Chapitre 1

Introduction aux Probabilités

Nous avons tous des notions intuitives de probabilités car la vie courante nous met en présence de phénomènes aléatoires : quel temps fera-t-il de main ? Quel sera le résultat du tirage du loto ? Quel est le temps de fonctionnement d'un lave-linge ? etc. Il s'agit de situations dont l'issue reste incertaine, mais au sujet desquelles nous disposons néanmoins d'une certaine information.

Le concept de probabilité a fait faire un pas considérable à « l'introduction de la rigueur dans le domaine de l'incertain ». Le calacul des probabilités, branche particulière des mathématiques, permet d'établir des lois pour mesurer l'incertain. Il a fallu, pour inventer ce type de concept, une très grande originalité de pensée.

Historiquement, le calcul des probabilités a démarré par l'étude des jeux de hasard : jeux de dés, de cartes, etc. Puis, il s'est développé au contact de nombreuses applications : prévisions météorologiques, économiques, besoin des banques et compagnies d'assurance, étude de phénomènes physiques et biologiques, et plus récemment, recommandations de produit, de musiques ou de vidéos (apprentissage automatique et profond).

Parmi les premiers problèmes de probabilités, on peut citer

- **le problème du prince de Toscane** : Le prince de Toscane demande à Galilée pourquoi, en lançant trois dés, on obtient plus souvent un total de 10 qu'un total de 9, alors qu'il y a six façons d'obtenir ces résultats ;
- **le problème du Chevalier de Méré** : Le Chevalier de Méré demande à Pascal s'il est plus facile d'obtenir un six au moins en lançant quatre fois de suite une seul dé, ou d'obtenir un double six au moins en lançant ving-quatr fois de suite deux dés.

Dans ce chapitre, nous commencerons par des rappels sur les ensembles, puis définirons la probabilité sur un univers et ses propriétés. Nous verrons comment calculer les probabilités d'événements grâce à l'équiprobabilité et au dénombrement. Nous finirons avec les probabilités conditionnelles et l'indépendance d'événements.

1.1 Rappel sur les Ensembles

Considérons un *ensemble* E, c'est-à-dire une collection d'objets appelés *éléments* ou *points* de E. Les *ensembles* seront principalement notés à l'aide de lettres *majuscules*, tandis que leurs éléments seront notés par des lettres *minuscules*. L'appartenance d'un élément i à l'ensemble E est notée $i \in E$, et $i \notin E$ signifie que l'élément i n'appartient pas à E.

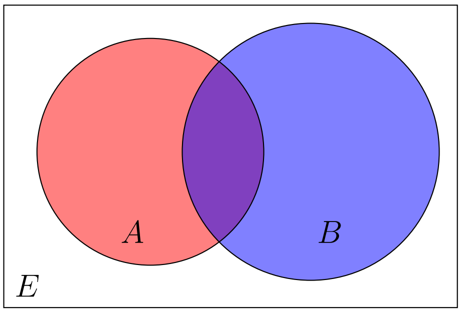
Une *partie A de E* est aussi un ensemble, appelé *sous-ensemble de E* : on écrit $A \subset E$ (on dit aussi que A est *inclus* dans E) lorsque A est un sous-ensemble de E.

On note $\mathrm{Card}(A)$ le *nombre d'éléments* contenus dans l'ensemble A si A est *fini*, c'est-à-dire que A contient un nombre fini d'éléments.

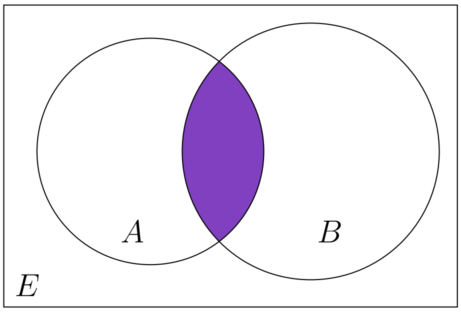
On note $\emptyset$ l'*ensemble vide*, c'est-à-dire l'ensemble ne contenant aucun élément. On a alors $\mathrm{Card}(\emptyset) = 0$.

Rappelons les opérations élémentaires sur les parties d'un ensemble :

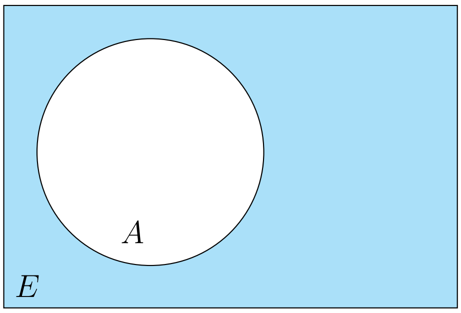
- **Réunion** : La réunion de deux ensembles A et B, notée $A \cup B$, est l'ensemble des éléments appartenant à au moins l'un de ces deux ensembles.



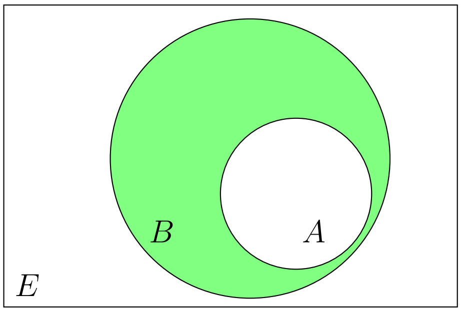
- **Intersection** : L'intersection de deux ensembles A et B, notée $A \cap B$, est l'ensemble des éléments appartenant à la fois à A et à B. Si $A \cap B = \emptyset$, on dit que les ensembles A et B sont *disjoints*.



