Chapitre 1

Les probabilités, mesure du hasard

Avant d'introduire, dans le chapitre suivant, la notion de mesure et d'espace mesurable, nous allons dans ce premier chapitre revoir les définitions et propriétés de base des probabilités, une probabilité n'étant en réalité qu'un cas particulier de mesure.

1.1 L'espace fondamental des probabilités

1.1.1 Espaces probabilisables

Éventualités

Dans une expérience ou une situation dont l'issue, pour le caractère auquel nous nous intéressons, semble indéterminée à l'avance, donc gouvernée par le hasard, on considère l'ensemble, noté Ω, de toutes les éventualités possibles. Cet ensemble se nomme univers des possibles.

Exemple 1 :
On lance un dé et on lit le chiffre lu sur la face supérieure. Dans ce cas, l'univers des possibles est :

$$\Omega = \{1,2,3,4,5,6\} \ .$$

On lance deux dés discernables, le résultat est un couple $\omega = (i,j)$, et dans ce cas :

$$\Omega = \{1,2,3,4,5,6\}^2 \ .$$

Si l'on s'intéresse à la somme des chiffres lus, nous obtenons :

$$\Omega = \{2,3,\ldots,12\} .$$

Exemple 2 :
On souhaite compter le nombre de pannes d'un appareil, pendant une période donnée. On ne sait pas borner, *a priori*, ce nombre. L'univers des possibles est donc :

$$\Omega = \mathbb{N}.$$

Exemple 3 :
On souhaite mesurer la durée de bon fonctionnement d'un appareil. Dans ce cas, nous obtenons :

$$\Omega = \mathbb{R}_+.$$

Exemple 4 :
On souhaite regarder le tracé d'un électro-encéphalogramme sur une période de temps donnée, entre l'instant $t = 0$ et l'instant $t = T$. Cette fois nous obtenons :

$$\Omega = \mathcal{C}[0,T] ,$$

où $\mathcal{C}[0,T]$ désigne l'ensemble des fonctions continues sur l'intervalle $[0,T]$.

Exemple 5 :
Répétons un jeu de pile ou face indéfiniment.
Pour cette expérience nous obtenons :

$$\Omega = \{0;1\}^{\mathbb{N}} .$$

À la i-ème expérience, nous avons :

$$\Omega_i = \{0;1\} .$$

Remarquons que, dans tous les cas, Ω résulte d'un choix délibéré, en se focalisant sur l'un des aspect d'une expérience aléatoire. Il constitue donc un compromis entre la finesse de description du phénomène observé et la lourdeur mathématique du modèle.

Événements

Un événement est un ensemble d'éventualités. On peut donc le représenter par un sous-ensemble A de Ω.
On dit que A se réalise dans l'éventualité ω si $\omega \in A$.

Si $A \subset B$, on dit que A implique B, car dans ce cas, $\omega \in A \Rightarrow \omega \in B$.
Exemple : Dans le cas du lancer d'un dé, l'ensemble $\{2,4,6\}$ est un ensemble d'éventualités qui peuvent être caractérisées par la propriété : le chiffre lu est pair.
Le plus souvent les événements qui nous intéressent sont ainsi caractérisés par une propriété que vérifient certaines éventualités.

Conséquences :

- Ω est un événement qui se réalise dans toutes les éventualités.
- $\emptyset$ est un événement qui ne se réalise jamais.
- A^c, noté parfois en probabilité $\overline{A}$, est l'événement contraire de A.
- $A \cap B$ représente la conjonction des événements A et B.
  Si $A \cap B = \emptyset$, A et B sont dits **incompatibles**.
- $A \cup B$ se réalise quand l'un ou l'autre, au moins, des événements A ou B se réalise.
- Dans le jeu de pile ou face, si A_i est un événement de Ω_i,

  $\bigcup_{i \in \mathbb{N}} A_i$ est un événement qui fait sens, intuitivement parlant.

Notations

Dans toute la suite de ce traité, $\mathcal{P}(\Omega)$ désigne l'ensemble des parties de Ω. Si Ω est fini ou dénombrable, $\mathcal{P}(\Omega)$ représente tous les événements possibles.
Lorsque Ω n'est, ni fini, ni dénombrable, $\mathcal{P}(\Omega)$ est une famille beaucoup trop riche, au sens où elle contient des ensembles qui représentent des événements auxquels l'être humain ne peut s'intéresser, car celui-ci ne dispose pas des moyens pour les observer ou même les caractériser. On se restreint donc souvent à une famille $\mathcal{A} \subset \mathcal{P}(\Omega)$, qui doit vérifier certaines propriétés de stabilité.

