Chapitre VIII

Probabilités

1 Espaces de probabilités finis

Définition 1 (Espace de probabilités fini)

Un espace de probabilités fini est un couple (Ω, P) où :

1. $\Omega = \{x_1, \ldots, x_n\}$ est un ensemble fini ;
2. P est une probabilité sur Ω, i.e. une application $P\colon \Omega \to [0,1]$ telle que :

$$P(x_1) + \ldots + P(x_n) = 1.$$

Définition 2 (Probabilité uniforme)

Soit $\Omega = \{x_1, \ldots, x_n\}$ un ensemble fini à $n \geq 1$ éléments. La probabilité uniforme sur Ω est l'application :

$$P\colon \Omega \to [0,1]\ ,\ x_i \mapsto \frac{1}{n}.$$

C'est l'unique probabilité sur Ω qui est une application constante, i.e. telle que tous les éléments de Ω ont la même image.

► **Exemple 1 :** L'ensemble $\Omega = [\![1,6]\!]$ muni de la probabilité uniforme

$$P\colon [\![1,6]\!] \to [0,1]\ ,\ i \mapsto \frac{1}{6}$$

est un espace de probabilités fini.

Définition 3 (Événement et probabilité d'un événement)

Soit $(\Omega = \{x_1, \ldots, x_n\}, P)$ un espace de probabilités fini.

1. Un événement est une partie de Ω.
2. Chacun des événements $\{x_i\}$, $i \in [\![1,n]\!]$, est appelé événement élémentaire.
3. Si $A = \{x_{i_1}, \ldots, x_{i_k}\}$ est un événement, alors la probabilité de A est le nombre $P(A)$ défini par :

$$P(A) = P(x_{i_1}) + \ldots P(x_{i_k}).$$

► **Exemple 2 :** On considère l'espace de probabités $([\![1,6]\!], P)$, où P est la loi uniforme sur $[\![1,6]\!]$. Alors l'événement $A = \{1,4\}$ a pour probabilité :

$$P(A) = P(1) + P(4) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3}.$$

On peut aussi calculer $P(A)$ à l'aide de la propriété suivante.

Proposition 1 (Probabilité d'un événement dans le cas uniforme)

Soit $(\Omega = \{x_1, \ldots, x_n\}, P)$ un espace de probabilités fini, où P est la probabilité uniforme. La probabilité d'un événement est donnée par la formule :

$$P(A) = \frac{\mathrm{Card}(A)}{\mathrm{Card}(\Omega)}.$$

Définition 4 (Expérience aléatoire à nombre d'issues fini et espace de probabilités fini associé)

Une expérience aléatoire est une expérience dont on ne peut pas prédire le résultat. Si le nombre d'issues possibles est fini, on associe à l'expérience aléatoire un espace de probabilités fini (Ω, P), où :

1. $\Omega = \{x_1, \ldots, x_n\}$ est l'ensemble de toutes les issues possibles ;
2. $P\colon \Omega \to [0,1]$ est l'application telle que pour tout $i \in [\![1, n]\!]$, $P(x_i)$ est la probabilité d'obtenir le résultat x_i.

Comme Ω contient tous les résultats possibles, on a bien :

$$P(x_1) + \ldots + P(x_n) = 1.$$

► **Exemple 3 :** On jette un dé non truqué à 6 faces numérotées de 1 à 6 et on lit le numéro obtenu. L'ensemble de toutes les issues possibles est $[\![1, 6]\!]$ et l'espace de probabilités associé est $([\![1, 6]\!], P)$, où P est la loi uniforme sur $[\![1, 6]\!]$ (le dé est non truqué et donc chacune des faces a la même probabilité d'apparaître).

► Remarque

Il existe de nombreuses expériences aléatoires dont le nombre d'issues n'est pas fini.

1. On se donne une pièce truquée telle que la probabilité que ≪ PILE ≫ apparaisse est p, avec $0 < p < 1$. On jette la pièce jusqu'à temps que l'on obtienne un ≪ PILE ≫ et on note n le nombre de lancers nécessaires pour obtenir un ≪ PILE ≫. L'ensemble de tous les résultats possibles est :
$$\Omega = \{1, 2, \ldots, n, \ldots\} \cup \{\infty\}.$$
Le résultat ∞ correspond au cas où ≪ PILE ≫ n'a été obtenu à aucun des lancers. Il est naturel de poser que pour tout $n \in \mathbb{N}^*$, la probabilité d'avoir n comme issue est :
$$P(n) = (1-p)^{n-1} p.$$
En revanche, il n'est pas évident, a priori, de donner une valeur pour $P(\infty)$.

   On aimerait pouvoir associer à cette expérience un espace de probabilités (infini) et pouvoir étendre les outils de la théorie des probabilités dans le cas fini à cette situation "infinie". Un problème se pose : on peut être amené à faire des sommes infinies...

2. On tire au hasard un nombre réel entre 0 et 1. L'ensemble de toutes les issues possibles de cette expérience aléatoire est alors $\Omega = [0, 1]$. Il est assez naturel de penser que :
   - la probabilité d'obtenir un nombre entre 0 et $\frac{1}{2}$ est $\frac{1}{2}$, nombre qui peut être vu comme le rapport entre la longueur de l'intervalle $\left[0, \frac{1}{2}\right]$ et la longueur de l'intervalle $[0, 1]$;
   - la probabilité d'obtenir un nombre entre $\frac{1}{3}$ et $\frac{2}{3}$ est $\frac{1}{3}$, nombre qui peut être vu comme le rapport entre la longueur de l'intervalle $\left[\frac{1}{3}, \frac{2}{3}\right]$ et la longueur de l'intervalle $[0, 1]$;
   - la probabilté d'obtenir un nombre entre a et b, avec $0 \leq a < b \leq 1$ est $b - a$, nombre qui peut être vu comme le rapport entre la longueur de l'intervalle $[a, b]$ et la longueur de l'intervalle $[0, 1]$.

   Deux problèmes se posent.
   (a) Quelles sont les parties de Ω pour lesquelles il est ≪ intéressant ≫ de définir une probabilité ? On aimerait que les intervalles réels inclus dans $[0, 1]$ soient des événements, mais poser que les seuls événements sont ces intervalles serait trop restrictif pour pouvoir étendre, plus tard, les formules de probabilités usuelles : e.g. la réunion de deux intervalles réels inclus dans $[0, 1]$ n'est, en général, pas un intervalle et le complémentaire d'un intervalle réel inclus dans $[0, 1]$ n'est, en général, pas un intervalle. L'ensemble des événements devrait contenir les intervalles réels inclus dans $[0, 1]$, mais pas seulement.
   (b) Comment définir la probabilité d'un événement ? Nous avons proposé une définition de probabilité pour certains intervalles réels inclus dans $[0, 1]$, mais il n'est pas évident, à priori, d'étendre cette définition.
   On remarque au passage que nous ne cherchons pas à définir la probabilité d'une issue (i.e. d'un nombre réel entre 0 et 1), mais la probabilité d'un événement (i.e. d'une certaine partie de $[0, 1]$).