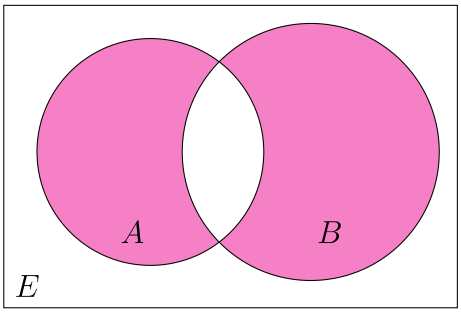
- **Complémentaire** : Soit E un ensemble et A une partie de E. Le complémentaire de A dans E, noté $E \setminus A$, ou $\overline{A}$ lorsqu'il n'y a pas d'ambiguïté sur E, est l'ensemble des éléments de E n'appartenant pas à A.



- **Différence** : Si A et B sont deux sous-ensembles de E, tels que $A \subset B$, on note $B \setminus A$ la *différence* entre A et B, c'est-à-dire l'ensemble des points qui sont dans B mais pas dans A. On a donc $B \setminus A = B \cap \overline{A}$.



- **Différence symétrique** : La *différence symétrique* de deux ensembles A et B, notée $A\Delta B$, est l'ensemble des éléments appartenant à l'un des deux ensembles A ou B, mais pas au deux. On a alors $A\Delta B = (A \setminus (A \cap B)) \cup (B \setminus (A \cap B)) = (A \cup B) \setminus (A \cap B)$.



Remarque 1.1. Les propriétés suivantes sont toujours vraies pour tous ensembles A et B :

$$A \subset A \cup B;\ B \subset A \cup B;\ A \cap B \subset A;\ A \cap B \subset B;\ A \cap B \subset A \cup B;\ \emptyset \subset A;\ \emptyset \subset B.$$

Par ailleurs, si $A \subset B$, alors $A \cap B = A$ et $A \cup B = B$.

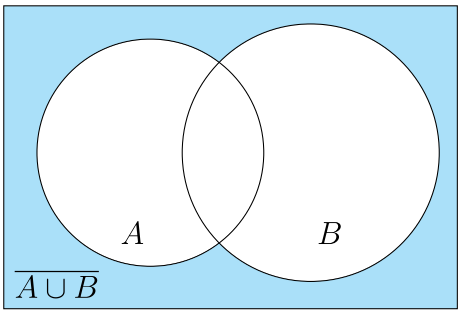
Propriété 1.1. *1. La réunion et l'intersection vérifient les propriétés suivantes*

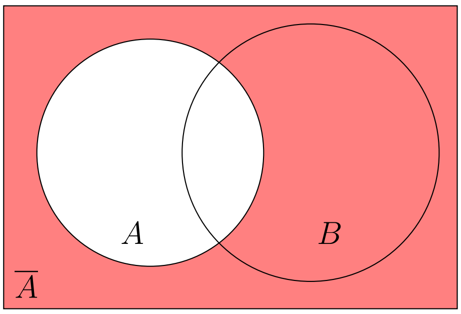
(a) **Commutativité** *:* $A \cap B = B \cap A$ *et* $A \cup B = B \cup A$.

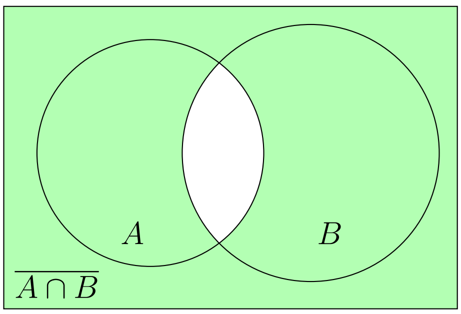
(b) **Associativité** *:* $A \cap (B \cap C) = (A \cap B) \cap C := A \cap B \cap C$ *et* $A \cup (B \cup C) = (A \cup B) \cup C := A \cup B \cup C$.

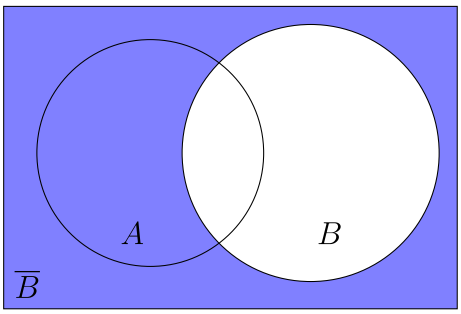
(c) **Distributivité** *:* $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$ *et* $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

2. **Lois de De Morgan** *:* $\overline{A \cup B} = \overline{A} \cap \overline{B}$ *et* $\overline{A \cap B} = \overline{A} \cup \overline{B}$.









1.2 Univers et Événements

L'ensemble des issues possibles d'une expérience aléatoire considérée forme ce que l'on nomme l'*univers des possibles*, ou tout simplement l'*univers*, et est noté Ω.

On appelle *événement élémentaire* un résultat possible d'une expérience aléatoire donnée, soit un *singleton* $\{\omega\}$, c'est-à-dire un ensemble ne contenant qu'un seul élément.

Un *événement* est une propriété dont on peut dire si elle est vraie ou non, une fois l'expérience réalisée. En termes mathématiques, un événement est une *partie de* Ω.

