Chapitre 02

Le formalisme de la théorie des probabilités

L'objet de la théorie des probabilités est de modéliser des phénomènes complexes dont il n'est pas en général possible de prédire avec certitude leur évolution ou les conséquences qu'ils peuvent engendrer. L'archétype d'un tel phénomème est le lancer d'une pièce à pile ou face : les mécanismes physiques à prendre en compte pour décrire l'expérience du lancer sont d'une telle complexité qu'il n'est pas envisageable de répondre de façon déterministe à la question la pièce va-t-elle tomber coté pile, face, sur la tranche ? On dit alors que le résultat de l'expérience est *aléatoire* ou encore *stochastique*. Voici d'autres exemples d'expériences usuelles dont le résultat est de nature aléatoire :

Expérience	Résultat observable
Lancer d'un dé	Un entier $k \in \{1, \ldots, 6\}$
Lancer d'une fléchette sur une cible	Point d'impact
Sondage à la sortie des urnes au cours d'un référendum	Nombre de Oui et de Non dans l'échantillon
Saut en longueur dans une compétition d'athlétisme	Saut éventuellement mordu, sinon un nombre $\ell \geqslant 0$
Mouvement d'un grain de pollen dans un liquide	Une trajectoire continue dans l'espace à trois dimensions

Pour modéliser ce type d'expériences, la démarche du probabiliste consiste tout d'abord à en préciser tous les résultats possibles. Ensuite, chaque résultat possible se voit attribuer un certain poids, une probabilité. Dans l'exemple du lancer à pile ou face, l'ensemble des résultat possibles est $\{\text{pile}, \text{face}, \text{tranche}\}$ ou plus simplement, si l'on néglige la possibilité que la pièce tombe sur la tranche, $\{\text{pile}, \text{face}\}$. Si la pièce est équilibrée, il est alors naturel de choisir les probabilités que la pièce tombe sur pile ou face égales à un demi. À la question de quel coté va tomber la pièce, la réponse du probabiliste n'est alors pas déterministe mais statistique : la pièce a une chance sur deux de tomber sur pile ou sur face.

Dans les prochains paragraphes, nous précisons le formalisme général de la théorie des probabilités, c'est-à-dire le cadre mathématique rigoureux dans lequel se formule cette théorie. Ce formalisme a été introduit au début du vingtième siècle par le mathématicien russe A. Kolmogorov.

1.1 Espace de probabilité

Avant toute chose, nous commençons par énoncer quelques rappels élémentaires de théorie des ensembles ainsi que de combinatoire (appelée aussi dénombrement) qui nous seront indispensables dans la suite.

1.1.1 Quelques rappels

Rappels de théorie des ensembles

Étant donnés deux ensembles A et B, on appelle l'*union* de A et B et on note $A \cup B$ l'ensemble formé des éléments qui appartiennent à l'ensemble A ou à l'ensemble B. On appelle *intersection* de A et B et on note $A \cap B$ l'ensemble formé des éléments qui appartiennent à l'ensemble A et l'ensemble B. Si l'intersection de A et B est vide, *i.e.* $A \cap B = \emptyset$, on dit que les ensembles A et B sont *disjoints*. Dans ce cas, l'union de A et B est dite *union disjointe*, et l'on note $A \sqcup B$.

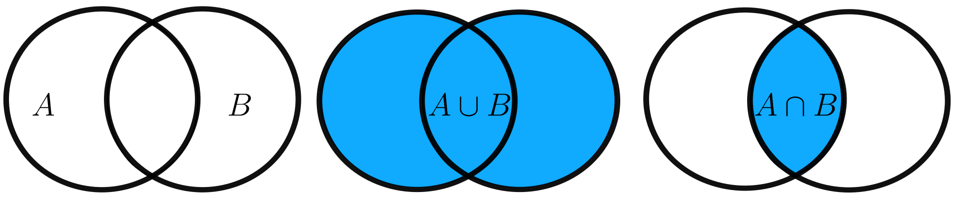

FIGURE 1.1 – Union et intersection de deux ensembles

Plus généralement, étant donnés des ensembles $(A_i)_{i \in I}$ indexés par un ensemble d'indice I, on note $\cup_{i \in I} A_i$ l'ensemble des éléments qui appartiennent à l'un des A_i et $\cap_{i \in I} A_i$ l'ensemble des éléments qui appartiennent à tous les A_i, de sorte que

— $x \in \bigcup_{i \in I} A_i$ signifie que x appartient à l'un des ensembles A_i ;

— $x \in \bigcap_{i \in I} A_i$ signifie que x appartient à tous les ensembles A_i.

Soient trois ensembles A, B et Ω tels que $A \subset \Omega$ et $B \subset \Omega$. On appelle *complémentaire* de A (dans Ω) et on note A^c l'ensemble des éléments de Ω qui ne sont pas dans A. On désigne par B *privé de* A et on note $B \backslash A$, l'ensemble des éléments de B qui ne sont pas dans A, c'est-à-dire $B \cap A^c$.
Si $(A_i)_{i \in I}$ est une famille d'ensembles inclus dans Ω, on a alors les relations :

$$\left(\bigcup_{i \in I} A_i \right)^c = \bigcap_{i \in I} A_i^c, \qquad \left(\bigcap_{i \in I} A_i \right)^c = \bigcup_{i \in I} A_i^c.$$

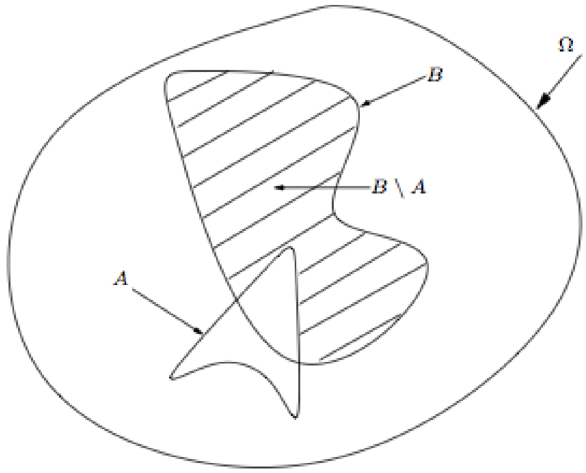

FIGURE 1.2 – Soustraction de deux ensembles.

