $(A_n)_{n \in \mathbb{N}}$ d'événements deux à deux disjoints, on ait $\lim\limits_{n \to +\infty} \mathbb{P}(A_n) = 0$.*

1.3 Propriétés des probabilités

On a déjà toutes les formules de maths sup qui se généralise à l'identique :

Théorème 2. Soit $(\Omega, \mathscr{A}, \mathbb{P})$ un espace probabilisé.

1) $\mathbb{P}(\varnothing) = 0$.

2) a) $\forall (A, B) \in \mathscr{A}^2$, $(A \cap B = \varnothing \Rightarrow \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B))$.
b) Plus généralement, pour tout $n \in \mathbb{N}^*$ puis tout n-uplet $(A_k)_{1 \leqslant k \leqslant n} \in \mathscr{A}^n$ d'événements deux à deux disjoints,

$$\mathbb{P}\left(\bigcup_{k=1}^{n} A_k\right) = \sum_{k=1}^{n} \mathbb{P}\left(A_k\right).$$

3) a) $\forall A \in \mathscr{A}$, $\mathbb{P}\left({}^cA\right) = 1 - \mathbb{P}(A)$.
b) $\forall A \in \mathscr{A}$, $\mathbb{P}(A) \in [0, 1]$.

4) a) $\forall (A, B) \in \mathscr{A}^2$, $A \subset B \Rightarrow \mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A)$.
b) $\forall (A, B) \in \mathscr{A}^2$, $A \subset B \Rightarrow \mathbb{P}(A) \leqslant \mathbb{P}(B)$ (croissance d'une probabilité).

5) $\forall (A, B) \in \mathscr{A}^2$, $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.

6) Pour tout $n \geqslant 2$ puis pour tout $(A_k)_{1 \leqslant k \leqslant n} \in \mathscr{A}^n$, $\mathbb{P}\left(\bigcup\limits_{k=1}^{n} A_k\right) \leqslant \sum\limits_{k=1}^{n} \mathbb{P}\left(A_k\right)$.

DÉMONSTRATION.

1) Pour $n \in \mathbb{N}$, on pose $A_n = \varnothing$.

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

et donc $\sum\limits_{n=1}^{+\infty} \mathbb{P}(\varnothing) = 0$ puis $\mathbb{P}(\varnothing) = 0$ (car $\mathbb{P}(\varnothing) \geqslant 0$).

2) a) Soit $(A, B) \in \mathscr{A}^2$ tel que $A \cap B = \varnothing$. On pose $A_0 = A$, $A_1 = B$ et pour $n \geqslant 2$, $A_n = \varnothing$. Les A_n, $n \in \mathbb{N}$, sont deux à deux disjoints et donc

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

b) Par récurrence sur n. Le cas $n = 2$ est a).

3) a) Soit $A \in \mathscr{A}$. Puisque $A \cap {}^cA = \varnothing$,

$$\mathbb{P}(A) + \mathbb{P}\left({}^cA\right) = \mathbb{P}\left(A \cup {}^cA\right) = \mathbb{P}(\Omega) = 1.$$

b) Soit $A \in \mathscr{A}$. $\mathbb{P}(A) = 1 - \mathbb{P}\left({}^cA\right) \leqslant 1$.

4) a) Soit $(A, B) \in \mathscr{A}^2$ tel que $A \subset B$. Alors $B = (B \setminus A) \cup A$ avec $(B \setminus A) \cap A = \varnothing$. Donc,

$$\mathbb{P}(A) + \mathbb{P}(B \setminus A) = \mathbb{P}(A \cup (B \setminus A)) = \mathbb{P}(B).$$

b) Soit $(A, B) \in \mathscr{A}^2$ tel que $A \subset B$. $\mathbb{P}(B) = \mathbb{P}(A) + \mathbb{P}(B \setminus A) \geqslant \mathbb{P}(A)$.

5) Soit $(A, B) \in \mathscr{A}^2$. $(B \setminus (A \cap B)) \cap A = \varnothing$ avec $A \cap B \subset B$ et donc

$$\mathbb{P}(A \cup B) = \mathbb{P}(A \cup (B \setminus (A \cap B))) = \mathbb{P}(A) + \mathbb{P}(B \setminus (A \cap B)) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B).$$

6) Montrons par récurrence que pour tout $n \geqslant 2$ puis pour tout $(A_k)_{1 \leqslant k \leqslant n} \in \mathscr{A}^n$, $\mathbb{P}\left(\bigcup_{k=1}^{n} A_k\right) \leqslant \sum_{k=1}^{n} \mathbb{P}(A_k)$.

- Soient A_1 et A_2 deux événements.

$$\mathbb{P}(A_1 \cup A_2) = \mathbb{P}(A_1) + \mathbb{P}(A_2) - \mathbb{P}(A_1 \cap A_2) \leqslant \mathbb{P}(A_1) + \mathbb{P}(A_2).$$

- Soit $n \geqslant 2$. Supposons que pour tout $(A_k)_{1 \leqslant k \leqslant n} \in \mathscr{A}^n$, $\mathbb{P}\left(\bigcup_{k=1}^{n} A_k\right) \leqslant \sum_{k=1}^{n} \mathbb{P}(A_k)$. Soit $(A_k)_{1 \leqslant k \leqslant n+1} \in \mathscr{A}^{n+1}$.

$$\begin{aligned}
\mathbb{P}\left(\bigcup_{k=1}^{n+1} A_k\right) &= \mathbb{P}\left(\left(\bigcup_{k=1}^{n} A_k\right) \cup A_{n+1}\right) \\
&\leqslant \mathbb{P}\left(\bigcup_{k=1}^{n} A_k\right) + \mathbb{P}(A_{n+1}) \text{ (d'après le cas } n = 2) \\
&\leqslant \sum_{k=1}^{n+1} \mathbb{P}(A_k) \text{ (par hypothèse de récurrence).}
\end{aligned}$$

On a montré par récurrence que pour tout $n \geqslant 2$ puis pour tout $(A_k)_{1 \leqslant k \leqslant n} \in \mathscr{A}^n$, $\mathbb{P}\left(\bigcup_{k=1}^{n} A_k\right) \leqslant \sum_{k=1}^{n} \mathbb{P}(A_k)$. ❏

En maths spé, on complète ces propriétés avec :

Théorème 3. Soit $(\Omega, \mathscr{A}, \mathbb{P})$ un espace probabilisé au plus dénombrable. Soit $(A_n)_{n \in \mathbb{N}}$ une suite d'événements.

