

Table des matières

1 Espaces probabilisés	9
1.1 Généralités	9
1.2 Espaces probabilisés finis et dénombrables	16
1.3 Probabilités conditionnelles	19
1.4 Evénements indépendants	23
1.5 Construction d'une probabilité : cas général	26
1.6 Fonction de répartition d'une probabilité sur \mathbb{R}	29
1.7 Exercices du Chapitre 1 .	39
1.8 Indications pour les exercices du Chapitre 1	44
1.9 Corrigés des exercices du Chapitre 1	49
2 Variables aléatoires	67
2.1 Généralités	67
2.2 Variables aléatoires à valeurs réelles .	71
2.3 Variables aléatoires indépendantes .	77
2.4 Exercices du Chapitre 2 .	83
2.5 Indications pour les exercices du Chapitre 2	89
2.6 Corrigés des exercices du Chapitre 2	94
3 Vecteurs aléatoires	121
3.1 Généralités	121
3.2 Espérance et matrice de covariance .	125
4 Fonctions caractéristiques	157
4.1 Cas d'une variable aléatoire	157
4.2 Cas d'un vecteur aléatoire	169

4.3 Vecteurs aléatoires gaussiens	170
4.4 Exercices du Chapitre 4 .	177
4.5 Indications pour les exercices du Chapitre 4	184
4.6 Corrigés des exercices du Chapitre 4	189
5 Convergence de variables aléatoires	211
5.1 Types de convergence	211
5.2 Relations entre convergences	214
5.3 Convergence presque sûre	219
5.4 Convergence en loi	221
5.5 Lois des grands nombres	224
5.6 Théorème Central Limite	227
5.7 Exercices du Chapitre 5	231
5.8 Indications pour les exercices du Chapitre 5	239
5.9 Corrigés des exercices du Chapitre 5	245
6 Annexe : L'intégrale de Lebesgue	267
6.1 Construction de l'intégrale de Lebesgue	267
6.2 Propriétés de l'intégrale de Lebesgue	269
6.3 Convergence monotone et dominée. Lemme de Fatou	270
6.4 Théorème de Fubini	271
7 Annexe : Théorème de Lévy	273
7.1 Familles tendues	273
7.2 Preuve du théorème de Lévy	273
Index	279
Index des symboles	283

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 : $\bigcup_{i \in \mathbb{N}} A_i \in \mathcal{F}$.

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 : $\bigcap_{i \in \mathbb{N}} A_i \in \mathcal{F}$.

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 $\bigcup_{i \in \mathbb{N}} D_i = \Omega$. 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} .

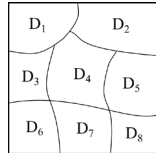


Figure 1. L'espace d'états $\Omega = [0, 1] \times [0, 1]$ ayant une partition de 8 éléments.

- 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 $\mathbf{P} : \mathcal{F} \rightarrow [0, 1]$ 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 côté 1 et comprendre la probabilité d'un sous-ensemble de Ω comme son aire.*

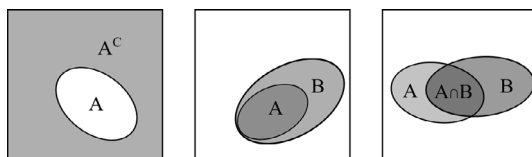


Figure 2. Illustration des propriétés élémentaires de probabilités.

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 $\mathbf{P} : \mathcal{F} \rightarrow [0, 1]$ 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

$$A = \bigcup_{i=0}^{+\infty} A_i,$$

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

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

$$B = \bigcap_{i=0}^{+\infty} B_i,$$

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

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, $\lim_{i \rightarrow +\infty} \mathbf{P}(B_i^c) = \mathbf{P}(B^c)$. 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 $A_{-1} = \emptyset$.

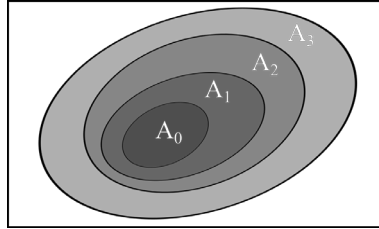


Figure 3. Construction des ensembles C_i .