On peut exprimer en termes d'opérations ensemblistes les *réunions, intersections* ou *passages au complémentaire* d'événements A et B (élémentaires ou non). Ces opérations correspondent également aux opérations logiques OU, ET et NON :

- L'événement consistant à obtenir A OU B est représenté par l'ensemble $A \cup B$, qui est la réunion de A et de B.
- L'événement consistant à obtenir A ET B sera représenté par l'intersection $A \cap B$.
- La négation de l'événement, c'est-à-dire le fait qu'il ne se produise pas, sera son complémentaire $\overline{A}$. Cette négation est donc l'événement de NE PAS obtenir A.

Deux événements A et B sont dits *incompatibles* ou *disjoints* si $A \cap B = \emptyset$, c'est-à-dire qu'ils ne peuvent pas être réalisés simultanément. L'ensemble Ω est qualifié d'*événement certain*. L'ensemble vide $\emptyset$ correspond à l'*événement impossible* (aucun élément de Ω n'est réalisé, ou un élément hors Ω est réalisé).

On note $\mathcal{A}$ l'ensemble de tous les événements. Souvent (mais pas toujours), on a $\mathcal{A} = \mathcal{P}(\Omega)$, l'ensemble de toutes les parties de Ω. En tous cas, $\mathcal{A}$ doit être *stable* par les opérations ensemblistes vues précédemment : si $A, B \in \mathcal{A}$, alors on doit avoir $\overline{A}, \overline{B} \in \mathcal{A}$, $A \cap B \in \mathcal{A}$, $A \cup B \in \mathcal{A}$, et aussi $\emptyset \in \mathcal{A}$ et $\Omega \in \mathcal{A}$.

Exemple 1.1. *Expérience aléatoire* : jet d'un dé à six faces.
L'univers est $\Omega = \{1, 2, 3, 4, 5, 6\}$.
Les événements élémentaires sont :

$$\text{« obtenir un 1 » noté } \{1\}, \text{ « obtenir un 2 » noté } \{2\}, \ldots, \text{ « obtenir un 6 » noté } \{6\}.$$

Tous les événements peuvent être obtenus par manipulations ensemblistes des événements élémentaires précédents. Par exemple, l'événement « obtenir un résultat pair » consiste à obtenir un 2, un 4 ou un 6, et sera noté au choix

$$\{2, 4, 6\} = \{2\} \cup \{4\} \cup \{6\}.$$

L'écriture en termes d'événements élémentaires sera primordiale pour le calcul des probabilités et permet de représenter un très grand nombre d'événements. On notera par exemple que

$$\text{« obtenir un résultat } \leq 3 \text{ »} = \{1, 2, 3\} = \{1\} \cup \{2\} \cup \{3\}.$$

De même, l'événement « obtenir un résultat pair inférieur ou égal à 3 » sera noté

$$\{2, 4, 6\} \cap \{1, 2, 3\} = \{2\}.$$

Le contraire d'un événement A correspond à son événement complémentaire noté $\overline{A}$. Pour l'exemple précédent, « ne pas obtenir un nombre pair », sera noté

$$\overline{\{2,4,6\}} = \Omega \setminus \{2,4,6\} = \{1,3,5\},$$

c'est-à-dire l'événement « obtenir un nombre impair ».

Vérifions sur un exemple les lois de De Morgan. Soient les événements A : « avoir une face supérieure ou égale à 4 » et B : « avoir une face inférieure ou égale à 5 ». Alors $A = \{4,5,6\}$, $B = \{1,2,3,4,5\}$, $A \cap B = \{4,5\}$, $A \cup B = \Omega$, $\overline{A} = \{1,2,3\}$, $\overline{B} = \{6\}$, et on trouve bien : $\overline{A \cap B} = \{1,2,3,6\} = \overline{A} \cup \overline{B}$. De même, on trouve $\overline{A \cup B} = \emptyset = \overline{A} \cap \overline{B}$.

Définition 1.1. *Soient E un ensemble et $(B_i)_{1 \le i \le n}$ une famille de parties non vides de E. On dit que $(B_i)_{1 \le i \le n}$ est une* partition de E *si*

$$\bigcup_{i=1}^{n} B_i = E \quad \textit{et} \quad B_i \cap B_j = \emptyset, \quad i \neq j.$$

Une partition est donc un ensemble de parties dont la réunion donne l'ensemble dont ils forment la partition et qui sont incompatibles. Ainsi, l'ensemble des événements élémentaires est une partition de l'univers Ω. De même, pour n'importe quel événement A, on a que A et $\overline{A}$ forment une partition de Ω (puisque $A \cap \overline{A} = \emptyset$ et $A \cup \overline{A} = \Omega$).

Exemple 1.2. *Expérience aléatoire* : jet d'un dé à six faces.
$\{\{1\},\{2\},\{3\},\{4\},\{5\},\{6\}\}$ est une partition de Ω, de même que $\{\{1,2\},\{3,4,5,6\}\}$.

Remarque 1.2. L'ensemble de toutes les parties de Ω est notée $\mathcal{P}(\Omega)$. Si $\mathrm{Card}(\Omega) = n$, alors $\mathrm{Card}\,(\mathcal{P}(\Omega)) = 2^n$.

Exemple 1.3. $\Omega = \{1,2,3\}$. $\mathcal{P}(\Omega) = \{\{\emptyset\},\{1\},\{2\},\{3\},\{1,2\},\{1,3\},\{2,3\},\Omega\}$.

