Chapitre 1

Espaces probabilisés

Après une brève introduction, nous donnerons des propriétés élémentaires des probabilités, puis nous étudierons la propriété plus délicate de σ -additivité. Ensuite, nous présenterons la construction d'une probabilité dans des espaces finis ou dénombrables, et quelques exemples de modélisation probabiliste. Ces exemples feront apparaître des lois classiques : la loi de Bernoulli, la loi binomiale, la loi de Poisson, la loi géométrique.

Puis, nous présenterons la construction d'une probabilité dans le cadre général. Pour rendre cette construction efficace, nous introduirons la notion indispensable de fonction de répartition.

Dans un espace particulier $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ nous donnerons la classification des fonctions de répartition : celles correspondant à des probabilités discrètes, à des probabilités continues et absolument continues par rapport à la mesure de Lebesgue. En particulier, des lois classiques à densité seront présentées : la loi uniforme, la loi normale, la loi gamma, la loi du khi-deux, la loi de Student et la loi de Cauchy.

1.1 Généralités

La théorie des probabilités vise à fournir des modèles mathématiques pour décrire des phénomènes aléatoires. On dit qu'un phénomène est aléatoire si, d'une part on ne peut pas prévoir à l'avance son résultat, mais que d'autre part, par « répétition », ce phénomène présente un certain caractère de régularité. Un jeu de roulette de « noir et rouge » au casino qui donne autant de chances de gagner en misant sur « noir » que « rouge » en est un exemple. Les jeux de Loto, de Keno et autres jeux, l'évolution du prix d'une action, du cours des changes, du prix d'un produit sont aussi des exemples de phénomènes aléatoires.

Un modèle probabiliste, d'après des travaux de Kolmogorov de 1933, est constitué de trois éléments :

- un espace d'états Ω,
- une σ-algèbre ou tribu $\mathcal{F}$,
- une probabilité $\mathbf{P}$.

L'espace d'états Ω est l'espace de tous les résultats possibles du phénomène aléatoire observé. Pour le lancement d'une pièce, on pourrait prendre $\Omega = \{p, f\}$, où p désigne « obtenir pile » et f désigne « obtenir face » ; pour n lancements d'une pièce, on pourrait prendre $\Omega = \{p, f\}^n$; pour une étude de durée de vie d'une ampoule électrique, on pourrait se servir de $\Omega = \mathbb{R}^+$. Il faut souligner que l'espace d'états Ω n'est pas unique. Cet espace doit être choisi de façon adéquate pour permettre d'effectuer le calcul des probabilités.

Le deuxième élément du modèle probabiliste est une tribu.

Définition 1.1. Un ensemble $\mathcal{F}$ de sous-ensembles de Ω est appelé **tribu** (ou σ-algèbre) s'il contient l'ensemble vide $\emptyset$ et s'il est stable par rapport au complémentaire et à l'union dénombrable d'éléments, c'est-à-dire

1. $\emptyset \in \mathcal{F}$,
2. pour tout $A \in \mathcal{F}$, on a : $A^c \in \mathcal{F}$,
3. pour toute suite $(A_i)_{i\in\mathbb{N}}$ telle que les $A_i \in \mathcal{F}$ on a : $\bigcup\limits_{i\in\mathbb{N}} A_i \in \mathcal{F}$.

Exemples 1.1. Donnons quelques exemples simples de tribus.

a) La tribu la plus pauvre $\mathcal{F} = \{\emptyset, \Omega\}$ s'appelle la tribu « grossière ».

b) La tribu la plus riche possible, notée habituellement par $\mathcal{P}(\Omega)$, est l'ensemble des toutes les parties de Ω.

Remarque 1.1. *Supposons que la propriété 2 de la définition d'une tribu soit vérifiée. Alors, la propriété 3 de stabilité par union dénombrable d'éléments de $\mathcal{F}$ est équivalente à la stabilité par rapport à l'intersection dénombrable d'éléments de $\mathcal{F}$, c'est-à-dire à la propriété :*

3′. pour toute suite $(A_i)_{i\in\mathbb{N}}$ telle que les $A_i \in \mathcal{F}$ on a : $\bigcap\limits_{i\in\mathbb{N}} A_i \in \mathcal{F}$.