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

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

De plus, par σ -additivité $\mathbf{P}(A) = \sum_{j=0}^{+\infty} \mathbf{P}(C_j)$. 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 $A_i = \bigcup_{j=0}^i C_j$ croissent vers $A = \bigcup_{j=0}^{+\infty} C_j$ et la propriété a) nous garantit que $\lim_{i \rightarrow +\infty} \mathbf{P}(A_i) = \mathbf{P}(A)$. 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 $i \rightarrow +\infty$, 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

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

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 $\lim_{i \rightarrow +\infty} \mathbf{P}(A \setminus A_i) = 0$. 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 $n \rightarrow +\infty$, 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 $\text{card}(\Omega) < +\infty$. 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 :

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

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,

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

□

Remarque 1.5. Si $\text{card}(\Omega) = N < +\infty$ et $\Omega = \{\omega_0, \omega_1, \omega_2, \dots, \omega_{N-1}\}$, le résultat similaire à la proposition 1.4 a lieu. Pour l'obtenir, il suffit de remplacer $+\infty$ par $N - 1$ dans la condition de la proposition.

Définition 1.5. L'espace fini Ω est dit **équiprobable** si les valeurs p_i de \mathbf{P} sur les singletons ω_i ne dépendent pas de i .

Dans ce cas là, $\mathbf{P}(\{\omega_i\}) = \frac{1}{\text{card}(\Omega)}$, et pour tout $A \in \mathcal{P}(\Omega)$

$$\mathbf{P}(A) = \frac{\text{card}(A)}{\text{card}(\Omega)}$$

Exemples 1.3. Donnons quelques exemples d'espaces équiprobables et non-équiprobables.

- a) L'espace d'états $\Omega = \{p, f\}$, modélisant le lancement au hasard d'une pièce équilibrée, est équiprobable :

$$\mathbf{P}(\{p\}) = 1/2, \quad \mathbf{P}(\{f\}) = 1/2.$$

- b) L'espace d'états $\Omega = \{p, f\}$, modélisant le lancement au hasard d'une pièce non-équilibrée, n'est pas équiprobable :

$$\mathbf{P}(\{p\}) = q, \quad \mathbf{P}(\{f\}) = 1 - q,$$

ici $q \neq 1/2$.

- c) L'espace d'états $\Omega = \{0, 1, 2\}$, modélisant le nombre de « piles » dans le lancement au hasard de deux pièces équilibrées identiques, n'est pas équiprobable :

$$\mathbf{P}(\{0\}) = 1/4, \quad \mathbf{P}(\{1\}) = 1/2, \quad \mathbf{P}(\{2\}) = 1/4.$$

Donnons quelques lois de probabilités classiques sur des espaces finis et dénombrables, ainsi que leurs utilisations en modélisation probabiliste.

Exemple 1.4. Pour modéliser une expérience avec seulement deux résultats possibles, on utilise l'espace d'états Ω à deux éléments, disons 0 et 1, muni de la tribu $\mathcal{P}(\Omega)$ et de la **loi de Bernoulli** $\mathcal{B}(p)$, $0 < p < 1$:

$$\mathbf{P}(\{0\}) = 1 - p, \quad \mathbf{P}(\{1\}) = p.$$

Cette loi est à utiliser dans la modélisation du lancement d'une pièce (avec les résultats possibles « pile » ou « face »), de la naissance d'un nourrisson (avec

les résultats possibles « fille » ou « garçon »), du traitement d'une maladie (avec les résultats possibles « succès » ou « échec »). Il faut remarquer que les résultats de toute expérience peuvent être réduits à deux résultats en faisant des regroupements ; simplement en procédant ainsi, on diminue, en général, l'information contenue dans des données. En effet, l'application, qui régit le regroupement, n'est pas toujours bijective, et ne permet pas, à partir de données regroupées, de reconstituer les données initiales.

Exemple 1.5. Pour modéliser le résultat égal au nombre de « succès » dans n expériences de Bernoulli, effectuées de manière indépendante, on fait intervenir la **loi binomiale** $\mathcal{B}(n, p)$ avec des paramètres $n \in \mathbb{N}^*$ et $0 < p < 1$. La loi binomiale est définie sur l'espace d'états $\Omega = \{0, 1, \dots, n\}$ muni de la tribu $\mathcal{P}(\Omega)$ et vérifie : pour $i \in \{0, 1, \dots, n\}$

$$\mathbf{P}(\{i\}) = \binom{n}{i} p^i (1-p)^{n-i},$$

où $\binom{n}{i} = \frac{n!}{i!(n-i)!}$. Par exemple, si on compte le nombre de « pile » obtenu dans n lancements d'une pièce, faits de manière indépendante, alors la probabilité que le nombre de « pile » soit égal à i est $\binom{n}{i} p^i (1-p)^{n-i}$. Pour plus de détails sur la modélisation voir l'exercice 1.6.