1.3 Probabilités d'événements

Définition 1.2. *La* probabilité *associée à une expérience aléatoire est une fonction qui à un événement associe un nombre réel compris entre 0 et 1, définie par :*

$$\mathbb{P} : \begin{array}{ccc} \mathcal{A} & \longrightarrow & [0\,;1] \\ A & \longmapsto & \mathbb{P}(A) \end{array},$$

où $\mathcal{A}$ est l'ensemble des événements de l'univers Ω.

La fonction est construite à partir des probabilités des événements élémentaires, dites *probabilités élémentaires*. Pour un univers Ω de cardinal n fini, on note chaque événement élémentaire A_i :« obtenir l'événement élémentaire i ». Rappelons que les événements élémentaires sont incompatibles et forment une partition de Ω. Les *probabilités élémentaires* (nombres réels compris entre 0 et 1) sont alors définies comme les probabilités associées à chacun des événements élémentaires, c'est-à-dire une famille de probabilités $(p_i)_{1 \le i \le n}$ telle que

$$\forall i \in \{1, \ldots, n\}, \quad p_i := \mathbb{P}(A_i) \in [0\,;1], \quad \text{avec} \quad p_1 + p_2 + \cdots + p_n = 1.$$

Les axiomes de Kolmogorov (1903–1987) suivants définissent les règles de construction des probabilités des événements possibles.

1.3.1 Axiomes des probabilités

Axiome I. *Événement certain :* $\mathbb{P}(\Omega) = 1$.

Axiome II. *Additivité :* si A et B sont deux événements *incompatibles* :

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) \quad \text{avec} \quad A \cap B = \emptyset.$$

Remarque 1.3. L'axiome II se généralise de la façon suivante : si $(B_i)_{i \in I}$ est une famille finie ou infinie d'événements mutuellement disjoints (c'est-à-dire $B_i \cap B_j = \emptyset$ pour tous $i \neq j$), alors $\mathbb{P}\left(\bigcup_{i \in I} B_i\right) = \sum_{i \in I} \mathbb{P}(B_i)$.

De ces axiomes, on déduit que

- la *probabilité du complémentaire* d'un événement A est donnée par

$$\mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A),$$

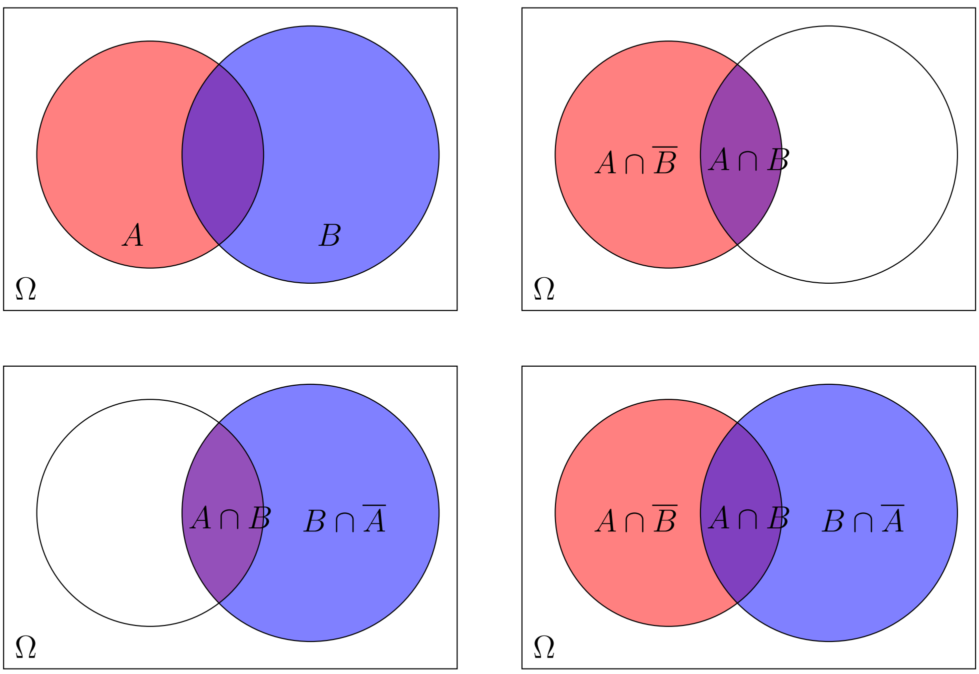
  puisque $A \cup \overline{A} = \Omega$ et $A \cap \overline{A} = \emptyset$.

- la *probabilité d'un événement impossible* est nulle : en effet, $\emptyset = \overline{\Omega}$ et on obtient le résultat en utilisant la propriété précédente.

- pour deux événements *quelconques* l'additivité devient :

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

  Ainsi, on peut montrer la formule d'additivité pour deux événements quelconques en écrivant les événements A, B et $A \cup B$ sous forme d'union disjointe (c'est-à-dire d'union d'événements incompatibles) : $A = (A \cap \overline{B}) \cup (A \cap B)$, $B = (A \cap B) \cup (B \cap \overline{A})$ et $A \cup B = (A \cap \overline{B}) \cup (A \cap B) \cup (B \cap \overline{A})$.