En effet, l'équivalence $3 \Leftrightarrow 3'$ découle de la relation :

$$\left(\bigcap_{i=0}^{+\infty} A_i\right)^c = \bigcup_{i=0}^{+\infty} A_i^c.$$

Définition 1.2. Soit $G \subseteq \mathcal{P}(\Omega)$ une famille de parties de Ω. On note $\mathcal{F}(G)$ la **tribu engendrée** par G. C'est la plus petite tribu contenant G, c'est-à-dire la tribu contenant G telle que pour toute autre tribu $\mathcal{F}'(G)$ contenant G, on a : $\mathcal{F}(G) \subseteq \mathcal{F}'(G)$.

On dit que l'ensemble de parties G est **générateur** de la tribu $\mathcal{F}(G)$ ou que G engendre la tribu $\mathcal{F}(G)$.

L'existence de $\mathcal{F}(G)$ découle de deux choses : tout d'abord $\mathcal{P}(\Omega)$ est une tribu contenant G, ensuite, on peut démontrer facilement que l'intersection des tribus est aussi une tribu. Donc, $\mathcal{F}(G)$ n'est rien d'autre que l'intersection de toutes les tribus contenant G.

Exemples 1.2. Donnons quelques exemples simples d'une tribu engendrée par un ensemble.

a) La tribu engendrée par une partition $\mathcal{D} = (D_i)_{i\in\mathbb{N}}$ de Ω. Rappelons que $\mathcal{D} = (D_i)_{i\in\mathbb{N}}$ est une partition de Ω, si les ensembles $D_i \neq \emptyset$, pour tout $i, j \in \mathbb{N}$ tels que $i \neq j$, $D_i \cap D_j = \emptyset$ et $\bigcup\limits_{i\in\mathbb{N}} D_i = \Omega$. Alors,

$$\mathcal{F}(\mathcal{D}) = \left\{ \bigcup_{i\in J} D_i \mid J \in \mathcal{P}(\mathbb{N}) \right\},$$

où $\mathcal{P}(\mathbb{N})$ est l'ensemble des toutes les parties de $\mathbb{N}$.

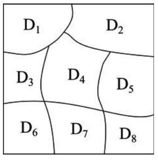

Figure 1. L'espace d'états $\Omega = [0,1] \times [0,1]$ ayant une partition de 8 éléments.

b) Les tribus boréliennes $\mathcal{B}(\mathbb{R})$, $\mathcal{B}(\mathbb{R}^+)$, $\mathcal{B}([0,1])$. Ce sont des tribus engendrées par les ouverts de $\mathbb{R}$, de $\mathbb{R}^+$ et de $[0,1]$ respectivement.

En théorie des probabilités, on utilise souvent du vocabulaire spécifique. On appelle les éléments de $\mathcal{F}$ des événements ; l'ensemble Ω lui même s'appelant l'événement certain et l'ensemble vide $\emptyset$, l'événement impossible. Des éléments ω de l'espace Ω, appartenant à $\mathcal{F}$, s'appellent des événements élémentaires ou des singletons.

Définition 1.3. Une application $\mathbf{P} : \mathcal{F} \to [0,1]$ est **une probabilité** si elle vérifie :

1. $\mathbf{P}(\Omega) = 1$,
2. $\mathbf{P}$ est σ- additive, c'est-à-dire pour toutes suites $(A_i)_{i\in\mathbb{N}}$ d'événements deux-à-deux disjoints : $A_i \cap A_j = \emptyset$ pour $i \neq j$, on a

$$\mathbf{P}\left(\bigcup_{i=0}^{+\infty} A_i\right) = \sum_{i=0}^{+\infty} \mathbf{P}(A_i).$$

Remarque 1.2. *La propriété de σ-additivité implique tout de suite que*

$$\mathbf{P}(\emptyset) = 0.$$

En effet, pour ce résultat il suffit de prendre une suite particulière $(A_i)_{i\in\mathbb{N}}$, notamment $A_i = \emptyset$ pour tout i.
La propriété de σ-additivité implique aussi la propriété d'additivité : pour tout $A \in \mathcal{F}$ et tout $B \in \mathcal{F}$ tels que $A \cap B = \emptyset$

$$\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B)$$

En effet, il suffit de prendre $A_0 = A$, $A_1 = B$, de compléter la suite par des ensembles vides, puis d'utiliser la propriété de σ-additivité. Notons aussi que la propriété d'additivité n'implique pas, en général, la σ-additivité (voir le théorème 1.2).

