Chapitre I : Espaces probabilisés

Les probabilités sont une branche des Mathématiques dont l'objet est l'étude des phénomènes aléatoires. Historiquement, il s'agissait essentiellement des jeux de hasard et des problèmes d'espérance de vie pour des calculs de rente.
L'approche mathématique de ces problèmes n'apparaît réellement qu'au **XVII**ᵉ siècle avec Pascal et Fermat, puis au siècle suivant avec Bayes et Laplace.
L'élaboration d'un cadre mathématique rigoureux est très récente et elle est due à Kolmogorov, qui a axiomatisé le calcul des probabilités (fondements du calcul des probabilités, 1933) et a permis en particulier l'utilisation de la théorie de la mesure.

I.1 - Expérience aléatoire et univers

Définition 1.1 *On appelle expérience aléatoire une expérience sur un système dont le résultat n'est pas connu d'avance et peut varier si on répète cette expérience.*

Exemple 1.1 *Jeter une pièce de monnaie, lancer des dés, prélever des boules dans une urne, lancer une fléchette en direction d'une cible, observer la position d'une particule dans un liquide...*

Le résultat, par hypothèse unique, de la réalisation de l'expérience aléatoire est noté ω.

Définition 1.2 *On appelle univers l'ensemble des résultats possibles. Il est noté* Ω.

Exemple 1.2 *On effectue deux jets successifs d'un dé :* $\Omega = [\![1;6]\!]^2$.

Exemple 1.3 *On lance indéfiniment une pièce de monnaie : un résultat possible est alors une suite de l'ensemble* $\{F;P\}$*, et donc* $\Omega = \{F;P\}^{\mathbb{N}}$.

La difficulté vient du fait qu'il est possible, pour une même expérience aléatoire, de définir plusieurs univers, suivant ce que l'on entend par le terme « résultat possible ».

Par exemple, pour une expérience aléatoire consistant à prélever une boule dans une urne contenant 2 boules rouges et 3 boules noires, on peut considérer qu'un résultat possible est une couleur ($\Omega = \{N; R\}$) ou l'une des 5 boules ($\Omega = \{R_1; R_2; N_1; N_2; N_3\}$).
De même, pour le lancer d'une fléchette contre une cible, on peut considérer comme résultat possible le point d'impact ($\Omega = \mathbb{R}^2$, après avoir muni le plan d'un repère), ou la trajectoire suivie par la fléchette ($\Omega = \mathcal{C}([0;1];\mathbb{R}^3)$, ensemble des applications continues de $[0;1]$ dans $\mathbb{R}^3$).

Remarque : si on répète la même expérience aléatoire d'univers Ω, on pourra choisir comme univers Ω^n dans le cas de n répétitions, et $\Omega^{\mathbb{N}}$ si on la répète indéfiniment.

I.2 - Événements

C'est une propriété $\mathcal{E}$ énonçable (c'est-à-dire accessible à l'expérience), vérifiée ou non selon le résultat obtenu. On l'identifie à l'ensemble des résultats de l'expérience pour lesquels elle est vérifiée :

$$\mathcal{E} \text{ est identifiée à } \{\omega \in \Omega / \mathcal{E} \text{ est vraie lorsque } \omega \text{ est réalisé}\}.$$

En pratique, on ne s'intéresse souvent qu'à un sous-ensemble d'événements. C'est le cas par exemple lorsque $\Omega = \mathbb{R}$: on considère les parties boréliennes, et pas toutes les parties de $\mathbb{R}$ (Voir annexe 1).

Notation : on note $\mathcal{A}$ l'ensemble des événements relatifs à l'expérience aléatoire. On a donc par définition $\mathcal{A} \subset \mathcal{P}(\Omega)$.

