

ESPACES PROBABILISÉS

Le but du chapitre est de généraliser la notion de loi de probabilité sur un ensemble fini Ω , vue en première année, au cas où l'ensemble Ω n'est pas nécessairement fini. On distinguera pour cela deux types d'ensembles infinis (cf. partie 1.1) :

- les ensembles *dénombrables* — infinis “pas trop gros” — pour lesquels les choses se passent de façon analogue au cas fini : par exemple, on peut y définir la probabilité de toute partie de Ω .
- les ensembles infinis *non-dénombrables* — infinis “trop gros” — pour lesquels les choses sont plus compliquées : par exemple, il n'est pas toujours possible (pour des raisons qui dépassent le cadre de ce cours) de définir la probabilité de toutes les parties de Ω .

Cette difficulté conduit à la notion de *tribu* (cf. partie 1.2), ensemble de parties de Ω (donc sous-ensemble de $\mathcal{P}(\Omega)$) sur lequel on pourra définir une loi de probabilité (cf. partie 1.3).

1 Espaces probabilisés

1.1 Préliminaires : ensembles dénombrables

1.1.1 Définition - Ensemble dénombrable.

On dit qu'un ensemble E est **dénombrable** s'il existe une bijection de $\mathbb{N}$ sur E .

1.1.2 Remarques. 1. Un ensemble dénombrable est donc un ensemble infini qui peut être décrit *en extension* sous la forme $\{x_n \mid n \in \mathbb{N}\}$.

2. La relation “être en bijection avec” est une relation d'équivalence sur la *classe* de tous les ensembles. Les ensembles dénombrables sont ceux de la classe d'équivalence de $\mathbb{N}$.

En particulier, tout ensemble en bijection avec un ensemble dénombrable est dénombrable.

1.1.3 Proposition - Exemples d'ensembles dénombrables.

1. Les ensembles $\mathbb{N}$, $\mathbb{Z}$ et $\mathbb{Q}$ sont dénombrables.
2. Toute partie infinie d'un ensemble dénombrable est dénombrable.
3. Le produit cartésien de deux ensembles dénombrables est dénombrable.

Démonstration.

1. • Le cas de $\mathbb{N}$ est évident, et l'application $\varphi : \mathbb{N} \rightarrow \mathbb{Z}, n \mapsto (-1)^n \lfloor \frac{n+1}{2} \rfloor$, est une bijection.
Rq. Cette bijection φ consiste à énumérer les entiers positifs avec les entiers naturels pairs, et les entiers négatifs avec les entiers naturels impairs.
 - L'application $\psi : \mathbb{Q} \rightarrow \mathbb{Z} \times \mathbb{N}^*, r \mapsto (a, b)$, où r s'écrit sous forme irréductible $r = \frac{a}{b}$, réalise une bijection de $\mathbb{Q}$ sur une partie infinie de $\mathbb{Z} \times \mathbb{N}^*$. Au vu des points 2 et 3, $\mathbb{Q}$ est dénombrable.
2. Si $\varphi : \mathbb{N} \rightarrow E$ est une bijection, et si F est une partie infinie de E , alors φ induit une bijection $\tilde{\varphi} : \mathbb{P} \rightarrow F$, où $\mathbb{P} = \varphi^{-1}(F)$ est une partie infinie de $\mathbb{N}$. Et une énumération des éléments de $\mathbb{P}$ dans l'ordre croissant fournit une bijection $\psi : \mathbb{N} \rightarrow \mathbb{P}$, d'où le résultat.
Rq. La bijection ψ se construit rigoureusement par récurrence, en posant $\psi(0) = \min \mathbb{P}$, et si $\psi(n)$ est construit, en posant $\psi(n+1) = \min\{x \in \mathbb{P} \mid x > \psi(n)\}$.
 - **Démo 1.** L'application $\chi : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}^*, (p, q) \mapsto 2^p(2q+1)$, est une bijection (tout entier naturel non nul s'écrit de façon unique comme produit d'une puissance de 2 et d'un nombre impair).
 - **Démo 2.** L'application $\psi : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}, (p, q) \mapsto p + \frac{1}{2}(p+q)(p+q+1)$ est une bijection (qui consiste à énumérer les couples d'entiers en suivant les diagonales d'équation $p+q=n$).
3. Si $\varphi_1 : \mathbb{N} \rightarrow E_1$ et $\varphi_2 : \mathbb{N} \rightarrow E_2$ sont deux bijections, alors $\varphi : \mathbb{N} \times \mathbb{N} \rightarrow E_1 \times E_2, (n, p) \mapsto (\varphi_1(n), \varphi_2(p))$, est une bijection. Il suffit donc de montrer que $\mathbb{N} \times \mathbb{N}$ est dénombrable.