Donnons des propriétés élémentaires de la probabilité $\mathbf{P}$.

Proposition 1.1. *La probabilité $\mathbf{P}$ vérifie :*

1. *pour $A \in \mathcal{F}$, on a : $\mathbf{P}(A^c) = 1 - \mathbf{P}(A)$, en particulier $\mathbf{P}(\emptyset) = 0$,*
2. *pour $A, B \in \mathcal{F}$ tels que $A \subseteq B$ nous avons*

$$\mathbf{P}(A) \leq \mathbf{P}(B),$$

3. *pour $A, B \in \mathcal{F}$ nous avons*

$$\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B) - \mathbf{P}(A \cap B),$$

en particulier

$$\mathbf{P}(A \cup B) \leq \mathbf{P}(A) + \mathbf{P}(B).$$

Preuve. Pour montrer la propriété 1, prenons $A \in \mathcal{F}$, alors $A^c \in \mathcal{F}$. De plus, $A \cap A^c = \emptyset$, $A \cup A^c = \Omega$ et $\mathbf{P}(\Omega) = 1$. Par σ-additivité $\mathbf{P}(A) + \mathbf{P}(A^c) = 1$, et, donc, la propriété 1 est vraie.

Pour établir la propriété 2, prenons $A, B \in \mathcal{F}$ tels que $A \subseteq B$. Alors, nous avons $(B \setminus A) \cup A = B$ et $(B \setminus A) \cap A = \emptyset$, d'où

$$\mathbf{P}(B) = \mathbf{P}(B \setminus A) + \mathbf{P}(A) \geq \mathbf{P}(A)$$

car $\mathbf{P}(B \setminus A) \geq 0$.

Pour démontrer la propriété 3, prenons $A, B \in \mathcal{F}$. Alors

$$A \cup B = (B \setminus A \cap B) \cup (A \setminus A \cap B) \cup A \cap B.$$

Les ensembles dans cette union étant disjoints, nous obtenons

$$\mathbf{P}(A \cup B) = \mathbf{P}(B \setminus A \cap B) + \mathbf{P}(A \setminus A \cap B) + \mathbf{P}(A \cap B).$$

Puisque $\mathbf{P}(B \setminus A \cap B) = \mathbf{P}(B) - \mathbf{P}(A \cap B)$ et $\mathbf{P}(A \setminus A \cap B) = \mathbf{P}(A) - \mathbf{P}(A \cap B)$, nous déduisons que l'égalité de la propriété 3 est vérifiée. Pour compléter la preuve, il nous reste à remarquer que $\mathbf{P}(A \cap B) \geq 0$. $\square$

Remarque 1.3. *Les propriétés élémentaires de* $\mathbf{P}$ *peuvent être interprétées à l'aide de dessins : prendre pour* Ω *un carré de coté 1 et comprendre la probabilité d'un sous-ensemble de* Ω *comme son aire.*

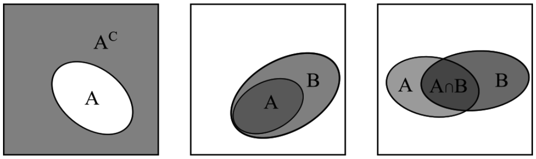

Figure 2. Illustration des propriétés élémentaires de probabilités.

Définition 1.4. Soit Ω un espace d'états, $\mathcal{F}$ est une tribu de Ω, et $\mathbf{P}$ est une probabilité sur $\mathcal{F}$. Le triplet $(\Omega, \mathcal{F}, \mathbf{P})$ est appelé **espace probabilisé**.

En se référant au cours de la théorie de la mesure, on pourrait aussi dire qu'un espace probabilisé n'est rien d'autre qu'un espace mesurable muni d'une mesure bornée positive, σ-additive, de masse totale 1.