On souhaite pouvoir définir certaines opérations sur les événements, et on utilise la correspondance suivante entre vocabulaires probabiliste et ensembliste :

Terminologie probabiliste	Terminologie ensembliste	Notation
Événement certain	Ensemble entier	Ω
Événement impossible	Ensemble vide	$\emptyset$
Événement élémentaire	Singleton	$\{\omega\}$
Événement contraire de A	Complémentaire de A	$\overline{A}$
A ou B	Réunion de A et B	$A \cup B$
A et B	Intersection de A et B	$A \cap B$
A implique B	A inclus dans B	$A \subset B$
A et B incompatibles	A et B disjoints	$A \cap B = \emptyset$
ω réalise A	ω appartient à A	$\omega \in A$

On exige que $\mathcal{A}$ contienne l'ensemble vide et l'univers, et qu'il soit stable *pour les opérations de réunion et d'intersection finies ou dénombrables, ainsi que par passage au complémentaire :*

i $\Omega \in \mathcal{A}$ et $\emptyset \in \mathcal{A}$

ii $A \in \mathcal{A} \Rightarrow \overline{A} \in \mathcal{A}$

iii $(\forall i \in I \subset \mathbb{N}, A_i \in \mathcal{A}) \Rightarrow \bigcup_{i \in I} A_i \in \mathcal{A}$

iv $(\forall i \in I \subset \mathbb{N}, A_i \in \mathcal{A}) \Rightarrow \bigcap_{i \in I} A_i \in \mathcal{A}$

c'est-à-dire que $\mathcal{A}$ soit une tribu de parties de Ω.

Remarque : ii et iii impliquent iv (utiliser les lois de De Morgan).
Si l'on s'intéresse à un ensemble d'événements $\mathcal{F}$ qui ne forme pas une tribu, on posera $\mathcal{A} = \mathcal{T}(\mathcal{F})$, tribu engendrée par $\mathcal{F}$ (c'est la plus petite tribu, au sens de l'inclusion, qui contient $\mathcal{F}$).

Exemple 1.4

1. *Si Ω est fini ou dénombrable, on s'intéresse naturellement aux événements élémentaires $\{\omega_i\}_{i \in I}$. La tribu engendrée est alors $\mathcal{P}(\Omega)$.*
2. *Si $\mathcal{F}$ est constitué d'un nombre fini ou dénombrable d'événements $(A_i)_{i \in I}$, qui forment une partition de Ω, la tribu engendrée est exactement l'ensemble des réunions quelconques d'événements A_i. Par exemple, en considérant une partition $\mathcal{F} = \{A, \overline{A}\}$ de Ω, on a $\mathcal{A} = \{\emptyset, \Omega, A, \overline{A}\}$.*
3. *Si $\Omega = \mathbb{R}$ ou $\Omega = \mathbb{R}^n$, on considère souvent comme tribu $\mathcal{A}$ la tribu des boréliens (engendrée par les ouverts, ou bien par les pavés $\prod_{i=1}^{n} [a_i; b_i]$, avec pour tout $i \in [\![1; n]\!], a_i \leq b_i$, ou bien, dans le cas de $\mathbb{R}$, par les ouverts $]-\infty; a[, a \in \mathbb{R}$). On ne peut pas prendre $\mathcal{P}(\mathbb{R})$ (voir annexe 1).*

Définition 1.3 *On appelle événement tout élément de la tribu $\mathcal{A}$.*

Définition 1.4 *On appelle événement élémentaire tout singleton (partie constituée d'un seul élément) de la tribu $\mathcal{A}$.*

Exemple 1.5 *On lance deux fois un dé à six faces. On pose $\Omega = [\![1; 6]\!]^2$ et $\mathcal{A} = \mathcal{P}(\Omega)$.*
On considère A : « la somme obtenue est supérieure ou égale à 11 ».
On a : $A = \{(5; 6); (6; 5); (6; 6)\}$ et A est un événement ($A \in \mathcal{A}$).
On considère B : « la somme obtenue est divisible par 3 et par 4 ».
On a : $B = \{(6; 6)\}$ et B est un événement élémentaire.

Attention : ne pas confondre résultat possible ($\omega \in \Omega$) et événement élémentaire ($\{\omega\} \in \mathcal{P}(\Omega)$). En particulier, on verra, lorsqu'on aura défini une probabilité P, que l'écriture $P(\omega)$ n'a aucun sens, et on prendra bien garde à écrire $P(\{\omega\})$.

