Chapitre 1

Espace probabilisé

1.1 Définitions

Un *espace probabilisé* est représenté par un triplet $(\Omega, \mathcal{F}, \mathbb{P})$, où Ω est un ensemble fondamental, $\mathcal{F}$ une collection d'événements et $\mathbb{P}$ une probabilité.

L'*ensemble fondamental* Ω est l'ensemble de tous les résultats élémentaires possibles d'une expérience aléatoire. Un *événement* est un sous-ensemble de l'ensemble fondamental Ω. Étant donné un *résultat élémentaire* $\omega \in \Omega$, on dit qu'un événement $E \subseteq \Omega$ se réalise si $\omega \in E$, et qu'il ne se réalise pas sinon.

La *collection des événements* $\mathcal{F}$ vérifie les conditions suivantes :

(1) $\Omega \in \mathcal{F}$, appelé *événement certain* ;

(2) $E^c = \{\omega \in \Omega : \omega \notin E\} \in \mathcal{F}$ si $E \in \mathcal{F}$, appelé *événement complémentaire* ;

(3) $\bigcup_{n=1}^{+\infty} E_n = \{\omega \in \Omega : \omega \in E_n \text{ pour un } n \geqslant 1\} \in \mathcal{F}$ si $E_1, E_2, \ldots \in \mathcal{F}$.

Les conditions (1), (2) et (3) font de $\mathcal{F}$ une σ-algèbre, appelée aussi *tribu*. Elles entraînent les conditions supplémentaires suivantes :

(4) $\varnothing = \Omega^c \in \mathcal{F}$, appelé *événement impossible* ;

(5) $\bigcup_{n=1}^{N} E_n = \bigcup_{n=1}^{+\infty} E_n \in \mathcal{F}$ si $E_1, \ldots, E_N \in \mathcal{F}$ et $E_{N+1} = E_{N+2} = \cdots = \varnothing$;

(6) $\bigcap_{n=1}^{N} E_n = (\bigcup_{n=1}^{N} E_n^c)^c \in \mathcal{F}$ si $E_1, \ldots, E_N \in \mathcal{F}$;

(7) $\bigcap_{n=1}^{+\infty} E_n = (\bigcup_{n=1}^{+\infty} E_n^c)^c \in \mathcal{F}$ si $E_1, E_2, \ldots \in \mathcal{F}$.

Les conditions (1), (2) et (5) définissent une *algèbre*. Les conditions ci-dessus font en sorte que pour un nombre fini ou infini dénombrable d'événements en incluant l'ensemble fondamental lui-même, on peut considérer les événements que tous se réalisent, qu'au moins un se réalise et qu'aucun ne se réalise.

La *probabilité* associe à chaque événement $E \subseteq \Omega$ un nombre réel $\mathbb{P}(E)$ compris entre 0 et 1 qui mesure les chances de réalisation de cet événement. Autrement dit, c'est une application

$$\mathbb{P} : \mathcal{F} \to [0, 1].$$

De plus, cette application vérifie les conditions suivantes :

(a) **(valeurs 0 et 1)** $\mathbb{P}(\varnothing) = 0$ et $\mathbb{P}(\Omega) = 1$.

(b) **($\boldsymbol{\sigma}$-additivité)** $\mathbb{P}(\bigcup_{n=1}^{+\infty} E_n) = \sum_{n=1}^{+\infty} \mathbb{P}(E_n)$ si $E_1, E_2, \ldots \in \mathcal{F}$ sont des *événements incompatibles*, c'est-à-dire si $E_i \cap E_j = \varnothing$ dès que $i \neq j$.

Les conditions (a) et (b) entraînent les autres conditions suivantes :

(c) **(additivité)** $\mathbb{P}(\bigcup_{n=1}^{N} E_n) = \sum_{n=1}^{N} \mathbb{P}(E_n)$ si $E_1, \ldots E_N \in \mathcal{F}$ et $E_i \cap E_j = \varnothing$ dès que $i \neq j$, en prenant $E_{N+1} = E_{N+2} = \cdots = \varnothing$ dans (b).

(d) **(complémentarité)** $\mathbb{P}(E^c) = 1 - \mathbb{P}(E)$ si $E \in \mathcal{F}$ par ce qui précède, car $\mathbb{P}(\Omega) = \mathbb{P}(E \cup E^c) = \mathbb{P}(E) + \mathbb{P}(E^c)$.