Théorème 1.2. *Supposons que l'application* $\mathbf{P} : \mathcal{F} \to [0, 1]$ *vérifie :*

1. $\mathbf{P}(\Omega) = 1$,
2. *pour tout* $A, B \in \mathcal{F}$, $A \cap B = \emptyset$, *nous avons la propriété d'additivité :*

$$\mathbf{P}(A \cup B) = \mathbf{P}(A) + \mathbf{P}(B).$$

Alors, la σ-additivité de $\mathbf{P}$ est équivalente à la continuité de $\mathbf{P}$ pour les suites croissantes et les suites décroissantes d'événements, c'est à dire

a) pour toute suite croissante d'événements $(A_i)_{i\in\mathbb{N}}$, $A_i \subseteq A_{i+1}$, telle que $A = \bigcup_{i=0}^{+\infty} A_i$,

$$\lim_{i\to+\infty} \mathbf{P}(A_i) = \mathbf{P}(A),$$

b) pour toute suite décroissante d'événements $(B_i)_{i\in\mathbb{N}}$, $B_i \supseteq B_{i+1}$, telle que $B = \bigcap_{i=0}^{+\infty} B_i$,

$$\lim_{i\to+\infty} \mathbf{P}(B_i) = \mathbf{P}(B).$$

Preuve. Tout d'abord établissons l'équivalence $a) \Leftrightarrow b)$ par le passage au complémentaire. En effet, si $a)$ est vérifiée, alors pour toute suite décroissante d'événements $(B_i)_{i\in\mathbb{N}}$, la suite $(B_i^c)_{i\in\mathbb{N}}$, forme une suite croissante d'événements. Donc, $\lim_{i\to+\infty} \mathbf{P}(B_i^c) = \mathbf{P}(B^c)$. Mais puisque pour tout i, $\mathbf{P}(B_i^c) = 1 - \mathbf{P}(B_i)$ et $\mathbf{P}(B^c) = 1 - \mathbf{P}(B)$, nous obtenons $b)$. De la même manière on montre que $b) \Rightarrow a)$.

Pour démontrer que σ-additivité implique $a)$, prenons une suite croissante d'événements $(A_i)_{i\in\mathbb{N}}$ et pour tout $i \in \mathbb{N}$, introduisons les ensembles $C_i = A_i \setminus A_{i-1}$, où par convention $A_{-1} = \emptyset$.

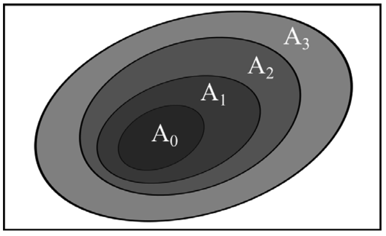

Figure 3. Construction des ensembles C_i.

Alors par construction, les ensembles C_i sont deux-à-deux disjoints, ils appartiennent à $\mathcal{F}$ et

$$A_i = \bigcup_{j=0}^{i} C_j, \qquad A = \bigcup_{j=0}^{+\infty} C_j.$$

De plus, par σ-additivité $\mathbf{P}(A) = \sum_{j=0}^{+\infty} \mathbf{P}(C_i)$. À son tour, la propriété d'additivité implique que $\mathbf{P}(A_i) = \sum_{j=0}^{i} \mathbf{P}(C_j)$. Puisque la somme partielle d'une série convergente converge vers la somme de cette série, la propriété $a)$ est établie.

Inversement, supposons que $a)$ est vérifiée et qu'on cherche à établir la σ-additivité. Soit $(C_i)_{i\in\mathbb{N}}$ une suite d'événements deux-à-deux disjoints. Alors, les ensembles $A_i = \bigcup\limits_{j=0}^{i} C_j$ croissent vers $A = \bigcup\limits_{j=0}^{+\infty} C_j$ et la propriété $a)$ nous garantit que $\lim_{i\to+\infty} \mathbf{P}(A_i) = \mathbf{P}(A)$. De plus, puisque les C_i sont deux-à-deux disjoints, on conclut par récurrence à partir de la propriété 2 que

$$\mathbf{P}(A_i) = \sum_{j=0}^{i} \mathbf{P}(C_j).$$

Mais la somme partielle d'une série convergente ne peut converger que vers sa somme, et donc en passant à la limite quand $i \to +\infty$, on trouve que

$$\mathbf{P}(A) = \mathbf{P}\left(\bigcup_{j=0}^{+\infty} C_j\right) = \sum_{j=0}^{+\infty} \mathbf{P}(C_j).$$