► Remarque

Soit $(\Omega = \{x_1, \ldots, x_n\}, P)$ un espace de probabilités fini. Alors en associant à un évènement A sa probabilité $P(A)$, on définit une application :

$$\widehat{P}\colon \mathcal{P}(\Omega) \to [0, 1]\ ,\ A \mapsto P(A)$$

qui vérifie :

1. $\widehat{P}(\Omega) = 1$;
2. $\widehat{P}(A \cup B) = \widehat{P}(A) + \widehat{P}(B)$, si A et B sont deux évènements disjoints (incompatibles).

Proposition 2 (définition équivalente d'une probabilité dans le cas fini)

Soit $(\Omega = \{x_1, \ldots, x_n\}, P)$ un espace de probabilités fini. Soit une application

$$\widehat{P} \colon \mathcal{P}(\Omega) \to [0,1] \ , \ A \mapsto \widehat{P}(A)$$

qui vérifie :

1. $\widehat{P}(\Omega) = 1$;
2. $\widehat{P}(A \cup B) = \widehat{P}(A) + \widehat{P}(B)$, si A et B sont deux évènements disjoints (incompatibles).

Alors l'application P définie par :

$$P \colon \Omega \to [0,1] \ , \ x_i \mapsto \widehat{P}(\{x_i\})$$

est une probabilité sur Ω.

▶ **Remarque**

La proposition ci-dessus fournit un autre moyen de construire une probabilité. C'est cette approche qui sera la plus à même d'être généralisée.

2 Introduction aux séries

Définition 5 (Série à termes positifs)

Soit $(u_n)_{n \geq n_0}$ une suite à termes positifs. Pour tout $n \geq n_0$, on pose :

$$S_n = \sum_{k=n_0}^{n} u_k.$$

Comme pour tout $n \geq n_0$, $S_{n+1} - S_n = u_{n+1} \geq 0$, la suite $(S_n)_{n \geq n_0}$ est croissante.

1. Si $(S_n)_{n \geq n_0}$ est bornée, alors la suite $(S_n)_{n \geq n_0}$ converge vers une limite finie $l \in \mathbb{R}$. Dans ce cas, on dit que la série de terme général u_n converge et on note
$$\sum_{n=n_0}^{+\infty} u_n$$
la limite l de la suite $(S_n)_{n \geq n_0}$.
2. Si $(S_n)_{n \geq n_0}$ n'est pas bornée, alors la suite $(S_n)_{n \geq n_0}$ tend vers $+\infty$ et on dit que la série de terme général diverge.

▶ **Remarque**

Soit $(u_n)_{n \geq n_0}$ une suite à termes positifs. La série de terme général u_n peut-être identifiée à la suite

$$\left(S_n = \sum_{k=n_0}^{n} u_k \right)_{n \geq n_0}$$

des sommes des premiers termes consécutifs de la suite $(u_n)_{n \geq n_0}$. En particulier, étudier la convergence ou la divergence de la série de terme général de terme général u_n revient à étudier la convergence ou la divergence de la suite $(S_n)_{n \geq n_0}$.

▶ **Exemple 4**

La série de terme général $\dfrac{1}{n(n+1)}$ converge. En effet, pour tout $n \in \mathbb{N}^*$, on a :

$$\frac{1}{n(n+1)} = \frac{1}{n} - \frac{1}{n+1}$$

et par suite, pour tout $n \in \mathbb{N}^*$:

$$\begin{aligned} S_n &= \sum_{k=1}^{n} \frac{1}{k(k+1)} \\ &= \sum_{k=1}^{n} \frac{1}{k} - \frac{1}{k+1} \\ &= 1 - \frac{1}{2} + \frac{1}{2} - \frac{1}{3} + \frac{1}{3} - \frac{1}{4} + \ldots + \frac{1}{n-1} - \frac{1}{n} + \frac{1}{n} - \frac{1}{n+1} \\ &= 1 - \frac{1}{n+1} \qquad \text{(somme télescopique).} \end{aligned}$$

On en déduit que la suite $(S_n)_{n\geq 1}$ converge vers 1. Par suite la série de terme général $\dfrac{1}{n(n+1)}$ converge et on a :

$$\sum_{n=1}^{+\infty} \frac{1}{n(n+1)} = 1.$$

▶ Exemple 5

La série de terme général n diverge. En effet, pour tout $n \in \mathbb{N}$, on a :

$$S_n = \sum_{k=0}^{n} k = 0 + 1 + 2 + 3 + \ldots + n = \frac{n(n+1)}{2}.$$

On en déduit que $\lim\limits_{n\to+\infty} S_n = +\infty$. Par suite la série de terme général n diverge.

Théorème 1 (Séries géométriques)

Soit $q \in \mathbb{R}^+$ et soit $(q^n)_{n\geq 0}$ la suite géométrique de raison q et de premier terme 1.

1. Si $q \in [0,1[$, alors la série de terme général q^n converge et on a :

$$\sum_{n=0}^{+\infty} q^n = \frac{1}{1-q}.$$

2. Si $q \in [1,+\infty[$, alors la série de terme général q^n diverge.

Preuve

- Si $q = 1$, alors on a : $S_n = \sum\limits_{k=0}^{n} q^k = \sum\limits_{k=0}^{n} 1 = n+1$. On en déduit que $\lim\limits_{n\to+\infty} S_n = +\infty$. Par suite, la série de terme général q^n diverge.
- Si q est différent de 1, alors on a (cf. cours sur les suites géométriques) :

$$(*) \qquad \forall\, n \in \mathbb{N} \quad S_n = \sum_{k=0}^{n} q^k = \frac{1-q^{n+1}}{1-q}.$$

- Si $q \in [0,1[$, on a $\lim\limits_{n\to+\infty} q^n = 0$ (cf. cours sur les suites géométriques) et donc $\lim\limits_{n\to+\infty} q^{n+1} = 0$. On en déduit (cf. $(*)$) que la suite $(S_n)_{n\geq 0}$ converge vers $\dfrac{1}{1-q}$. Par suite, la série de terme général q^n converge et $\sum\limits_{n=0}^{+\infty} q^n = \dfrac{1}{1-q}$.
- Si $q \in]1,+\infty[$, alors $\lim\limits_{n\to+\infty} q^n = +\infty$ (cf. cours sur les suites géométriques) et donc $\lim\limits_{n\to+\infty} q^{n+1} = +\infty$. On en déduit (cf. $(*)$) que $\lim\limits_{n\to+\infty} S_n = +\infty$ et donc que la série de terme général q^n diverge.