Par exemple, si l'on considère les ensembles G et A des germanophones et des anglophones dans la population francaise, le complémentaire de $G \cap A$ est $G^c \cup A^c$, *i.e.* le contraire de "parler allemand et anglais" et "ne pas parler allemand ou ne pas parler anglais.

On appelle cardinal de A et on note Card(A) ou encore $\#A$ le nombre d'éléments qu'il contient. Si A et B sont des ensembles finis, on a la relation

$$\mathrm{Card}(A) + \mathrm{Card}(B) = \mathrm{Card}(A \cup B) + \mathrm{Card}(A \cap B).$$

Étant donné un ensemble Ω, on désigne par $\mathcal{P}(\Omega)$ l'ensemble de ses parties. Par exemple, si $\Omega = \{0, 1\}$, alors $\mathcal{P}(\Omega) = \{\emptyset, \{0\}, \{1\}, \{0, 1\}\}$. Si l'ensemble Ω est fini de cardinal n, alors on a $\mathrm{Card}(\mathcal{P}(\Omega)) = 2^n$.

On rappelle les notations usuelles concernant les sommes et les produits, si $a_1, a_2, \ldots, a_n$ sont des nombres réels :

$$\sum_{i=1}^{n} a_i = a_1 + a_2 + \ldots + a_n, \quad \prod_{i=1}^{n} a_i = a_1 \times a_2 \times \ldots \times a_n.$$

Rappels de combinatoire

Soit A un ensemble à n éléments. Le nombre de permutations des éléments de A est appelé *factorielle* n, que l'on note $n!$. Ce nombre est égal à

$$n! := n \times (n-1) \times (n-2) \times \cdots \times 2 \times 1.$$

Par exemple, il y a $6 = 3!$ permutations possibles de 3 symboles a, b, c : (a, b, c), (a, c, b), (b, a, c), (b, c, a), (c, a, b), (c, b, a).

Remarque 1.1.1. Tous les élements sont ici supposés distinguables et on tient compte de l'ordre des éléments.

On peut aussi définir la factorielle grâce à la fonction Γ : $\Gamma(x) = \int_0^\infty u^{x-1}e^{-u}du$ qui a les propriétés suivantes : $\Gamma(n+1) = n!$ pour n entier et $\Gamma(x+1) = x\Gamma(x)$. La formule de Stirling permet de construire une estimation asymptotique de la factorielle

$$n! \approx n^n e^{-n}\sqrt{2\pi n}(1 + \frac{1}{12n} + \frac{1}{288n^2} + \ldots).$$

Le nombre de façons de choisir p éléments de A parmi les n est appelé *arrangement de p objets parmi n*. Il est souvent noté A_n^p et vaut :

$$A_n^p := \frac{n!}{(n-p)!} = n \times (n-1) \times (n-2) \times \cdots \times (n-p+1).$$

Remarque 1.1.2. Ici encore, on tient compte de l'ordre des éléments.

Le nombre de façons de choisir p éléments de A parmi les n éléments sans tenir compte de l'ordre est appelé *combinaison de p objets parmi n*. Il est noté C_n^p ou encore $\binom{n}{p}$ et vaut :

$$C_n^p := \frac{n!}{p!(n-p)!} = \frac{A_n^p}{p!}.$$

On a les propriétés suivantes :

$$C_n^0 = C_n^n = 1, \quad C_n^p = C_n^{n-p}, \quad C_n^p = C_{n-1}^{p-1} + C_{n-1}^p, \quad \sum_{p=1}^n C_n^p = 2^n.$$

Quelques exercices d'entrainement

Exercice 1 :
Combien existe-t-il de plaques minéralogiques à 7 caractères (les 2 premiers étant des lettres et les 5 autres des chiffres) ? Même question si l'on impose que les répétitions de lettres ou de chiffres sont exclues.

Correction : Si on autorise les répétitions, on a 26×26 choix pour les lettres, et $10 \times 10 \times ... \times 10 = 10^5$ pour les chiffres, soit au total : $N = 26^2 \times 10^5$ possibilités. Si les répétitions sont proscrites, alors on a 26×25 choix pour les lettres et $10 \times 9 \times 8 \times 7 \times 6$ choix pour les chiffres, soit au total : $N = 26 \times 25 \times 10 \times 9 \times 8 \times 7 \times 6 = 19656000$ possibilités.

Exercice 2 :
On doit asseoir sur un même rang 4 allemands, 3 français, et 3 anglais ; les gens de même nationalité devant rester groupés. Combien de dispositions sont possibles ?

Correction : Les personnes de même nationalité devant rester groupées, on peut tout d'abord choisir l'ordre des 3 nationalités sur le rang : pour cela on $N = 3! = 6$ configurations possibles :

D	F	GB

D	GB	F

F	D	GB

F	GB	D

GB	D	F

GB	F	D

Ensuite, on peut permuter les personnes au sein d'une même nationalité, au total il y a donc $N = 6 \times 4! \times 3! \times 3!$ configurations.

Exercice 3 :
Combien existe-t-il d'arrangements différents avec les lettres des mots suivants : a) pinte ; b) proposition ; c) Mississipi ; d) arrangement ?