1) (continuité croissante) Si la suite $(A_n)_{n \in \mathbb{N}}$ est croissante pour l'inclusion, alors la suite numérique $(\mathbb{P}(A_n))_{n \in \mathbb{N}}$ converge et de plus,

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

2) (continuité décroissante) Si la suite $(A_n)_{n \in \mathbb{N}}$ est décroissante pour l'inclusion, alors la suite numérique $(\mathbb{P}(A_n))_{n \in \mathbb{N}}$ converge et de plus,

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

Démonstration .

1) Soit $(A_n)_{n \in \mathbb{N}}$ une suite d'événements, croissante pour l'inclusion. Donc, $\forall k \in \mathbb{N}$, $A_k \subset A_{k+1}$.
On se ramène au cas d'événements deux à deux disjoints : on pose $B_0 = A_0$ et pour $k \in \mathbb{N}^*$, on pose $B_k = A_k \setminus A_{k-1}$. Les B_k sont deux à disjoints et, par récurrence,

$$\forall n \in \mathbb{N}^*,\ A_n = A_0 \cup \left(\bigcup_{k=1}^{n} (A_k \setminus A_{k-1})\right) = \bigcup_{k=0}^{n} B_k.$$

Ceci fournit déjà $\bigcup_{k=0}^{+\infty} B_k = \bigcup_{k=0}^{+\infty} A_k$ et d'autre part, pour $n \in \mathbb{N}^*$,

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

Puisque $(B_k)_{k\in\mathbb{N}}$ est une suite d'événements deux à deux disjoints, la suite $(\mathbb{P}(A_n))_{n\in\mathbb{N}} = \left(\sum_{k=0}^{n} \mathbb{P}(B_k)\right)_{n\in\mathbb{N}}$ converge et

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

2) Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'événements, décroissante pour l'inclusion. Si pour $n \in \mathbb{N}$, on pose $A'_n = {}^cA_n$, la suite $(A'_n)_{n\in\mathbb{N}}$ est une suite d'événements, croissante pour l'inclusion. D'après 1), la suite $(\mathbb{P}(A'_n))$ converge et donc la suite $(\mathbb{P}(A_n))_{n\in\mathbb{N}} = (1 - \mathbb{P}(A'_n))_{n\in\mathbb{N}}$ converge. De plus,

$$\begin{aligned}
\lim_{n\to+\infty} \mathbb{P}(A_n) &= 1 - \lim_{n\to+\infty} \mathbb{P}(A'_n) \\
&= 1 - \mathbb{P}\left(\bigcup_{k=0}^{+\infty} A'_k\right) = 1 - \mathbb{P}\left(\bigcup_{k=0}^{+\infty} ({}^cA_k)\right) = 1 - \mathbb{P}\left({}^c\left(\bigcap_{k=0}^{+\infty} A_k\right)\right) \\
&= \mathbb{P}\left(\bigcap_{k=0}^{+\infty} A_k\right).
\end{aligned}$$

❑

⇛ **Commentaire** . *Dans la démonstration précédente, dans le cas d'une suite d'événements $(A_n)_{n\in\mathbb{N}}$ croissante pour l'inclusion, on aurait également pu constater que la suite $(\mathbb{P}(A_n))_{n\in\mathbb{N}}$ était une suite numérique croissante, majorée par 1 et donc convergente.*

Une conséquence du théorème 3 est :

Théorème 3 bis. Soit $(\Omega, \mathscr{A}, \mathbb{P})$ un espace probabilisé. Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'événements. Alors,

- La suite $\left(\mathbb{P}\left(\bigcup_{k=0}^{n} A_k\right)\right)_{n\in\mathbb{N}}$ converge et

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

- La suite $\left(\mathbb{P}\left(\bigcap_{k=0}^{n} A_k\right)\right)_{n\in\mathbb{N}}$ converge et

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

Démonstration . On applique le théorème 3 à la suite $(B_n)_{n\in\mathbb{N}} = \left(\bigcup_{k=0}^{n} A_k\right)_{n\in\mathbb{N}}$ (resp. $(B_n)_{n\in\mathbb{N}} = \left(\bigcap_{k=0}^{n} A_k\right)_{n\in\mathbb{N}}$) qui est croissante (resp. décroissante) pour l'inclusion.

❑

Théorème 4. Soit $(\Omega, \mathscr{A}, \mathbb{P})$ un espace probabilisé. Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'événements. Alors,

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

Démonstration . Pour $n \in \mathbb{N}$, posons $A'_n = \bigcup_{k=0}^{n} A_k$ de sorte que pour $n \in \mathbb{N}$, $\bigcup_{k=0}^{n} A'_k = \bigcup_{k=0}^{n} A_k$.

La suite $(A'_n)_{n\in\mathbb{N}}$ est une suite croissante pour l'inclusion. Donc, la suite $(\mathbb{P}(A'_n))_{n\in\mathbb{N}}$ converge vers $\mathbb{P}\left(\bigcup_{k=0}^{+\infty} A'_k\right) = \mathbb{P}\left(\bigcup_{k=0}^{+\infty} A_k\right)$ d'après le théorème 3. D'autre part, pour $n \in \mathbb{N}$, d'après le théorème 2,

$$\mathbb{P}(A'_n) = \mathbb{P}\left(\bigcup_{k=0}^{n} A_k\right) \leqslant \sum_{k=0}^{n} \mathbb{P}(A_k) \quad (*).$$

La série numérique de terme général $\mathbb{P}(A_k)$, $n \in \mathbb{N}$, à termes positifs, converge dans $\overline{\mathbb{R}}$ vers un réel positif ou vers $+\infty$. Quand n tend vers $+\infty$ dans l'inégalité $(*)$, on obtient

$$\mathbb{P}\left(\bigcup_{k=0}^{+\infty} A_k\right) \leqslant \sum_{k=0}^{+\infty} \mathbb{P}(A_k), \text{ (le second membre étant réel ou infini).}$$

❏

Remarque. Dans la planche d'exercices n° 41 de maths sup (« Dénombrements »), on détaille la *formule du crible* qui donne le calcul de card $\left(\bigcup_{k=0}^{n} A_k\right)$. Cette démonstration et la formule qui en résulte, peuvent être reproduite à l'identique pour les probabilités en remplaçant card par $\mathbb{P}$.