Les conditions (a) et (b) garantissent en fait qu'une probabilité est une mesure fini. Dans ce contexte, un événement est un ensemble mesurable.

Plus généralement, un *espace mesuré* est un triplet (S, Σ, μ), où S est un ensemble, Σ est une σ-algèbre de *sous-ensembles mesurables* de S, et enfin $\mu : \Sigma \to [0, +\infty]$ est une *mesure*, c'est-à-dire une application σ-additive avec $\mu(\varnothing) = 0$. L'espace mesuré est *fini* si $\mu(S) < +\infty$. Il est *σ-fini* si $S = \bigcup_{n=1}^{+\infty} S_n$ avec $\mu(S_n) < +\infty$ pour tout $n \geqslant 1$.

1.2 Schéma de Bernoulli

Voici l'exemple le plus important d'un espace probabilisé fini. L'expérience aléatoire consiste à observer la réalisation d'un événement ou de son complémentaire n fois. Imaginons, par exemple, qu'on lance une pièce de monnaie n fois avec comme résultat face ou pile à chacune des fois. En représentant face par 1 et pile par 0, l'ensemble des résultats élémentaires est donné par

$$\Omega = \{\omega = (\omega_1, \ldots, \omega_n) : \omega_i = 1 \text{ ou } 0 \text{ pour } i = 1, \ldots, n\}.$$

Comme collection d'événements, on considère l'ensemble des parties de Ω, soit

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

C'est la *tribu discrète*. Enfin, la probabilité de $E \subseteq \Omega$ est définie par

$$\mathbb{P}(E) = \sum_{\omega \in E} p_\omega,$$

où

$$p_\omega = p^{\sum_{i=1}^n \omega_i} (1-p)^{n-\sum_{i=1}^n \omega_i}$$

pour $0 \leqslant p \leqslant 1$.

Ainsi, on a $\mathbb{P}(\varnothing) = 0$ et

$$\mathbb{P}(\Omega) = \sum_{k=0}^{n} \left(\begin{array}{c} n \\ k \end{array} \right) p^k (1-p)^{n-k} = (p + (1-p))^n = 1$$

par la *formule du binôme de Newton*, où

$$\left(\begin{array}{c} n \\ k \end{array} \right) = \frac{n!}{k!(n-k)!}$$

représente le nombre de $\omega = (\omega_1, \ldots, \omega_n) \in \Omega$ avec $\sum_{i=1}^n \omega_i = k$. De plus,

$$\mathbb{P}(\bigcup_{i=1}^N E_i) = \sum_{i=1}^{N} \sum_{\omega \in E_i} p_\omega = \sum_{i=1}^{N} \mathbb{P}(E_i)$$

pour des événements incompatibles $E_1, \ldots, E_N$. Ce n'est pas le cas pour les événements $E_1, \ldots, E_n$ si E_i est l'événement *obtenir face au i-ième jet*, qui est représenté par

$$E_i = \{\omega \in \Omega : \omega_i = 1\}$$

et dont la probabilité est

$$\mathbb{P}(E_i) = \sum_{k=0}^{n-1} \left(\begin{array}{c} n-1 \\ k \end{array} \right) p^{k+1} (1-p)^{n-k-1} = p\,(p + (1-p))^{n-1} = p.$$

En revanche, c'est le cas pour les événements $F_1, \ldots, F_n$ si F_i est l'événement *obtenir face pour la première fois au i-ième jet*, qui est représenté par

$$F_i = \{\omega \in \Omega : \omega_1 = \cdots = \omega_{i-1} = 0, \omega_i = 1\}$$

et dont la probabilité est

$$\begin{aligned}
\mathbb{P}(F_i) &= \sum_{k=0}^{n-i} \left(\begin{array}{c} n-i \\ k \end{array} \right) p^{k+1} (1-p)^{n-k-1} \\
&= (1-p)^{i-1} p \sum_{k=0}^{n-i} \left(\begin{array}{c} n-i \\ k \end{array} \right) p^k (1-p)^{n-i-k} \\
&= (1-p)^{i-1} p.
\end{aligned}$$

On remarque cependant que

$$\bigcup_{i=1}^{n} E_i = \bigcup_{i=1}^{n} F_i = B,$$

dont le complémentaire

$$B^c = \{\omega \in \Omega : \omega_1 = \cdots = \omega_n = 0\}$$

est l'événement *n'obtenir aucune face* en n jets. Par additivité, on a

$$\mathbb{P}(B) = \mathbb{P}(\bigcup_{i=1}^{n} F_i) = \sum_{i=1}^{n} \mathbb{P}(F_i) = \sum_{i=1}^{n} (1-p)^{i-1} p = 1 - (1-p)^n.$$