On en déduit que

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

Ainsi

$$\mathbb{P}(A \cap \overline{B}) = \mathbb{P}(A) - \mathbb{P}(A \cap B), \quad \mathbb{P}(B \cap \overline{A}) = \mathbb{P}(B) - \mathbb{P}(A \cap B),$$

d'où

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

- *Propriété de monotonie par inclusion* :

$$A \subset B \Longrightarrow \mathbb{P}(A) \leq \mathbb{P}(B).$$

Dans ce cas $A \cup B = B$ et $A \cap B = A$. Or $B = A \cup B = (A \cap B) \cup (B \cap \overline{A})$ et par la propriété d'additivité, on a

$$\mathbb{P}(B) = \mathbb{P}(A \cup B) = \mathbb{P}(A \cap B) + \mathbb{P}(B \cap \overline{A}) = \mathbb{P}(A) + \mathbb{P}(B \cap \overline{A}) \geq \mathbb{P}(A),$$

puisque $0 \leq \mathbb{P}(B \cap \overline{A}) \leq 1$.

Si A est inclus dans B (on dit parfois que A *implique* B), il est alors intuitif que la probabilité de A soit inférieure à celle de B (A sera toujours réalisé lorsque B le sera, et sa probabilité ne pourra être que plus petite). À l'inverse, on trouve $\overline{A} \supset \overline{B} \Longrightarrow \mathbb{P}(\overline{A}) \geq \mathbb{P}(\overline{B})$.

1.3.2 Formule des probabilités totales

Pour toute partition d'un ensemble E, en appliquant la règle d'additivité, on obtient la **formule des probabilités totales** : pour une partition $(B_i)_{i\in I}$ de E, on a

$$\mathbb{P}\left(\bigcup_{i\in I} B_i\right) = \sum_{i\in I} \mathbb{P}(B_i) = \mathbb{P}(E).$$

Si E est l'univers des possibles Ω, sachant que la probabilité de l'événement certain est fixée à 1, on trouve $\sum_{i\in I} \mathbb{P}(B_i) = 1$. On a donc que la somme des probabilités élémentaires est égale à 1 puisque les événements élémentaires forment une partition de l'univers : si $(A_i)_{1\leq i\leq n}$ représente les événements élémentaires de Ω, en posant $p_i = \mathbb{P}(A_i)$, on a alors

$$\sum_{i=1}^{n} p_i = p_1 + p_2 + \cdots + p_n = 1.$$

Exemple 1.4. Dans le cas d'un jet de dé, la partition élémentaire de l'univers en

$$\Omega = \bigcup_{i=1}^{6} \{i\} = \{1\} \cup \{2\} \cup \cdots \cup \{6\},$$

nous permet d'avoir dans tous les cas :

$$p_1 + p_2 + p_3 + p_4 + p_5 + p_6 = 1.$$

Un autre exemple de partition est donné par $\{A, B\}$ avec les événements A : « obtenir un résultat pair » et B : « obtenir un résultat impair ». On a en effet $A \cup B = \Omega$ et $A \cap B = \emptyset$, ce qui permet de vérifier la formule des probabilités totales dans ce cas particulier :

$$\mathbb{P}(A) + \mathbb{P}(B) = 1.$$

1.4 Équiprobabilité

Les événements élémentaires sont dits *équiprobables* si les probabilités élémentaires p_i sont *toutes identiques*. Cette hypothèse peut être émise à partir de statistiques descriptives ou simplement par souci de bon sens.

En cas d'équiprobabilité, **et seulement dans ce cas**, on pourra évaluer la probabilité d'un événement A par :

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

C'est loin d'être le cas en général mais c'est souvent pris par hypothèse pour les exemples de base.

La somme des probabilités élémentaires étant égale à 1, dans le cas d'équiprobabilité, chaque événement élémentaire à donc la même probabilité. Si $\text{Card}(\Omega) = n$, alors

$$p_1 = p_2 = \cdots = p_n,$$

et ainsi

$$\forall i = 1, \ldots n, \quad p_i = \frac{1}{n}.$$

Exemple 1.5. Dans le cas du jet d'un dé à six faces, l'hypothèse d'équiprobabilité peut-être émise lorsque le dé n'est pas truqué (on parle d'un dé *équilibré* dans ce cas). Les probabilités élémentaires, en notant A_i chaque événement élémentaire, sont alors :

$$p_i = \mathbb{P}(A_i) = \mathbb{P}(\text{« obtenir un } i \text{ »}) = \frac{1}{6},$$

puisque $\text{Card}(\Omega) = 6$ et $\text{Card}(A_i) = 1$. On trouve bien $\displaystyle\sum_{i=1}^{6} p_i = 1$.

1.5 Dénombrement

Dans le cas d'un expérience aléatoire où l'univers Ω est *fini* et où l'on a *équiprobabilité* des résultats possibles, on a pour tout événement $A \in \mathcal{A}$, $\mathbb{P}(A) = \dfrac{\text{Card}(A)}{\text{Card}(\Omega)}$. Ainsi, le calcul d'une probabilité se résume à des calculs de dénombrement pour déterminer le cardinal de l'univers Ω et celui de l'événement A.

1.5.1 Permutations

Définition 1.3. *Le* nombre de permutations *d'un ensemble à n éléments est égal à $n! = 1 \times 2 \times 3 \times \cdots \times (n-1) \times n$.*

Exemple 1.6. On permute au hasard les n tomes d'une encyclopédie. Quelle est la probabilité que les tomes 1 et 2 se retrouvent côte à côte dans cet ordre ?