1.4 Evénements négligeables, événements presque sûrs

$\varnothing$ est l'événement impossible. Sa probabilité est nulle. Mais il dorénavant possible de trouver des événements non vides de probabilité nulle. Par exemple, si on choisit un entier au hasard, la probabilité que cet entier soit 0 est nulle (cette phrase mériterait d'être largement détaillée et précisée).
De même, $\mathbb{P}(\Omega) = 1$ mais $\mathbb{P}(A) = 1 \not\Rightarrow A = \Omega$. Ceci nous conduit à poser la définition suivante :

DÉFINITION 4. Soit $(\Omega, \mathscr{A}, \mathbb{P})$ un espace probabilisé au plus dénombrable. Soit $A \in \mathscr{A}$.

A est un **événement négligeable** si et seulement si $\mathbb{P}(A) = 0$.
A est un **événement presque sûr** si et seulement si $\mathbb{P}(A) = 1$.

Le vocabulaire mis en place montre que la notion est délicate. Il est pénible de se dire qu'un événement a 100 chances sur 100 de se réaliser et pourtant que cet événement n'est que presque sûr mais n'est pas certain.

Théorème 5. Soit $(\Omega, \mathscr{A}, \mathbb{P})$ un espace probabilisé. Soit $(A, B) \in \mathscr{A}^2$.

- Si B est négligeable, alors $\mathbb{P}(A \cap B) = 0$ et $\mathbb{P}(A \cup B) = \mathbb{P}(A)$.
- Si B est presque sûr, alors $\mathbb{P}(A \cap B) = \mathbb{P}(A)$ et $\mathbb{P}(A \cup B) = 1$.

DÉMONSTRATION .

- Supposons $\mathbb{P}(B) = 0$. Puisque $A \cap B \subset B$, $\mathbb{P}(A \cap B) \leqslant \mathbb{P}(B) = 0$ puis $\mathbb{P}(A \cap B) = 0$. Mais alors,

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

- Supposons $\mathbb{P}(B) = 1$. Puisque $B \subset A \cup B$, $1 = \mathbb{P}(B) \leqslant \mathbb{P}(A \cup B)$ puis $\mathbb{P}(A \cup B) = 1$. Mais alors,

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

❏

Théorème 6. Soit $(\Omega, \mathscr{A}, \mathbb{P})$ un espace probabilisé au plus dénombrable.

Une réunion finie ou dénombrable d'événements négligeables est un événement négligeable.
Une intersection finie ou dénombrable d'événements presque sûrs est un événement presque sûr.

DÉMONSTRATION .

- Soit $(A_n)_{n\in\mathbb{N}} \in \mathscr{A}^{\mathbb{N}}$. Supposons chaque A_n, $n \in \mathbb{N}$, négligeable. D'après le théorème 4 de la page précédente,

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

et donc $\mathbb{P}\left(\bigcup_{n=0}^{+\infty} A_n\right) = 0$. Ceci démontre le résultat pour une réunion dénombrable d'événements négligeables. Le cas d'une famille finie d'événements négligeables s'obtient en prenant les A_n vides à partir d'un certain rang.

- Soit $(A_n)_{n\in\mathbb{N}} \in \mathscr{A}^{\mathbb{N}}$. Supposons chaque A_n, $n \in \mathbb{N}$, presque sûr. On en déduit que chaque ${}^cA_n$ est négligeable et donc que ${}^c\left(\bigcap_{n=0}^{+\infty} A_n\right) = \bigcup_{n=0}^{+\infty} {}^cA_n$ est négligeable puis que $\bigcap_{n=0}^{+\infty} A_n$ est un événement presque sûr.

❏

Exercice 1. Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'événements d'un espace probabilisé $(\Omega, \mathscr{A}, \mathbb{P})$ tel que la série de terme général $\mathbb{P}(A_n)$, $n \in \mathbb{N}$, converge.

Montrer que l'événement $\bigcap_{n=0}^{+\infty}\left(\bigcup_{p=n}^{+\infty} A_p\right)$ est négligeable.

Solution 1. Pour $n \in \mathbb{N}$, posons $B_n = \bigcup_{p=n}^{+\infty} A_p$. $(B_n)_{n\in\mathbb{N}}$ est une suite d'événements, décroissante pour l'inclusion. Par continuité décroissante,

$$\mathbb{P}\left(\bigcap_{n=0}^{+\infty}\left(\bigcup_{p=n}^{+\infty} A_p\right)\right) = \mathbb{P}\left(\bigcap_{n=0}^{+\infty} B_n\right) = \lim_{n\to+\infty} \mathbb{P}(B_n).$$

Maintenant, pour $n \in \mathbb{N}$, $\mathbb{P}(B_n) = \mathbb{P}\left(\bigcup_{p=n}^{+\infty} A_p\right) \leqslant \sum_{p=n}^{+\infty} \mathbb{P}(A_n) = R_n$. Pour $n \geqslant 1$, R_n est le reste à l'ordre $n-1$ de la série de terme général $\mathbb{P}(A_n)$, $n \in \mathbb{N}$, qui est convergente. On sait alors que pour chaque $n \in \mathbb{N}^*$, R_n existe dans $\mathbb{R}$ puis que $\lim_{n\to+\infty} R_n = 0$. On en déduit que $\lim_{n\to+\infty} \mathbb{P}(B_n) = 0$ puis que

$$\mathbb{P}\left(\bigcap_{n=0}^{+\infty}\left(\bigcup_{p=n}^{+\infty} A_p\right)\right) = 0.$$

L'événement $\bigcap_{n=0}^{+\infty}\left(\bigcup_{p=n}^{+\infty} A_p\right)$ est donc négligeable.