Définition 1.5 *Le couple* $(\Omega; \mathcal{A})$*, où* Ω *est l'univers et* $\mathcal{A}$ *une tribu de parties de* Ω*, est appelé espace probabilisable.*

I.3 - Probabilité

Il s'agit d'affecter à chaque événement A un poids $P(A)$ indiquant sa « chance » d'être réalisé si l'on effectue l'expérience aléatoire.

Des considérations relatives à l'expérience peuvent conduire, dans le cas où Ω est fini, à affecter des poids de probabilité identiques à chaque événement élémentaire : ce sera l'hypothèse d'équiprobabilité, par exemple lorsqu'on jette un dé équilibré, que l'on tire des cartes dans un jeu bien battu, que l'on prélève des boules indiscernables...

Une autre approche est l'approche « fréquentiste » : on se donne un événement fixé A, on répète n fois l'expérience aléatoire et on note n_A le nombre de fois où l'événement A a été réalisé. La fréquence de réalisation de A au cours de ces n répétitions est donc : $\dfrac{n_A}{n}$. On démontrera que, lorsque n tend vers l'infini, $\dfrac{n_A}{n}$ fluctue de moins en moins, autour d'une valeur limite f_A (c'est la loi faible des grands nombres, voir chapitre VIII), appelée fréquence de réalisation de A.
La limite principale de cette approche est qu'elle ne concerne que des expériences aléatoires aisément répétables, et dans des conditions identiques.
Cependant, l'application f possède les propriétés élémentaires suivantes :

1. $\forall A \in \mathcal{A}, f_A \geq 0$ (positivité)
2. $f_\Omega = 1$ (totalité)
3. $\forall (A; B) \in \mathcal{A}^2, A \cap B = \emptyset \Rightarrow f_{A \cup B} = f_A + f_B$ (additivité)

Par analogie avec ces trois propriétés de la fréquence, on définit ce qu'est une probabilité de manière purement axiomatique (présentation due à Kolmogorov), sans référence à une quelconque observation de la réalisation de l'expérience aléatoire :

Définition 1.6 *Étant donné un espace probabilisable* $(\Omega; \mathcal{A})$*, on appelle probabilité sur* $(\Omega; \mathcal{A})$ *toute application* $P : \mathcal{A} \to \mathbb{R}$ *satisfaisant aux trois axiomes suivants :*

1. *$\forall A \in \mathcal{A}, P(A) \geq 0$ (positivité)*
2. *$P(\Omega) = 1$ (totalité)*
3. *Pour toute suite (A_n) d'éléments deux à deux disjoints,*
$P\left(\bigcup_n A_n\right) = \sum_{n=0}^{+\infty} P(A_n)$ *(additivité)*

Remarque : avec le vocabulaire de la théorie de la mesure, P est donc une mesure positive finie, de masse totale égale à 1.

Définition 1.7 *Soit $A \in \mathcal{A}$.*
Si $P(A) = 0$, on dit que A est un événement presque impossible.
Si $P(A) = 1$, on dit que A est un événement presque certain.

Remarque : on se place dans le cas d'un univers fini, $\Omega = \{\omega_1; \cdots; \omega_n\}$, et on s'intéresse aux événements élémentaires en posant comme tribu $\mathcal{A} = \mathcal{P}(\Omega)$. La donnée de n nombres réels positifs $p_1, \cdots, p_n$ de somme égale à 1 permet de définir une probabilité P sur $\mathcal{A}$ en posant :

i $\forall i \in [\![1; n]\!], P(\{\omega_i\}) = p_i$

ii $\forall A \in \mathcal{P}(\Omega), P(A) = \sum_{i/\omega_i \in A} p_i$ (puisque P doit être additive)

Prenons le cas du jet d'un dé, où l'on pose $\Omega = [\![1; 6]\!]$.
La donnée de

$$p_1 = p_2 = p_3 = p_4 = p_5 = \frac{1}{10} \text{ et } p_6 = \frac{5}{10}$$

permet de définir une probabilité P comme décrit ci-dessus.
Si on note A l'événement « le résultat est pair », alors :

$$P(A) = \sum_{i/\omega_i \in A} p_i = p_2 + p_4 + p_6 = \frac{7}{10}$$

On procède de la même façon lorsque Ω est dénombrable, en se donnant une suite de réels positifs $(p_i)_{i \in \mathbb{N}}$ tels que la série $\sum_{i \geq 0} p_i$ converge vers 1.