1.1.4 Proposition - Exemples d'ensembles infinis non dénombrables.

1. Un intervalle non trivial de $\mathbb{R}$ (donc en particulier $\mathbb{R}$) et $\mathbb{C}$ sont non dénombrables.
2. L'ensemble des suites à valeurs dans un ensemble de cardinal ≥ 2 est non dénombrable.
3. L'ensemble des parties d'un ensemble dénombrable est non dénombrable.

Démonstration.

1. • Tout intervalle non trivial $|a; b|$ de $\mathbb{R}$ est en bijection avec $|0; 1|$ (crochets ouverts ou fermés), donc il suffit de montrer qu'une application $\varphi : \mathbb{N}^* \rightarrow |0; 1|$ ne peut pas être surjective. En effet, un $x \in |0; 1|$ dont, pour tout $n \in \mathbb{N}^*$, la n -ième décimale du développement décimal propre est différente de celle de $\varphi(n)$, ne peut pas avoir d'antécédent par φ .
• Puisque $\mathbb{C}$ contient un ensemble infini non dénombrable, il est lui même non dénombrable.
2. L'argument est analogue à celui du premier point : si E est un ensemble de cardinal ≥ 2 , alors une application $\varphi : \mathbb{N} \rightarrow E^{\mathbb{N}}$ ne peut pas être surjective. En effet, une suite $(u_n)_{n \in \mathbb{N}} \in E^{\mathbb{N}}$ telle que, pour tout $n \in \mathbb{N}$, u_n est distinct du n -ième terme de $\varphi(n)$, ne peut pas avoir d'antécédent par φ .
3. Résulte du fait qu'une application $\varphi : E \rightarrow \mathcal{P}(E)$ ne peut pas être surjective. En effet la partie $F = \{x \in E \mid x \notin \varphi(x)\}$ ne saurait avoir d'antécédent par φ .

1.2 Tribu sur un ensemble**1.2.1 Définitions - Tribu - Espace probabilisable.**

- Soit Ω un ensemble. On appelle **tribu** sur Ω tout sous-ensemble $\mathcal{T}$ de $\mathcal{P}(\Omega)$ contenant Ω , et *stable* par passage au complémentaire et par union dénombrable, i.e. tel que :
 - a. $\Omega \in \mathcal{T}$,
 - b. Pour tout $A \in \mathcal{T}$, $\bar{A} = \Omega \setminus A \in \mathcal{T}$,
 - c. Pour toute suite $(A_n)_{n \in \mathbb{N}}$ d'éléments de $\mathcal{T}$, $\bigcup_{n=0}^{+\infty} A_n \in \mathcal{T}$.
- On dit qu'un ensemble Ω muni d'une tribu $\mathcal{T}$ sur Ω , i.e. formellement le couple $(\Omega, \mathcal{T})$, est un **espace probabilisable**.

1.2.2 Exemples. Soit Ω un ensemble.

- Les ensembles $\{\emptyset, \Omega\}$ et $\mathcal{P}(\Omega)$ sont deux tribus sur Ω (respectivement la plus petite et la plus grosse, au sens de l'inclusion). Lorsque Ω est fini ou dénombrable, on le munit a priori, sauf mention du contraire, de sa tribu maximale $\mathcal{P}(\Omega)$.
- Pour tout $A \in \mathcal{P}(\Omega)$, l'ensemble $\{\emptyset, A, \bar{A}, \Omega\}$ est une tribu sur Ω , dite *engendrée* par A .

1.2.3 Théorème - Propriétés des tribus.

Soit $(\Omega, \mathcal{T})$ un espace probabilisable. Alors $\mathcal{T}$ contient l'ensemble vide, est stable par union finie, par intersection finie ou dénombrable, et par différence.

Démonstration.

- On a $\Omega \in \mathcal{T}$ donc $\emptyset = \bar{\Omega} \in \mathcal{T}$.
- La stabilité par union finie résulte de ce que si l'on complète une famille finie $(A_1, \dots, A_p)$ d'éléments de $\mathcal{T}$ en une suite $(A_n)_{n \in \mathbb{N}}$, en posant $A_n = \emptyset$ si $n \notin [1; p]$, alors $\bigcup_{n=1}^p A_n = \bigcup_{n=0}^{+\infty} A_n$.
- La stabilité par intersection finie ou dénombrable résulte de ce que pour tout ensemble I et toute famille $(A_i)_{i \in I}$ de parties de Ω , on a $\bigcap_{i \in I} A_i = \overline{\bigcup_{i \in I} \bar{A}_i}$.
- La stabilité par différence résulte de ce que pour toutes parties A et B de Ω , $A \setminus B = A \cap \bar{B}$.

1.2.4 Définitions - Vocabulaire des probabilités.

Soit $(\Omega, \mathcal{T})$ un espace probabilisable.