► **Exemple 6**

La série de terme général $\left(\frac{1}{2}\right)^n$ converge et on a :

$$\sum_{n=0}^{+\infty} \left(\frac{1}{2}\right)^n = \frac{1}{1-\frac{1}{2}} = 2.$$

Théorème 2 (Convergence et « linéarité »)

Soit $(u_n)_{n\geq n_0}$ et $(v_n)_{n\geq n_0}$ deux suites à termes positifs telle que les séries de termes généraux u_n et v_n convergent. Soient $\lambda, \mu \in \mathbb{R}^+$. Alors la série de terme général $\lambda u_n + \mu v_n$ converge et on a :

$$\sum_{n=n_0}^{+\infty} \lambda u_n + \mu v_n = \lambda \left(\sum_{n=n_0}^{+\infty} u_n\right) + \mu \left(\sum_{n=n_0}^{+\infty}\right) v_n.$$

► **Exemple 7**

Soit $q \in [0,1[$. On va démontrer que la série de terme général $n\,q^n$ converge et que :

$$\sum_{n=1}^{+\infty} n\,q^n = \frac{q}{(1-q)^2}.$$

1. Soit $n \in \mathbb{N}$. On définit la fonction f_n par :

$$f_n \colon \mathbb{R}\setminus\{1\} \to \mathbb{R}\ ,\ x \mapsto \sum_{k=0}^{n} x^k.$$

Cette fonction est dérivable sur $\mathbb{R}\setminus\{1\}$ et on a :

$$(*) \qquad \forall\, x \in \mathbb{R}\setminus\{1\} \quad f_n'(x) = \sum_{k=1}^{n} k\,x^{k-1}.$$

Or, d'après le cours sur les suites géométriques, on a :

$$\forall\, x \in \mathbb{R}\setminus\{1\} \quad f_n(x) = \frac{1-x^{n+1}}{1-x}.$$

À l'aide de cette observation, on obtient une autre expression pour la dérivée de f_n :

$$(**) \qquad \forall\, x \in \mathbb{R}\setminus\{1\} \quad f_n'(x) = \frac{n\,x^{n+1}-(n+1)x^n+1}{(1-x)^2}.$$

En comparant $(*)$ et $(**)$, il vient :

$$(***) \qquad \forall\, x \in \mathbb{R}\setminus\{1\} \quad \sum_{k=1}^{n} k\,x^{k-1} = \frac{n\,x^{n+1}-(n+1)x^n+1}{(1-x)^2}.$$

2. En appliquant $(***)$ avec $x = q$, on a donc :

$$(****) \qquad \sum_{k=1}^{n} k\,q^{k-1} = \frac{n\,q^{n+1}-(n+1)q^n+1}{(1-q)^2}.$$

D'après le cours sur les suites géométriques, on a $\lim_{n\to+\infty} q^n = 0$ et donc $\lim_{n\to+\infty} q^{n+1} = 0$. Par croissances comparées, on a donc : $\lim_{n\to+\infty} n\,q^{n+1} = 0$ et $\lim_{n\to+\infty} (n+1)q^n = 0$. On en déduit, grâce à $(****)$, que la suite $\left(\sum_{k=1}^{n} k\,q^{k-1}\right)_{n\geq 1}$ converge vers $\frac{1}{(1-q)^2}$. La série de terme général $n\,q^{n-1}$ est donc convergente et :

$$\sum_{n=1}^{+\infty} n\,q^{n-1} = \frac{1}{(1-q)^2}.$$

3. Comme pour tout $n \in \mathbb{N}^*$, $n\,q^n = q \times (n\,q^{n-1})$, on déduit de ce qui précède, en appliquant le théorème 2, que la série de terme général $n\,q^n$ converge et que :

$$\sum_{n=1}^{+\infty} n\,q^n = q \times \frac{1}{(1-q)^2} = \frac{q}{(1-q)^2}.$$

Théorème 3 (Convergence d'une série associée à une sous-suite)

Soit $(u_n)_{n \geq n_0}$ une suite à termes positifs telle que la série $\sum u_n$ converge. Soit $i_0 < i_1 < \ldots < i_n < \ldots$ une suite strictement croissante d'entiers tous supérieurs ou égaux à n_0. Alors la série de terme général u_{i_n} converge et :

$$\sum_{n=0}^{+\infty} u_{i_n} \leq \sum_{n=n_0}^{+\infty} u_k.$$

3 Espaces de probabilités dénombrables

Définition 6 (Ensemble dénombrable)

Un ensemble Ω est dit dénombrable si et seulement si il existe une bijection

$$f\colon \mathbb{N} \to \Omega\ ,\ i \mapsto x_i$$

i.e. si l'on peut faire la liste

$$x_0, x_1, x_2, \ldots, x_n, \ldots$$

des éléments de Ω en les numérotant à l'aide des entiers naturels.

▶ **Remarque**

Un espace dénombrable est donc infini, mais il existe des ensembles infinis non dénombrables.

▶ **Exemples 8**

Les ensembles $\mathbb{N}$, $\mathbb{N}^*$, $\mathbb{Z}$, $\mathbb{Q}$ et $\mathbb{N} \times \mathbb{N}$ sont dénombrables. En revanche, les ensembles $\mathbb{R}$, $\mathbb{C}$ et $\mathcal{P}(\mathbb{N})$ (ensemble des parties de $\mathbb{N}$) ne sont pas dénombrables.

Définition 7 (Espace de probabilités dénombrable)

Un espace de propabilités dénombrable est un couple (Ω, P) tel que :

1. $\Omega = \{x_0, x_1, \ldots, x_n, \ldots\}$ est un ensemble dénombrable ;
2. $P\colon \Omega \to [0,1]$ est une application telle que la série de terme général $P(x_n)$ converge et vérifie :

$$\sum_{n=0}^{+\infty} P(x_n) = 1.$$

Définition 8 (Événement et probabilité d'un évènement)

Soit $(\Omega = \{x_0, x_1, \ldots, x_n, \ldots\}, P)$ un espace de probabilités dénombrable.