Posons A :« les tomes 1 et 2 sont côte à côte dans cet ordre » et Ω :« toutes les permutations possibles des tomes ». Alors $\text{Card}(\Omega) = n!$ et nous sommes en équiprobabilité.

Soit B_i :« le tome 1 occupe la position i et le tome 2 occupe la position $i+1$ », pour $i \in \{1, \ldots, n-1\}$. Alors $A = \bigcup_{i=1}^{n-1} B_i$ et les B_i sont deux à deux disjoints. D'où $\mathbb{P}(A) = \sum_{i=1}^{n-1} \mathbb{P}(B_i)$ et $\mathbb{P}(B_i) = \dfrac{\text{Card}(B_i)}{\text{Card}(\Omega)}$.

Or $\text{Card}(B_i) = (n-2)!$, soit toutes les façons de placer les $n-2$ tomes différents de 1 et 2 en dehors des emplacements i et $i+1$. Donc

$$\mathbb{P}(A) = \sum_{i=1}^{n-1} \mathbb{P}(B_i) = \sum_{i=1}^{n-1} \frac{(n-2)!}{n!} = \sum_{i=1}^{n-1} \frac{1}{n(n-1)} = (n-1)\frac{1}{n(n-1)} = \frac{1}{n}.$$

1.5.2 Arrangements

Définition 1.4. *Le* nombre d'arrangements *de p éléments parmi n (avec $0 \leq p \leq n$) est le nombre de parties ordonnées de p éléments distincts dans un ensemble à n éléments. On le note A_n^p et il est défini par* $A_n^p = \dfrac{n!}{(n-p)!}$.

Lorsque l'on choisit p objets parmi n objets et que l'ordre dans lequel les objets sont sélectionnés revêt une importance, on peut les représenter par un p-uplet d'éléments distincts et on en constitue une liste ordonnée sans répétition possible, c'est-à-dire dans laquelle l'ordre des éléments est pris en compte (si l'on permute deux éléments de la liste, on a une liste différente, et un élément ne peut être présent qu'une seule fois). Une telle liste ordonnée est un arrangement.

Exemple 1.7. Quelle est la probabilité que, dans une assemblée de n personnes, tous les participants aient une date d'anniversaire différente ?

Soient A :« Toutes les dates d'anniversaire sont différentes » et Ω : « ensemble de toutes les dates d'anniversaire possibles ». En supposant que toutes les années ne font que 365 jours, on a $\text{Card}(\Omega) = 365^n$ et par équiprobabilité, on a

— si $n > 365$, $\text{Card}(A) = 0$, donc $\mathbb{P}(A) = 0$;

— si $n \leq 365$, $\text{Card}(A) = A_{365}^n$, donc $\mathbb{P}(A) = \dfrac{A_{365}^n}{365^n}$.

1.5.3 Combinaisons

Définition 1.5. *Le* nombre de combinaisons *de p éléments parmi n (avec $0 \leq p \leq n$) est le nombre de parties de p éléments distincts dans un ensemble à n éléments sans que l'ordre des éléments ne compte. Autrement dit, les combinaisons de taille p d'un ensemble à n éléments sont les sous-ensembles qui ont pour taille p et dont les éléments sont différents. On le note C_n^p et il est défini par* $C_n^p = \dfrac{n!}{p!(n-p)!}$.

Contrairement aux arrangements, les combinaisons s'intéressent uniquement aux éléments choisis parmi l'ensemble, et non à l'ordre dans lequel ils sont tirés.

On a par ailleurs toujours $C_n^0 = C_n^n = 1$ et $C_n^k = C_n^{n-k}$. On peut calculer tous les coefficients (en utilisant éventuellement un triangle de Pascal) par l'intermédiaire de la **formule de Pascal**

$$C_n^k = C_{n-1}^k + C_{n-1}^{k-1}.$$

n											
0						$C_0^0 = 1$					
1					$C_1^0 = 1$		$C_1^1 = 1$				
2				$C_2^0 = 1$		$C_2^1 = 2$		$C_2^2 = 1$			
3			$C_3^0 = 1$		$C_3^1 = 3$		$C_3^2 = 3$		$C_3^3 = 1$		
4		$C_4^0 = 1$		$C_4^1 = 4$		$C_4^2 = 6$		$C_4^3 = 4$		$C_4^4 = 1$	
5	$C_5^0 = 1$		$C_5^1 = 5$		$C_5^2 = 10$		$C_5^3 = 10$		$C_5^4 = 5$		$C_5^5 = 1$
$\vdots$											

Triangle de Pascal.

Exemple 1.8. Quelle est la probabilité de gagner au loto en jouant une grille ?

Soient A :« Gagner au loto en jouant une grille » et Ω : « Ensemble de toutes les grilles possibles ».

Au loto, il faut choisir 6 numéros distincts entre 1 et 49, l'ordre ne compte pas. On a alors $\mathrm{Card}(\Omega) = C_{49}^6$ et $\mathrm{Card}(A) = 1$. D'où $\mathbb{P}(A) = \dfrac{1}{C_{49}^6}$.

1.6 Probabilités Conditionnelles

On cherche ici à « mesurer » l'influence d'un événement sur la probabilité de réalisation d'un autre. Par exemple, lorsque l'on tire une carte dans un jeu, la valeur de la seconde carte tirée n'aura pas la même probabilité selon que l'on remet la première carte ou non.

