Probabilités

Chapitre 1 : Fondement des probabilités

Lucie Le Briquer

1 Introduction

Objectif. Trouver un cadre pour modéliser l'aléa, le hasard.

- Jeux de hasards, dés, roulettes
- Sondages, échantillons (on capture 1000 tortues, il y en a 543 bleues, est-ce assez significatif pour conclure qu'il y a une proportion $> \frac{1}{2}$ de tortues bleues)
- Situation où on ne maîtrise pas toutes les interactions (météo, cours de la Bourse, physique statistique)

Cadre. (Axiomatique de Kolmogorov) À un problème impliquant l'aléa, on va construire un espace de probabilité $(\Omega, \mathcal{A}, \mathbb{P})$ pour modéliser cet aléa.

Remarque. On ne travaille *jamais* avec plusieurs espaces de probabilités et plus tard on ne précisera pas lequel (sauf pour les chaînes de Markov où Ω est explicite et il y a tout une famille de $\mathbb{P}$).

1. Ω est un ensemble non vide qui représente l'ensemble des possibles. Un élément $\omega \in \Omega$ est l'aléa. La donnée de ω donne toute l'information sur le modèle. Un seul ω est réalisé, mais on ne sait pas lequel.

   Exemple. On jette 3 dés, $\Omega = \{1, ..., 6\}^3$. Tout ce qui dépend de ω est aléatoire, le reste est déterministe.

2. $\mathcal{A} \subseteq \mathcal{P}(\Omega)$ qui est une tribu ($\sigma-$algèbre) :
   - non vide : $\Omega \in \mathcal{A}$
   - stable pour $\bigcup_{\mathbb{N}}$: si $A_n \in \mathcal{A} \forall n \Rightarrow \bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$
   - stable par complémentarité : $A \in \mathcal{A} \Rightarrow \bar{A} \in \mathcal{A}$

   Les éléments A,B,C,... sont des *évènements* et représentent des propriétés que l'on peut observer sur l'aléa.

Exemple. $A=$"Les 3 dés sont pairs" $=\{2,4,6\}^3$

3. $\mathbb{P}:\mathcal{A}\to[0,1]$ est une mesure de probabilité :
   - si les A_n sont disjoints :
   $$\mathbb{P}\left(\bigsqcup_{\mathbb{N}} A_n\right)=\sum_{\mathbb{N}}\mathbb{P}(A_n)$$
   - $\mathbb{P}(\Omega)=1$

   $\mathbb{P}(A)$ représente les chances qu'a ω l'aléa d'être dans A ($\mathbb{P}(A)=1\to$ sur toutes les observations $\omega\in A$, $\mathbb{P}(A)=0\to$ sur toutes les observations $\omega\notin A$, $\mathbb{P}(A)=\frac{1}{3}\to$ une fois sur trois on observera $\omega\in A$).

On pose un problème :

1. on *modélise* par le choix d'un $(\Omega,\mathcal{A},\mathbb{P})$ adapté
2. on travaille dans cet espace pour résoudre

2 Exemples

2.1 Un premier exemple dans le cadre discret

Problème. on jette 2 dés, quelle est la probabilité que la somme soit paire ?

Modélisation. $\Omega=\{1,...,6\}^2$, $\mathcal{A}=\mathcal{P}(\Omega)$ (toujours si Ω est fini), $\mathbb{P}$=probabilité uniforme ($\mathbb{P}(A)=\frac{\mathrm{Card}(A)}{36}$).

Résolution. $A=$"La somme des dés est paire" $=\{2,4,6\}^2\cup\{1,3,5\}^2$, donc :

$$\mathbb{P}(A)=\frac{\mathrm{Card}(A)}{36}=\frac{9+9}{36}=\frac{1}{2}$$

2.2 Un second exemple dans le cas continu : l'aiguille de Buffon

Problème. On lance une aiguille de 1cm sur un parquet de lattes de 1cm. Quelle est la probabilité que l'aiguille soit à cheval sur 2 lattes ?