Correction : Dans le mot "pinte" chaque lettre apparaît une seule fois, le nombre d'arrangements de lettres distincts que l'on peut former est donc $5! = 120$. Dans le mot "proposition", il y a 11 lettres dont 2 "p", 3 "o", 2 "i". Pour ne pas compter plusieurs fois le même arrangement (par exemple, si on ne regarde que les "p", "pproosition" apparaît deux fois, si on ne regarde que les "o", "oooprpsitin" apparaît $3! = 6$ fois...) on est amené à diviser le nombre des permutations possibles des lettres par $2! \times 3! \times 2! = 24$. Le nombre d'arrangements distincts est donc

$$N = \frac{11!}{2! \times 3! \times 2!} = 1663200.$$

De même pour "Mississipi", il y a 10 lettres dont 4 "i" et 4 "s", le nombre de possibilités est alors $N = \frac{10!}{4! \times 4!} = 6300$. Pour "arrangement", on trouve

$$N = \frac{11!}{2! \times 2! \times 2! \times 2!} = 2494800.$$

Exercice 4 :
On veut former un comité de 7 personnes, constitué de 2 démocrates, 2 républicains, et 3 indépendants. On a le choix parmi 6 démocrates, 5 républicains, et 4 indépendants. Combien de choix sont possibles ?

Correction : On détermine le nombre de possibilités dans chacune des 3 obédiences, le nombre total de choix possibles est alors le produit de ces trois nombres. Pour les démocrates, on a C_6^2 choix, pour les républicains C_5^2, et pour les indépendants C_4^3. Le nombre comités distincts que l'ont peut ainsi former est :

$$N = C_6^2 \times C_5^2 \times C_4^3 = 600.$$

Nous pouvons maintenant introduire la notion fondamentale d'espace de probabilité. Il s'agit d'un triplet $(\Omega, \mathcal{F}, \mathbb{P})$, où Ω est un ensemble, $\mathcal{F}$ une tribu, et $\mathbb{P}$ une mesure de probabilité. Nous donnons la définition et le rôle de ces trois objets dans les prochains paragraphes.

1.1.2 Univers des possibles

Comme indiqué ci-dessus, le premier élément Ω d'un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ est un ensemble. Plus précisément, on a la définition suivante :

Définition 1.1.3. Étant donnée une expérience aléatoire, on appelle *univers des possibles*, et l'on note souvent Ω, l'ensemble des résultats possibles de l'expérience.

La description explicite de l'ensemble Ω est la première étape fondamentale dans la modélisation d'un phénomène aléatoire. Comme nous le verrons plus loin, le choix de Ω n'est pas toujours unique. Les pseudo-paradoxes qui apparaissent parfois entre deux protagonistes concernant une expérience où intervient le hasard relèvent le plus souvent de deux choix distincts d'ensembles des possibles. Aussi est-t-il important de bien choisir l'ensemble Ω avec lequel on travaille, et de se tenir à ce choix.

Exemple 1.1.4. Voici quelques expériences aléatoires et les ensembles des possibles correspondants :

1. On jette un dé. L'ensemble Ω est alors l'ensemble $\{1,2,3,4,5,6\}$ à 6 éléments. Ici, l'élément $\omega = 2 \in \Omega$ signifie que la face visible du dé après le lancer est 2.
2. On jette deux dés. L'ensemble Ω est alors l'ensemble $\{1,2,3,4,5,6\}^2$ c'est-à-dire $\Omega = \{(i,j),\ i,j \in \{1,2,3,4,5,6\}\} = \{(1,1),(2,1),(3,6),\ldots\}$. L'élément $\omega = (3,5) \in \Omega$ correspond à un lancer où le premier dé donne 3 et le second dé donne 5 ;
3. On joue dix fois à pile ou face. On a alors $\Omega = \{\text{pile}, \text{face}\}^{10}$. On peut aussi choisir pour ensemble des possibles $\Omega' = \{\text{pile}, \text{face}, \text{tranche}\}^{10}$ si l'on veut tenir compte du fait que la pièce peut tomber sur la tranche ;
4. On fait un sondage auprès de 1000 personnes à la sortie d'un référendum. On a alors $\Omega = \{\text{oui}, \text{non}, \text{blanc}\}^{1000}$;
5. On distribue une main au poker. L'ensemble des possibles correspondant à cette expérience est alors $\Omega = \{\text{choix de 5 cartes parmi 52}\}$ qui a pour cardinal le coefficient binomial $\binom{52}{5}$.

Remarque 1.1.5. Il n'est pas toujours possible de décrire de façon rigoureuse l'univers des possibles. On peut penser par exemple à l'expérience aléatoire de la météo du lendemain ! Néanmoins, dans les cas simples que nous envisagerons dans la suite, on peut la plupart du temps décrire explicitement l'ensemble Ω.

1.1.3 Tribu et évènements

Dans la suite, on va vouloir calculer la probabilité de certaines parties de l'ensemble des possibles Ω. Par exemple, lorsque l'on jette deux dés, l'ensemble des possibles Ω est $\{1,2,3,4,5,6\}^2$, et l'on voudrait calculer la probabilité que le premier dé donne 2 et le second est impair, c'est-à-dire la probabilité de l'ensemble :

$$\{(2,j), j = 1,3,5\}.$$

Définition 1.1.6. On appelle *tribu* et on note $\mathcal{F}$ l'ensemble des parties de Ω dont on pourra calculer la probabilité. Lorsque l'ensemble Ω est fini ou dénombrable, on choisira pour $\mathcal{F}$ l'ensemble de toutes les parties de Ω c'est-à-dire :

$$\mathcal{F} = \mathcal{P}(\Omega).$$

Définition 1.1.7. Les éléments de $\mathcal{F} = \mathcal{P}(\Omega)$ sont appelés des *évènements.* On dit encore que ce sont des ensembles *mesurables* par rapport à la tribu $\mathcal{F}$.