1. Un événement est une partie de Ω.
2. Si $A = \{x_{i_1}, \ldots, x_{i_n}\}$ est un événement fini, alors la probabilité de A est le nombre $P(A)$ défini par :
$$P(A) = P(x_{i_1}) + \ldots P(x_{i_n}).$$
3. Si A est un évènement infini, alors il existe une unique suite strictement croissante d'entiers naturels $i_0 < i_1 < \ldots < i_n < \ldots$ telle que :
$$A = \{x_{i_0}, x_{i_1}, \ldots, x_{i_n}, \ldots\}.$$
La probabilité de A est définie par :
$$P(A) = \sum_{n=0}^{+\infty} P(x_{i_n}).$$
D'après le théorème 3, cette somme est bien définie et on sait que :
$$P(A) \leq \sum_{n=0}^{+\infty} P(x_n) = 1.$$

▶ **Exemple 9 :** L'ensemble $\Omega = \mathbb{N}$ muni de l'application
$$P\colon \mathbb{N} \to [0,1]\ ,\ n \mapsto \left(\frac{1}{2}\right)^{n+1}$$
est un espace de probabilités dénombrable. En effet, on a vu (cf. exemple 6) que la série de terme général $\left(\frac{1}{2}\right)^{n}$ converge et que $\displaystyle\sum_{n=0}^{+\infty} \left(\frac{1}{2}\right)^{n} = 2$. On en déduit, grâce au théorème 2, que la série de terme général $\left(\frac{1}{2}\right)^{n+1} = \frac{1}{2} \times \left(\frac{1}{2}\right)^{n}$ converge et que : $\displaystyle\sum_{n=0}^{+\infty} \left(\frac{1}{2}\right)^{n+1} = \frac{1}{2} \times 2 = 1$.

Soit A l'événement « l'entier est pair », i.e. $A = \{0, 2, 4, 6, 8, 10, 12, \ldots\}$. On a donc :
$$\begin{aligned}
P(A) &= P(0) + P(2) + P(4) + P(6) + P(8) + P(10) + P(12) + \ldots \\
&= \sum_{n=0}^{+\infty} P(2n) \\
&= \sum_{n=0}^{+\infty} \left(\frac{1}{2}\right)^{2n+1} \\
&= \sum_{n=0}^{+\infty} \frac{1}{2} \times \left(\left(\frac{1}{2}\right)^{2}\right)^{n} \\
&= \frac{1}{2} \times \left(\sum_{n=0}^{+\infty} \left(\frac{1}{4}\right)^{n}\right) \qquad \text{(cf. théorème 2)} \\
&= \frac{1}{2} \times \frac{1}{1 - \frac{1}{4}} \qquad \text{(cf. théorème 1)} \\
&= \frac{2}{3}.
\end{aligned}$$

▶ **Remarque**

Soit $(\Omega = \{x_0, x_1, \ldots, x_n, \ldots\}, P)$ un espace de probabilités dénombrable. Alors en associant à un évènement A sa probabilité $P(A)$, on définit une application :

$$\widehat{P} \colon \mathcal{P}(\Omega) \to [0,1] \ , \ A \mapsto P(A)$$

qui vérifie :

1. $\widehat{P}(\Omega) = 1$;
2. si $(A_n)_{n\in\mathbb{N}}$ une suite d'événements deux à deux disjoints (incompatibles), alors la série de termes général $\widehat{P}(A_n)$ converge et on a :

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

Nous admettrons la propriété 2 dont la démonstration n'est pas immédiate et requiert des résultats supplémentaires sur les séries à termes généraux positifs, résultats que nous ne discutons pas ici.

Proposition 3 (définition équivalente d'une probabilité dans le cas dénombrable)

Soit $(\Omega = \{x_0, x_1, \ldots, x_n, \ldots\}, P)$ un espace de probabilités dénombrable. Soit une application

$$\widehat{P} \colon \mathcal{P}(\Omega) \to [0,1] \ , \ A \mapsto \widehat{P}(A)$$

qui vérifie :

1. $\widehat{P}(\Omega) = 1$;
2. Si $(A_n)_{n\in\mathbb{N}}$ une suite d'événements deux à deux disjoints, alors la série de terme général $\widehat{P}(A_n)$ converge et on a :

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

Alors l'application P définie par :

$$P \colon \Omega \to [0,1] \ , \ x_i \mapsto \widehat{P}(\{x_i\})$$

est une probabilité sur Ω.

▶ **Remarque**

Nous admettrons cette propriété. À nouveau, il nous faudrait des outils supplémentaires sur les séries à termes généraux positifs pour pouvoir en donner une preuve. C'est cette dernière définition de probabilité qui sera prise comme définition de probabilité dans le cas général.

4 Espaces de probabilités généraux

Soit Ω un ensemble quelconque, appelé univers.

Contrairement aux deux cas précédents (Ω fini, Ω dénombrable), un événement ne sera pas nécessairement une sous-partie quelconque de Ω dans le cas général. Sans parler des problèmes théoriques sous-jacents, on ne sera pas forcément intéressé, dans le cas général, par connaître la probabilité de toutes les parties de Ω.

Pour indiquer quelles sont les parties de Ω ≪ intéressantes ≫, on va ajouter à Ω un sous-ensemble $\mathcal{T}$ de $\mathcal{P}(\Omega)$, appelé ensemble des événements ou tribu, possédant ≪ certaines propriétés ≫.

Définition 9 (tribu)

Une tribu $\mathcal{T}$ sur Ω est un ensemble de parties de Ω tel que :

1. $\Omega \in \mathcal{T}$;
2. $\mathcal{T}$ est stable par passage au complémentaire, i.e. :
$$\forall\, A \in \mathcal{T} \quad \overline{A} \in \mathcal{T}.$$
3. $\mathcal{T}$ est stable par réunion dénombrable, i.e. pour tout suite $(A_n)_{n\in\mathbb{N}}$ d'éléments de $\mathcal{T}$, on a :
$$\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{T}.$$

Définition 10 (espace probabilisable, événement d'un espace probabilisable)

1. Un espace probabilisable est un couple $(\Omega, \mathcal{T})$ où Ω est un ensemble et $\mathcal{T}$ est une tribu sur Ω.
2. Un événement d'un espace probabilisable $(\Omega, \mathcal{T})$ est une sous-partie de Ω qui appartient à la tribu $\mathcal{T}$.

► **Remarque**

Dans le cas où Ω est fini ou dénombrable, on choisira toujours $\mathcal{T} = \mathcal{P}(\Omega)$. Toute partie de Ω sera donc, dans ces cas, un événement, ce qui est cohérent avec ce qui précède. Dans le cas où Ω n'est ni fini, ni dénombrable, la tribu $\mathcal{T}$ sera, en pratique, toujours différente de $\mathcal{P}(\Omega)$.

Définition 11 (espace de probabilités)

Un espace de probabilités est un triplet $(\Omega, \mathcal{T}, P)$ où :

1. $(\Omega, \mathcal{T})$ est un espace probabilisable ;
2. $P\colon \mathcal{T} \to [0,1]$ est une application, appelée probabilité, telle que :
   (a) $P(\Omega) = 1$;
   (b) Pour toute suite d'événements $(A_n)_{n\in\mathbb{N}}$ deux à deux disjoints (incompatibles), i.e.
   $$\forall\, n \in \mathbb{N} \quad \forall\, m \in \mathbb{N} \quad n \neq m \Longrightarrow A_n \cap A_m = \emptyset,$$
   la série de terme général $P(A_n)$ converge et :
   $$\sum_{n=0}^{+\infty} P(A_n) = P\left(\bigcup_{n\in\mathbb{N}} A_n\right).$$
   On remarque que d'après la propriété 3 de la définition d'une tribu, $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{T}$. On peut donc considérer la probabilité de l'événement $\bigcup_{n\in\mathbb{N}} A_n$.