□

Remarque 1.4. *Les conditions a) et b) sont équivalentes à une seule condition souvent plus facile à vérifier : pour toute suite d'événements $(B_i)_{i\in\mathbb{N}}$ décroissante vers l'ensemble vide $\emptyset$*

$$\lim_{i\to+\infty} \mathbf{P}(B_i) = 0$$

En effet, d'une part cette condition est un cas particulier de b) avec $B = \emptyset$. D'autre part, si $(A_i)_{i\in\mathbb{N}}$ est une suite croissante d'événements, alors $(A\setminus A_i)_{i\in\mathbb{N}}$ est une suite décroissante vers l'ensemble vide, et $\lim\limits_{i\to+\infty} \mathbf{P}(A \setminus A_i) = 0$. Mais $\mathbf{P}(A \setminus A_i) = \mathbf{P}(A) - \mathbf{P}(A_i)$, et donc a) a lieu.

Corollaire 1.3. *Pour toute suite $(A_i)_{i\in\mathbb{N}}$ d'événements, on a :*

$$\mathbf{P}\left(\bigcup_{i=0}^{+\infty} A_i\right) \leq \sum_{i=0}^{+\infty} \mathbf{P}(A_i).$$

Preuve. La propriété 3 de la proposition 1.1 combinée avec un raisonnement par récurrence nous donne : pour tout $n \in \mathbb{N}$

$$\mathbf{P}\left(\bigcup_{i=0}^{n} A_i\right) \leq \sum_{i=0}^{n} \mathbf{P}(A_i).$$

En passant à la limite à l'aide du théorème 1.2. quand $n \to +\infty$, nous arrivons au résultat souhaité. □

1.2 Espaces probabilisés finis et dénombrables

Les espaces probabilisés finis et dénombrables sont des cas très particuliers d'espaces probabilisés. Rappelons qu'un espace d'états Ω est appelé espace d'états fini si $\mathrm{card}(\Omega) < +\infty$. Un espace d'états Ω est appelé espace d'états dénombrable si Ω contient un nombre dénombrable d'éléments. Dans ces cas là, on peut numéroter tous les éléments de l'espace Ω et on obtient :

$$\Omega = \{\omega_0, \omega_1, \omega_2, \ldots, \omega_n, \ldots\}.$$

On peut aussi prendre comme tribu $\mathcal{P}(\Omega)$, ce qui évite toutes les questions difficiles de mesurabilité. De plus, sur l'espace fini la propriété de σ- additivité devient équivalente à l'additivité, ce qui s'avère faux dans le cas général. Ensuite, la probabilité $\mathbf{P}$ sur un espace fini ou dénombrable peut être donnée de façon simple : en précisant les valeurs de l'application $\mathbf{P}$ sur les singletons. Notons que cela s'avère faux dans le cas général.

Proposition 1.4. *Une probabilité sur un espace d'états Ω dénombrable est entièrement caractérisée par ses valeurs sur les singletons: $p_i = \mathbf{P}(\{\omega_i\})$, $i \in \mathbb{N}$, qui vérifient*

$$\sum_{i=0}^{+\infty} p_i = 1.$$

Preuve. Pour tout $A \in \mathcal{F}$ on pose

$$\mathbf{P}(A) = \sum_{\omega_i \in A} \mathbf{P}(\{\omega_i\}).$$

Puisque la série de terme général p_i converge, la définition de l'application $\mathbf{P}$ ci-dessus est correcte. Il reste à vérifier la propriété de σ- additivité : pour toute suite d'événements $(A_i)_{i\in\mathbb{N}}$ deux-à deux disjoints

$$\mathbf{P}\left(\bigcup_{j=0}^{+\infty} A_j\right) = \sum_{j=0}^{+\infty} \mathbf{P}(A_j).$$

Cette propriété, en effet, découle du théorème de sommation par paquets des séries absolument convergentes. Notons $I_j = \{i \in \mathbb{N} \mid \omega_i \in A_j\}$ et $I = \bigcup_{j=0}^{+\infty} I_j$. Alors,

$$\mathbf{P}\left(\bigcup_{j=0}^{+\infty} A_j\right) = \sum_{i\in I} \mathbf{P}(\{\omega_i\}) = \sum_{j=0}^{+\infty}\sum_{i\in I_j} p_i = \sum_{j=0}^{+\infty} \mathbf{P}(A_j).$$

□