Le texte qui suit, en miniature, pourra être ommis par le lecteur, il concerne la "vraie" définition de la notion de tribu. En effet, sauf lorsque Ω est fini ou dénombrable, on ne peut pas s'intéresser à l'ensemble $\mathcal{P}(\Omega)$ de toutes les parties de Ω, celui-ci étant en quelque sorte "trop gros". On se restreindra donc à un sous-ensemble $\mathcal{F}$ de $\mathcal{P}(\Omega)$, qui constituera l'ensemble des parties dont on peut calculer la probabilité. Afin d'obtenir un modèle aussi cohérent que possible, il importe néanmoins d'imposer certaines conditions de stabilité à l'ensemble $\mathcal{F}$: par union, intersection, passage au complémentaire, etc. Aussi, voici la "vraie" notion de tribu.

Définition 1.1.8. Soit Ω un ensemble et $\mathcal{F}$ un sous-ensemble de parties de Ω, *i.e.* $\mathcal{F} \subset \mathcal{P}(\Omega)$. On dit que $\mathcal{F}$ est une tribu si elle vérifie les 3 conditions suivantes :

1. $\Omega \in \mathcal{F}$;
2. si A appartient à $\mathcal{F}$, alors son complémentaire A^c appartient aussi à $\mathcal{F}$;
3. si $(A_n)_{n\in\mathbb{N}}$ est une suite d'éléments de $\mathcal{F}$, alors $\bigcup_{n=0}^{\infty} A_n$ appartient à $\mathcal{F}$.

On vérifie sans problème à partir des trois axiomes ci-dessus que toute tribu $\mathcal{F}$ contient l'ensemble vide $\emptyset$, est stable par union finie, intersection finie ou dénombrable. Ainsi, on retiendra qu'une tribu est stable par combinaisons au plus dénombrables d'opérations usuelles sur les ensembles, bref par toutes les manipulations classiques.

Exemple 1.1.9. Voici trois exemples classiques de tribus :

— La tribu triviale : $\mathcal{F} = \{\emptyset, \Omega\}$;
— La tribu engendreée par une partie A de Ω : $\mathcal{F} = \{\emptyset, A, A^c, \Omega\}$;
— La tribu pleine : $\mathcal{F} = \mathcal{P}(\Omega)$.

Exemple 1.1.10. On jette deux dés discernables. L'ensemble des résultats possibles est alors

$$\Omega = \{(i,j),\ i,j \in \{1,2,3,4,5,6\}\}.$$

La tribu engendrée par le singleton $\{(1,1)\}$ est composée des quatre évènements $\{\emptyset, (1,1), \Omega\backslash(1,1), \Omega\}$. Si on choisit la tribu pleine $\mathcal{F} = \mathcal{P}(\Omega)$, l'évènement "la somme des deux dés est supérieure ou égale à dix" correspond à l'ensemble $\{(5,5),(5,6),(6,5)\}$; si on introduit les deux ensembles

$$A = \{\text{les deux dés sont pairs}\}, \text{ et } B = \{\text{les deux sont distincts}\},$$

alors $A \cap B$ correspond à l'évènement $\{(2,4),(4,2),(2,6),(6,2),(4,6),(6,4)\}$.

En pratique, lorsque Ω est fini ou dénombrable, on considère donc en général la tribu pleine $\mathcal{P}(\Omega)$. En revanche, si Ω n'est pas dénombrable, comme c'est le cas dans l'exemple d'une suite infinie de lancers ($\Omega = \{\text{pile}, \text{face}\}^{\mathbb{N}}$), on ne considérera pas la tribu $\mathcal{F} = \mathcal{P}(\Omega)$, mais une tribu plus petite.

Le couple $(\Omega, \mathcal{F})$ est appelée espace mesurable ou encore espace probabilisable. Pour compléter la description de la notion d'espace de probabilité, il nous reste à introduire la notion de mesure de probabilité. C'est l'objet du prochain paragraphe.

1.1.4 Probabilité

Une fois fixés un univers Ω et une tribu $\mathcal{F}$, on peut définir proprement ce qu'est une probabilité $\mathbb{P}$ sur $(\Omega, \mathcal{F})$ et par suite un espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$: à chaque évènement, on associe un nombre positif compris entre 0 et 1, sa probabilité.

Définition 1.1.11. On appelle probabilité sur $(\Omega, \mathcal{F})$ une application $\mathbb{P}$ de $\mathcal{F}$ dans l'intervalle $[0, 1]$ telle que :

1. $\mathbb{P}(\Omega) = 1$;
2. pour toute famille au plus dénombrable d'évènements deux à deux disjoints $(A_n)_{n\geqslant 0}$ on a $\mathbb{P}(\bigcup_{n\geqslant 0} A_n) = \sum_{n\geqslant 0} \mathbb{P}(A_n)$.

Le triplé $(\Omega, \mathcal{F}, P)$ est alors appelé *espace de probabilité* ou encore *espace probabilisé*.