- Les éléments de $\mathcal{T}$ sont appelés des **événements**.
 Ω est l'**univers**, ou l'**événement certain**, et $\emptyset$ est l'**événement impossible**.
- Deux événements A et B sont dits **incompatibles** si $A \cap B = \emptyset$, et **contraires** si $B = \bar{A}$.
- Un **système complet d'événements** est une *partition* de Ω constituée d'éléments de $\mathcal{T}$.

1.2.5 Remarque (Modélisation mathématique d'une expérience aléatoire). Les issues possibles d'une expérience aléatoire constituent un ensemble Ω , l'univers de cette expérience aléatoire.

- Si l'on souhaite considérer plus particulièrement les issues vérifiant une certaine assertion $\mathcal{A}$, il convient de choisir une tribu $\mathcal{T}$ sur Ω contenant la partie $A = \{\omega \in \Omega \mid \omega \text{ vérifie } \mathcal{A}\}$ de Ω constituée des issues vérifiant l'assertion $\mathcal{A}$.

Dans ce contexte, on dit que la partie A est l'événement « l'assertion $\mathcal{A}$ est vérifiée ».

- Les propriétés des tribus listées en 1.2.1 et 1.2.3 permettent alors les considérations usuelles sur les événements : si A, B et A_i (où $i \in I$, ensemble fini ou dénombrable) sont des événements définis par des assertions $\mathcal{A}, \mathcal{B}$ et $\mathcal{A}_i$ (où $i \in I$) respectivement, alors :
 - ★ $\bar{A}$ est l'événement « l'assertion $\mathcal{A}$ n'est pas vérifiée ».
 - ★ $A \cup B, A \cap B$ et $A \setminus B$ sont respectivement les événements « au moins l'une des deux assertions $\mathcal{A}$ et $\mathcal{B}$ est vérifiée », « $\mathcal{A}$ et $\mathcal{B}$ sont vérifiées », et « seule $\mathcal{A}$ est vérifiée ».
 - ★ $\bigcup_{i \in I} A_i$ et $\bigcap_{i \in I} A_i$ sont respectivement les événements « l'une (au moins) des assertions $\mathcal{A}_i$ est vérifiée » et « toutes les assertions $\mathcal{A}_i$ sont vérifiées ».

1.2.6 Exemples. 1. Si l'expérience aléatoire consiste à lancer une fois un dé à 6 faces, alors l'univers est $\Omega = \llbracket 1; 6 \rrbracket$, que l'on munit usuellement de la tribu maximale $\mathcal{T} = \mathcal{P}(\Omega)$.

Les événements « le résultat obtenu est pair » et « le résultat obtenu est supérieur ou égal à 3 » sont respectivement les parties $A = \{2, 4, 6\}$ et $B = \{3, 4, 5, 6\}$ de Ω .

Alors $\bar{A}, A \cup B, A \cap B$ et $A \setminus B$ sont respectivement les événements « le résultat obtenu est impair », « le résultat obtenu n'est pas 1 », « on a obtenu 4 ou 6 » et « on a obtenu 2 ».

2. Si l'expérience aléatoire consiste à lancer indéfiniment un dé à 6 faces, alors l'univers est $\Omega = \llbracket 1; 6 \rrbracket^{\mathbb{N}^*}$, que l'on munit usuellement d'une tribu $\mathcal{T}$ sur Ω contenant, pour tous $k \in \llbracket 1; 6 \rrbracket$ et $p \in \mathbb{N}$, la partie $A_{k,p} = \{(u_n)_{n \in \mathbb{N}^*} \in \Omega \mid u_p = k\}$ de Ω , à savoir l'événement « on a obtenu k au p -ième lancer ».

Alors $\bigcup_{p=1}^{+\infty} A_{k,p}$ et $\bigcap_{p=1}^{+\infty} A_{k,p}$ sont respectivement les événements « on a obtenu (au moins) une fois k » et « on n'a obtenu que des k ».

1.3 Loi de probabilité sur un espace probabilisable**1.3.1 Définitions - Loi de probabilité - Espace probabilisé.**

- Soit $(\Omega, \mathcal{T})$ un espace probabilisable. On appelle (**loi de**) **probabilité** sur $(\Omega, \mathcal{T})$ toute application $P : \mathcal{T} \rightarrow [0; 1]$ de *poinds total* 1 et σ -*additive*, i.e. telle que :
 - a. $P(\Omega) = 1$,
 - b. Pour toute suite $(A_n)_{n \in \mathbb{N}}$ d'événements deux à deux incompatibles, la série $\sum P(A_n)$ converge et $P\left(\bigcup_{n=0}^{+\infty} A_n\right) = \sum_{n=0}^{+\infty} P(A_n)$.
- On dit qu'un ensemble Ω , muni d'une tribu $\mathcal{T}$ sur Ω et d'une probabilité P sur $(\Omega, \mathcal{T})$, i.e. formellement le triplet $(\Omega, \mathcal{T}, P)$, est un **espace probabilisé**.

