

Chapitre 2

Introduction aux probabilités

L'étude des probabilités est très récente dans l'histoire des mathématiques. Les débuts se font avec l'étude des jeux de hasard notamment jeu de dés et le problème des partis, correspondance entre Pascal et Fermat (bien que Cardan ait fait des écrits antérieurs) où l'on s'intéressait à la répartition des gains lors de l'interruption d'une partie. La raison pour laquelle cet intérêt a été tardif, alors que les mathématiques présentaient tous les outils (fractions...) pour traiter des problèmes simples d'équiprobabilité est certainement lié à la religion et la superstition, il ne fallait pas défier le destin ni étudier la volonté d'une puissance divine. Huygens introduit quelques temps après le droit d'espérer gagner pour un joueur bien que le jeu ne soit que le fruit du hasard. Puis arrivent Bernouilli, Bayes, Laplace, Poisson, Bienaymé, Tchebychev, Markov, Kolmogorov...

Laplace donne comme définition de la probabilité le rapport du nombre de cas favorables à un événement à celui de tous les cas possibles, "mais cela suppose les divers cas également possibles" Cette définition qui a longtemps été reprise amène un problème, comment reconnaître que tous les cas sont également possibles?

On lance un dé non truqué. On ne peut pas savoir à l'avance la face qui va apparaître. On sait seulement que ce sera un chiffre entre 1 et 6. Le résultat d'une telle expérience est donc aléatoire, du latin alea qui signifie dé.

2.1 Définitions et propriétés de l'application P

Dans toute la suite, Ω désignera un ensemble que l'on appellera univers, ce sera l'ensemble de toutes les issues qu'il peut y avoir lors d'une expérience aléatoire.

Exemple d'univers possibles :

On lance 2 dés; $\Omega = \{1..6\}^2$ est l'ensemble des issues possibles de cette expérience. Son côté aléatoire est traduit par le fait qu'avant de l'avoir réalisée, on connaît les possibles $((i, j), i, j \in \{1..6\})$ mais on ne connaît pas le couple (i, j) réalisé à la fin de l'expérience.

On tire n fois à pile ou face avec une pièce. L'univers Ω peut s'écrire $\Omega =$

$\{P, F\}^n$ ou si l'on code par 0 l'apparition de pile et 1 l'apparition de face, $\Omega = \{0, 1\}^n$. Si on fait varier le nombre de lancers et qu'on suppose qu'il peut être infini, on forme alors une suite $(U_n)_{n \in \mathbb{N}}$ où U_i est l'issue du i ème tirage. L'univers Ω est alors l'ensemble de ces suites $(U_n)_{n \in \mathbb{N}}$ correspondant à toutes les issues possibles de nos n tirages $\Omega = \{0, 1\}^{\mathbb{N}}$.

Toute expérience aléatoire possède un univers Ω de possibles qui décrivent mathématiquement les issues de cette expérience aléatoire.

Le choix de Ω sera déterminant dans la résolution des exercices. Une partie de Ω est appelée événement, un élément de Ω événement élémentaire. Ainsi, pour $\Omega = \{1 \dots 6\}$ l'univers associé à l'expérience aléatoire du tirage d'un dé, obtenir un nombre pair se traduit par l'événement $\{2, 4, 6\}$ partie de Ω , obtenir un 6 se traduit par l'événement élémentaire $\{6\}$, partie également de Ω de cardinal 1.

Definition 1: On appelle $\wp(\Omega)$ l'ensemble des parties de Ω . $A \in \wp(\Omega) \iff A \subset \Omega$

C'est l'ensemble dont les éléments sont des sous-ensembles de Ω que nous avons dénombré dans le chapitre 1.

Travaillant par la suite sur des ensembles finis, nous ferons le choix de ne pas introduire la notion de tribu dans ce cours et nous nous restreindrons à l'ensemble $\wp(\Omega)$ des parties de Ω comme espace probabilisable $(\Omega, \wp(\Omega))$.

On veut ainsi définir une probabilité sur l'ensemble, en associant à chaque événement A un réel positif $p(A)$, compris entre 0 et 1 et qui nous donne le "degré plus ou moins probable de réalisation de cet événement" sur une échelle telle que 0 soit le degré quasi impossible et 1 le degré quasi certain.

A chaque expérience aléatoire est associé un univers.