Cas particulier important : **le cas d'équiprobabilité (Ω fini, de cardinal n).**

Des considérations relatives à l'expérience (dé équilibré, boules indiscernables, ...) peuvent conduire à définir une probabilité P sur $\mathcal{P}(\Omega)$ telle que :

$$P(\{\omega_1\}) = P(\{\omega_2\}) = \cdots = P(\{\omega_n\})$$

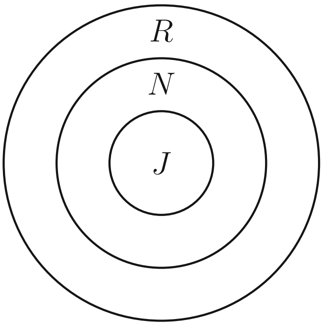
On a alors : $\forall i \in [\![1;n]\!], P(\{\omega_i\}) = \dfrac{1}{n}$, et P est entièrement déterminée, puisqu'elle est connue sur les n événements élémentaires $\{\omega_i\}$. En effet :

$$\begin{aligned}
\forall A \in \mathcal{P}(\Omega), P(A) &= P\left(\bigcup_{\omega \in A} \{\omega\}\right) \\
&= \sum_{\omega \in A} P(\{\omega\}) \text{ (additivité de } P\text{)} \\
&= \sum_{\omega \in A} \frac{1}{n} \\
&= \frac{\text{Card}(A)}{n}
\end{aligned}$$

Définition 1.8 *L'application P définie ci-dessus est appelée la probabilité uniforme.*

Dans ce cas, les calculs de probabilité se ramènent à des problèmes de dénombrement, puisque $P(A) = \dfrac{\text{Card}(A)}{\text{Card}(\Omega)}$.

Exemple de probabilité non uniforme : on lance une fléchette contre la cible suivante :



On note la couleur obtenue : on pose donc $\Omega = \{J; N; R\}$.
Si la probabilité de chacun des trois événements élémentaires est proportionnelle à l'aire de chacune des trois parties colorées, on a :

$$P(\{J\}) = \frac{1}{9}; P(\{N\}) = \frac{3}{9}; P(\{R\}) = \frac{5}{9}$$

L'exemple 1.6 constitue aussi un cas où P n'est pas uniforme.

Propriété 1.1 *Toute probabilité P possède les propriétés suivantes :*

1. $P(\emptyset) = 0$

2. *$\forall (A;B) \in \mathcal{A}^2, A \subset B \Rightarrow P(A) \leq P(B)$*
3. *$\forall A \in \mathcal{A}, P(A) \in [0;1]$ et $P(A) = 1 - P(\overline{A})$*
4. *Formule du crible de Poincaré : $\forall (A_1; \cdots; A_n) \in \mathcal{A}^n$,*

$$\begin{aligned}
P\left(\bigcup_{i=1}^{n} A_i\right) &= \sum_{1\leq i\leq n} P(A_i) - \sum_{1\leq i_1<i_2\leq n} P(A_{i_1} \cap A_{i_2}) \\
&\quad + \cdots \\
&\quad +(-1)^{k-1} \sum_{1\leq i_1<i_2<\cdots<i_k\leq n} P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}) \\
&\quad + \cdots \\
&\quad +(-1)^{n-1} P(A_1 \cap \cdots \cap A_n)
\end{aligned}$$