Par complémentarité, on a aussi

$$\mathbb{P}(B) = 1 - \mathbb{P}(B^c) = 1 - (1-p)^n.$$

1.3 Mesure de Lebesgue sur $(0,1]$

Voici maintenant l'exemple d'un espace probabilisé continu qui s'avère le plus important. Il modélise l'expérience qui consiste à choisir un nombre au hasard dans l'intervalle $(0,1]$. L'ensemble fondamental est alors le sous-ensemble des nombres réels

$$\Omega = (0,1] = \{x \in \mathbb{R} : 0 < x \leqslant 1\}.$$

On considère la collection des sous-intervalles

$$\Pi(0,1] = \{(0,b] : 0 < b \leqslant 1\}.$$

Cette collection est fermée sous l'opération d'intersection de deux quelconques de ses éléments, ce qui signifie que $I \cap J \in \Pi(0,1]$ si $I, J \in \Pi(0,1]$. C'est ce qui est appelé un *π-système*.

Comme σ-algèbre d'événements, on prend la plus petite σ-algèbre de sous-ensembles de $(0,1]$ qui contient $\Pi(0,1]$. C'est en fait l'intersection de toutes les tribus contenant $\Pi(0,1]$, dont en particulier la tribu discrète $\mathcal{P}(0,1]$ qui est définie comme l'ensemble de toutes les parties de $(0,1]$. On obtient ainsi la *σ-algèbre engendrée par le π-système* $\Pi(0,1]$, qu'on représente par

$$\mathcal{F} = \sigma(\Pi(0,1]).$$

Cette σ-algèbre est la *tribu borélienne* de $(0,1]$ qui est défini comme la σ-algèbre engendrée par les ouverts de $(0,1]$, c'est-à-dire

$$\mathcal{B}(0,1] = \sigma(\text{ouverts de } (0,1]).$$

En effet, tout ouvert de $(0,1]$ est une union d'un nombre fini ou infini dénombrable d'intervalles disjoints de forme

$$\begin{aligned}(a,b)\cap(0,1] &= (0,b)\cap(0,a]^c\cap(0,1]\\ &= \textstyle\bigcup_{n=1}^{+\infty}\{(0,b-1/n]\cap(0,a]^c\cap(0,1]\}\in\sigma(\Pi(0,1])\end{aligned}$$

pour $0\leqslant a<b<+\infty$. Par conséquent, on a $\mathcal{B}(0,1]\subseteq\mathcal{F}$. D'autre part, on a

$$(0,b]=\textstyle\bigcap_{n=1}^{+\infty}\{(0,b+1/n)\cap(0,1]\}\in\mathcal{B}(0,1]$$

pour $0<b\leqslant 1$, d'où $\mathcal{F}\subseteq\mathcal{B}(0,1]$. On conclut que $\mathcal{F}=\mathcal{B}(0,1]$. On a bien sûr l'inclusion $\mathcal{B}(0,1]\subseteq\mathcal{P}(0,1]$. En revanche, l'inclusion inverse n'est pas vraie.

Enfin, on considère une probabilité

$$\mathbb{P}:\sigma(\Pi(0,1])\to[0,1]$$

qui est telle que $\mathbb{P}((0,b])=b$ pour $0<b\leqslant 1$. Par additivité, on a alors

$$\mathbb{P}((a,b])=\mathbb{P}((0,b]\cap(0,a]^c)=\mathbb{P}((0,b])-\mathbb{P}((0,a])=b-a$$

pour $0\leqslant a<b\leqslant 1$. Une probabilité qui vérifie cette condition existe. De plus, elle est unique. C'est la *mesure de Lebesgue* sur $(0,1]$, qu'on représente par λ. Son existence et son unicité découlent de deux grands résultats en théorie de la mesure qui sont énoncés ci-dessous sans démonstration.