1.3.2 Proposition - Loi de probabilité sur un univers fini ou dénombrable.

Soit Ω un ensemble non vide, fini ou dénombrable. Alors pour toute famille $(p_\omega)_{\omega \in \Omega}$ de réels positifs dont la somme vaut 1, il existe une unique loi de probabilité P sur $(\Omega, \mathcal{P}(\Omega))$ telle que $\forall \omega \in \Omega, P(\{\omega\}) = p_\omega$. Elle est définie, pour tout $A \in \mathcal{P}(\Omega)$, par $P(A) = \sum_{\omega \in A} p_\omega$.

Rq. Dans le cas où A est infini, cette somme est la somme d'une série convergente.

Démonstration.

- Montrons que l'application P définie dans l'énoncé est une loi de probabilité.
Par hypothèse sur la famille $(p_\omega)_{\omega \in \Omega}$, on a $P(\Omega) = \sum_{\omega \in \Omega} p_\omega = 1$. Et si $(A_n)_{n \in \mathbb{N}}$ est une suite d'événements deux à deux incompatibles, alors en sommant par paquets (formule admise dans le cas dénombrable), on a $P(\bigcup_{n=0}^{+\infty} A_n) = \sum_{\omega \in \bigcup_{n=0}^{+\infty} A_n} p_\omega = \sum_{n=0}^{+\infty} \sum_{\omega \in A_n} p_\omega = \sum_{n=0}^{+\infty} P(A_n)$.
- L'unicité résulte de la σ -additivité et de ce que toute partie de Ω est finie ou dénombrable (1.1.3), donc est la réunion finie ou dénombrable de ses singletons.

1.3.3 Remarques. 1. Définir une loi de probabilité sur un univers fini ou dénombrable Ω se résume donc à définir la probabilité p_ω de chaque singleton de Ω , et à vérifier que $\sum_{\omega \in \Omega} p_\omega = 1$. C'est ce qui est fait plus loin, en partie 3, pour définir les lois de probabilités usuelles.

2. Ce résultat est faux si Ω est infini non dénombrable, la σ -additivité ne permettant alors plus de définir la probabilité de toute partie de Ω à partir des probabilités des singletons.

1.3.4 Exemples. 1. Dans le cas du lancer d'un dé à 6 faces, on munira $\Omega = \llbracket 1; 6 \rrbracket$ de la loi de probabilité P suivante, selon l'hypothèse faite sur le dé lancé :

- si le dé est *équilibré*, alors P est la loi *uniforme*, définie par $P(\{k\}) = \frac{1}{6}$ pour tout $k \in \llbracket 1; 6 \rrbracket$.
- si le dé a été truqué de façon à donner $k \in \llbracket 1; 6 \rrbracket$ avec une probabilité proportionnelle à k , alors P est la loi définie par $P(\{k\}) = \frac{k}{21}$, pour tout $k \in \llbracket 1; 6 \rrbracket$.

La probabilité que le résultat obtenu soit pair est alors $\frac{1}{2}$ ou $\frac{4}{7}$ selon le cas envisagé.

- La suite $(\frac{1}{2^n})_{n \in \mathbb{N}^*}$ définit une loi de probabilité P sur $\mathbb{N}^*$ pour laquelle la probabilité de l'ensemble $\mathbb{P} = \{2k \mid k \in \mathbb{N}^*\}$ des entiers naturels pairs non nuls est $P(\mathbb{P}) = \sum_{k=1}^{+\infty} \frac{1}{2^{2k}} = \frac{1}{3}$.
- Dans le cas des lancers successifs d'un dé à 6 faces *équilibré*, on admettra l'existence d'une tribu $\mathcal{T}$ sur $\Omega = \llbracket 1; 6 \rrbracket^{\mathbb{N}^*}$ contenant les $A_{k,p}$ (définis en 1.2.6), et d'une loi de probabilité P sur $(\Omega, \mathcal{T})$, telles que pour tous $k \in \llbracket 1; 6 \rrbracket$ et $p \in \mathbb{N}^*$, $P(A_{k,p}) = \frac{1}{6}$.

1.3.5 Théorème - Propriétés des lois de probabilités.

Soient $(\Omega, \mathcal{T}, P)$ un espace probabilités.