De l'axiomatique de Kolmogorov, on déduit aisément les propriétés suivantes :

Proposition 1.1.12. *Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité. Alors on a*

1. $\mathbb{P}(\emptyset) = 0$;
2. *pour tout $A \in \mathcal{F}$, $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$;*
3. *pour tout $A, B \in \mathcal{F}$ tels que $A \subset B$, $\mathbb{P}(A) \leqslant \mathbb{P}(B)$;*
4. *pour tout $A, B \in \mathcal{F}$, $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$;*
5. *Si $(A_n)_{n\geqslant 0}$ est une suite d'évènements, alors $\mathbb{P}(\bigcup_{n\geqslant 0} A_n) \leqslant \sum_{n\geqslant 0} \mathbb{P}(A_n)$. Il n'y a égalité que si les évènements A_n sont deux à deux disjoints.*

Proposition 1.1.13 (Continuité monotone séquentielle). *Soient $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité et $(A_n)_{n\geqslant 0}$ une suite d'évènements.*

1. *Si la suite A_n est croissante, c'est-à-dire si $A_0 \subset A_1 \subset \ldots \subset A_n \subset \ldots$, alors*
$$\mathbb{P}\left(\bigcup_{n\geqslant 0} A_n\right) = \lim_{n\to\infty} \mathbb{P}(A_n) \ ;$$
2. *Si la suite A_n est décroissante, c'est-à-dire si $A_0 \supset A_1 \supset \ldots \supset A_n \supset \ldots$, alors*
$$\mathbb{P}\left(\bigcap_{n\geqslant 0} A_n\right) = \lim_{n\to\infty} \mathbb{P}(A_n).$$

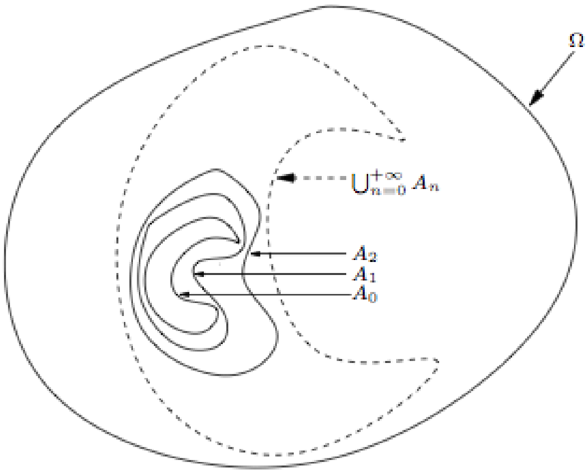

FIGURE 1.3 – Une famille croissante d'ensembles.

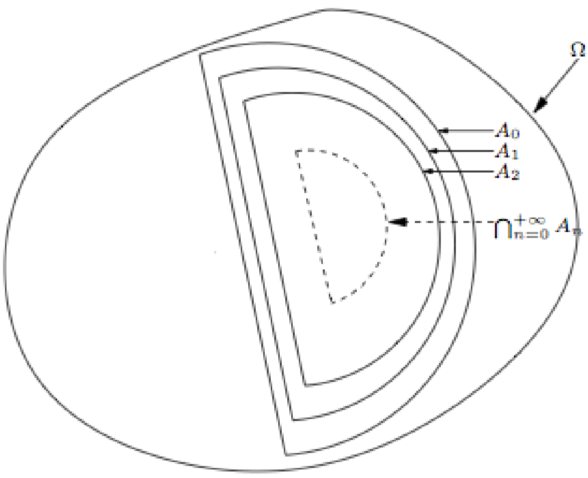

FIGURE 1.4 – Suite décroissante d'ensembles.

Quelques exemples de calculs de probabilités

Exercice 1 :
Dix athlètes participent à une course que chacun a la même chance d'emporter (pas d'ex aequo). Ils portent des dossards numérotés de 1 à 10. Quelle est la probabilité que l'un des coureurs portant les numéros 1, 2 ou 3 l'emporte ?

Correction : On note A_i l'évènement "le coureur au dossard i l'emporte" et A l'évènement "un des coureurs portant les numéros 1, 2 ou 3 l'emporte". L'évènement A s'écrit simplement comme l'union $A = A_1 \cup A_2 \cup A_3$ et les trois évènements sont disjoints donc

$$\mathbb{P}(A) = \mathbb{P}(A_1 \cup A_2 \cup A_3) = \mathbb{P}(A_1) + \mathbb{P}(A_2) + \mathbb{P}(A_3) = \frac{1}{10} + \frac{1}{10} + \frac{1}{10} = \frac{3}{10}.$$

Exercice 2 :
Un sac contient des billes noires et rouges, portant une marque ou non. La probabilité d'observer une bille rouge et marquée est de 2/10, une bille marquée de 3/10 et une bille noire de 7/10. Quelle est la probabilité d'observer une bille rouge ou marquée ?

Correction : On note R pour rouge, N pour noire, M pour marquée et M^c pour non marquée. On cherche la probabilité de l'évènement $R \cup M$. On a

$$\mathbb{P}(R \cup M) = \mathbb{P}(R) + \mathbb{P}(M) - \mathbb{P}(R \cap M) = \frac{3}{10} + \frac{3}{10} - \frac{2}{10} = \frac{4}{10}.$$

Exercice 3 :
Lors d'une loterie de Noël, 300 billets sont vendus aux parents d'élèves d'une école ; 4 billets sont gagnants. J'achète 10 billets, quelle est la probabilité pour que je gagne au moins un lot ?