Par exemple l'univers Ω associé :

- au tirage de 2 boules sans remise dans une urne où se trouvent 2 boules rouges et une boule noire est : $\Omega = \{(R1, N), (R2, N), (R1, R2)\}$ soit C_3^2 possibles (l'ordre d'apparition n'a pas d'importance).
- tirage avec remise : $\Omega = \{(R1, R1), (R2, R2), (N, N), (R1, N), (R2, N), (R1, R2)\}$ soit Γ_3^2 possibles
- tirage d'une seule boule : $\Omega = \{R1, R2, N\}$ soit C_3^1 possibles
- tirage de 3 boules : $\Omega = \{(R1, R2, N)\}$ soit C_3^3 possibles
- tirage de 3 boules avec remise : $\Omega = \{(R1, R2, N), (R1, R1, N), (R1, R1, R2), (R2, R2, N), (R2, R2, R1), (N, N, R1), (N, N, R2), (N, N, N), (R2, R2, R2), (R1, R1, R1)\}$ soit Γ_3^3 possibles

Definition 2: une probabilité sur un espace $(\Omega, \wp(\Omega))$ est une application $p : \wp(\Omega) \rightarrow \mathbb{R}_+$ qui vérifie :

(i) Axiome de dénombrable additivité si $(A_n)_{n \in \mathbb{N}}$ est une suite d'éléments deux à deux disjoints de $\wp(\Omega)$, $p(\cup A_n) = \sum p(A_n)$

(ii) $p(\Omega) = 1$

On en déduit que $p(\Omega) \rightarrow [0, 1]$ car $(i) \Rightarrow p(A) + p(\bar{A}) = p(\Omega) = 1$ ainsi $p(A) \in [0, 1]$ pour tout $A \in \Omega$.

Definition 3: $(\Omega, \wp(\Omega), p)$ est un espace probabilisable.

Definition 4: deux événements tels que $A \cap B = \emptyset$ sont dits incompatibles.

Proposition 1: Si p est une probabilité sur Ω alors

i $\forall A \in \wp(\Omega), p(\bar{A}) = 1 - p(A)$

ii $\forall A \in \wp(\Omega), p(A) \in [0, 1]$

iii $p(\emptyset) = 0$

iv $\forall A, B \in \wp(\Omega), p(A \setminus B) = p(A) - p(A \cap B)$

v $\forall A, B \in \wp(\Omega), A \cap B = \emptyset, p(A \cup B) = p(A) + p(B)$ (axiome d'additivité faible)

Notez que $A \setminus B$ signifie A privé de B

- $\Rightarrow (i)$ d'après l'axiome de dénombrable additivité,
 $p(\Omega) = p(A \cup \bar{A}) = p(A) + p(\bar{A}) = 1$ d'où $p(\bar{A}) = 1 - p(A)$
- $\Rightarrow (ii) \forall A \in \wp(\Omega), p(A) + p(\bar{A}) = 1$, comme $p(A)$ et $p(\bar{A}) \geq 0$ on a donc $p(A) \leq 1$.
- $\Rightarrow (iii)$ en remarquant que $\Omega = \bar{\emptyset}$ on a $p(\emptyset) = 1 - p(\Omega) = 0$
- $\Rightarrow (iv)$ $A \setminus B$ et $(A \cap B)$ sont deux événements incompatibles
 ainsi, d'après la dénombrable additivité, on a: $p(A \setminus B \cup (A \cap B)) = p(A \setminus B) + p(A \cap B)$
 or $A \setminus B \cup (A \cap B) = A$
 ainsi $p(A) = p(A \setminus B) + p(A \cap B)$
- $\Rightarrow (v)$ Comme A et B sont incompatibles, l'axiome de dénombrable additivité s'applique d'où le résultat attendu.

Théorème 1 : $\forall A, B \in \wp(\Omega), p(A \cup B) = p(A) + p(B) - p(A \cap B)$

$\Rightarrow$ On se ramène à des intervalles qu'on sait disjoints pour appliquer l'axiome d'additivité :

$A \cup B = (A \setminus B) \cup (B \setminus A) \cup (A \cap B)$ réunion disjointe

Ainsi, $p(A \cup B) = p(A \setminus B) + p(B \setminus A) + p(A \cap B)$
Or d'après (iv) de Prop 1, $p(A \setminus B) = p(A) - p(A \cap B)$
et $p(B \setminus A) = p(B) - p(A \cap B)$
ainsi :

$$\begin{aligned} p(A \cup B) &= p(A) - p(A \cap B) + p(B) - p(A \cap B) + p(A \cap B) \\ &= p(A) + p(B) - p(A \cap B) \end{aligned}$$

Ce qui amène la généralisation avec la formule de Poincaré ou du crible :
Soit $n \in \mathbb{N}$, pour tout $A_1 \dots A_n \in \wp(\Omega)$,