► **Exemple 10 :** Il existe une « plus petite tribu » $\mathcal{B}$ sur $\Omega = [0,1]$ contenant tous les intervalles réels inclus dans $[0,1]$. Cette tribu est appelée tribu borélienne (borélien fait référence au mathématicien Émile Borel) et n'est pas égale à $\mathcal{P}([0,1])$, i.e. certaines parties de $[0,1]$ n'appartiennent pas à la tribu borélienne. Sur l'espace probabilisable $([0,1], \mathcal{B})$, il existe une unique probabilité P telle que pour tout $a, b \in \mathbb{R}$, $0 \leq a < b \leq 1$, on a :

$$P(]a,b[) = P([a,b[) = P([a,b]) = P(]a,b]) = b - a.$$

On dit que P est la probabilité uniforme sur $[0,1]$.
Les résultats cités ci-dessous sont admis. Leurs démonstrations dépassent le cadre du programme.

Théorème 4 (propriétés élémentaires des espaces de probabilités)

Soit $(\Omega, \mathcal{T}, P)$ un espace de probabilités. On a les propriétés suivantes.

1. $P(\emptyset) = 0$.
2. Si A et B sont deux événements disjoints (incompatibles), alors $P(A \cup B) = P(A) + P(B)$.
3. Si $A_1, \ldots, A_n$ sont des événements deux à deux disjoints (incompatibles), i.e. :
$$\forall\, i \in [\![1, n]\!] \qquad \forall\, j \in [\![1, n]\!] \qquad i \neq j \Longrightarrow A_i \cap A_j = \emptyset,$$
alors $P(A_1 \cup \ldots \cup A_n) = P(A_1) + \ldots + P(A_n)$.
4. Soit A un événement. Alors $P(\overline{A}) = 1 - P(A)$.
5. Soient A et B deux événements tels que $A \subset B$. Alors $P(A) \leq P(B)$.
6. Soient A et B deux événements. Alors $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Preuve

1. Pour tout $n \in \mathbb{N}$, on pose $A_n = \emptyset$. Alors la suite $(A_n)_{n\in\mathbb{N}}$ est une suite d'événements deux à deux disjoints. La convergence de la série de terme général $P(A_n) = P(\emptyset)$ force l'égalité $P(\emptyset) = 0$.

2. Soient A et B deux événements incompatibles. On pose $A_0 = A$, $A_1 = B$ et $A_n = \emptyset$ pour tout entier $n \geq 2$. La suite $(A_n)_{n\in\mathbb{N}}$ est alors une suite d'événements deux à deux disjoints. Alors la série de terme général $P(A_n)$ converge et :
$$(*) \qquad \sum_{n=0}^{+\infty} P(A_n) = P\left(\bigcup_{n\in\mathbb{N}} A_n\right).$$
Or
$$(**) \qquad \begin{aligned} \sum_{n=0}^{+\infty} P(A_n) &= P(A_0) + P(A_1) + P(A_2) + P(A_3) + P(A_4) + P(A_5) + \ldots \\ &= P(A) + P(B) + P(\emptyset) + P(\emptyset) + P(\emptyset) + P(\emptyset) + \ldots \\ &= P(A) + P(B) \qquad (P(\emptyset) = 0) \end{aligned}$$
De plus on a :
$$(***) \qquad \begin{aligned} \bigcup_{n\in\mathbb{N}} A_n &= A_0 \cup A_1 \cup A_2 \cup A_3 \cup A_4 \cup A_5 \cup \ldots \\ &= A \cup B \cup \emptyset \cup \emptyset \cup \emptyset \cup \emptyset \cup \ldots \\ &= A \cup B \end{aligned}$$
De $(*)$, $(**)$ et $(***)$, on déduit que $P(A \cup B) = P(A) + P(B)$.

3. On peut prouver la propriété 3. en généralisant la démonstration donnée pour la propriété 2. Les détails sont laissés en exercice.

4. Soit A un événement. Alors A et $\overline{A}$ sont deux événements disjoints. On a donc, d'après 2. :
$$(*) \qquad P(A \cup \overline{A}) = P(A) + P(\overline{A}).$$
Or $A \cup \overline{A} = \Omega$ et $P(\Omega) = 1$. De $(*)$, on déduit donc $P(A) + P(\overline{A}) = 1$ et par suite $P(\overline{A}) = 1 - P(A)$.

5. Soient A et B deux événements tels que $A \subset B$. Alors $B = A \cup (B \cap \overline{A})$ (faire un diagramme de Venn). Comme les deux événements A et $B \cap \overline{A}$ sont disjoints, on a d'après 2. :
$$(*) \qquad P(B) = P(A) + P(B \cap \overline{A}).$$
Or $P(B \cap \overline{A}) \geq 0$ (car l'ensemble d'arrivée de P est $[0, 1]$). On a donc $P(A) \leq P(B)$.

6. Soient A et B deux événements. Alors $A \cup B = (A \cap B) \cup (\overline{A} \cap B) \cup (A \cap \overline{B})$ (faire un diagramme de Venn). Les trois événements $A \cap B$, $\overline{A} \cap B$ et $A \cap \overline{B}$ étant deux à deux disjoints, on a, d'après 3. :
$$(*) \qquad P(A \cup B) = P((A \cap B) \cup (\overline{A} \cap B) \cup (A \cap \overline{B})) = P(A \cap B) + P(\overline{A} \cap B) + P(A \cap \overline{B}).$$
Les événements $A \cap B$ et $A \cap \overline{B}$ sont disjoints et leur réunion est A (faire un diagramme de Venn). On a donc, d'après 2., $P(A) = P(A \cap B) + P(A \cap \overline{B})$. On en déduit :
$$(**) \qquad P(A \cap \overline{B}) = P(A) - P(A \cap B).$$
De même, on montre :
$$(***) \qquad P(\overline{A} \cap B) = P(B) - P(A \cap B).$$
De $(*)$, $(**)$ et $(***)$, on déduit :
$$P(A \cup B) = P(A) + P(B) - P(A \cap B).$$

► **Remarque**

On considère à nouveau l'espace de probabilités $(\Omega = [0,1], \mathcal{B}, P)$ introduit dans l'exemple 10. On remarque, dans cet exemple, deux phénomènes que l'on n'a pas rencontrés dans le cas fini.