Proposition 1.3.1 (théorème d'extension de Carathéodory). *Soit*

$$\mu_0:\mathcal{A}\to[0,+\infty]$$

une application σ-additive sur une algèbre $\mathcal{A}$ de sous-ensembles de S avec $\mu_0(\varnothing)=0$. Alors, l'application

$$\mu:\sigma(\mathcal{A})\to[0,+\infty]$$

définie pour tout A dans la σ-algèbre engendrée par $\mathcal{A}$ par

$$\mu(A)=\inf\left\{\sum_{n=1}^{+\infty}\mu_0(A_n):A_n\in\mathcal{A}\text{ pour }n\geqslant 1, A\subseteq\textstyle\bigcup_{n=1}^{+\infty}A_n\right\}$$

est une mesure, appelée mesure extérieure*, qui est telle que $\mu=\mu_0$ sur $\mathcal{A}$.*

Proposition 1.3.2 (théorème d'unicité d'une mesure). *Soient*

$$\mu_1, \mu_2 : \sigma(\Pi) \to [0, +\infty]$$

deux mesures définies sur la σ-algèbre engendrée par un Π-système de sous-ensembles de S. Si $\mu_1(S) = \mu_2(S) < +\infty$ et $\mu_1(I) = \mu_2(I)$ pour tout $I \in \Pi$, alors $\mu_1 = \mu_2$ sur $\sigma(\Pi)$.

Pour montrer l'existence de la mesure de Lebesgue sur $(0, 1]$, il suffit alors de montrer que l'algèbre engendrée par $\Pi(0, 1]$ est

$$\mathcal{A}(\Pi(0,1]) = \{(a_1, b_1] \cup \cdots \cup (a_r, b_r] : 0 < a_1 < b_1 < \cdots < a_r < b_r \leqslant 1\},$$

et que l'application

$$\lambda : \mathcal{A}(\Pi(0,1]) \to [0,1]$$

définie par

$$\lambda((a_1, b_1] \cup \cdots \cup (a_r, b_r]) = \sum_{k=1}^{r} (b_k - a_k)$$

avec $\lambda(\varnothing) = 0$ est σ-additive. L'extension sur $\sigma(\mathcal{A}(\Pi(0,1])) = \sigma(\Pi(0,1])$ est une probabilité sur $(0,1]$. En ajoutant le point 0 ou en retirant le point 1, tous deux de mesure 0, on obtient une probabilité sur $\mathcal{B}[0,1]$ ou sur $\mathcal{B}(0,1)$, respectivement.

Par complémentarité et σ-additivité, on a en particulier

$$\begin{aligned}
\lambda(\text{nombres irrationnels dans } (0,1]) &= 1 - \lambda(\text{nombres rationnels dans } (0,1]) \\
&= 1 - \lambda\left(\textstyle\bigcup_{a \in \mathbb{Q} \cap (0,1]} \{a\}\right) \\
&= 1 - \sum_{a \in \mathbb{Q} \cap (0,1]} \lambda(\{a\}) \\
&= 1,
\end{aligned}$$

où $\mathbb{Q}$ représente l'ensemble des nombres rationnels, à la condition cependant que

$$\begin{aligned}
\lambda(\{a\}) &= \lim_{n \to +\infty} \lambda\left((a - 1/n, a]\right) \\
&= \lim_{n \to +\infty} (a - (a - 1/n)) \\
&= 0
\end{aligned}$$

pour $0 < a \leqslant 1$. C'est une propriété de continuité qui découle des autres propriétés d'une probabilité tel que cela est montré dans la section suivante.

L'unicité de la mesure de Lebesgue sur $(0, 1]$ est une conséquence directe du théorème d'unicité d'une mesure étant donné ses valeurs sur un π-système. Ce théorème est en fait un corollaire du résultat plus général énoncé ci-dessous.

Proposition 1.3.3 (théorème $\boldsymbol{\lambda}$-$\boldsymbol{\pi}$ de Dynkin). *Soit $\mathcal{C}$ une collection de sous-ensembles de S qui contient un π-système Π et qui est un λ-système, c'est-à-dire telle que :*

(a) $S \in \mathcal{C}$ *;*

(b) $B^c \in \mathcal{C}$ *si* $B \in \mathcal{C}$ *;*

(c) $\bigcup_{n=1}^{+\infty} B_n \in \mathcal{C}$ *si* $B_n \in \mathcal{C}$ *pour* $n \geqslant 1$ *avec* $B_n \cap B_m = \varnothing$ *dès que* $n \neq m$.

Alors, $\mathcal{C}$ contient nécessairement $\sigma(\Pi)$, la σ-algèbre engendrée par Π.