1.5 Systèmes complets ou quasi-complets d'événements

DÉFINITION 5. Soit $(\Omega, \mathscr{A}, \mathbb{P})$ un espace probabilisé au plus dénombrable. Soit $(A_n)_{n\in\mathbb{N}} \in \mathscr{A}^{\mathbb{N}}$.

$(A_n)_{n\in\mathbb{N}}$ est un **système complet d'événements** si les événements A_n, $n \in \mathbb{N}$, sont deux à deux disjoints et $\bigcup_{n=0}^{+\infty} A_n = \Omega$.

DÉFINITION 5 BIS. Soit $(\Omega, \mathscr{A}, \mathbb{P})$ un espace probabilisé au plus dénombrable. Soit $(A_n)_{n\in\mathbb{N}} \in \mathscr{A}^{\mathbb{N}}$.

$(A_n)_{n\in\mathbb{N}}$ est un **système quasi-complet d'événements** si les événements A_n, $n \in \mathbb{N}$, sont deux à deux disjoints et $\sum_{n=0}^{+\infty} \mathbb{P}(A_n) = 1$ (ou encore $\mathbb{P}\left(\bigcup_{n=0}^{+\infty} A_n\right) = 1$ ou encore $\bigcup_{n=0}^{+\infty} A_n$ est un événement presque sûr).

Un système complet d'événements est en particulier un système quasi-complet d'événements. L'intérêt d'un système quasi-complet d'événements réside dans le théorème suivant :

Théorème 7. Soit $(\Omega, \mathscr{A}, \mathbb{P})$ un espace probabilisé. Soit $(A_n)_{n\in\mathbb{N}} \in \mathscr{A}^{\mathbb{N}}$ un système quasi-complet d'événements. Soit $B \in \mathscr{A}$. Alors,

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

DÉMONSTRATION. Puisque $\bigcup_{n=0}^{+\infty} A_n$ est presque sûr, on a $\mathbb{P}\left(B \cap \bigcup_{n=0}^{+\infty} A_n\right) = \mathbb{P}(B)$ d'après le théorème 5. Mais $B \cap \bigcup_{n=0}^{+\infty} A_n = \bigcup_{n=0}^{+\infty} (B \cap A_n)$ et de plus, les événements $B \cap A_n$, $n \in \mathbb{N}$, sont deux à deux disjoints. On en déduit que

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

❑

1.6 Distribution de probabilités discrète sur Ω

Dans ce le théorème qui suit, on se place dans le cas particulier où $\mathscr{A} = \mathscr{P}(\Omega)$.

Théorème 8. Soit Ω un univers non vide.

1) Soit $\mathbb{P}$ une probabilité sur l'espace probabilisable $(\Omega, \mathscr{P}(\Omega))$. Pour $\omega \in \Omega$, on pose $p_\omega = \mathbb{P}(\{\omega\})$. Alors, $(p_\omega)_{\omega\in\Omega}$ est une famille de réels positifs, sommable, de somme 1. De plus, pour $A \in \mathscr{P}(\Omega)$,

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

2) On suppose de plus Ω est au plus dénombrable. Soit $(p_\omega)_{\omega\in\Omega}$ une famille de réels positifs indexée par Ω, sommable, de somme 1. Il existe une probabilité $\mathbb{P}$ et une seule sur l'espace probabilisable $(\Omega, \mathscr{P}(\Omega))$ telle que

$$\forall \omega \in \Omega,\ \mathbb{P}(\{\omega\}) = p_\omega.$$

Démonstration .

1) Les p_ω, $\omega \in \Omega$, sont des réels positifs. Ensuite, $\Omega = \bigcup_{\omega\in\Omega} \{\omega\}$ et donc, puisque les $\{\omega\}$, $\omega \in \Omega$, sont deux à deux disjoints,

$$1 = \mathbb{P}(\Omega) = \sum_{\omega\in\Omega} \mathbb{P}(\{\omega\}) = \sum_{\omega\in\Omega} p_\omega.$$

Ceci montre que la famille $(p_\omega)_{\omega\in\Omega}$ est sommable, de somme 1. Dit autrement, $(\{\omega\})_{\omega\in\Omega}$ est un système complet d'événements. D'après le théorème 7, pour $A \in \mathscr{P}(\Omega)$,

$$\mathbb{P}(A) = \sum_{\omega\in\Omega} \mathbb{P}(A \cap \{\omega\}) = \sum_{\omega\in A} \mathbb{P}(\{\omega\}) = \sum_{\omega\in A} p_\omega.$$

2) Soit $(p_\omega)_{\omega\in\Omega}$ une famille de réels positifs indexée par Ω au plus dénombrable, sommable, de somme 1. D'après 1), si $\mathbb{P}$ est une probabilité sur $(\Omega, \mathscr{P}(\Omega))$ telle que $\forall \omega \in \Omega$, $\mathbb{P}(\{\omega\}) = p_\omega$, on a nécessairement

$$\forall A \in \mathscr{P}(\Omega),\ \mathbb{P}(A) = \sum_{\omega\in A} p_\omega,$$

ce qui montre l'unicité de $\mathbb{P}$. Il s'agit alors de vérifier que $\mathbb{P}$ ainsi définie, est bien une probabilité.

Pour chaque $A \in \mathscr{P}(\Omega)$, la famille $(p_\omega)_{\omega\in A}$ est sommable en tant que sous-famille de la famille sommable $(p_\omega)_{\omega\in\Omega}$ et de plus

$$0 \leqslant \mathbb{P}(A) = \sum_{\omega\in A} p_\omega \leqslant \sum_{\omega\in\Omega} p_\omega = 1.$$

Donc, $\mathbb{P}$ est une application de $\mathscr{P}(\Omega)$ dans $[0, 1]$. Ensuite,

$$\mathbb{P}(\Omega) = \sum_{\omega\in\Omega} p_\omega = 1.$$

Soit enfin $(A_n)_{n\in\mathbb{N}}$ une suite d'éléments de $\mathscr{P}(\Omega)$ deux à deux disjoints. L'événement $A = \bigcup_{n=0}^{+\infty} A_n$ est une partie de l'ensemble dénombrable Ω et donc $\bigcup_{n=0}^{+\infty} A_n$ est au plus dénombrable. On peut poser $\bigcup_{n=0}^{+\infty} A_n = \{\omega_k,\ k \in I\}$ où l'ensemble des indices I est soit vide, soit un ensemble du type $[\![1, p]\!]$, $p \in \mathbb{N}^*$, soit $\mathbb{N}$. Pour chaque n, on note I_n l'ensemble des indices $k \in I$ tels que $\omega \in A_n$. Les ensembles I_n sont deux à deux disjoints et leur réunion est I.