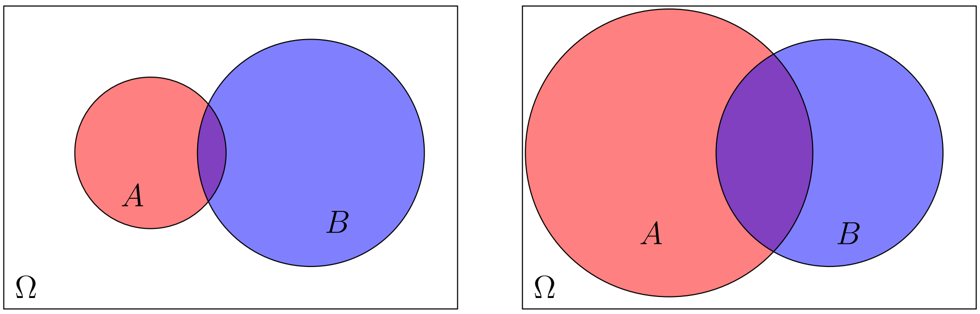
Ce constat a conduit à la notion de *probabilité conditionnelle* de B *sachant* A.

Définition 1.6. *Soient deux événements A et B tels que $\mathbb{P}(A) \neq 0$. La* probabilité conditionnelle *de B* sachant A *est alors définie par*

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

Il convient alors de ne pas confondre $\mathbb{P}(B|A)$ et $\mathbb{P}(A \cap B)$: $\mathbb{P}(B|A)$ évalue les chances d'obtenir B lorsque A est déjà réalisé (de manière certaine) alors que la $\mathbb{P}(A \cap B)$ évalue celles de voir A et B réalisés simultanément.

Ici, on évalue donc les chances de B sur un sous-ensemble de l'univers – celui pour lequel A est réalisé – et on pondère la probabilité de l'intersection selon la « taille » de A : plus A est important, c'est-à-dire plus $\mathbb{P}(A)$ est grand, plus $A \cap B$ a des chances de se réaliser.



On obtient un nombre $\mathbb{P}(B|A)$ entre 0 et 1, toujours plus grand que $\mathbb{P}(A \cap B)$ car on divise cette dernière par le nombre $\mathbb{P}(A)$ plus petit que 1.

Lorsque que A est fixé, cela détermine une nouvelle probabilité

$$\mathbb{P}_A : \begin{array}{ccc} \mathcal{A} & \longrightarrow & [0\,;1] \\ B & \longmapsto & \mathbb{P}(B|A) \end{array}.$$

On note parfois la probabilité conditionnelle $\mathbb{P}_A(B)$.

Remarque 1.4. Il s'agit bien d'une probabilité car elle vérifie les axiomes constitutifs (I-II) et les propriétés qui en découlent. En particulier, $\mathbb{P}(\overline{B}|A) = 1 - \mathbb{P}(B|A)$. Mais par contre, $\mathbb{P}(B|A) \neq 1 - \mathbb{P}(B|\overline{A})$.

Les probabilités conditionnelles permettent également d'obtenir deux expressions de la probabilité de l'intersection :

$$\mathbb{P}(A \cap B) = \mathbb{P}(A|B) \cdot \mathbb{P}(B) = \mathbb{P}(B|A) \cdot \mathbb{P}(A).$$

On en déduit aisément la **formule de Bayes I** :

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(A|B) \cdot \mathbb{P}(B)}{\mathbb{P}(A)}. \tag{1.1}$$

Les probabilités conditionnelles permettent aussi d'écrire, en utilisant la formule des probabilités totales :

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

car $A \cap B$ et $A \cap \overline{B}$ forment une partition de A.

De là, on obtient une nouvelle version de (1.1), souvent plus utile, appelée **formule de Bayes II** :

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(A|B) \cdot \mathbb{P}(B)}{\mathbb{P}(A|B)\mathbb{P}(B) + \mathbb{P}(A|\overline{B})\mathbb{P}(\overline{B})}. \tag{1.2}$$

Cette formule est importante et souvent utile en probabilités pour corriger de mauvaises intuitions dues à une vision trop équiprobable des événements. La formule (1.2) permet en effet d'évaluer le rapport entre le « nombre de cas favorables » et le « nombre de cas possibles ».

Exemple 1.9. Dans la région Provence-Alpes-Côte d'Azur (PACA), 5% des PME font faillite dans une année. Ce pourcentage est de 1% pour les grandes entreprises, qui ne représentent que 30% des entreprises de la région. Une entreprise de Marseille fait faillite. Quelle est la probabilité que ce soit une PME ?

Soient les événements P : « L'entreprise est une PME » et F : « L'entreprise fait faillite ». D'après l'énoncé, on a $\mathbb{P}(F|P) = 0,05$, $\mathbb{P}(F|\overline{P}) = 0,01$ et $\mathbb{P}(\overline{P}) = 0,3$. Alors, en utilisant (1.2),

$$\mathbb{P}(P|F) = \frac{\mathbb{P}(F|P)\mathbb{P}(P)}{\mathbb{P}(F|P)\mathbb{P}(P) + \mathbb{P}(F|\overline{P})\mathbb{P}(\overline{P})} = \frac{0,05 \times 0,7}{0,05 \times 0,7 + 0,01 \times 0,3} \approx 0,9211.$$

On peut généraliser les formules précédentes de la façon suivante

Proposition 1.1. *Soit $(B_i)_{1 \leq i \leq n}$ une partition de Ω.*

1. **Formules des probabilités totales** *: $\forall A \in \mathcal{A}$, $\mathbb{P}(A) = \displaystyle\sum_{i=1}^{n} \mathbb{P}(A \cap B_i)$.*

   Si, de plus, $\mathbb{P}(B_i) \neq 0$ pour tout i, alors : $\forall A \in \mathcal{A}$, $\mathbb{P}(A) = \displaystyle\sum_{i=1}^{n} \mathbb{P}(A|B_i)\mathbb{P}(B_i)$.