1.4 Propriétés d'une probabilité

Les conditions (a), (b), (c) et (d) pour une probabilité $\mathbb{P} : \mathcal{F} \to [0,1]$ entraînent les propriétés supplémentaires suivantes pour lesquelles il est sous-entendu que tous les sous-ensembles considérés sont dans la σ-algèbre $\mathcal{F}$:

(e) (formule d'inclusion-exclusion) On a

$$\begin{aligned}\mathbb{P}(\textstyle\bigcup_{n=1}^{N} E_n) = &\sum_{1\leqslant n \leqslant N} \mathbb{P}(E_n) - \sum_{1\leqslant n_1 < n_2 \leqslant N} \mathbb{P}\left(E_{n_1} \cap E_{n_2}\right) \\ &+ \cdots + (-1)^{r-1} \sum_{1\leqslant n_1 < \cdots < n_r \leqslant N} \mathbb{P}\left(E_{n_1} \cap \cdots \cap E_{n_r}\right) \\ &+ \cdots + (-1)^{N-1} \mathbb{P}\left(E_1 \cap \cdots \cap E_N\right).\end{aligned}$$

Cela se démontre par récurrence sur N à partir de l'identité

$$\mathbb{P}(E_1 \cup E_2) = \mathbb{P}(E_1) + \mathbb{P}(E_2) - \mathbb{P}(E_1 \cap E_2),$$

qui découle de

$$\begin{aligned}\mathbb{P}(E_1 \cup E_2) &= \mathbb{P}(E_1 \cap E_2^c) + \mathbb{P}(E_1^c \cap E_2) + \mathbb{P}(E_1 \cap E_2) \\ &= \mathbb{P}(E_1) - \mathbb{P}(E_1 \cap E_2) + \mathbb{P}(E_2) - \mathbb{P}(E_1 \cap E_2) + \mathbb{P}(E_1 \cap E_2)\end{aligned}$$

par la propriété d'additivité.

(f) (monotonicité) On a

$$\mathbb{P}(F) \leqslant \mathbb{P}(E) \text{ si } F \subseteq E.$$

En effet, on a alors l'identité

$$\mathbb{P}(E) = \mathbb{P}(F) + \mathbb{P}(E \cap F^c)$$

par additivité.

(g) (sous-σ-additivité) On a

$$\mathbb{P}(\bigcup_{n=1}^{+\infty} E_n) \leqslant \sum_{n=1}^{+\infty} \mathbb{P}(E_n).$$

En effet, les événements

$$F_n = E_n \cap (\bigcup_{i=1}^{n-1} E_i)^c \subseteq E_n$$

pour $n \geqslant 1$ sont incompatibles et vérifient $\bigcup_{n=1}^{+\infty} F_n = \bigcup_{n=1}^{+\infty} E_n$. On a alors

$$\mathbb{P}(\bigcup_{n=1}^{+\infty} E_n) = \mathbb{P}(\bigcup_{n=1}^{+\infty} F_n) = \sum_{n=1}^{+\infty} \mathbb{P}(F_n) \leqslant \sum_{n=1}^{+\infty} \mathbb{P}(E_n)$$

par σ-additivité et monotonicité.

(h) (sous-additivité) On a

$$\mathbb{P}(\bigcup_{n=1}^{N} E_n) \leqslant \sum_{n=1}^{N} \mathbb{P}(E_n)$$

en prenant $E_{N+1} = E_{N+2} = \cdots = \varnothing$ dans la propriété précédente.

(i) (continuité vers le haut) On a

$$\mathbb{P}(\bigcup_{n=1}^{N} E_n) \uparrow \mathbb{P}(\bigcup_{n=1}^{+\infty} E_n) \text{ lorsque } N \uparrow +\infty.$$

En effet, utilisant les événements F_n pour $n \geqslant 1$ tels que définis dans la démonstration de la propriété (g), on a

$$\begin{aligned} \mathbb{P}(\bigcup_{n=1}^{N} E_n) = \mathbb{P}(\bigcup_{n=1}^{N} F_n) &= \sum_{n=1}^{N} \mathbb{P}(F_n) \\ &\uparrow \sum_{n=1}^{+\infty} \mathbb{P}(F_n) = \mathbb{P}(\bigcup_{n=1}^{+\infty} F_n) = \mathbb{P}(\bigcup_{n=1}^{+\infty} E_n) \end{aligned}$$

par additivité et σ-additivité. En prenant des événements E_n pour $n \geqslant 1$ incompatibles, on voit que les propriétés d'additivité et de continuité vers le haut entraînent la propriété de σ-additivité.