Puisque la famille $(p_\omega)_{\omega\in A}$ est sommable, le théorème de sommation par paquets permet d'écrire

$$\sum_{n=0}^{+\infty} \mathbb{P}(A_n) = \sum_{n=0}^{+\infty} \left(\sum_{k\in I_n} p_{\omega_k} \right) = \sum_{k\in I} p_{\omega_k} = \mathbb{P}(A).$$

Ainsi, $\mathbb{P}$ est une probabilité sur l'espace $(\Omega, \mathscr{P}(\Omega))$ et $\forall \omega \in \Omega$, $\mathbb{P}(\{\omega\}) = p_\omega$ et donc $\mathbb{P}$ convient.

❏

Ceci nous invite à poser la définition suivante :

Définition 6. Soit Ω un ensemble non vide.

Une **distribution de probabilités discrète** sur Ω est une famille $(p_\omega)_{\omega\in\Omega}$ de réels positifs indexée par Ω, sommable de somme 1.

On rappelle que le support d'une famille sommable est au plus dénombrable et donc le support d'une distribution de probabilités discrète $(p_\omega)_{\omega\in\Omega}$, c'est-à-dire l'ensemble des $\omega \in \Omega$ tels que $p_\omega > 0$, est au plus dénombrable.

Le théorème 8 montre que si Ω est au plus dénombrable, la donnée d'une probabilité sur $\mathscr{P}(\Omega)$ équivaut à la donnée d'une distribution de probabilités discrète sur Ω. On obtient donc toutes les probabilités sur $\mathscr{P}(\Omega)$.

2 Probabilités conditionnelles. Indépendance

2.1 Probabilités conditionnelles

2.1.1 Définition

Définition 7. Soit $(\Omega, \mathscr{A}, \mathbb{P})$ un espace probabilisé au plus dénombrable. Soit $(A, B) \in \mathscr{A}^2$.

Si $\mathbb{P}(A) \neq 0$, la **probabilité de** B **sachant** A, notée $\mathbb{P}_A(B)$, est

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

Théorème 9. Soit $(\Omega, \mathscr{A}, \mathbb{P})$ un espace probabilisé. Soit $A \in \mathscr{A}$ tel que $\mathbb{P}(A) \neq 0$.

$\mathbb{P}_A$ est une probabilité sur l'espace probabilisable $(\Omega, \mathscr{A})$.

Démonstration .

- Puisque $\mathbb{P}(A) \neq 0$, pour tout $B \in \mathscr{A}$, $\mathbb{P}_A(B)$ existe dans $\mathbb{R}$ et de plus $\mathbb{P}_A(B) \geqslant 0$.
- $\mathbb{P}_A(\Omega) = \dfrac{\mathbb{P}(A \cap \Omega)}{\mathbb{P}(A)} = \dfrac{\mathbb{P}(A)}{\mathbb{P}(A)} = 1$.
- Soit $(B_n)_{n\in\mathbb{N}}$ une suite d'événements deux à deux disjoints.

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

On a montré que $\mathbb{P}_A$ est une probabilité sur l'espace probabilisable $(\Omega, \mathscr{A})$.

❏

On note que si A est un événement tel que $\mathbb{P}(A) \neq 0$ alors $\mathbb{P}_A(A) = 1$.

2.1.2 Formule des probabilités composées

Un résultat immédiat est :

Théorème 10. Soit $(\Omega, \mathscr{A}, \mathbb{P})$ un espace probabilisé. Soit $(A, B) \in \mathscr{A}^2$.

Si $\mathbb{P}(A) \neq 0$, $\mathbb{P}(A \cap B) = \mathbb{P}(A) \times \mathbb{P}_A(B)$.
Si $\mathbb{P}(B) \neq 0$, $\mathbb{P}(A \cap B) = \mathbb{P}(B) \times \mathbb{P}_B(A)$.

Plus généralement,

Théorème 11. (formule des probabilités composées)

Soit $(\Omega, \mathscr{A}, \mathbb{P})$ un espace probabilisé. Soit $(A_n)_{n\in\mathbb{N}} \in \mathscr{A}^{\mathbb{N}}$ une suite d'événements tels que $\forall n \in \mathbb{N}$, $\mathbb{P}(A_0 \cap \ldots \cap A_n) \neq 0$.

Pour tout $n \in \mathbb{N}^*$,

$$\mathbb{P}(A_0 \cap \ldots \cap A_n) = \mathbb{P}_{A_0\cap\ldots\cap A_{n-1}}(A_n) \times \mathbb{P}_{A_0\cap\ldots\cap A_{n-2}}(A_{n-1}) \times \ldots \times \mathbb{P}_{A_0}(A_1) \times \mathbb{P}(A_0).$$

Démonstration . Pour $n \in \mathbb{N}$, posons $p_n = \mathbb{P}(A_0 \cap \ldots \cap A_n)$. Par hypothèse, pour tout $n \in \mathbb{N}$, $p_n \neq 0$.