Correction : L'univers des possibles est ici l'ensemble des combinaisons de 10 billets parmi les 300 ; il y en a $\binom{300}{10}$. Je ne 10 gagne rien si les 10 billets achetés se trouvent parmi les 296 billets perdants, ceci arrive avec la probabilité :

$$q = \frac{\binom{296}{10}}{\binom{300}{10}}.$$

La probabilité p cherchée est celle de l'évêment complémentaire :

$$p = 1 - q = 1 - \frac{\binom{296}{10}}{\binom{300}{10}} \approx 0.127.$$

Proposition 1.1.14. *Soit $(A_n)_{n\in\mathbb{N}}$ une suite d'évènements qui constituent une partition de l'ensemble Ω c'est-à-dire $\Omega = \bigsqcup_{n\in\mathbb{N}} A_n$. Alors pour tout $B \in \mathcal{F}$, on a*

$$\mathbb{P}(B) = \sum_n \mathbb{P}(B \cap A_n).$$

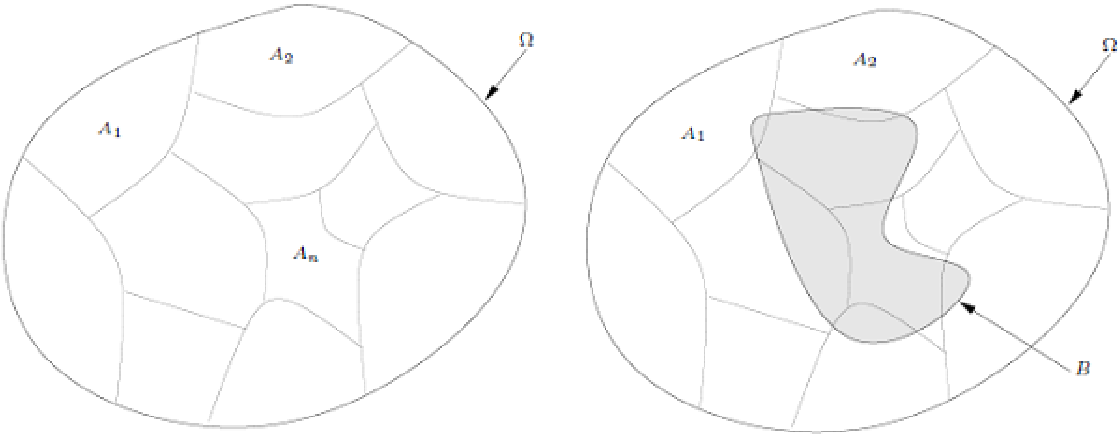

FIGURE 1.5 – Probabilité et partition.

Exemple 1.1.15. Un étudiant a les probabilités suivantes d'avoir la note i à un module, le module étant noté sur 10 *i.e.* $i = 1 \dots 10$. Quelle est la probabilté qu'il valide son module, c'est-à-dire qu'il obtienne une note supérieure ou égale à 5 ?

Note	0	1	2	3	4	5	6	7	8	9	10
Proba	1/11	0	0	1/11	1/11	2/11	2/11	2/11	1/11	1/11	0

On note B l'évèment "il valide son module" et A_i l'évèment "il obtient la note i. Les A_i forment une partition de l'ensemble des notes possibles et l'on a donc :

$$\mathbb{P}(B) = \sum_{i=0}^{10} \mathbb{P}(B \cap A_i) = 0 + \sum_{i=5}^{10} \mathbb{P}(A_i) = 8/11.$$

Remarque 1.1.16. Étant donné un espace probabilisable $(\Omega, \mathcal{F})$, le choix de la probabilité $\mathbb{P}$ n'est bien sûr pas unique. Ce choix doit se faire en accord avec l'expérience aléatoire que l'on souhaite modéliser. Par exemple, si on joue à pile ou face et que l'on précise que la pièce est équilibrée, on choisira naturellement $\mathbb{P}$ de sorte que

$$\mathbb{P}(\text{pile}) = \mathbb{P}(\text{face}) = 1/2.$$

En revanche, si l'on précise que la pièce est truquée, on préférera choisir $\mathbb{P}$ de sorte que $\mathbb{P}(\text{pile}) \neq \mathbb{P}(\text{face})$.

Remarque 1.1.17. Au risque de se répéter, insistons sur le fait que dans la modélisation d'une expérience aléatoire, l'espace de probabilité $(\Omega, \mathcal{F}, \mathbb{P})$ avec lequel on travaille n'est a priori pas unique. Il résulte d'un choix, et que ce choix doit pouvoir être justifié :

— le choix de l'ensemble des possibles Ω n'est pas unique, pensez au jeu de pile ou face avec $\Omega_1 = \{\text{pile}, \text{face}\}$ et $\Omega_2 = \{\text{pile}, \text{face}, \text{tranche}\}$;

— le choix de la tribu n'est pas unique, on peut choisir la tribu pleine, la tribu engendrée par un évènement, etc. ;

— le choix de la probabilité $\mathbb{P}$ n'est pas unique comme indiqué dans la remarque précédente.

L'exemple suivant est caractéristique. L'énoncé est n'est pas assez précis, de sorte que plusieurs choix de modélisations sont possibles et donc plusieurs réponses à la question posée sont envisageables. Il n'y a pas une réponse meilleure que l'autre : elles répondent à des questions différentes !

Exemple 1.1.18. On tire une corde au hasard dans un disque de rayon R. Quelle est la probabilité que la longueur ℓ de la corde soit supérieure à R ?

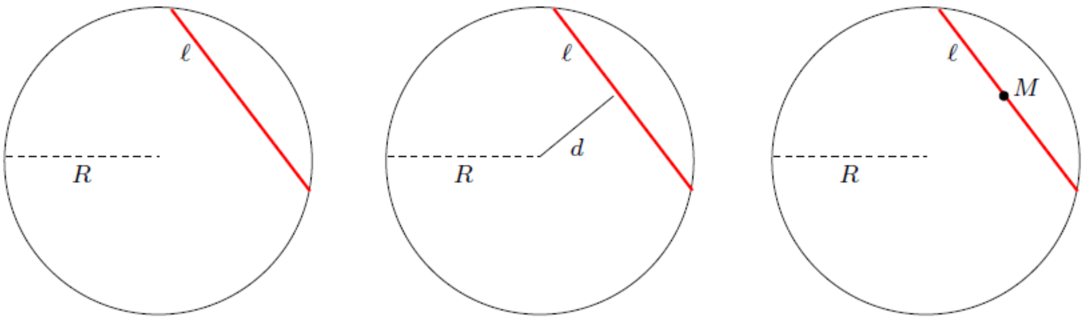

FIGURE 1.6 – Dans les trois exemples, on tire uniformément selon la longueur, la distance au centre, le milieu de la corde.