$p(A_1 \cup \dots \cup A_n) = \sum_{k=1}^n (-1)^{k-1} (\sum_{1 \leq i_1 < i_2 < \dots < i_k \leq n} p(A_{i_1} \cap \dots \cap A_{i_k}))$
même formule que celle appliquée aux cardinaux, la probabilité de l'événement A_i remplace son cardinal. Autre manière de l'écrire plus intuitivement (voir Chap 1) :

$$p(\cup_{i=1}^n A_i) = \sum_{k=1}^n (-1)^{k-1} \sum_{I \subset \{1..n\}, \text{Card} I = k} p(\bigcap_{I} A_I)$$

La démonstration est la même que celle abordée pour les cardinaux, par récurrence sur n.

Pour le rang 1 c'est trivial. Le rang 2 est donné par le théorème 1.

On suppose la propriété vérifiée pour un rang n quelconque. Au rang n+1 :
 $p(A_1 \cup \dots \cup A_{n+1}) = p(A_1 \cup \dots \cup A_n) + p(A_{n+1}) - p(\cup_{1..n} A_i \cap A_{n+1}) (*)$.

or, d'après l'hypothèse de récurrence,

$$\begin{aligned} p(\bigcup_{i=1}^n A_i) &= \sum_{k=1}^n (-1)^{k-1} \sum_{I \subset \{1..n\}, \text{Card} I = k} p(\bigcap_{i \in I} A_i) \text{ ET} \\ p(\bigcup_{i=1}^n A_i \cap A_{n+1}) &= p(\bigcup_{i=1}^n (A_i \cap A_{n+1})) \\ &= \sum_{k=1}^n (-1)^{k-1} \sum_{I \subset \{1..n\}, \text{Card} I = k} p(\bigcap_{i \in I} (A_i \cap A_{n+1})) \\ &= \sum_{k=1}^n (-1)^{k-1} \sum_{I \subset \{1..n\}, \text{Card} I = k} p(\bigcap_{i \in I} A_i \cap A_{n+1}) \end{aligned}$$

On remplace dans (*) :

$$p\left(\bigcup_{i=1}^{n+1} A_i\right) = \sum_{k=1}^n (-1)^{k-1} \sum_{I \subset \{1..n\}, \text{Card} I=k} p\left(\bigcap_{i \in I} A_i\right) + p(A_{n+1}) - \sum_{k=1}^n (-1)^{k-1} \sum_{I \subset \{1..n\}, \text{Card} I=k} p\left(\bigcap_{i \in I} A_i \cap A_{n+1}\right)$$

Or

$$\begin{aligned} \sum_{k=1}^n (-1)^{k-1} \sum_{I \subset \{1..n\}, \text{Card} I=k} p\left(\bigcap_{i \in I} A_i \cap A_{n+1}\right) &= \sum_{k=1}^n (-1)^{k-1} \sum_{n+1 \in J \subset \{1..n+1\}, \text{Card} J=k+1} p\left(\bigcap_{i \in J} A_i\right) \\ &= \sum_{k=2}^{n+1} (-1)^{k-2} \sum_{n+1 \in J \subset \{1..n+1\}, \text{Card} J=k} p\left(\bigcap_{i \in J} A_i\right) \\ &= \sum_{k=1}^{n+1} (-1)^{k-2} \sum_{n+1 \in J \subset \{1..n+1\}, \text{Card} J=k} p\left(\bigcap_{i \in J} A_i\right) - (-1)^0 p(A_{n+1}) \end{aligned}$$

D'où

$$\begin{aligned} p\left(\bigcup_{i=1}^{n+1} A_i\right) &= \sum_{k=1}^n (-1)^{k-1} \sum_{I \subset \{1..n\}, \text{Card} I=k} p\left(\bigcap_{i \in I} A_i\right) \\ &- (-1)^{n-1} p\left(\bigcap_{i \in [1..n+1]} A_i\right) + \sum_{k=1}^n (-1)^{k-1} \sum_{n+1 \in J \subset \{1..n+1\}, \text{Card} J=k} p\left(\bigcap_{i \in J} A_i\right) \\ &= \sum_{k=1}^n (-1)^{k-1} \left(\sum_{I \subset \{1..n\}, \text{Card} I=k} p\left(\bigcap_{i \in I} A_i\right) + \sum_{n+1 \in J \subset \{1..n+1\}, \text{Card} J=k} p\left(\bigcap_{i \in J} A_i\right) \right) \\ &+ (-1)^{n+1-1} p\left(\bigcap_{i \in [1..n+1]} A_i\right) \\ &= \sum_{k=1}^{n+1} (-1)^{k-1} \sum_{I \subset \{1..n+1\}, \text{Card} I=k} p\left(\bigcap_{i \in I} A_i\right) \blacksquare \end{aligned}$$