Soit $n \in \mathbb{N}^*$. Pour $k \in [\![0, n-1]\!]$,

$$\mathbb{P}_{A_0\cap\ldots\cap A_k}(A_{k+1}) = \frac{\mathbb{P}(A_0 \cap \ldots \cap A_{k+1})}{\mathbb{P}(A_0 \cap \ldots \cap A_k)} = \frac{p_{k+1}}{p_k}.$$

Par suite, pour $n \in \mathbb{N}^*$,

$$\begin{aligned}\mathbb{P}(A_0 \ldots \cap A_n) &= p_n \\ &= p_0 \prod_{k=0}^{n-1} \frac{p_{k+1}}{p_k} \text{ (produit télescopique)} \\ &= \mathbb{P}(A_0) \times \mathbb{P}_{A_0}(A_1) \times \ldots \times \mathbb{P}_{A_0 \cap \ldots \cap A_{n-2}}(A_{n-1}) \times \mathbb{P}_{A_0 \cap \ldots \cap A_{n-1}}(A_n).\end{aligned}$$

❑

Exercice 2. (d'après CCP 2017, MP, Maths 2)

Une particule possède deux états possibles, l'état A et l'état B et passe d'un état à l'autre de façon aléatoire. L'état de la particule au temps $n+1$ dépend uniquement de son état au temps n de la façon suivante :

- si au temps n, la particule est dans l'état A, au temps $n+1$, elle passe à l'état B avec une probabilité $\frac{1}{2}$.
- si au temps n, la particule est dans l'état B, au temps $n+1$, elle passe à l'état A avec une probabilité $\frac{1}{4}$.

A l'étape 0, la particule est dans l'un des deux états A ou B avec la même probabilité égale à $\frac{1}{2}$.

Pour $n \in \mathbb{N}$, on note T_n l'événement : « la première étape où la particule est à l'état A est l'étape n ».

Calculer $\mathbb{P}(T_n)$ en fonction de n.

Solution 2. Pour $n \in \mathbb{N}$, on note A_n l'événement : « à l'étape n, la particule est dans l'état A » et B_n l'événement : « à l'étape n, la particule est dans l'état B » (donc, $B_n = \overline{A_n}$).

T_0 est l'événement A_0 et donc $\mathbb{P}(T_0) = \mathbb{P}(A_0) = \frac{1}{2}$. Ensuite

$$\mathbb{P}(T_1) = \mathbb{P}(B_0 \cap A_1) = \mathbb{P}_{B_0}(A_1) \times \mathbb{P}(B_0) = \frac{1}{4} \times \frac{1}{2} = \frac{1}{8}$$

Soit $n \geqslant 1$. T_n est l'événement $B_0 \cap B_1 \cap \ldots \cap B_{n-1} \cap A_n$. D'après la formule des probabilités composées et puisque l'état de la particule à chaque étape dépend uniquement de l'état de la particule à l'état précédent,

$$\begin{aligned}\mathbb{P}(T_n) &= \mathbb{P}_{\bigcap_{k=0}^{n-1} B_k}(A_n) \times \mathbb{P}_{\bigcap_{k=0}^{n-2} B_k}(B_{n-1}) \times \ldots \times \mathbb{P}_{B_0}(B_1) \times \mathbb{P}(B_0) \\ &= \mathbb{P}_{B_{n-1}}(A_n) \times \mathbb{P}_{B_{n-2}}(B_{n-1}) \times \ldots \times \mathbb{P}_{B_0}(B_1) \times \mathbb{P}(B_0) \\ &= \frac{1}{4} \times \underbrace{\frac{3}{4} \times \ldots \times \frac{3}{4}}_{n-1} \times \frac{1}{2} = \frac{1}{8}\left(\frac{3}{4}\right)^{n-1} \\ &= \frac{1}{6}\left(\frac{3}{4}\right)^{n}\end{aligned}$$

ce qui reste vrai quand $n = 1$. Donc, $\mathbb{P}(T_0) = \frac{1}{2}$ et $\forall n \geqslant 1$, $\mathbb{P}(T_n) = \frac{1}{6}\left(\frac{3}{4}\right)^{n}$.

2.1.3 Formule des probabilités totales

Le théorème 7 fournit la formule des probabilités totales :

Théorème 12. (formule des probabilités totales)

Soit $(\Omega, \mathscr{A}, \mathbb{P})$ un espace probabilisé. Soit $(A_i)_{i \in I} \in \mathscr{A}^I$ un système quasi-complet d'événements, tels que $\forall i \in I$, $\mathbb{P}(A_i) > 0$.

Pour tout $B \in \mathscr{A}$,

$$\mathbb{P}(B) = \sum_{i \in I} \mathbb{P}(B \cap A_i) = \sum_{i \in I} \mathbb{P}_{A_i}(B) \times \mathbb{P}(A_i).$$

2.1.4 Formule de Bayes

Théorème 13. (formule de Bayes ou formule de probabilité des causes)

Soit $(\Omega, \mathscr{A}, \mathbb{P})$ un espace probabilisé. Soit $(A_i)_{j \in I} \in \mathscr{A}^I$ un système quasi-complet d'événements, tels que $\forall i \in I$, $\mathbb{P}(A_i) > 0$.

Soit $B \in \mathscr{A}$ tel que $\mathbb{P}(B) \neq 0$. Alors, pour tout $i \in I$,

$$\mathbb{P}_B(A_i) = \frac{\mathbb{P}(A_i) \times \mathbb{P}_{A_i}(B)}{\displaystyle\sum_{j \in I} \mathbb{P}(A_j) \times \mathbb{P}_{A_j}(B)}$$

Démonstration . D'après la formule des probabilités totales, pour $i \in I$,

$$\mathbb{P}_B(A_i) = \frac{\mathbb{P}(B \cap A_i)}{\mathbb{P}(B)} = \frac{\mathbb{P}(A_i) \times \mathbb{P}_{A_i}(B)}{\displaystyle\sum_{j \in I} \mathbb{P}(A_j) \times \mathbb{P}_{A_j}(B)}$$

❏

Exercice 3. (exercice d'oral de la banque CCP)

1) Enoncer la formule de Bayes.

2) On dispose de 100 dés dont 25 sont pipés.

Pour chaque dé pipé, la probabilité d'obtenir le chiffre 6 lors d'un lancer vaut $\frac{1}{2}$.

a) On tire un dé au hasard parmi les 100 dés. On lance ce dé et on obtient le chiffre 6.
Quelle est la probabilité que ce dé soit pipé ?

b) Soit $n \in \mathbb{N}^*$.
On tire un dé au hasard parmi les 100 dés. On lance ce dé n fois et on obtient n fois le chiffre 6.
Quelle est la probabilité p_n que ce dé soit pipé ?

c) Déterminer $\lim\limits_{n \to +\infty} p_n$. Interpréter ce résultat.

Solution 3.

1) Formule de Bayes.