- On a $P(\emptyset) = 0$, et plus généralement pour tout $A \in \mathcal{T}$, $P(\bar{A}) = 1 - P(A)$.
- Soient $A, B \in \mathcal{T}$.
 - On a $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
En particulier, $P(A \cup B) \leq P(A) + P(B)$, avec égalité si A et B sont incompatibles.
 - Croissance.* Si $A \subset B$, alors $P(A) \leq P(B)$, et plus précisément $P(B) - P(A) = P(B \setminus A)$.
- Soit $(A_n)_{n \in \mathbb{N}}$ une suite d'événements.
 - Sous-additivité.* On a $P\left(\bigcup_{n=0}^{+\infty} A_n\right) \leq \sum_{n=0}^{+\infty} P(A_n)$, le terme de droite pouvant valoir $+\infty$.
 - Continuité croissante.* Si $\forall n \in \mathbb{N}, A_n \subset A_{n+1}$, alors $P\left(\bigcup_{n=0}^{+\infty} A_n\right) = \lim_{n \rightarrow +\infty} P(A_n)$.
 - Continuité décroissante.* Si $\forall n \in \mathbb{N}, A_{n+1} \subset A_n$, alors $P\left(\bigcap_{n=0}^{+\infty} A_n\right) = \lim_{n \rightarrow +\infty} P(A_n)$.

Démonstration.

1. La σ -additivité appliquée à la suite $(A, \bar{A}, \emptyset, \emptyset, \dots)$ montre que $P(\emptyset) = 0$ et que $P(A) + P(\bar{A}) = 1$.
2. a. Par σ -additivité et 1, on a $P(A \cup B) = P(A) + P(B \setminus A)$ et $P(B) = P(B \setminus A) + P(A \cap B)$.
b. Les deux égalités de 2a donnent, dans le cas $A \subset B$, $P(B) = P(B \setminus A) + P(A)$.
3. Posons $B_0 = A_0$ et pour tout $n \geq 1$, $B_n = A_n \setminus \bigcup_{k=0}^{n-1} A_k$, de sorte que les B_n sont des événements deux à deux disjoints, avec $B_n \subset A_n$ et $\bigcup_{n=0}^{+\infty} B_n = \bigcup_{n=0}^{+\infty} A_n$.
a. Par σ -additivité et croissance, on a alors $P(\bigcup_{n=0}^{+\infty} A_n) = \sum_{n=0}^{+\infty} P(B_n) \leq \sum_{n=0}^{+\infty} P(A_n)$.
b. Si $\forall k \in \mathbb{N}$, $A_k \subset A_{k+1}$, alors $\forall n \geq 1$, $B_n = A_n \setminus A_{n-1}$, donc par 2b, $P(B_n) = P(A_n) - P(A_{n-1})$.
Par σ -additivité et télescopage, on a alors $P(\bigcup_{n=0}^{+\infty} A_n) = \sum_{n=0}^{+\infty} P(B_n) = \lim_{n \rightarrow +\infty} P(A_n)$.
c. Si $\forall k \in \mathbb{N}$, $A_{k+1} \subset A_k$, alors $\bar{A}_k \subset \bar{A}_{k+1}$, et donc par 3b, $P(\bigcup_{n=0}^{+\infty} \bar{A}_n) = \lim_{n \rightarrow +\infty} P(\bar{A}_n)$.
Par 1, on a alors, puisque $\bigcup_{n=0}^{+\infty} \bar{A}_n = \overline{\bigcap_{n=0}^{+\infty} A_n}$, $P(\bigcap_{n=0}^{+\infty} A_n) = \lim_{n \rightarrow +\infty} P(A_n)$.

1.3.6 Remarques. 1. Puisque $P(\emptyset) = 0$, la σ -additivité et la sous-additivité sont également valables pour une famille finie d'événements.

2. La continuité (dé)croissante donne, pour une suite d'événements $(A_n)_{n \in \mathbb{N}}$ quelconque :

$$P\left(\bigcup_{k=0}^{+\infty} A_k\right) = \lim_{n \rightarrow +\infty} P\left(\bigcup_{k=0}^n A_k\right) \quad \text{et} \quad P\left(\bigcap_{k=0}^{+\infty} A_k\right) = \lim_{n \rightarrow +\infty} P\left(\bigcap_{k=0}^n A_k\right).$$

1.3.7 Exemples. Considérons les lancers successifs d'un dé équilibré à 6 faces (1.2.6 et 1.3.4).

- L'événement $A = \bigcup_{p=1}^{+\infty} A_{6,p}$ « on obtient (au moins) une fois 6 » est la réunion croissante des $A_n = \bigcup_{k=1}^n A_{6,k}$ « on obtient (au moins) une fois 6 au cours des n premiers lancers ».
On verra en 2.2.6 que $\forall n \geq 1$, $P(A_n) = 1 - (\frac{5}{6})^n$, donc par continuité croissante, $P(A) = 1$.
- L'événement $B = \bigcap_{p=1}^{+\infty} A_{6,p}$ « on n'obtient que des 6 » est l'intersection décroissante des $B_n = \bigcap_{k=1}^n A_{6,k}$ « on n'obtient que des 6 au cours des n premiers lancers ».
On verra en 2.2.6 que $\forall n \geq 1$, $P(B_n) = (\frac{1}{6})^n$, donc par continuité décroissante, $P(B) = 0$.

1.3.8 Définitions - Vocabulaire des probabilités (suite).

Soit $(\Omega, \mathcal{T}, P)$ un espace probablisé.