Il est normal que la formule de Poincaré valable en probabilité le soit en cardinal car il suffit de remarquer qu'en cas d'équiprobabilité (voir plus loin) on a bien :

$$p\left(\bigcap_{i=1}^n A_i\right) = \frac{|\bigcap A_i|}{|\Omega|}$$

2.2 Cas particulier de l'équiprobabilité

Definition 5: Considérons un univers fini $\Omega = \{\omega_1 \dots \omega_n\}$

On dit qu'il y a équiprobabilité si et seulement si

$$\forall (i, j) \in [1..n]^2, p(\omega_i) = p(\omega_j)$$

Comme $\sum_{i=1}^n p(\omega_i) = 1$ il vient alors $\forall i \in [1..n], p(\omega_i) = \frac{1}{n} = \frac{1}{\text{Card}(\Omega)}$

La probabilité d'un événement élémentaire est $\frac{1}{\text{Card} \Omega}$. On appelle cette probabilité une probabilité uniforme sur Ω .

Proposition 2: Soit Ω un univers fini et p probabilité uniforme sur Ω .

La probabilité d'un événement quelconque $A \in \wp(\Omega)$ est donnée par

$$p(A) = \frac{\text{card}A}{\text{Card}\Omega} \text{ ou bien par } p(A) = \frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}}.$$

Exemple : Exercice 2 TD2

Proposition 3: Si $(A_n)_{n \in \mathbb{N}}$ est une suite croissante au sens de l'inclusion d'éléments de $\wp(\Omega)$ (i.e. $A_i \subset A_{i+1}$) alors la suite $(p(A_n))_{n \in \mathbb{N}}$ est croissante et on a le résultat:

$$\lim_{n \rightarrow \infty} p(A_n) = p(\cup_{n \in \mathbb{N}} A_n)$$

$\Rightarrow$ Soit $n \in \mathbb{N}$, $p(A_{n+1}) - p(A_n) = p((A_{n+1} \setminus A_n) \cup A_n) - p(A_n)$
 $A_{n+1} \setminus A_n$ et A_n étant disjoints, l'axiome de l'additivité nous donne :
 $p(A_{n+1}) - p(A_n) = p(A_{n+1} \setminus A_n) + p(A_n) - p(A_n) = p(A_{n+1} \setminus A_n) \geq 0$

La suite $(p(A_n))_{n \in \mathbb{N}}$ est donc croissante et majorée par $p(\Omega) = 1$ donc elle converge. On définit la suite $(B_n)_{n \in \mathbb{N}}$ telle que :

$$\begin{cases} B_0 = A_0 \\ \forall n \in \mathbb{N}^*, B_n = (A_n \setminus A_{n-1}) \end{cases}$$

Les événements B_n sont disjoints deux à deux et par construction $\cup A_n = \cup B_n$.

On a, d'après l'axiome d'additivité, $p(\bigcup_{n \in \mathbb{N}} B_n) = \sum_{n \in \mathbb{N}} p(B_n) = \lim_{N \rightarrow \infty} \sum_{n=0}^N p(B_n)$

Or $\forall N \in \mathbb{N}$, $p(A_N) = \sum_{n=0}^N p(B_n)$ par récurrence évidente

D'où $\lim_{N \rightarrow \infty} p(A_N) = \lim_{N \rightarrow \infty} \sum_{n=0}^N p(B_n) = p(\bigcup_{n \in \mathbb{N}} B_n) = p(\bigcup_{n \in \mathbb{N}} A_n)$ ■

Proposition 3bis : De même, si $(A_n)_{n \in \mathbb{N}}$ est une suite décroissante au sens de l'inclusion d'éléments de $\wp(\Omega)$ i.e. $\forall n \in \mathbb{N}, A_{n+1} \subset A_n$ alors la suite $(p(A_n))_{n \in \mathbb{N}}$ est décroissante et

$$\lim_{n \rightarrow \infty} p(A_n) = p(\bigcap_{n \in \mathbb{N}} A_n)$$

$\Rightarrow$ On se sert de Proposition 3 en passant aux complémentaires soit $\forall n \in \mathbb{N}, B_n = \overline{A_n}$