Modélisation. On repère l'aiguille par :

- x la distance de son centre au bord de la latte en dessous $0\leqslant x<1$
- θ son angle $\theta\in[0,\pi[$

On pose alors comme cadre :

- $\Omega=\{(x,\theta)\in[0,1[\times[0,\pi[\}$
- $\mathcal{A}=\mathcal{B}([0,1[\times[0,\pi[)$
- $\mathbb{P}=\frac{1}{\pi}\mathrm{Lebesgue}|_{[0,1[\times[0,\pi[}$

Résolution. $A =$ "L'aiguille touche 2 lattes"

$$\begin{aligned}\bar{A} &= \text{"L'aiguille ne touche pas 2 lattes"}\\ &= \left\{(x,\theta) \in [0,1[\times[0,\pi[\ \middle| \ \left(x \leqslant \frac{1}{2} \text{ et } x - \frac{1}{2}\sin\theta \geqslant 0\right) \text{ ou } \left(x > \frac{1}{2} \text{ et } x + \frac{1}{2}\sin\theta < 1\right)\right\}\end{aligned}$$

$$\begin{aligned}\mathbb{P}(\bar{A}) &= \mathbb{P}\left(\left\{(x,\theta)\in\Omega \mid x \leqslant \frac{1}{2} \text{ et } x - \frac{1}{2}\sin\theta \geqslant 0\right\}\right) + \mathbb{P}\left(\left\{(x,\theta)\in\Omega \mid x > \frac{1}{2} \text{ et } x + \frac{1}{2}\sin\theta < 1\right\}\right)\\ &\underset{\text{symétrie}}{=} 2\mathbb{P}\left(\left\{(x,\theta)\in\Omega \mid x \leqslant \frac{1}{2} \text{ et } x - \frac{1}{2}\sin\theta \geqslant 0\right\}\right)\\ &= 2\int_{x\in[0,1[}\int_{0\leqslant\theta\leqslant\pi} \mathbb{1}_{\frac{1}{2}\sin\theta\leqslant x\leqslant\frac{1}{2}} \frac{dxd\theta}{\pi}\\ &= 2\int_0^\pi \frac{1}{2}(1-\sin\theta)\frac{d\theta}{\pi}\\ &= 1 - \frac{2}{\pi}\end{aligned}$$

Donc,

$$\mathbb{P}(A) = \frac{2}{\pi}$$

2.3 Un troisième exemple plus complexe

Problème. On jette des dés à 6 faces (une infinité de fois), quelle est la probabilité que le premier 6 apparaisse au bout d'un nombre pair de lancers ?

Modélisation. $\Omega = \{1,...,6\}^{\mathbb{N}}$. $\omega = (\omega_1, \omega_2, ...)$ avec ω_i le résultat du $i-$ème lancer. Soit :

$$A_{i_1,...,i_k} = \{\omega = (\omega_1,\omega_2,...) \in \Omega \mid \omega_1 = i_1, ..., \omega_k = i_k\}$$

où $k \in \mathbb{N}$ et $(i_1,...,i_k) \in \{1,...,6\}^k$. On veut que les $A_{i_1,...,i_k}$ soient des évènements. On pose donc :

$$\mathcal{A} = \sigma(A_{i_1,...,i_k} \mid k \in \mathbb{N}^*, (i_1,...,i_k) \in \{1,...,6\}^k)$$

($\sigma =$ plus petite tribu contenant ...).
et $\mathbb{P}$ la mesure de probabilité telle que $\mathbb{P}(A_{i_1,...,i_k}) = \frac{1}{6^k}$.

Difficultés.

- Pourquoi ce choix de tribu ?
- Existe-t-il une telle probabilité $\mathbb{P}$?
- Est-elle unique ?

Résolution.

$$\begin{aligned}
&\mathbb{P}(\text{premier 6 au bout d'un nombre pair de lancers})\\
&= \mathbb{P}\left(\bigsqcup_{p\in\mathbb{N}^*} \text{le premier 6 est au bout de } 2p \text{ lancers}\right)\\
&= \sum_{p\in\mathbb{N}^*} \sum_{i_1,\dots,i_{2p-1}\in\{1,\dots,5\}^{2p-1}} \mathbb{P}(A_{i_1,\dots,i_{2p-1},6})\\
&= \sum_{p\in\mathbb{N}^*} \sum_{i_1,\dots,i_{2p-1}\in\{1,\dots,5\}^{2p-1}} \frac{1}{6^{2p}}\\
&= \sum_{p\in\mathbb{N}^*} \left(\frac{5}{6}\right)^{2p} \frac{1}{5}\\
&= \frac{5}{6^2}\frac{1}{1-\left(\frac{5}{6}\right)^2}\\
&= \frac{5}{11}
\end{aligned}$$

Existence de $\mathbb{P}$. Construire une infinité d'évènements indépendants est parfois difficile (on y reviendra). Ici on a une construction directe.
$x \in [0,1[$ a un développement 6-adique unique :

$$x = \sum_{i=1}^{+\infty} \frac{x_i}{6^i} \quad \text{avec } x_i \in \{0,1,\dots,5\} \text{ non tous égaux à 5 à partir d'un certain rang}$$