1. Si $a \in [0,1]$, alors $P(\{a\}) = P([a,a]) = a - a = 0$. Il existe donc des événements différents de l'ensemble vide et qui ont une probabilité nulle.
2. On a $P(]0,1[) = 1 - 0 = 1$, alors que $]0,1[$ n'est pas égal à $\Omega = [0,1]$.

Ceci conduit à introduire deux définitions nouvelles.

Définition 12 (événement négligeable et événement presque sûr)

Soit $(\Omega, \mathcal{T}, P)$ un espace de probabilités et soit A un événement.

1. A est un événement négligeable si $P(A) = 0$.
2. A est un événement presque sûr si $P(A) = 1$.

► **Remarque**

Si A et B sont deux événements d'un espace de probabilités fini, on sait que pour calculer $P(A \cap B)$, on peut utiliser l'indépendance de A et B (quand cette propriété est satisfaite, ce qui n'est pas toujours le cas) ou plus généralement la notion de probabilité conditionnelle. Ces deux notions s'étendent naturellement au cas d'un espace de probabilités général.

Définition 13 (indépendance)

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

1. Deux événements A et B sont dits indépendants si $P(A \cap B) = P(A) \times P(B)$.
2. Des événements $A_1, A_2, \ldots, A_n$ sont dits deux à deux indépendants si :
$$\forall\, i \in [\![1,n]\!] \qquad \forall\, j \in [\![1,n]\!] \qquad i \neq j \Longrightarrow P(A_i \cap A_j) = P(A_i) \times P(A_j).$$
3. Des événements $A_1, A_2, \ldots, A_n$ sont dits mutuellement indépendants si pour tout sous-ensemble I de $[\![1,n]\!]$ on a :
$$P\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} P(A_i).$$

► **Exemple 11 :** Soit $(\Omega, \mathcal{T}, P)$ un espace de probabilités. On explicite la condition d'indépendance mutuelle dans le cas de 4 événements A_1, A_2, A_3 et A_4. Les événements A_1, A_2, A_3 et A_4 sont mutuellement indépendants si les propriétés suivantes sont satisfaites :

$$P(A_1 \cap A_2) = P(A_1) \times P(A_2) \qquad P(A_1 \cap A_3) = P(A_1) \times P(A_3) \qquad P(A_1 \cap A_4) = P(A_1) \times P(A_4)$$
$$P(A_2 \cap A_3) = P(A_2) \times P(A_3) \qquad P(A_2 \cap A_4) = P(A_2) \times P(A_4)$$
$$P(A_3 \cap A_4) = P(A_3) \times P(A_4)$$

$$P(A_1 \cap A_2 \cap A_3) = P(A_1) \times P(A_2) \times P(A_3) \qquad P(A_1 \cap A_2 \cap A_4) = P(A_1) \times P(A_2) \times P(A_4)$$
$$P(A_1 \cap A_3 \cap A_4) = P(A_1) \times P(A_3) \times P(A_4) \qquad P(A_2 \cap A_3 \cap A_4) = P(A_2) \times P(A_3) \times P(A_4)$$

$$P(A_1 \cap A_2 \cap A_3 \cap A_4) = P(A_1) \times P(A_2) \times P(A_3) \times P(A_4)$$

► **Remarque**

Soit $(\Omega, \mathcal{T}, P)$ un espace de probabilités. Si les événements $A_1, A_2, \ldots, A_n$ sont mutuellement indépendants, alors ils sont deux à deux indépendants, mais la réciproque est ≪ en général ≫ fausse.

Définition 14 (probabilité conditionnelle)

Soit $(\Omega, \mathcal{T}, P)$ un espace de probabilités. Soient A et B deux événements. Si A *n'est pas négligeable* (i.e. $P(A) \neq 0$), alors la probabilité de B sachant A, notée $P(B \,/\, A)$ ou $P_A(B)$, est définie par :

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

La condition $P(A) \neq 0$ est imposée pour pouvoir diviser par $P(A)$.

5 Les théorèmes fondamentaux

Les théorèmes fondamentaux :

- formule du crible ou de Poincaré,
- formule des probabilités composées,
- formule des probabilités totales,
- théorème de Bayes,

vus dans le cas d'un espace de probabilités fini, s'étendent au cas d'un espace de probabilités général. Les preuves ne sont pas rappelées, car elles sont analogues voire quasi identiques à celles données dans le cas fini. La seule nouveauté réside dans une extension de la formule des probabilités totales, dont une démonstration est proposée.

Théorème 5 (formule de Poincaré)

Soit $(\Omega, \mathcal{T}, P)$ un espace de probabilités. Soient $A_1, A_2, \ldots, A_n$ des événements. Alors on a :

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{k=1}^{n} \left((-1)^{k-1} \sum_{1 \leq i_1 < i_2 < \ldots < i_k \leq n} P(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}) \right).$$

► **Exemple 12 :** On explicite la formule de Poincaré pour une réunion de 4 événements A_1, A_2, A_3 et A_4 d'un espace de probabilités $(\Omega, \mathcal{T}, P)$.

$$\begin{aligned} P(A_1 \cup A_2 \cup A_3 \cup A_4) \quad = \quad & P(A_1) + P(A_2) + P(A_3) + P(A_4) \\ - \; & (P(A_1 \cap A_2) + P(A_1 \cap A_3) + P(A_1 \cap A_4) + P(A_2 \cap A_3) + P(A_2 \cap A_4) + P(A_3 \cap A_4)) \\ + \; & P(A_1 \cap A_2 \cap A_3) + P(A_1 \cap A_2 \cap A_4) + P(A_1 \cap A_3 \cap A_4) + P(A_2 \cap A_3 \cap A_4) \\ - \; & P(A_1 \cap A_2 \cap A_3 \cap A_4) \end{aligned}$$

On remarque la présentation du membre de droite en probabilités des événements $P(A_i)$, probabilités des intersections doubles, probabilités des intersections triples, et probabilité de l'intersection quadruple (correspondant aux différents termes d'indices k de la somme $\displaystyle\sum_{k=1}^{4}$ de la formule de Poincaré de l'encadré). On note également l'alternance des signes.