- Un événement $A \in \mathcal{T}$ tel que $P(A) = 0$ est dit **presque impossible**, ou **négligeable**.
- Un événement $A \in \mathcal{T}$ tel que $P(A) = 1$ est dit **presque sûr**.
- Un **système presque complet d'événements** est une famille d'événements deux à deux disjoints et de réunion presque sûre.

2 Conditionnement et indépendance

2.1 Probabilités conditionnelles

2.1.1 Définition.

Soient $(\Omega, \mathcal{T}, P)$ un espace probablisé et $B \in \mathcal{T}$ tel que $P(B) \neq 0$. Alors pour tout $A \in \mathcal{T}$, la **probabilité conditionnelle** de A sachant B est le réel, noté $P_B(A)$ ou $P(A|B)$, défini par

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

2.1.2 Remarque. Sous les hypothèses de 2.1.1, l'application $P_B : \mathcal{T} \rightarrow [0; 1]$, $A \mapsto P_B(A)$, ainsi définie est une loi de probabilité sur $(\Omega, \mathcal{T})$.

Démonstration.

On a $P_B(\Omega) = P(B)/P(B) = 1$. Et si $(A_n)_{n \in \mathbb{N}}$ est une suite d'événements 2 à 2 incompatibles, alors $(A_n \cap B)_{n \in \mathbb{N}}$ est une suite d'événements 2 à 2 incompatibles et $(\bigcup_{n=0}^{+\infty} A_n) \cap B = \bigcup_{n=0}^{+\infty} (A_n \cap B)$, donc $P_B(\bigcup_{n=0}^{+\infty} A_n) = \frac{1}{P(B)} P(\bigcup_{n=0}^{+\infty} (A_n \cap B)) = \frac{1}{P(B)} \sum_{n=0}^{+\infty} P(A_n \cap B) = \sum_{n=0}^{+\infty} P_B(A_n)$.

2.1.3 Théorème - Les trois formules fondamentales du calcul des probabilités.

Soient $(\Omega, \mathcal{T}, P)$ un espace probabilisé.

1. **Formule de Bayes.** Pour tous $A, B \in \mathcal{T}$ de probabilité non nulle, $P_B(A) = \frac{P_A(B)P(A)}{P(B)}$.
2. **Formule des probabilités totales.** Si $(A_n)_{n \in \mathbb{N}}$ est un système (presque) complet d'événements, alors pour tout $B \in \mathcal{T}$, on a, en posant $P(A_n)P_{A_n}(B) = 0$ si $P(A_n) = 0$:

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

3. **Formule des probabilités composées.** Si $A_1, A_2, \dots, A_n$ sont des événements tels que $P(A_1 \cap \dots \cap A_{n-1}) \neq 0$, alors pour tout $k \in \llbracket 1; n-2 \rrbracket$, $P(A_1 \cap \dots \cap A_k) \neq 0$ et

$$P\left(\bigcap_{k=1}^n A_k\right) = P(A_1)P_{A_1}(A_2)P_{A_1 \cap A_2}(A_3) \cdots P_{A_1 \cap \dots \cap A_{n-1}}(A_n).$$

Démonstration.

1. Par définition, les deux termes $P(B)P_B(A)$ et $P(A)P_A(B)$ valent $P(A \cap B)$.
2. Quitte à compléter $(A_n)_{n \in \mathbb{N}}$ avec le complémentaire de $\bigcup_{n=0}^{+\infty} A_n$, qui est négligeable, on peut supposer que $(A_n)_{n \in \mathbb{N}}$ est un système complet d'événements.
Alors $B = \bigcup_{n=0}^{+\infty} (B \cap A_n)$, et les $B \cap A_n$ sont 2 à 2 disjoints, donc par σ -additivité, $P(B) = \sum_{n=0}^{+\infty} P(B \cap A_n)$. La dernière égalité résulte de la définition des probabilités conditionnelles.
3. Par récurrence à partir de la définition des probabilités conditionnelles.

- 2.1.4 Remarques.** • La formule de Bayes est parfois donnée sous la forme suivante : si $(A_n)_{n \in \mathbb{N}}$ est un système (presque) complet d'événements, alors pour tout $B \in \mathcal{T}$ et tout $k \in \mathbb{N}$,

$$P_B(A_k) = \frac{P_{A_k}(B)P(A_k)}{\sum_{n=0}^{+\infty} P_{A_n}(B)P(A_n)}.$$

- La formule des probabilités totales est également valable pour un système (presque) complet d'événements fini (en complétant la famille finie d'événements en une suite avec des $\emptyset$).

2.1.5 Exemple. Une urne contient initialement 7 boules noires et 3 boules blanches. On en tire successivement 3 boules selon le protocole suivant : si on obtient une boule noire alors on la remet, et si on obtient une boule blanche alors on la remplace par une noire.