Soit $(\Omega, \mathscr{A}, \mathbb{P})$ un espace probabilisé.
Soit $(A_i)_{i \in I}$ un système quasi-complet d'événements, de cet espace tel que pour tout $i \in I$, $\mathbb{P}(A_i) > 0$.
Soit B un événement tel que $\mathbb{P}(B) \neq 0$. Alors,

$$\forall i \in I,\ \mathbb{P}_B(A_i) = \frac{\mathbb{P}(A_i) \times \mathbb{P}_{A_i}(B)}{\displaystyle\sum_{j \in I} \mathbb{P}(A_j) \times \mathbb{P}_{A_j}(B)}.$$

2) a) Notons A l'événement « le dé est pipé » et B l'événement « on obtient le chiffre 6 ». La probabilité demandée est $\mathbb{P}_B(A)$.

$(A, \overline{A})$ est un système complet d'événements. On a $\mathbb{P}(A) = \frac{25}{100} = \frac{1}{4} \neq 0$ et $\mathbb{P}(\overline{A}) = 1 - \frac{1}{4} = \frac{3}{4}$. Ensuite $\mathbb{P}_A(B) = \frac{1}{2}$ et $\mathbb{P}_{\overline{A}}(B) = \frac{1}{6}$. Donc,

$$P(B) = P(A) \times P_A(B) + P(\overline{A}) \times P_{\overline{A}}(B) = \frac{1}{4} \times \frac{1}{2} + \frac{3}{4} \times \frac{1}{6} = \frac{1}{4} \neq 0.$$

D'après la formule de Bayes,

$$P_B(A) = \frac{P(A) \times P_A(B)}{P(A) \times P_A(B) + P(\overline{A}) \times P_{\overline{A}}(B)} = \frac{\frac{1}{4} \times \frac{1}{2}}{\frac{1}{4}} = \frac{1}{2}.$$

La probabilité que ce dé soit pipé est $\frac{1}{2}$.

b) Notons A l'événement « le dé est pipé » et B_n l'événement « on obtient n fois le chiffre 6 ». La probabilité demandée est $p_n = \mathbb{P}_{B_n}(A)$.

$(A, \overline{A})$ est un système complet d'événements. On a toujours $\mathbb{P}(A) = \dfrac{1}{4} \neq 0$ et $\mathbb{P}\left(\overline{A}\right) = \dfrac{3}{4}$. Ensuite $\mathbb{P}_A(B_n) = \dfrac{1}{2^n}$ et $\mathbb{P}_{\overline{A}}(B_n) = \dfrac{1}{6^n}$. Donc,

$$\mathbb{P}(B_n) = \mathbb{P}(A) \times \mathbb{P}_A(B_n) + \mathbb{P}\left(\overline{A}\right) \times \mathbb{P}_{\overline{A}}(B_n) = \frac{1}{4} \times \frac{1}{2^n} + \frac{3}{4} \times \frac{1}{6^n} \neq 0.$$

D'après la formule de Bayes,

$$p_n = \mathbb{P}_{B_n}(A) = \frac{P(A) \times \mathbb{P}_A(B_n)}{\mathbb{P}(A) \times \mathbb{P}_A(B_n) + \mathbb{P}\left(\overline{A}\right) \times \mathbb{P}_{\overline{A}}(B_n)} = \frac{\dfrac{1}{4} \times \dfrac{1}{2^n}}{\dfrac{1}{4} \times \dfrac{1}{2^n} + \dfrac{3}{4} \times \dfrac{1}{6^n}} = \frac{1}{1 + \dfrac{1}{3^{n-1}}}.$$

La probabilité que ce dé soit pipé est $p_n = \dfrac{1}{1 + \dfrac{1}{3^{n-1}}}$.

c) $\displaystyle\lim_{n \to +\infty} \frac{1}{3^{n-1}} = 0$ et donc $\displaystyle\lim_{n \to +\infty} p_n = 1$. Ceci signifie que si au bout d'un grand nombre de lancers, on a obtenu à chaque fois le 6, il est quasiment sûr que le dé est pipé.

2.2 Evénements indépendants

2.2.1 Cas de deux événements

Définition 8. Soit $(\Omega, \mathscr{A}, \mathbb{P})$ un espace probabilisé. Soit $(A, B) \in \mathscr{A}^2$.

A et B sont **indépendants** si et seulement si $\mathbb{P}(A \cap B) = \mathbb{P}(A) \times \mathbb{P}(B)$.

Un résultat immédiat est :

Théorème 14. Soit $(\Omega, \mathscr{A}, \mathbb{P})$ un espace probabilisé. Soit $(A, B) \in \mathscr{A}^2$.

Si $\mathbb{P}(A) \neq 0$, A et B sont indépendants si et seulement si $\mathbb{P}_A(B) = \mathbb{P}(B)$.

Théorème 15. Soit $(\Omega, \mathscr{A}, \mathbb{P})$ un espace probabilisé. Soit $(A, B) \in \mathscr{A}^2$. Si A et B sont indépendants, alors $\overline{A}$ et B sont indépendants, A et $\overline{B}$ sont indépendants et $\overline{A}$ et $\overline{B}$ sont indépendants.

Démonstration. Supposons A et B indépendants. Donc $\mathbb{P}(A \cap B) = \mathbb{P}(A) \times \mathbb{P}(B)$. D'après la formule des probabilités totales, $\mathbb{P}(B) = \mathbb{P}(A \cap B) + \mathbb{P}\left(\overline{A} \cap B\right)$ et donc

$$\begin{aligned} \mathbb{P}\left(\overline{A} \cap B\right) &= \mathbb{P}(B) - \mathbb{P}(A \cap B) = \mathbb{P}(B) - \mathbb{P}(A) \times \mathbb{P}(B) = \mathbb{P}(B)\left(1 - \mathbb{P}\left(\overline{A}\right)\right) \\ &= \mathbb{P}\left(\overline{A}\right) \times \mathbb{P}(B), \end{aligned}$$

et donc $\overline{A}$ et B sont indépendants. En appliquant à B et A, on a aussi A et $\overline{B}$ sont indépendants. Enfin, en appliquant à A et $\overline{B}$, on a aussi $\overline{A}$ et $\overline{B}$ sont indépendants.