Définition 1 *On appelle* **tribu** *d'événements ou $\sigma-$algèbre, sur Ω, toute partie $\mathcal{A} \subset \mathcal{P}(\Omega)$ telle que :*

1. $\Omega \in \mathcal{A}$.
2. $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$.
3. *si $(A_n)_{n \geq 1}$ est une suite d'éléments de $\mathcal{A}$ alors* $\bigcap_{n \geq 1} A_n \in \mathcal{A}$.

Le couple $(\Omega,\mathcal{A})$ s'appelle **espace probabilisable** ou plus généralement **espace mesurable** .

L'importance de la notion de tribu est due au fait que les tribus permettent de définir rigoureusement la notion d'ensemble mesurable. Le terme lui même a été introduit en 1932 par un mathématicien français du groupe Bourbaki.

Progressivement formalisées pendant le premier tiers du XX^e siècle, les tribus constituent le cadre dans lequel s'est développée la théorie de la mesure. Les exemples les plus connus sont les tribus sur $\mathbb{R}$. Les plus utilisées, et nous y reviendrons par la suite, sont les tribus boréliennes, du nom d'Émile Borel qui construisit la tribu borélienne de la droite réelle en 1898, et la tribu de Lebesgue, formée des ensembles mesurables définis par Henri Lebesgue en 1901. En conséquence, les tribus sont fondamentales en théorie des probabilités, car celle-ci s'appuie sur la théorie de la mesure. Dans ce domaine des probabilités, les tribus sont un outil puissant car elles sont à la base de la définition de concepts parmi les plus importants : espérance conditionnelle, martingales, etc. Cette notion n'a cependant réellement d'intérêt que si l'univers des possibles Ω n'est pas de cardinal fini, ce qui est le cas dans la majorité des phénomènes aleatoires complexes que l'être humain souhaite étudier.

Exemples de tribus

1. La plus petite : $\{\emptyset,\Omega\}$ appelée tribu triviale.
2. La plus grande : $\mathcal{P}(\Omega)$.
3. La plus petite qui contienne un événement A donné : $\{\emptyset, A, A^c, \Omega\}$.

Propriétés des tribus

Nous énumérons ici quelques propriétés fondamentales des tribus qui découlent immédiatement de la définition. Soit $\mathcal{A}$ une tribu, nous avons :

$$\emptyset \in \mathcal{A} \text{ car } \emptyset = \Omega^c \, .$$

$$A,B \in \mathcal{A} \Rightarrow A \setminus B = A \cap B^c \in \mathcal{A} \, .$$

$$A,B \in \mathcal{A} \Rightarrow A \triangle B \in \mathcal{A} \, .$$

$$(A_n)_{n\geq 1} \subset \mathcal{A} \Rightarrow \bigcup_{n\geq 1} A_n \in \mathcal{A} \text{ car } \left(\bigcap_{n\geq 1} A_n^c\right)^c = \bigcup_{n\geq 1} A_n \, .$$

Il y a donc équivalence, compte tenu de la stabilité par complémentation, entre la stabilité par union dénombrable et la stabilité par intersection dénombrable.
Enfin,

$$\liminf_n A_n = A_* = \bigcup_{n\geq 1} \bigcap_{k\geq n} A_k \in \mathcal{A} .$$

$$\limsup_n A_n = A^* = \bigcap_{n\geq 1} \bigcup_{k\geq n} A_k \in \mathcal{A} .$$

L'interprétation de ces ensembles sera vue à l'exercice 6, page 16.

Remarque : Le plus souvent, on impose que $\mathcal{A}$ contienne une famille donnée d'événements. En effet, dans la mesure où l'ensemble $\mathcal{P}(\Omega)$ ne peut être écrit en extension, l'idée est d'obtenir une famille $\mathcal{A}$, plus accessible que $\mathcal{P}(\Omega)$, et qui pourra être caractérisée par une propriété, une représentation, celles-ci étant accessibles au cerveau humain. Ceci nous conduit naturellement à la définition suivante.