1. Probabilité que la deuxième boule obtenue soit blanche ?
2. Probabilité d'avoir tiré une boule blanche en premier, sachant que la seconde était blanche ?
3. Probabilité de tirer 3 boules blanches à la suite ?

2.2 Indépendance**2.2.1 Définitions - Indépendance - Indépendance mutuelle.**

Soit $(\Omega, \mathcal{T}, P)$ un espace probabilité.

- On dit que deux événements A et B sont **indépendants** si $P(A \cap B) = P(A)P(B)$, i.e. dans le cas où $P(B) \neq 0$, si $P_B(A) = P(A)$.
- Plus généralement, des événements A_i , pour $i \in I$, sont dits **mutuellement indépendants** si pour toute partie finie $J \subset I$, $P\left(\bigcap_{j \in J} A_j\right) = \prod_{j \in J} P(A_j)$.

2.2.2 Remarque. L'indépendance mutuelle ne se résume pas à la propriété sur l'intersection de tous les A_i : $P(\bigcap_{i \in I} A_i) = \prod_{i \in I} P(A_i)$ (ce produit n'a pas forcément de sens si I est infini) ; ni à l'indépendance deux à deux : $\forall i, j \in I$ avec $i \neq j$, $P(A_i \cap A_j) = P(A_i)P(A_j)$.

2.2.3 Exemples. On lance successivement deux dés équilibrés à 6 faces.

- On considère les événements A_1 « le résultat du premier dé est pair », A_2 « le résultat du second dé est impair » et A_3 « les deux résultats sont de même parité ».

Alors $P(A_i) = \frac{1}{2}$ et $P(A_i \cap A_j) = \frac{1}{4}$ si $i \neq j$, mais $P(A_1 \cap A_2 \cap A_3) = 0 \neq P(A_1)P(A_2)P(A_3)$.

- On considère les événements A_1 « le résultat du second dé est ≥ 4 », A_2 « le résultat du second dé est impair » et A_3 « la somme des deux résultats vaut 9 ».

Alors $P(A_1) = P(A_2) = \frac{1}{2}$ et $P(A_3) = \frac{1}{9}$, puis $P(A_1 \cap A_2 \cap A_3) = \frac{1}{36} = P(A_1)P(A_2)P(A_3)$, mais $P(A_1 \cap A_2) = \frac{1}{6} \neq P(A_1)P(A_2)$ et $P(A_1 \cap A_3) = \frac{1}{12} \neq P(A_1)P(A_3)$.

2.2.4 Proposition - Propriétés de l'indépendance mutuelle.

Soient A_i , pour $i \in I$, des événements mutuellement indépendants.

- Pour tout $J \subset I$, les événements A_j , pour $j \in J$, sont mutuellement indépendants.
- Posons $B_i = A_i$ ou $B_i = \overline{A_i}$ pour tout $i \in I$. Alors les événements B_i , pour $i \in I$, sont mutuellement indépendants.

Démonstration.

- Toute partie finie d'une partie de I est une partie finie de I .
- Soit J une partie finie de I .

Montrons que $P(\bigcap_{j \in J} B_j) = \prod_{j \in J} P(B_j)$, par récurrence sur $n = \text{card}\{j \in J \mid B_j = \overline{A_j}\}$.

Le cas $n = 0$ est trivial. Supposons $n \geq 1$ et le résultat établi pour toute partie finie K de I telle que $\text{card}\{j \in K \mid B_j = \overline{A_j}\} = n - 1$. Soient alors $j_0 \in J$ tel que $B_{j_0} = \overline{A_{j_0}}$, et $J' = J \setminus \{j_0\}$. Alors par formule des probabilités totales associée au système complet d'événements $(A_{j_0}, \overline{A_{j_0}})$, puis par hypothèse de récurrence (appliquée deux fois), on a

$$\begin{aligned} P(\bigcap_{j \in J} B_j) &= P(\overline{A_{j_0}} \cap (\bigcap_{j \in J'} B_j)) = P(\bigcap_{j \in J'} B_j) - P(A_{j_0} \cap (\bigcap_{j \in J'} B_j)) \\ &= \prod_{j \in J'} P(B_j) - P(A_{j_0}) \prod_{j \in J'} P(B_j) = P(\overline{A_{j_0}}) \prod_{j \in J'} P(B_j) = \prod_{j \in J} P(B_j). \end{aligned}$$

2.2.5 Remarque. L'indépendance modélise le fait que la réalisation d'un événement n'a pas d'influence sur la réalisation des autres, comme pour des événements liés à des tirages successifs avec remise (urne/jeu de cartes), ou à des lancers successifs (dé/pièce de monnaie).

2.2.6 Exemple. Dans le cas des lancers successifs d'un dé équilibré à 6 faces, on suppose l'indépendance mutuelle des événements de type $A_{\cdot, n}$, pour $n \in \mathbb{N}^*$ (définis en 1.2.6).