❏

On note enfin de manière un peu anecdotique que :

Théorème 16. Soit $(\Omega, \mathscr{A}, \mathbb{P})$ un espace probabilisé au plus dénombrable. Soit $(A, B) \in \mathscr{A}^2$.

Si A est négligeable, alors A et B sont indépendants.
Si A est presque sûr, alors A et B sont indépendants.

Démonstration. Si A est négligeable, alors $\mathbb{P}(A) = 0$ puis $\mathbb{P}(A) \times \mathbb{P}(B) = 0$. D'autre part, $A \cap B \subset A$ et donc $\mathbb{P}(A \cap B) \leqslant \mathbb{P}(A) = 0$ puis $\mathbb{P}(A \cap B) = 0 = \mathbb{P}(A) \times \mathbb{P}(B)$. Ceci montre que A et B sont indépendants.

Si A est presque sûr, alors $\overline{A}$ est négligeable puis $\overline{A}$ et B sont indépendants d'après ce qui précède. On en déduit que A et B sont indépendants d'après le théorème précédent.

❏

2.2.2 Familles d'événements mutuellement indépendants

DÉFINITION 9. Soit $(\Omega, \mathscr{A}, \mathbb{P})$ un espace probabilisé. Soit $(A_i)_{i\in I} \in \mathscr{A}^I$ une famille non vide d'événements.

La famille $(A_i)_{i\in I} \in \mathscr{A}^I$ est une famille d'**événements mutuellement indépendants** (ou plus simplement indépendants) si et seulement si pout toute partie finie non vide J de I,

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

La famille $(A_i)_{i\in I} \in \mathscr{A}^I$ est une famille d'**événements deux à deux indépendants** si et seulement si pour tout $(j,k) \in I^2$,

$$j \neq k \Rightarrow \mathbb{P}(A_j \cap A_k) = \mathbb{P}(A_j) \times \mathbb{P}(A_k).$$

Un résultat immédiat est :

Théorème 17. Soit $(\Omega, \mathscr{A}, \mathbb{P})$ un espace probabilisé au plus dénombrable. Soit $(A_i)_{i\in I} \in \mathscr{A}^I$ une famille non vide d'événements au plus dénombrable.

Si les événements A_i, $i \in I$, sont mutuellement indépendants, alors les événements A_i, $i \in I$, sont deux à deux indépendants.

Il faut prendre garde au fait que la réciproque est fausse ou encore deux à deux indépendants $\not\Rightarrow$ mutuellement indépendants. Considérons par exemple un univers à quatre éléments $\Omega = \{a, b, c, d\}$. On munit $(\Omega, \mathscr{P}(\Omega))$ de la probabilité uniforme. Soient $A = \{a, b\}$, $B = \{b, c\}$ et $C = \{c, a\}$.
$\mathbb{P}(A \cap B) = \mathbb{P}(\{b\}) = \dfrac{1}{4} = \dfrac{1}{2} \times \dfrac{1}{2} = \mathbb{P}(A) \times \mathbb{P}(B)$. Puis, par symétrie des rôles,
$\mathbb{P}(A \cap C) = \mathbb{P}(A) \times \mathbb{P}(C)$ et $\mathbb{P}(B \cap C) = \mathbb{P}(B) \times \mathbb{P}(C)$. Donc, les événements A, B et C sont deux à deux indépendants.
Maintenant, $\mathbb{P}(A \cap B \cap C) = \mathbb{P}(\varnothing) \neq \left(\dfrac{1}{2}\right)^3 = \mathbb{P}(A) \times \mathbb{P}(B) \times \mathbb{P}(C)$. Donc, les événements A, B et C ne sont pas mutuellement indépendants.

Théorème 18. Soit $(\Omega, \mathscr{A}, \mathbb{P})$ un espace probabilisé. Soit $(A_i)_{i\in I} \in \mathscr{A}^I$ une famille non vide d'événements.

Si les événements A_i, $i \in I$, sont mutuellement indépendants, alors les événements B_i, $i \in I$, sont mutuellement indépendants où, pour tout $i \in I$, B_i est soit A_i, soit $\overline{A_i}$.

DÉMONSTRATION. Il suffit de démontrer le résultat quand un et un seul des B_i est $\overline{A_i}$ et tous les autres sont A_i. Le résultat s'en déduit alors par récurrence sur le cardinal d'une partie finie J de I.

Soit $i_0 \in I$. Pour $i \in I$, on pose $B_i = \left\{\begin{array}{l}\overline{A_{i_0}} \text{ si } i = i_0\\ A_i \text{ si } i \neq i_0\end{array}\right.$.

Soit J une partie finie non vide de I ayant au moins deux éléments. Si $i_0 \notin J$, on a immédiatement $\mathbb{P}\left(\bigcap_{i\in J} B_i\right) = \prod_{i\in I} \mathbb{P}(B_i)$. Si $i_0 \in J$, les événements A_{i_0} et $\bigcap_{i\in J\setminus\{i_0\}} A_i$ sont indépendants car

$$\begin{aligned}\mathbb{P}\left(A_{i_0} \cap \left(\bigcap_{i\in J\setminus\{i_0\}} A_i\right)\right) &= \mathbb{P}\left(\bigcap_{i\in J} A_i\right) = \prod_{i\in J} \mathbb{P}(A_i) = \mathbb{P}(A_{i_0}) \times \prod_{i\in J\setminus\{i_0\}} \mathbb{P}(A_i)\\ &= \mathbb{P}(A_{i_0}) \times \mathbb{P}\left(\bigcap_{i\in J\setminus\{i_0\}} A_i\right)\end{aligned}$$

Mais alors, les événements B_{i_0} et $\bigcap_{i\in J\setminus\{i_0\}} B_i$ sont indépendants d'après le théorème 15 et donc

$$\mathbb{P}\left(\bigcap_{i\in J} B_i\right) = \mathbb{P}(B_{i_0}) \times \mathbb{P}\left(\bigcap_{i\in J\setminus\{i_0\}} B_i\right) = \prod_{i\in J} \mathbb{P}(B_i).$$

On a montré que les B_i, $i \in I$, sont mutuellement indépendants.

❑