5. *Si $(A_i)_{i\in\mathbb{N}}$ est une suite croissante d'événements (au sens de l'inclusion), alors la suite de réels $(P(A_i))_{i\in\mathbb{N}}$ est croissante, et :*

$$\lim_{n\to+\infty} P(A_n) = P\left(\bigcup_{i=0}^{+\infty} A_i\right)$$

6. *Si $(A_i)_{i\in\mathbb{N}}$ est une suite décroissante d'événements (au sens de l'inclusion), alors la suite de réels $(P(A_i))_{i\in\mathbb{N}}$ est décroissante, et :*

$$\lim_{n\to+\infty} P(A_n) = P\left(\bigcap_{i=0}^{+\infty} A_i\right)$$

Remarque : pour $n = 2$, la formule de Poincaré s'écrit :

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$$

Pour $n = 3$, elle s'écrit :

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

Démonstration des propriétés de P :

1. On utilise l'additivité de P :

$$\begin{aligned}
P(\Omega \cup \emptyset) &= P(\Omega) + P(\emptyset) \\
\Rightarrow P(\Omega) &= P(\Omega) + P(\emptyset) \\
\Rightarrow P(\emptyset) &= 0
\end{aligned}$$

2. On décompose B : $A \subset B \Rightarrow B = (B \cap \overline{A}) \cup A$ (union disjointe), d'où :

$$\begin{aligned} P(B) &= P(B \cap \overline{A}) + P(A) \\ \Rightarrow P(B) - P(A) &= P(B \cap \overline{A}) \\ \Rightarrow P(B) - P(A) &\geq 0 \text{ (par positivité de } P) \end{aligned}$$

3. D'après la propriété précédente :

$$\begin{aligned} A \subset \Omega &\Rightarrow P(A) \leq P(\Omega) \\ &\Rightarrow P(A) \leq 1 \end{aligned}$$

et d'autre part :

$$A \cup \overline{A} = \Omega \Rightarrow P(A) + P(\overline{A}) = 1$$

4. On démontre la formule par récurrence sur n (voir l'exercice 3.1 pour une autre démonstration, utilisant les propriétés de l'espérance).
Soit P_n la propriété : $\forall (A_1; \cdots; A_n) \in \mathcal{A}^n$,

$$\begin{aligned} P\left(\bigcup_{i=1}^{n} A_i\right) = \ & \sum_{1 \leq i \leq n} P(A_i) - \sum_{1 \leq i_1 < i_2 \leq n} P(A_{i_1} \cap A_{i_2}) \\ & + \cdots \\ & + (-1)^{k-1} \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq n} P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}) \\ & + \cdots \\ & + (-1)^{n-1} P(A_1 \cap \cdots \cap A_n) \end{aligned}$$

Pour $n = 1$: il est clair que P_1 est vraie.
Avant de montrer que : $\forall n \in \mathbb{N}, P_n \Rightarrow P_{n+1}$, on va démontrer que P_2 est vraie car la méthode sera identique pour montrer l'hérédité :

$$\begin{aligned} A_1 \cup A_2 &= A_1 \cup (A_2 \cap \overline{A}_1) \text{ (union disjointe)} \\ \Rightarrow P(A_1 \cup A_2) &= P(A_1) + P(A_2 \cap \overline{A}_1) \\ \text{Or : } A_2 &= (A_2 \cap A_1) \cup (A_2 \cap \overline{A}_1) \text{ (union disjointe)} \\ \Rightarrow P(A_2) &= P(A_2 \cap A_1) + P(A_2 \cap \overline{A}_1) \\ \text{d'où : } P(A_1 \cup A_2) &= P(A_1) + P(A_2) - P(A_1 \cap A_2) \end{aligned}$$

On démontre à présent que : $\forall n \in \mathbb{N}, P_n \Rightarrow P_{n+1}$.

$$\begin{aligned} \bigcup_{i=1}^{n+1} A_i &= \left(\bigcup_{i=1}^{n} A_i\right) \cup \left(A_{n+1} \cap \overline{\bigcup_{i=1}^{n} A_i}\right) \text{ (union disjointe)} \\ \Rightarrow P\left(\bigcup_{i=1}^{n+1} A_i\right) &= P\left(\bigcup_{i=1}^{n} A_i\right) + P\left(A_{n+1} \cap \overline{\bigcup_{i=1}^{n} A_i}\right) \end{aligned}$$