2. **Formule de Bayes II** *: $\forall A \in \mathcal{A}$ tel que $\mathbb{P}(A) \neq 0$, $\forall i = 1, \ldots, n$,*

$$\mathbb{P}(B_i|A) = \frac{\mathbb{P}(A|B_i)\mathbb{P}(B_i)}{\sum_{j=1}^{n} \mathbb{P}(A|B_j)\mathbb{P}(B_j)}.$$

1.7 Indépendance

1.7.1 Indépendance de deux événements

Définition 1.7. *Deux événements sont dits* indépendants *si et seulement si*

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

On le note $A \perp\!\!\!\perp B$.

Deux événements sont indépendants lorsque la probabilité de l'un est la même que l'autre soit réalisé ou non, c'est-à-dire

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(A) \cdot \mathbb{P}(B)}{\mathbb{P}(B)} = \mathbb{P}(A) \quad \text{et de même} \quad \mathbb{P}(B|A) = \mathbb{P}(B).$$

Remarque 1.5. $A \perp\!\!\!\perp B \Longleftrightarrow A \perp\!\!\!\perp \overline{B} \Longleftrightarrow \overline{A} \perp\!\!\!\perp B \Longleftrightarrow \overline{A} \perp\!\!\!\perp \overline{B}$.

Exemple 1.10. Soient deux événements A et B avec $\mathbb{P}(A) = 0,7$, $\mathbb{P}(B) = 0,8$ et $\mathbb{P}(A \cup B) = 0,9$. Les événements A et B son-ils indépendants ?

On sait que

$$\begin{aligned} & \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B) \\ \Longleftrightarrow \quad & \mathbb{P}(A \cap B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cup B) = 0,7 + 0,8 - 0,9 = 0,6. \end{aligned}$$

On voit alors que $\mathbb{P}(A \cap B) \neq \mathbb{P}(A) \cdot \mathbb{P}(B) = 0,7 \times 0,8 = 0,56$. Donc A et B ne sont pas indépendants. On remarque par ailleurs qu'ils ne sont pas incompatible puisque $\mathbb{P}(A \cap B) \neq 0$.

Une autre situation usuelle, où l'hypothèse d'indépendance est étudiée, est celle des tirages au sort successifs avec ou sans remise. Lorsque le tirage est effectué *avec remise* du premier élément tiré au sort, on se retrouve dans une situation identique lors du second tirage au sort, de sorte que le résultat du premier n'influence en rien le résultat du second : *deux tirages successifs avec remise sont indépendants*. Lorsque le tirage est au contraire effectué *sans remise*, l'élément tiré au sort au premier tirage ne peut plus être tiré au second, modifiant donc les probabilités des événements du second tirage par rapport au premier, de sorte que les résultats des deux tirages sont liés : *deux tirages successifs sans remise ne sont pas indépendants*.

Exemple 1.11. Soient deux jets consécutifs d'un même dé équilibré à 6 faces. Une hypothèse naturelle consiste à considérer ces événements comme étant indépendants. Pour les événements A : « obtenir un six au 1$^{\text{er}}$ jet » et B : « obtenir un six au 2$^{\text{nd}}$ jet », on a :

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36},$$

de sorte que la probabilité d'obtenir un double six est évaluée à environ $2,78\%$.

1.7.2 Indépendance mutuelle et indépendance deux à deux

Définition 1.8. *Soit $(A_i)_{1 \leq i \leq n}$ une famille d'événements.*

1. *Les événements* $(A_i)_{1\leq i\leq n}$ *sont dits* mutuellement indépendants *(ou* indépendants dans leur ensemble*) si*
$$\forall I \subset \{1,\ldots,n\}, \quad \mathbb{P}\left(\bigcap_{i\in I} A_i\right) = \prod_{i\in I} \mathbb{P}(A_i).$$
2. *Les événements* $(A_i)_{1\leq i\leq n}$ *sont dits* deux à deux indépendants *si*
$$\forall i,j \in \{1,\ldots,n\},\ i \neq j, \quad \mathbb{P}(A_i \cap A_j) = \mathbb{P}(A_i)\cdot\mathbb{P}(A_j).$$

Remarque 1.6. L'indépendance mutuelle est plus forte que l'indépendance deux à deux.

Exemple 1.12. On lance simultanément deux dés discernables et on considère les événements suivants : A_1 : « le premier dé donne un nombre pair », A_2 : « le second dé donne un nombre pair » et A_3 : « la somme des dés est paire ».

Ces événements sont deux à deux indépendants, mais pas mutuellement indépendants. En effet, $\mathbb{P}(A_1) = \mathbb{P}(A_2) = \mathbb{P}(A_3) = 1/2$, $\mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_1 \cap A_3) = \mathbb{P}(A_2 \cap A_3) = 1/4$ et $\mathbb{P}(A_1 \cap A_2 \cap A_3) = 1/4 \neq \mathbb{P}(A_1)\cdot\mathbb{P}(A_2)\cdot\mathbb{P}(A_3)$.