Définition 2 *Soit $\mathcal{F}$ une famille non vide de parties de Ω. On appelle* **tribu engendrée** *par $\mathcal{F}$ la plus petite tribu $\sigma(\mathcal{F})$, contenant $\mathcal{F}$.*

Comme l'intersection de deux tribus est une tribu, on montrera à la proposition 6 que $\sigma(\mathcal{F})$ est l'intersection de toutes les tribus contenant $\mathcal{F}$.

Remarque : La réunion de deux tribus n'est, en général, pas une tribu.

Dans le cadre de ce livre, les tribus d'intérêt sont les tribus sur $\mathbb{R}$ ou $\mathbb{R}^k$.

Définition 3 *On appelle* **tribu borélienne** *de $\mathbb{R}$ (E. Borel 1871-1956) et on note $\mathcal{B}_{\mathbb{R}}$ ou $\mathcal{B}(\mathbb{R})$ la tribu engendrée par la famille des intervalles*

$$\mathcal{B}_{\mathbb{R}} = \sigma\left(\{\ [a,b],[a,b[,]a,b],]a,b[\ ;\ a \in \mathbb{R}\,,\, b \in \mathbb{R}\ \}\right)\ ,$$

et **tribu borélienne sur** $\mathbb{R}^k$ *celle engendrée par les produits d'intervalles*

$$\mathcal{B}_{\mathbb{R}^k} = \sigma\left(\{\ \prod_{i=1}^{k}]a_i,b_i]\ ;\ a_i \in \mathbb{R}\,,\, b_i \in \mathbb{R}\ \}\right)\ .$$

Ce sont les tribus usuelles sur $\mathbb{R}$ ou sur $\mathbb{R}^k$.
Elles seront revues ultérieurement à la définition 7 du chapitre suivant, puis complétées.
En pratique, lorsque l'univers des possibles Ω est fini ou dénombrable on considère $\mathcal{P}(\Omega)$ comme tribu usuelle, et, dans ce cas, la notion de tribu n'a pas lieu d'être vue.

1.1.2 Probabilités, espaces probabilisés

Le calcul des probabilités est parfois connu du lecteur, mais la théorie des probabilités ne peut être construite rigoureusement sans la théorie de la mesure. Une probabilité est en effet un cas particulier de mesure, et la théorie de la mesure va, de fait, généraliser ce qui a pu être vu à propos des probabilités. Il semble donc utile de rappeler ces notions.

Définition d'une probabilité, notée $\mathbb{P}$

Définition 4 *Soit Ω l'univers des possibles et $\mathcal{A}$ une tribu sur Ω. On appelle* **probabilité sur** $(\Omega,\mathcal{A})$*, toute application* $\mathbb{P} \ : \ \mathcal{A} \longrightarrow [0,1]$ *, vérifiant :*

1. $\mathbb{P}(\Omega) = 1$.
2. *si* $(A_n)_{n\geq 1}$ *est une suite finie ou dénombrable d'éléments de* $\mathcal{A}$*, et si* $\forall\, i \neq j$ *on a* $A_i \cap A_j = \emptyset$ *alors :*

$$\mathbb{P}\left(\bigcup_{n=1}^{+\infty} A_n\right) = \sum_{n=1}^{+\infty} \mathbb{P}(A_n) \ .$$

Le triplet $(\Omega,\mathcal{A},\mathbb{P})$ *s'appelle* **espace probabilisé** *ou espace fondamental.*

Deux propriétés seulement caractérisent donc une probabilité, ce qui est très peu. La première est un choix arbitraire et correspond à une normalisation qui consiste à mesurer toute probabilité entre 0 et 1. La deuxième correspond à une propriété de cohérence, que l'esprit humain applique bien évidemment aux notions de dénombrement, de longueur, de surface ou de volume. Cette deuxième propriété, qui fait intervenir la notion d'union dénombrable correspond donc à une logique intuitive.