Un autre exemple classique, où la loi binomiale intervient, concerne le tirage de boules. Si d'une urne contenant N_1 boules blanches et N_2 boules noires on tire au hasard n boules avec remise, alors la probabilité d'obtenir i boules blanches parmi n boules tirées est égale à $\binom{n}{i} p^i (1-p)^{n-i}$, où $p = \frac{N_1}{N_1 + N_2}$. Pour plus de détails sur la modélisation voir l'exercice 1.4.

Exemple 1.6. Pour modéliser les résultats de tirages de boules sans remise, on utilise souvent la **loi hypergéométrique** $\mathcal{H}(n, N_1, N_2)$. Ici $n, N_1, N_2 \in \mathbb{N}^*$, $n \leq \min(N_1, N_2)$. La loi hypergéométrique $\mathcal{H}(n, N_1, N_2)$ est définie sur l'espace d'états $\Omega = \{0, 1, \dots, n\}$ muni de la tribu $\mathcal{P}(\Omega)$ et vérifie : pour $i \in \{0, 1, \dots, n\}$

$$\mathbf{P}(\{i\}) = \frac{\binom{N_1}{i} \binom{N_2}{n-i}}{\binom{N_1+N_2}{n}}.$$

Par exemple, si d'une urne contenant N_1 boules blanches et N_2 boules noires on tire n boules sans remise, la probabilité que le nombre de boules blanches obtenu dans n tirages soit égale à i , est $\frac{\binom{N_1}{i} \binom{N_2}{n-i}}{\binom{N_1+N_2}{n}}$. Pour plus de détails voir l'exercice 1.5.

Exemple 1.7. La loi de Poisson $\mathcal{P}(\lambda)$, $\lambda > 0$, apparaît, en général, dans des situations où il s'agit de modéliser le nombre de « succès » dans une suite d'expériences de Bernoulli, faites de manière indépendante. Mais les circonstances sont particulières : la probabilité de « succès » est petite et le nombre d'expériences n est grand, de sorte que le produit np reste stable. La **loi de Poisson** $\mathcal{P}(\lambda)$, $\lambda > 0$, est définie sur un espace d'états $\Omega = \mathbb{N}$ muni de la tribu $\mathcal{P}(\mathbb{N})$ et vérifie : pour $i \in \mathbb{N}$

$$\mathbf{P}(\{i\}) = \frac{\lambda^i \exp(-\lambda)}{i!}.$$

La loi de Poisson intervient comme loi du nombre des événements rares, tels que des naissances, des catastrophes, des accidents...

Exemple 1.8. La loi géométrique $\mathcal{G}(p)$, $0 < p < 1$, apparaît tout naturellement comme la loi du nombre d'expériences de Bernoulli, effectuées de manière indépendante, qui soient nécessaires jusqu'au premier « succès ». La **loi géométrique** $\mathcal{G}(p)$, $0 < p < 1$ est définie sur un espace d'états $\Omega = \mathbb{N}^*$ muni de la tribu $\mathcal{P}(\mathbb{N}^*)$ et vérifie : pour $i \in \mathbb{N}^*$

$$\mathbf{P}(\{i\}) = p(1-p)^{i-1}.$$

Par exemple, si on lance une pièce jusqu'à l'obtention de « pile », alors la probabilité que le nombre de lancements nécessaires soit égal à i est exactement $p(1-p)^{i-1}$. Pour plus de détails voir l'exercice 1.6.

1.3 Probabilités conditionnelles

Les probabilités conditionnelles jouent un rôle important comme outil théorique ainsi qu'un outil dans des applications. Par exemple, pour apprécier l'efficacité de traitement d'une maladie pour différents groupes de malades, différenciés par âge, conditions de vie, maladies intercurrentes, etc. on utilise les probabilités conditionnelles. Donnons la définition de la probabilité conditionnelle.

Soit $(\Omega, \mathcal{F}, \mathbf{P})$ un espace probabilisé. Soit $A \in \mathcal{F}$ tel que $\mathbf{P}(A) \neq 0$. On pose pour tout $B \in \mathcal{F}$

$$\mathbf{P}(B | A) = \frac{\mathbf{P}(B \cap A)}{\mathbf{P}(A)}. \quad (1.1)$$

La probabilité $\mathbf{P}(B | A)$ est appelée probabilité conditionnelle de B sachant A .

Proposition 1.5. *L'application $P(\cdot | A) : \mathcal{F} \rightarrow [0, 1]$ est une probabilité.*

Preuve. Soit $(\Omega, \mathcal{F}, \mathbf{P})$ un espace probabilisé. Soit $A \in \mathcal{F}$ tel que $\mathbf{P}(A) > 0$. La probabilité est une application de \mathcal{F} dans $[0, 1]$ qui vérifie $\mathbf{P}(\Omega) = 1$ et qui est σ -additive. Vérifions ces deux propriétés.