1. Dans ce premier cas, on choisit ici de modéliser le hasard en supposant que longueur de la corde est choisie uniformément parmi toutes les longueurs possibles. La longueur ℓ varie ici continûment dans $[0, 2R]$, de sorte que la probabilité cherchée vaut $1/2$;
2. On décide maintenant de modéliser le hasard en supposant que c'est la distance au centre de la corde qui est choisie uniformément au hasard. La longueur ℓ est donc déterminée par la distance d de la corde au centre du disque. Ici, d varie continûment dans $[0, R]$, et $\ell = \sqrt{R^2 - d^2} \geqslant R \Leftrightarrow d \leqslant \sqrt{3}/2R$, de sorte que la probabilité cherchée vaut $\sqrt{3}/2$;
3. Enfin, on décide de modéliser l'expérience en supposant que c'est la milieu M de la corde qui est choisi uniformément dans le disque. Dans ce cas, $\ell \geqslant R$ a lieu *ssi* M est dans le disque concentrique de rayon $\sqrt{3}/2$ de sorte que la probabilité cherchée vaut $3/4$.

1.2 Exemples d'espaces de probabilités

Afin de se familiariser avec les notions introduites ci-dessus, on donne maintenant des exemples d'expériences aléatoires et les espaces de probabilités correspondants.

1.2.1 Probabilité uniforme sur un ensemble fini

Loi uniforme pour un dé

Reprenons l'exemple du lancer de dé. On a vu que l'univers des possibles est est $\Omega = \{1, 2, 3, 4, 5, 6\}$ de cardinal 6. On munit Ω de la tribu des parties $\mathcal{F} = \mathcal{P}(\Omega)$. On vérifie alors que l'application

$$\mathbb{P} : \mathcal{F} \to [0, 1], \qquad A \mapsto \mathbb{P}(A) := \frac{\mathrm{Card}(A)}{6},$$

est bien une mesure de probabilité. Ainsi, dans cette modélisation, la probabilité d'obtenir un chiffre plus grand que 5 avec un lancer est

$$\mathbb{P}(\{5\} \cup \{6\}) = \mathbb{P}(\{5, 6\}) = \frac{2}{6} = \frac{1}{3}.$$

Loi uniforme sur un ensemble fini

Plus généralement, dès qu'on considère une expérience aléatoire où $\mathrm{Card}(\Omega)$, le nombre de résultats possibles, est fini, et que parmi ces résultats aucun n'est privilégié, on choisira naturellement la tribu des parties $\mathcal{F} = \mathcal{P}(\Omega)$ et la probabilité dite *uniforme* définie de la façon suivante :

$$\mathbb{P}(A) = \frac{\mathrm{Card}(A)}{\mathrm{Card}(\Omega)}, \quad \text{pour tout } A \in \mathcal{F}.$$

Par exemple, si on joue trois fois de suite à pile ou face (on note $\mathtt{p}, \mathtt{f}$ pour simplifier) avec une pièce équilibrée, l'ensemble des possibles est $\Omega = \{\mathtt{p}, \mathtt{f}\}^3$ qui a pour cardinal $2^3 = 8$. Notons A l'évènement le premier et le troisième lancer donnent pile, c'est-à-dire $A = \{(\mathtt{p}, \mathtt{p}, \mathtt{p}), (\mathtt{p}, \mathtt{f}, \mathtt{p})\}$. Alors

$$\mathbb{P}(A) = \frac{\mathrm{Card}(A)}{\mathrm{Card}(\Omega)} = \frac{2}{2^3} = \frac{1}{4}.$$

Remarque 1.2.1. La probabilité uniforme sur un ensemble fini est encore appelée *équiprobabilité*. On dit alors que tous les évènements élémentaires $\omega \in \Omega$ sont équiprobables.

1.2.2 Probabilité sur un ensemble au plus dénombrable

Loi générale sur un ensemble fini

On a vu que lorsqu'on a équiprobabilité sur un univers fini, la mesure de probabilité $\mathbb{P}$ est celle qui à tout évènement A associe le rapport de son cardinal au cardinal de Ω. En d'autres termes $\Omega = \{\omega_1, \dots, \omega_n\}$ et pour tout $i = 1, \dots, n$: $p_i = \mathbb{P}(\{\omega_i\}) = 1/n$. Supposer que l'on n'a pas équiprobabilité des évènements élémentaires ω_i revient à considérer une suite $(p_1, \dots, p_n)$ de nombres positifs et sommant à 1, mais dont tous les coefficients p_i ne sont pas égaux. On définit alors encore une mesure de probabilité sur $\mathcal{P}(\Omega)$ en considérant pour tout évènement $A \in \mathcal{P}(\Omega)$:

$$\mathbb{P}(A) = \sum_{i, \omega_i \in A} p_i$$

où la notation "$i, \omega_i \in A$" signifie que la somme est effectuée sur l'ensemble des indices i pour lesquels ω_i appartient à A.

Exemple 1.2.2. On lance 3 fois de suite une pièce équilibrée et on compte le nombre de fois où pile est apparu. On a donc $\Omega = \{0, 1, 2, 3\}$, mais il n'y a pas équiprobabilité puisque les probabilités élémentaires sont $(1/8, 3/8, 3/8, 1/8)$.