Soit :

$$\varphi : \begin{cases} [0,1[& \to & \Omega \\ x = \sum \frac{x_i}{6^i} & \to & (x_1+1, x_2+2, \dots) \end{cases}$$

φ est mesurable car :

$$\varphi^{-1}(A_{i_1,\dots,i_k}) = \left[\sum_{j=1}^{k} \frac{i_j-1}{6^j} ; \sum_{j=1}^{k} \frac{i_j+1}{6^j} + \frac{1}{6^k}\right[$$

(il suffit de vérifier la mesurabilité sur une partie génératrice car $\varphi^{-1}(\bar{A}) = \overline{\varphi^{-1}(A)}$)
On pose $\mathbb{P}$ la mesure image de Lebesgue par φ :

$$\mathbb{P}(A) = \text{Lebesgue}(\varphi^{-1}(A))$$

et $\mathbb{P}(A_{i_1,\dots,i_k}) = \text{Lebesgue}\left(\left[\sum \frac{i_j-1}{6^j}; \sum \frac{i_j-1}{6^j} + \frac{1}{6^k}\right[\right)$. Donc $\mathbb{P}$ existe.

Choix de la tribu. $\mathcal{A} = \sigma(A_{i_1,\dots,i_k} \mid \dots)$. Déraisonnable de ne pas prendre les $A_{i_1,\dots,i_k}$. Choisir une tribu plus grosse mène aux mêmes difficultés que définir Lebesgue sur + que les boréliens. Ce $\mathcal{A}$ est la tribu “cylindrique”.

Unicité de $\mathbb{P}$: $\exists$ une unique probabilité sur $(\Omega, \mathcal{A})$ telle que

$$\mathbb{P}(A_{i_1,\dots,i_k}) = \frac{1}{6^k} \quad \forall i_1, \dots, i_k$$

La raison vient des classes monotones.

Rappels 1 (classes)

1. une classe $\mathcal{C}$ est une partie de $\mathcal{A}$: $\mathcal{C} \subseteq \mathcal{A}$
2. c'est une tribu si elle est stable par complémentaire et union dénombrable
3. on appelle

$$\sigma(\mathcal{C}) = \text{plus petite tribu qui la contient} = \bigcap_{\mathcal{B} \text{ tribu}, \ \mathcal{C} \subseteq \mathcal{B} \subseteq \mathcal{A}} \mathcal{B}$$

Définition 1 (classe monotone)

$\mathcal{C}$ est une classe monotone si :

- $\Omega \in \mathcal{C}$
- si $A \subseteq B$ et A, B dans $\mathcal{C}$ alors $B \backslash A$
- si A_n est une suite croissante dans $\mathcal{C}$ ($A_n \subseteq A_{n+1}$) alors $\bigcup_{\mathbb{N}} A_n \in \mathcal{C}$

Remarque. Une tribu est une classe monotone.

On définit pour une classe $\mathcal{C}$:

$$\mathcal{M}(\mathcal{C}) = \text{plus petite classe monotone contenant } \mathcal{C} = \bigcap_{\mathcal{C} \subseteq \mathcal{M} \subseteq \mathcal{A}, \ \mathcal{M} \text{ cl. monotone}} \mathcal{M}$$

Intérêt de la notion :

- simple à vérifier
- si μ et ν sont deux mesures de probabilité, $\{A \in \Omega \mid \mu(A) = \nu(A)\}$ est une classe monotone (si $A \subseteq B$, $\mu(B \backslash A) = \mu(B) - \mu(A)$; si A_n suite croissante, $\mu(\bigcup A_n) = \lim \mu(A_n)$)

S'il existe une autre mesure $\mathbb{Q}$ telle que $\mathbb{Q}(A_1, ..., i_k) = \frac{1}{6^k}$ on a :

$$\mathbb{P}|_{\mathcal{M}(A_{i_1,...,i_k}|..)} = \mathbb{Q}|_{\mathcal{M}(A_{i_1,...,i_k})|..}$$

Il manque un résultat pour dire $\mathcal{M}(A_{i_1,...,i_k}) = \sigma(A_{i_1,...,i_k}) = \mathcal{A}$

Lemme 1 (lemme des classes monotones)

Si $\mathcal{C}$ est une classe stable par $\bigcap$ finie alors :

$$\mathcal{M}(\mathcal{C}) = \sigma(\mathcal{C})$$

Ici avec $\mathcal{C} = \left\{A_{i_1,...,i_k} | i_1, ..., i_k \in \{1, ..., 6\}^k, k \in \mathbb{N}^*\right\} \bigcup \{\emptyset\}$ on vérifie que le lemme s'applique et on montre alors que $\mathbb{P} = \mathbb{Q}$.