Exemple des éventualités équiprobables

L'exemple des éventualités équiprobables, souvent évoqué dans les classes d'enseignement du secondaire, correspond à une expérience aléatoire relativement simple où l'ensemble des éventualités possibles est de cardinal fini, et où toutes les éventualités sont équiprobables.
Nous avons alors :

$$\Omega \text{ fini}, \mathcal{A} = \mathcal{P}(\Omega)\,,\ \text{et } \forall\, \omega, \omega' \in \Omega\,, \mathbb{P}(\{\omega\}) = \mathbb{P}(\{\omega'\})\ .$$

Par définition, nous avons :

$$\Omega = \bigcup_{\omega \in \Omega} \{\omega\}, \text{ et } \mathbb{P}(\Omega) = 1\ .$$

En appliquant la deuxième propriété d'une probabilité, nous obtenons :

$$\mathbb{P}(\{\omega\}) = \frac{1}{Card(\Omega)} \quad \text{et} \quad \forall\, A \subset \Omega\ ,\ \mathbb{P}(A) = \frac{Card(A)}{Card(\Omega)}\ .$$

Cette situation correspond, par exemple, au lancer d'un dé non pipé à six faces. Il peut être intéressant d'apprendre à programmer, donc à simuler ce type d'expérience.

Espaces probabilisés discrets

Plus généralement, si Ω est fini ou dénombrable (c'est-à-dire en bijection avec $\mathbb{N}$), toute probabilité est entièrement définie par la donnée des probabilités des événements élémentaires réduits à une seule éventualité.
Notons : $p_\omega = \mathbb{P}(\{\omega\})$. Nous avons alors :

$$\sum_{\omega\in\Omega} p_\omega = 1 \quad \text{et} \quad A \in \mathcal{P}(\Omega) \Rightarrow\ \mathbb{P}(A) = \sum_{\omega\in A} p_\omega\ .$$

Exemple : La probabilité de Poisson est définie sur $(\mathbb{N},\mathcal{P}(\mathbb{N}))$ par :

$$\forall\, n \geq 0\,, \mathbb{P}(\{n\}) = \frac{e^{-\lambda}\lambda^n}{n!} \text{ où } \lambda \in \mathbb{R}\,, \lambda \geq 0\ .$$

Cette probabilité, définie sur un ensemble non fini, est très utile en modélisation et permet de donner un premier exemple de probabilité sur $\mathbb{N}$. Par exemple, cela peut modéliser le nombre de requêtes qui arrivent à un serveur du réseau informatique pendant une durée de temps donnée.

Propriétés élémentaires des probabilités

Nous rappelons ici les propriétés élémentaires et utiles des probabilités. L'événement $\overline{A}$ désigne ici l'événement complémentaire de A.

Proposition 1 *Nous avons :*

1. $\mathbb{P}(\emptyset) = 0$
2. $\mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A)$.
3. *Soient* $A,B \in \mathcal{A}$, *alors* $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.
4. $A \subset B \Rightarrow \mathbb{P}(A) \leq \mathbb{P}(B)$.
5. *Soit* $(A_n)_{n\geq 1}$ *une suite croissante (* i.e. $A_n \subset A_{n+1}$*) d'éléments de* $\mathcal{A}$. *Celle-ci vérifie la propriété dite de monotonie croissante :*

$$\mathbb{P}\left(\bigcup_{n=1}^{+\infty} A_n\right) = \lim_{n\to+\infty} \mathbb{P}(A_n) \ .$$

6. *Soit* $(A_n)_{n\geq 1}$ *une suite décroissante (* i.e. $A_n \supset A_{n+1}$*) d'éléments de* $\mathcal{A}$. *Celle-ci vérifie la propriété de monotonie décroissante :*

$$\mathbb{P}\left(\bigcap_{n=1}^{+\infty} A_n\right) = \lim_{n\to+\infty} \mathbb{P}(A_n) \ .$$

Démonstration

- Les propriétés 1), 2) sont aisées à démontrer et découlent immédiatement de la définition.
  Démontrons la propriété 3).
  Nous avons :

  $$A \cup B = A \cup (B \cap \overline{A}) \ ,$$

  D'où,

  $$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B \cap \overline{A}) \ .$$

  Or,

  $$\mathbb{P}(B) = \mathbb{P}(A \cap B) + \mathbb{P}(B \cap \overline{A}) \ .$$

  D'où le résultat : $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$. ■
- Démontrons la propriété 4). Puisque $A \subset B$, nous obtenons :

  $$\mathbb{P}(B) = \mathbb{P}(A) + \mathbb{P}(B \cap \overline{A})) \geq \mathbb{P}(A) \ . \quad ■$$