Théorème 6 (formule des probabilités composées)

Soit $(\Omega, \mathcal{T}, P)$ un espace de probabilités. Soient $A_1, A_2, \ldots, A_n$ des événements tels que $A_1 \cap A_2 \cap \ldots \cap A_{n-1}$ *n'est pas négligeable* (i.e. $P(A_1 \cap A_2 \cap \ldots \cap A_{n-1}) \neq 0$). Alors on a :

$$P(A_1 \cap A_2 \cap \ldots \cap A_n) = P(A_1) \times P(A_2 \,/\, A_1) \times P(A_3 \,/\, A_1 \cap A_2) \times \ldots \times P(A_n \,/\, A_1 \cap A_2 \cap \ldots \cap A_{n-1}).$$

► **Remarque**

Dans le théorème précédent, toutes les probabilités conditionnelles sont bien définies. En effet, comme $P(A_1 \cap A_2 \cap \ldots \cap A_{n-1}) \neq 0$, on a :

$P(A_1) \neq 0$ $\quad (A_1 \cap A_2 \cap \ldots \cap A_{n-1} \subset A_1$ et propriété 5 du théorème 4)
$P(A_1 \cap A_2) \neq 0$ $\quad (A_1 \cap A_2 \cap \ldots \cap A_{n-1} \subset A_1 \cap A_2$ et propriété 5 du théorème 4)
$\vdots$
$P(A_1 \cap A_2 \cap \ldots \cap A_{n-2}) \neq 0$ $\quad (A_1 \cap A_2 \cap \ldots \cap A_{n-1} \subset A_1 \cap A_2 \cap \ldots \cap A_{n-2}$ et propriété 5 du théorème 4)

▶ **Exemple 13 :** On explicite la formule des probabilités composées pour une intersection de 4 événements A_1, A_2, A_3 et A_4 d'un espace de probabilités $(\Omega, \mathcal{T}, P)$.

$$P(A_1 \cap A_2 \cap A_3 \cap A_4) = P(A_1) \times P(A_2 \,/\, A_1) \times P(A_3 \,/\, A_1 \cap A_2) \times P(A_4 \,/\, A_1 \cap A_2 \cap A_3)$$

▶ **Remarque**

Cette formule des probabilités composées peut être appliquée pour rédiger soigneusement ce que l'on ≪ voit ≫ sur un arbre de probabilités.

Définition 15 (partition finie et partition dénombrable d'un espace probabilisable)

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

1. Une partition finie de $(\Omega, \mathcal{T})$ est une famille $(A_1, \ldots, A_n)$ d'événements telle que :
   - les événements $A_1, \ldots, A_n$ sont deux à deux disjoints,
   - la réunion des événements A_1, …, A_n est égale à Ω, i.e. : $\displaystyle\bigcup_{i=1}^{n} A_i = \Omega$.
2. Une partition dénombrable de $(\Omega, \mathcal{T})$ est une suite $(A_n)_{n \in \mathbb{N}}$ d'événements telle que :
   - les événements A_1, …, A_n, … sont deux à deux disjoints,
   - la réunion des événements A_1, …, A_n, … est égale à Ω, i.e. : $\displaystyle\bigcup_{n \in \mathbb{N}} A_n = \Omega$.

Théorème 7 (formules des probabilités totales)

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

1. Soit A un événement qui n'est ni négligeable, ni presque sûr (i.e. $P(A) \neq 0$ et $P(A) \neq 1$). (On a ainsi $P(\overline{A}) = 1 - P(A) \neq 0$; $\overline{A}$ n'est donc pas négligeable.) Alors pour tout événement B, on a :

$$P(B) = P(B \,/\, A) \times P(A) + P(B \,/\, \overline{A}) \times P(\overline{A}).$$

2. Soit $(A_1, \ldots, A_n)$ une partition finie de $(\Omega, \mathcal{T})$ telle que :

$$\forall\, i \in [\![1, n]\!] \quad P(A_i) \neq 0.$$

   Alors pour tout événement B, on a :

$$P(B) = \sum_{i=1}^{n} P(B \,/\, A_i) \times P(A_i).$$

3. Soit $(A_n)_{n \in \mathbb{N}}$ une partition dénombrable de $(\Omega, \mathcal{T})$ telle que :

$$\forall\, n \in \mathbb{N} \quad P(A_n) \neq 0.$$

   Alors pour tout événement B, la série de terme général $P(B \,/\, A_n) \times P(A_n)$ converge et on a :

$$P(B) = \sum_{n=0}^{+\infty} P(B \,/\, A_n) \times P(A_n).$$

Preuve

La démonstration des propriétés 1. et 2. est analogue à celle connue dans le cas fini. Nous ne donnons de preuve que pour la troisième propriété.

Soit $(A_n)_{n\in\mathbb{N}}$ une partition finie de $(\Omega, \mathcal{T})$ telle que : $\forall\, n \in \mathbb{N} \quad P(A_n) \neq 0$. Pour tout $n \in \mathbb{N}$, soit B_n l'événement défini par :

$$B_n = B \cap A_n.$$

Alors $(B_n)_{n\in\mathbb{N}}$ est une suite d'événements deux à deux disjoints. En effet, si n et m sont deux entiers naturels distincts, alors :

$$B_n \cap B_m = (B \cap A_n) \cap (B \cap A_m) = B \cap (\underbrace{A_n \cap A_m}_{\emptyset}) = \emptyset.$$

La propriété $A_n \cap A_m = \emptyset$ vient du fait que les événements $(A_n)_{n\in\mathbb{N}}$ sont deux à deux disjoints, d'après la définition d'une partition dénombrable.

D'après la définition d'un espace de probabilités, la série de terme général $P(B_n)$ converge et on a :

$$P\left(\bigcup_{n\in\mathbb{N}} B_n\right) = \sum_{n=0}^{+\infty} P(B_n).$$

Comme pout tout $n \in \mathbb{N}$, $P(B_n) = P(B \cap A_n) = P(B \,/\, A_n) \times P(A_n)$ (on peut introduire la probabilité conditionnelle $P(B \,/\, A_n)$ car $P(A_n) \neq 0$), on en déduit que la série de terme général $P(B \,/\, A_n) \times P(A_n)$ converge et que :

$$(*) \qquad P\left(\bigcup_{n\in\mathbb{N}} B_n\right) = \sum_{n=0}^{+\infty} P(B \,/\, A_n) \times P(A_n).$$

Or on a :

$$(**) \qquad \bigcup_{n\in\mathbb{N}} B_n = \bigcup_{n\in\mathbb{N}} B \cap A_n = B \cap \underbrace{\left(\bigcup_{n\in\mathbb{N}} A_n\right)}_{\Omega} = B.$$

La propriété $\bigcup_{n\in\mathbb{N}} A_n = \Omega$ est conséquence du fait que $(A_n)_{n\in\mathbb{N}}$ est une partition dénombrable.

De $(*)$ et $(**)$, on déduit que :

$$P(B) = \sum_{n=0}^{+\infty} P(B \,/\, A_n) \times P(A_n).$$

Théorème 8 (théorème de Bayes)

Soit $(\Omega, \mathcal{T}, P)$ un espace de probabilités. Soient A et B deux événements non négligeables (i.e. $P(A \neq 0)$ et $P(B \neq 0)$). Alors on a :

$$P(A \,/\, B) = \frac{P(B \,/\, A) \times P(A)}{P(B)}.$$

► Remarque

En appliquant une des trois formules des probabilités totales au dénominateur $P(B)$ de la fraction apparaissant dans le théorème de Bayes, on obtient de nouvelles versions de ce théorème, pour des événements A intervenant dans la formule des probabilités totales considérée.