Ainsi, pour tout $n \in \mathbb{N}^*$, on a $P(\bigcap_{k=1}^n A_{6,k}) = (\frac{1}{6})^n$, et en passant aux complémentaires, $P(\bigcup_{k=1}^n \overline{A_{6,k}}) = 1 - P(\bigcap_{k=1}^n A_{6,k}) = 1 - (\frac{5}{6})^n$. Cela justifie les calculs de 1.3.7.

3 Lois de probabilité usuelles

3.1 Probabilités usuelles sur un univers fini (rappels)

3.1.1 Définition - Loi uniforme.

Soit Ω un ensemble fini non vide. La **loi uniforme** sur Ω , en abrégé $\mathcal{U}(\Omega)$, est la loi de probabilité P sur Ω définie par $P(\{\omega\}) = \frac{1}{\text{card}(\Omega)}$, pour tout $\omega \in \Omega$.

3.1.2 Remarques. • Plus généralement, pour tout $A \subset \Omega$, $P(A) = \frac{\text{Card}(A)}{\text{Card}(\Omega)}$.

- Situation type :** résultat d'un tirage *au hasard* dans un ensemble fini (urne, jeu de cartes) - résultat d'un lancer de dé ou d'une pièce *non truquée*, ou *équilibrée*.

3.1.3 Définition - Loi de Bernoulli.

Soit $p \in [0; 1]$. La **loi de Bernoulli** de paramètre p , en abrégé $\mathcal{B}(p)$, est la loi de probabilité P sur $\{0; 1\}$ définie par $P(\{1\}) = p$ et $P(\{0\}) = 1 - p$.

3.1.4 Remarque. Situation type : succès/échec d'une expérience dont la probabilité de succès (encodé par 1) est p , et celle de l'échec (encodé par 0) est donc $1 - p$.

3.1.5 Définition - Loi Binomiale.

Soient $n \in \mathbb{N}^*$ et $p \in [0; 1]$. La **loi binomiale** de paramètres n et p , en abrégé $\mathcal{B}(n, p)$, est la loi de probabilité P sur $\llbracket 0; n \rrbracket$ définie par $P(\{k\}) = \binom{n}{k} p^k (1 - p)^{n-k}$, pour tout $k \in \llbracket 0; n \rrbracket$.

3.1.6 Remarque. Situation type : nombre de succès obtenus lors de la répétition de n expériences indépendantes les unes des autres, et dont la probabilité de succès est p .

3.2 Probabilités usuelles sur un univers dénombrable**3.2.1 Définition - Loi géométrique.**

Soit $p \in]0; 1]$. La **loi géométrique** de paramètre p , en abrégé $\mathcal{G}(p)$, est la loi de probabilité sur $\mathbb{N}^*$ définie par $P(\{k\}) = (1 - p)^{k-1} p$, pour tout $k \in \mathbb{N}^*$.

3.2.2 Remarques. 1. **Situation type :** rang d'apparition du premier succès obtenu lors d'une suite d'expériences indépendantes les unes des autres, et dont la probabilité de succès est p .
 2. La loi géométrique est l'unique loi de probabilité *sans mémoire* sur $\mathbb{N}^*$, i.e. telle que, en notant $\forall i \in \mathbb{N}, A_i = \llbracket i; +\infty \rrbracket$, on a $\forall k, n \in \mathbb{N}, P_{A_n}(A_{n+k}) = P(A_k)$.
 Ainsi dans une suite d'expériences indépendantes les unes des autres et de même probabilité de succès, le temps d'attente d'un premier succès après n échecs ne dépend pas de n .

3.2.3 Définition - Loi de Poisson.

Soit $\lambda \in \mathbb{R}_+^*$. La **loi de Poisson** de paramètre λ , en abrégé $\mathcal{P}(\lambda)$, est la loi de probabilité P sur $\mathbb{N}$ définie par $P(\{k\}) = \frac{\lambda^k}{k!} e^{-\lambda}$, pour tout $k \in \mathbb{N}$.

3.2.4 Remarques. 1. La loi de Poisson est la *loi des événements rares* : si P est une loi de Poisson de paramètre λ et si P_n est une loi binomiale de paramètres n et $\frac{\lambda}{n}$, alors pour tout $k \in \mathbb{N}$, $P_n(\{k\}) \xrightarrow{n \rightarrow +\infty} P(\{k\})$. L'adjectif *rare* vient du fait que $\frac{\lambda}{n} \xrightarrow{n \rightarrow +\infty} 0$.
 2. **Situation type :** nombre d'événements qui se produisent dans un intervalle de temps donné (où λ est le nombre d'événements qui se produisent en moyenne dans un tel intervalle de temps) : nombre de pannes d'un appareil en un an, nombre d'accidents sur une portion de route donnée en un mois, nombre de clients entrant dans un magasin en une journée, etc.