1) Pour tout $B \in \mathcal{F}$

$$\begin{aligned}\mathbf{P}(B | A) &= \frac{\mathbf{P}(B \cap A)}{\mathbf{P}(A)} \geq 0, \\ \mathbf{P}(B | A) &= \frac{\mathbf{P}(B \cap A)}{\mathbf{P}(A)} \leq \frac{\mathbf{P}(A)}{\mathbf{P}(A)} = 1.\end{aligned}$$

2) Pour une suite d'ensembles $(B_k)_{k \in \mathbb{N}^*}$ appartenant à \mathcal{F} deux-à-deux disjoints, nous avons

$$\begin{aligned}\mathbf{P}\left(\bigcup_{k=1}^{+\infty} B_k | A\right) &= \frac{\mathbf{P}\left(\left(\bigcup_{k=1}^{+\infty} B_k\right) \cap A\right)}{\mathbf{P}(A)} \\ &= \frac{\mathbf{P}\left(\bigcup_{k=1}^{+\infty} (B_k \cap A)\right)}{\mathbf{P}(A)} = \frac{\sum_{k=1}^{+\infty} \mathbf{P}(B_k \cap A)}{\mathbf{P}(A)}\end{aligned}$$

car $(B_k \cap A) \cap (B_j \cap A) = (B_k \cap B_j) \cap A = \emptyset$ si $k \neq j$. Donc,

$$\mathbf{P}\left(\bigcup_{k=1}^{+\infty} B_k | A\right) = \sum_{k=1}^{+\infty} \mathbf{P}(B_k | A)$$

□

Nous allons donner quelques formules importantes liées aux probabilités conditionnelles.

On peut réécrire la définition de la probabilité conditionnelle comme suit :

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

C'est un cas particulier de la formule des probabilités conditionnelles en cascade, appelée aussi formule des probabilités composées.

Proposition 1.6. *Soit $n \geq 2$ et soit A_1, A_2, \dots, A_n des ensembles appartenant à \mathcal{F} tels que $\mathbf{P}\left(\bigcap_{k=1}^{n-1} A_k\right) \neq 0$.*

$$\mathbf{P}\left(\bigcap_{k=1}^n A_k\right) = \mathbf{P}(A_1) \mathbf{P}(A_2 | A_1) \mathbf{P}(A_3 | A_1 \cap A_2) \cdots \mathbf{P}\left(A_n | \bigcap_{k=1}^{n-1} A_k\right) \quad (1.2)$$

Preuve. Nous démontrons la formule (1.2) par récurrence.

Pour $n = 2$, par (1.1) on a $P(A_1 \cap A_2) = P(A_1)P(A_2 | A_1)$. Supposons que la formule (1.2) est vraie pour un $n \geq 2$. Alors en prenant $n + 1$ ensembles mesurables A_1, A_2, \dots, A_{n+1} , nous avons

$$\mathbf{P} \left(\bigcap_{k=1}^{n+1} A_k \right) = \mathbf{P} \left(\bigcap_{k=1}^n A_k \cap A_{n+1} \right) = \mathbf{P} \left(A_{n+1} \mid \bigcap_{k=1}^n A_k \right) \mathbf{P} \left(\bigcap_{k=1}^n A_k \right).$$

En appliquant la formule (1.2) pour les ensembles A_1, A_2, \dots, A_n , on obtient la formule (1.2) pour $n + 1$ événements A_1, A_2, \dots, A_{n+1} . □

Exemple 1.9. D'une urne contenant 5 boules noires et 5 boules rouges on tire successivement 3 boules : si on tire une boule noire, on l'enlève, si on tire une boule rouge, on la retire et on rajoute une boule noire à la place. Quelle est la probabilité de tirer 3 boules rouges ?

Écrivons 1 si la boule tirée est rouge et 0 si elle est noire. Alors, notre espace probabilisé est constitué de

$$\Omega = \{\omega = (\omega_1, \omega_2, \omega_3) \mid \omega_i \in \{0, 1\} \text{ pour } i \in \{1, 2, 3\}\},$$

$\mathcal{F} = \mathcal{P}(\Omega)$ et une probabilité \mathbf{P} . La probabilité \mathbf{P} est définie en utilisant la formule des probabilités conditionnelles en cascade, qui sont définies naturellement avant de définir \mathbf{P} . Nous l'expliquons pour l'événement A .