Une application typique du théorème de Bayes est le ≪ renversement du conditionnement ≫, i.e. le calcul de $P(A \,/\, B)$ connaissant la valeur de $P(B \,/\, A)$ (ainsi que les valeurs de $P(A)$ et $P(B)$).

Théorème 9 (théorème de Bayes et première formule des probabilités totales)

Soit $(\Omega, \mathcal{T}, P)$ un espace de probabilités. Soit A un événement qui n'est ni négligeable, ni presque sûr (i.e. $P(A) \neq 0$ et $P(A) \neq 1$). Alors pour tout événement B non négligeable (i.e. $P(B) \neq 0$), on a :

$$P(A \,/\, B) = \frac{P(B \,/\, A) \times P(A)}{P(B \,/\, A) \times P(A) + P(B \,/\, \overline{A}) \times P(\overline{A})}.$$

Théorème 10 (théorème de Bayes et deuxième formule des probabilités totales)

Soit $(\Omega, \mathcal{T}, P)$ un espace de probabilités. Soit $(A_1, \ldots, A_n)$ une partition finie de $(\Omega, \mathcal{T})$ telle que :

$$\forall\, i \in [\![1, n]\!] \quad P(A_i) \neq 0.$$

Alors pour tout événement B non négligeable (i.e. $P(B) \neq 0$), on a :

$$\forall\, i \in [\![1, n]\!] \quad P(A_i \,/\, B) = \frac{P(B \,/\, A_i) \times P(A_i)}{\displaystyle\sum_{j=1}^{n} P(B \,/\, A_j) \times P(A_j)}.$$

Théorème 11 (théorème de Bayes et troisième formule des probabilités totales)

Soit $(\Omega, \mathcal{T}, P)$ un espace de probabilités. Soit $(A_n)_{n\in\mathbb{N}}$ une partition dénombrable de $(\Omega, \mathcal{T})$ telle que :

$$\forall\, n \in \mathbb{N} \quad P(A_n) \neq 0.$$

Alors pour tout événement B non négligeable (i.e. $P(B) \neq 0$), on a :

$$\forall\, n \in \mathbb{N} \quad P(A_n \,/\, B) = \frac{P(B \,/\, A_n) \times P(A_n)}{\displaystyle\sum_{k=0}^{+\infty} P(B \,/\, A_k) \times P(A_k)}.$$

6 Espaces de probabilités produits

Nous admettrons le théorème suivant, utile par exemple pour modéliser des épreuves répétées.

Théorème 12 (espace de probabilités produit)

Soient $(\Omega_1, \mathcal{P}(\Omega_1), P_1)$ et $(\Omega_2, \mathcal{P}(\Omega_2), P_2)$ deux espaces de probabilités finis. Alors il existe une unique probabilité P sur l'espace probabilisable $(\Omega_1 \times \Omega_2, \mathcal{P}(\Omega_1 \times \Omega_2))$ telle que :

$$\forall\, A_1 \in \mathcal{P}(\Omega_1) \quad \forall\, A_2 \in \mathcal{P}(\Omega_2) \quad P(A_1 \times A_2) = P_1(A_1) \times P_2(A_2).$$

Cette probabilité P est appelée probabilité produit de P_1 et P_2.

► **Remarque**

Ce théorème admet une généralisation au cas où l'on a n espaces de probabilités finis $(\Omega_1, \mathcal{P}(\Omega_1), P_1), \ldots, (\Omega_n, \mathcal{P}(\Omega_n), P_n)$, où $n \geq 2$ est un entier quelconque. Alors il existe une unique probabilité P sur l'espace probabilisable $(\Omega_1 \times \ldots \times \Omega_n, \mathcal{P}(\Omega_1 \times \ldots \times \Omega_n))$ telle que :

$$\forall\, A_1 \in \mathcal{P}(\Omega_1) \quad \ldots \quad \forall\, A_n \in \mathcal{P}(\Omega_n) \quad P(A_1 \times \ldots \times A_n) = P_1(A_1) \times \ldots \times P_n(A_n).$$

Cette probabilité P est appelée probabilité produit des $P_1, \ldots, P_n$.

► **Exemple 14 :** On considère le jeu de « Pile » ou « Face ». On dispose d'une pièce de monnaie que l'on lance n fois de suite. On suppose que les résultats des différents lancers sont indépendants.

On note $p \in [0, 1]$ la probabilité d'obtenir « Pile » lors d'un lancer. Un lancer sera alors modélisé par l'espace de probabilité $(\Omega = \{\text{« Pile »}, \text{« Face »}\}, \mathcal{P}(\Omega), P)$ où P est la probabilité définie par :

$$P\colon \begin{array}{rcl} \{\text{« Pile »}, \text{« Face »}\} & \to & [0,1] \\ \text{« Pile »} & \mapsto & p \\ \text{« Face »} & \mapsto & 1-p \end{array}.$$

La suite des n lancers sera modélisée par l'espace de probabilités $(\Omega^n, \mathcal{P}(\Omega^n), Q)$, où Q est la probabilité produit de $\underbrace{P, \dots, P}_{n \text{ fois}}$.

Par exemple, la probabilité de l'événement ≪ on obtient n ≪ Pile ≫ ≫ sera :

$$Q(\underbrace{\{\text{≪ Pile ≫}\} \times \ldots \times \{\text{≪ Pile ≫}\}}_{n \text{ termes}}) = \underbrace{P(\{\text{≪ Pile ≫}\}) \times \ldots \times P(\{\text{≪ Pile ≫}\})}_{n \text{ termes}} = P(\{\text{≪ Pile ≫}\})^n = p^n$$

et la probabilité de l'événement ≪ on obtient k ≪ Pile ≫ puis $n-k$ ≪ Face ≫ ≫ $(0 \leq k \leq n)$ sera :

$$\begin{aligned}
& Q(\underbrace{\{\text{≪ Pile ≫}\} \times \ldots \times \{\text{≪ Pile ≫}\}}_{k \text{ termes}} \times \underbrace{\{\text{≪ Face ≫}\} \times \ldots \times \{\text{≪ Face ≫}\}}_{n-k \text{ termes}}) \\
= \quad & \underbrace{P(\{\text{≪ Pile ≫}\}) \times \ldots \times P(\{\text{≪ Pile ≫}\})}_{k \text{ termes}} \times \underbrace{P(\{\text{≪ Face ≫}\}) \times \ldots \times P(\{\text{≪ Face ≫}\})}_{n-k \text{ termes}} \\
= \quad & P(\{\text{≪ Pile ≫}\})^k \times P(\{\text{≪ Face ≫}\})^{n-k} \\
= \quad & p^k(1-p)^{n-k}.
\end{aligned}$$