Exemple 1.2.3. On lance deux dés équilibrés et on note S la somme des deux lancers. L'ensemble des valeurs possibles pour S est $\Omega = \{2, 3, \dots, 11, 12\}$. Les probabilités pour les valeurs possibles de S sont alors :

k	2	3	4	5	6	7	8	9	10	11	12
$\mathbb{P}(S = k)$	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

Loi sur un ensemble dénombrable

Si on veut construire une probabilité $\mathbb{P}$ sur un ensemble infini dénombrable, typiquement sur $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$, on ne peut plus avoir équiprobabilité des évènements élémentaires $\{n\}$. Supposons en effet que pour tout $n \in \mathbb{N}$ on ait $\mathbb{P}(\{n\}) = p > 0$, alors l'additivité de $\mathbb{P}$ imposerait que :

$$\mathbb{P}(\mathbb{N}) = \sum_{n \geqslant 0} \mathbb{P}(\{n\}) = \sum_{n \geqslant 0} p = +\infty$$

ce qui est en contradiction avec la condition $\mathbb{P}(\mathbb{N}) = 1$. Une façon de construire une probabilité sur $(\mathbb{N}, \mathcal{P}(\mathbb{N}))$ est de généraliser le procédé que l'on vient de voir pour les ensembles finis : considérer une suite $(p_n)_{n \geqslant 0}$ de nombres positifs telle que la série $\sum_{n \geqslant 0} p_n$ soit convergente et de somme 1. Comme précédemment, on définit alors pour tout événement $A \in \mathcal{P}(\mathbb{N})$:

$$\mathbb{P}(A) = \sum_{n, n \in A} p_n.$$

Exemple 1.2.4. On lance une pièce équilibrée jusqu' à ce que pile apparaisse (toujours en excluant le cas improbable où pile n'apparaît jamais). On a donc $\Omega = \{1, 2, \dots\} = \mathbb{N}^*$. On a clairement $p_1 = \mathbb{P}(\{1\}) = 1/2$, $p_2 = \mathbb{P}(\{2\}) = 1/4$ et de façon générale $p_n = \mathbb{P}(\{n\}) = 1/2^n$. On reconnaît dans les p_n les termes d'une suite géométrique dont la somme vaut bien 1 :

$$\sum_{n \geqslant 1} 2^{-n} = 1.$$

1.2.3 Espace de probabilité continu

Donnons à présent quelques exemples de probabilités sur des espaces continus. Considérons ainsi un intervalle $\Omega =]a, b[\subset \mathbb{R}$.

On est ici dans un cas où l'ensemble Ω n'est pas dénombrable et où la tribu $\mathcal{P}(\Omega)$ est "trop grosse". Aussi, on considère une tribu $\mathcal{F}$ plus "petite", celle formée des intersections / unions dénombrables d'intervalles du type $[c, d[$ (on parle de tribu *borelienne*).

Supposons que l'on dispose d'une fonction positive f définie sur l'intervalle $[a, b]$ et telle que

$$\int_a^b f(x)dx = 1.$$

On peut alors définir une probabilité $\mathbb{P}$ sur $\mathcal{F}$ de la façon suivante : pour tout intervalle $A = [c, d[$ dans $[a, b]$

$$\mathbb{P}(A) = \int_A f(x)dx = \int_c^d f(x)dx.$$

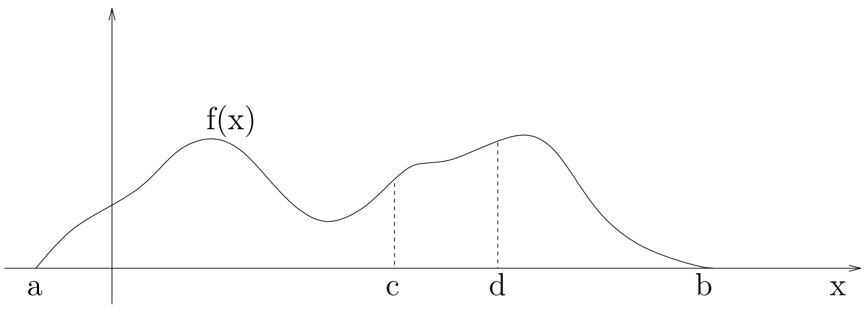

FIGURE 1.7 – Probabilité sur un intervalle via une densité.

Exemple 1.2.5 (Probabilité uniforme continue). Un bus est censé passer toutes les dix minutes à République pour se rendre à Beaulieu. Un passager arrive à l'arrêt de bus. On cherche à modéliser son temps d'attente T. A priori, on peut supposer que ce temps d'attente est dans l'intervalle $\Omega = [0, 10]$. On munit cet ensemble de la tribu borélienne. N'ayant pas d'information sur l'heure théorique de passage du bus et l'heure d'arrivée du passager, on peut supposer que le temps d'attente est uniforme, *i.e.* pour tout $0 < c < d < 10$:

$$\mathbb{P}(T \in [c, d[) = \frac{1}{10}|d - c| = \int_c^d f(x)dx$$

où la fonction f est constante égale à $1/10$ sur l'intervalle $[0, 10]$ de sorte que $\int_0^{10} f(x)dx = 1$.

Exemple 1.2.6. On cherche à modéliser le temps de demi-vie d'un atome radioactif. Ce temps T est aléatoire et l'expérience montre qu'il peut être très grand. On supposera que T est à valeurs dans $\Omega = \mathbb{R}^+ = [0, +\infty[$. Là encore, on suppose Ω muni de sa tribu borélienne. Des considérations physiques montrent que la probabilité ci-dessous décrit bien le temps de demi-vie T :

$$\mathbb{P}(T \in [c, d[) = \int_c^d e^{-x} dx, \quad \text{pour } 0 < c < d < +\infty.$$