Alors (B_n) est croissante. Ainsi, d'après Proposition 3, $(p(B_n))_{n \in \mathbb{N}}$ est croissante et $\lim_{n \rightarrow \infty} p(B_n) = p(\bigcup_{n \in \mathbb{N}} B_n)$

Comme $\forall n \in \mathbb{N}, p(A_n) = 1 - p(B_n)$ et que $(p(B_n))_{n \in \mathbb{N}}$ est croissante, $(p(A_n))_{n \in \mathbb{N}}$ est décroissante et $\lim_{n \rightarrow \infty} p(A_n) = 1 - \lim_{n \rightarrow \infty} p(B_n) = 1 - p(\bigcup_{n \in \mathbb{N}} B_n)$

$$\text{Or } p(\bigcup_{n \in \mathbb{N}} B_n) = p(\bigcup_{n \in \mathbb{N}} \overline{A_n}) = p(\overline{\bigcap_{n \in \mathbb{N}} A_n}) = 1 - p(\bigcap_{n \in \mathbb{N}} A_n) \quad \blacksquare$$

Exemple d'application de la propriété 3 : On lance une infinité de fois un dé. On étudie la probabilité de l'événement "la face 6 n'apparaît pas". Ici $\Omega = \{1 \dots 6\}^{\mathbb{N}}$

Soit A_n l'événement "la face 6 n'apparaît pas au cours des n premiers lancers". On peut écrire : $p(A_1) = \frac{5}{6}$ (équiprobabilité cas favorables rapportés aux cas possibles)

$p(A_2) = \frac{5}{6} * \frac{5}{6}$ (le 6 ne sort ni au premier essai ni au second)

De manière générale, soit $n \in \mathbb{N}, p(A_n) = (\frac{5}{6})^n$

On remarque que $\forall n \in \mathbb{N}, A_n \subset A_{n+1}$ (si la face 6 n'est pas apparue au cours des $n+1$ premiers lancers, alors elle n'est pas apparue avant, soit au cours des n premiers).

Ainsi, nous sommes dans les conditions d'application de la proposition 3, $\bigcup_{n \in \mathbb{N}} A_n$

représente l'événement "la face 6 n'apparaît jamais au cours d'une infinité de lancers" alors :

$$p(\bigcup_{n \in \mathbb{N}} A_n) = \lim_{n \rightarrow \infty} p(A_n) = \lim_{n \rightarrow \infty} (\frac{5}{6})^n = 0 \text{ (suite géométrique de raison inférieure à 1)}$$

Ainsi, la probabilité de l'événement "la face 6 n'apparaît jamais suite à une infinité de lancers" est nulle.

Ce qui nous amène à la définition :

Definition 5: Si $p(A) = 0$ (resp. $p(A) = 1$) on dit que A est un événement quasi (ou presque) impossible (resp. quasi (ou presque) certain).

L'exemple illustre bien la raison pour laquelle on ne dit pas impossible, alors que les possibles de cette dernière expérience étaient nombreux, (toutes les suites u_n prenant leur valeur dans $1..5$), pourtant leur probabilité est nulle.

2.3 Un exemple fréquent de loi de probabilité : la loi binomiale de Newton

Différents exemples de son application :

- Exemple 1: on jette 50 fois une pièce, on s'intéresse à la probabilité d'obtenir 37 fois pile.
- Exemple 2: dans une urne où se trouvent 5 boules blanches et 3 noires, on procède à 10 tirages avec remise. On s'intéresse à la probabilité d'obtenir 5 boules noires sur nos 10 tirages.
- Exemple 3: TD2 exercice 4
- Exemple 4: dans un QCM il y a 20 questions à chacune 5 réponses possibles dont une seule est correcte. On s'intéresse à la probabilité d'avoir au moins 8 bonnes réponses de la part d'un candidat répondant au hasard à chacune des 20 questions.

Definition 6 : On appelle épreuve de Bernoulli de paramètre p une expérience aléatoire qui ne possède que deux issues contraires : une réussite de probabilité p et un échec de probabilité $1-p$.

Remarque : L'univers des issues d'une épreuve de Bernoulli est une paire :
exemple 1 : $\Omega = \{P, F\}$
exemple 2 : $\Omega = \{B, N\}$
exemple 4 : $\Omega = \{V, F\}$

Definition 7: un schéma de Bernoulli est une expérience qui consiste en la répétition de n épreuves de Bernoulli identiques et indépendantes.

Definition 8: la loi binomiale de Newton est une loi de probabilité qui comptabilise le nombre de réussites sur n épreuves d'un schéma de Bernoulli de paramètre p

On a alors $p(A) = C_n^k p^k (1-p)^{n-k}$