Notons

$$\begin{aligned} A &= \{\omega \in \Omega \mid \omega_1 = \omega_2 = \omega_3 = 1\}, & A_1 &= \{\omega \in \Omega \mid \omega_1 = 1\}, \\ A_2 &= \{\omega \in \Omega \mid \omega_2 = 1\}, & A_3 &= \{\omega \in \Omega \mid \omega_3 = 1\}. \end{aligned}$$

Alors par la formule des probabilités conditionnelles en cascade

$$\mathbf{P}(A) = \mathbf{P}(A_1 \cap A_2 \cap A_3) = \mathbf{P}(A_1)\mathbf{P}(A_2 \mid A_1)\mathbf{P}(A_3 \mid A_1 \cap A_2).$$

Il est naturel de définir $\mathbf{P}(A_1)$, $\mathbf{P}(A_2 \mid A_1)$ et $\mathbf{P}(A_3 \mid A_1 \cap A_2)$ par des modèles avec probabilités équiprobables.

Puisque l'urne contient 5 boules noires et 5 boules rouges, $\mathbf{P}(A_1) = \frac{1}{2}$. Si la boule rouge est tirée au premier tirage, on a 6 boules noires et 4 boules rouges, d'où $\mathbf{P}(A_2 \mid A_1) = \frac{4}{10} = \frac{2}{5}$. Si la boule rouge est tirée au second tirage, on a 7 boules noires et 3 boules rouges, et $\mathbf{P}(A_3 \mid A_1 \cap A_2) = \frac{3}{10}$. Donc,

$$\mathbf{P}(A) = \frac{1}{2} \times \frac{2}{5} \times \frac{3}{10} = \frac{3}{50}.$$

Donnons maintenant la formule des probabilités totales.

Proposition 1.7. *Soit $(A_k)_{k \in \mathbb{N}}$ une partition mesurable de Ω en événements de probabilités $\mathbf{P}(A_k) > 0$ pour tout $k \in \mathbb{N}$. Alors pour tout événement $A \in \mathcal{F}$*

$$\mathbf{P}(A) = \sum_{k=0}^{\infty} \mathbf{P}(A | A_k) \mathbf{P}(A_k)$$

Preuve. Soit $(\Omega, \mathcal{F}, \mathbf{P})$ un espace probabilisé. Soit $(A_k)_{k \in \mathbb{N}}$ une partition mesurable de Ω , c'est-à-dire une suite d'événements deux-à-deux disjoints tels que $\bigcup_{k=0}^{+\infty} A_k = \Omega$. On suppose que $\mathbf{P}(A_k) > 0$ pour tout $k \in \mathbb{N}$.

Pour tout événement $A \in \mathcal{F}$

$$\mathbf{P}(A) = \mathbf{P}(A \cap \Omega) = \mathbf{P}\left(A \cap \left(\bigcup_{k=0}^{+\infty} A_k\right)\right) = \mathbf{P}\left(\bigcup_{k=0}^{+\infty} (A \cap A_k)\right)$$

Puisque les ensembles $(A \cap A_k)_{k \in \mathbb{N}}$ sont deux-à-deux disjoints, par σ -additivité on a :

$$\mathbf{P}\left(\bigcup_{k=0}^{+\infty} (A \cap A_k)\right) = \sum_{k=0}^{+\infty} \mathbf{P}(A \cap A_k) = \sum_{k=0}^{+\infty} \mathbf{P}(A | A_k) \mathbf{P}(A_k)$$

□

Présentons maintenant la formule de Bayes, très utile dans les calculs de probabilités conditionnelles.

Proposition 1.8. *1. Soient deux événements A, B tels que $\mathbf{P}(A) \neq 0$ et $\mathbf{P}(B) \neq 0$. Alors*

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

2. Soit $(A_k)_{k \in \mathbb{N}}$ une partition mesurable de Ω . Pour tout événement A de probabilité $\mathbf{P}(A) \neq 0$, et pour tout $k \in \mathbb{N}$

$$\mathbf{P}(A_k | A) = \frac{\mathbf{P}(A | A_k) \mathbf{P}(A_k)}{\mathbf{P}(A)} = \frac{\mathbf{P}(A | A_k) \mathbf{P}(A_k)}{\sum_{i=0}^{\infty} \mathbf{P}(A | A_i) \mathbf{P}(A_i)}.$$

Preuve. La partie 1 découle immédiatement de la définition (1.1) appliquée à $\mathbf{P}(B | A)$ et à $\mathbf{P}(A | B)$. La partie 2 découle de la partie 1 avec $B = A_k$ et de la formule des probabilités totales. □