

Probabilités sur un univers fini

I Espaces probabilisés finis

1) Introduction

La théorie des probabilités consiste en la modélisation des phénomènes dans lesquels intervient le hasard.

Définition. Une expérience aléatoire est une expérience renouvelable dont le résultat ne peut être prévu à l'avance et qui, renouvelée dans des conditions identiques, ne donne pas forcément le même résultat à chaque fois.


Voici quelques exemples d'expériences aléatoires simples : lancer de dé, lancer de pièce, tirage de boules dans une urne, tirage d'une carte dans un jeu de cartes, etc.

La théorie des probabilités est apparue historiquement pour étudier les jeux de hasard. Elle est désormais utilisée pour décrire des phénomènes trop complexes pour être analysés en détail (le hasard est donc une hypothèse simplificatrice).

Par exemple : l'écoulement à travers un matériau poreux (percolation), le comportement des molécules dans un aimant (modèle d'Ising) ou dans un gaz, l'évolution d'une grande population de consommateur, l'évolution d'une espèce qui se reproduit, etc.

Mathématiquement, pour décrire une expérience aléatoire, on se donne :

- Un ensemble Ω , appelé univers, qui contient toutes les résultats possibles de l'expérience.
Pour un lancer de dé, on se donne $\Omega = \llbracket 1 ; 6 \rrbracket$. Pour un lancer de pièce, on se donne $\Omega = \{P; F\}$ ou $\Omega = \{0; 1\}$. Pour un lancer de 3 pièces, on se donne $\Omega = \{P; F\}^3$.
- Un ensemble de phénomènes, appelés événements, qui peuvent se produire ou ne pas se produire lors de l'expérience. Un événement est décrit par une partie de Ω (généralement l'ensemble des résultats de l'expérience menant à la réalisation de ce phénomène).
Si on lance un dé, alors l'événement « obtenir un chiffre pair » est décrit par la partie $\{2; 4; 6\}$. Si le dé tombe sur 2, 4 ou 6, alors cet événement est réalisé.
- Une fonction qui à tout événement associe la probabilité d'occurrence de A , c'est-à-dire un réel de $[0; 1]$ qui représente la chance que A soit réalisé.
Si le dé est bien équilibré, alors on peut prendre la fonction qui, à l'événement « obtenir la face numérotée k », associe $1/6$ pour tout $k \in \llbracket 1 ; 6 \rrbracket$.

 Il se peut que l'ensemble Ω des résultats possibles d'une expérience aléatoire soit un ensemble infini.

Par exemple l'ensemble des résultats possibles de l'expérience consistant à compter le nombre de lancers de pièce nécessaires pour obtenir un Pile est \mathbb{N}^ . L'ensemble des résultats possibles de l'expérience consistant à lancer une flèche sur une cible est un disque.*

Officiellement le programme de première année se limite au cas des univers finis. Le cas d'un univers dénombrable (c'est par exemple le cas si on considère une infinité de lancers de dés) sera vu en deuxième année. Le cas d'un univers non dénombrable (c'est par exemple le cas si on lance une fléchette sur une cible) ne sera pas étudié en CPGE scientifiques.

Il va sans dire que $\Omega \neq \emptyset$. Et c'est ce que l'on suppose désormais dans tous les chapitres de probabilités.

En général on définit les événements à l'aide de phrases et on ne cherche pas à les décrire mathématiquement.

Parfois, en exercice, nous « tricherons » un peu en nous limitant artificiellement à un univers fini... cette limitation du programme est avant tout technique.

2) Univers et événements

Nous allons à présent définir les notions d'univers, d'événements et de probabilités proprement.

Définition. On appelle univers (des possibles) l'ensemble Ω des résultats possibles de l'expérience aléatoire étudiée.

Les éléments de Ω sont appelés les éventualités ou les issues de l'expérience aléatoire.

Une partie A de Ω (c'est-à-dire $A \subset \mathcal{P}(\Omega)$) est appelé événement. On dit qu'il est réalisé si le résultat de l'expérience est un élément de A .

Si $\omega \in \Omega$, le singleton $\{\omega\}$ est appelé événement élémentaire.

L'événement \emptyset est l'événement impossible. L'événement Ω est l'événement certain.

Pour le moment, on ne parle pas encore de probabilités mais seulement de considérations ensemblistes afin de décrire des phénomènes.

Remarques :

- ⚠ Une éventualité ne doit pas être confondue avec un événement (une éventualité est un élément de Ω , alors qu'un événement est une partie de Ω). En particulier, il ne faut pas confondre $\omega \in \Omega$ et l'événement $\{\omega\}$.

- Le choix de Ω dépend de l'expérience que l'on cherche à modéliser.

Par exemple, si on lance deux dés indiscernables, on peut prendre $\Omega = \llbracket 1; 6 \rrbracket^2$ l'ensemble des couples (a, b) avec a et b dans $\llbracket 1; 6 \rrbracket$ mais on peut également prendre

$$\Omega = \{ \{a; b\} \mid (a, b) \in \llbracket 1; 6 \rrbracket^2 \}$$

En effet, les deux dés étant indiscernables, on sait qu'on a obtenu les nombres a et b mais on ne sait pas dans quel ordre, on peut donc se contenter de la partie contenant a et b . Le premier univers a 36 éléments, le second en a $\binom{6}{1} + \binom{6}{2} = 21$.

- À part dans les situations dites « d'équiprobabilité » (cf. paragraphe 1.5.d), on construit rarement l'univers Ω . En général Ω est de toute façon trop compliqué pour être décrit mathématiquement et la plupart des énoncés se contentent d'admettre qu'il existe un univers décrivant l'expérience étudiée.

Dans la pratique, on prendra toujours le premier (on expliquera pourquoi dans la suite).

Exemples :

- Si on lance n dés successivement (avec $n \in \mathbb{N}^*$)...

- Si on tire simultanément p boules dans une urne contenant n boules numérotées de 1 à n (avec p et n des entiers tels que $0 \leq p \leq n$)...

- Si cette fois on tire les boules successivement et sans remise...

Dans ces exemples, comme dans le chapitre 33, on se demande si l'ordre compte et si on peut avoir plusieurs fois le même élément.

Dans toute la suite de ce chapitre, Ω désigne un ensemble fini et non vide. L'étape suivante est de dresser une liste d'événements « simples ». En général, on définit les événements à l'aide de phrases et on ne cherche pas à les décrire mathématiquement (puisque l'on ne construit pas Ω en général).

« simple » au sens où l'on arrive facilement à calculer leur probabilité.

Exemples :

- Si on lance un dé, on peut poser, pour tout $k \in \llbracket 1; 6 \rrbracket$, A_k : « obtenir la face numérotée k ».
- Si on tire une boule dans une urne contenant des boules bleues, rouges et vertes, on peut poser R : « tirer une boule rouge », B : « tirer une boule bleue » et V : « tirer une boule verte ».
- Si on lance successivement n pièces de monnaie, on peut poser, pour tout $k \in \llbracket 1; n \rrbracket$, A_k : « obtenir Pile au $k^{\text{ième}}$ lancer ».



Réflexe en début d'exercice de probabilités : définir des événements (tous ceux auxquels on peut penser). Si un événement dépend d'un paramètre k (un instant, un numéro par exemple), on l'indexe avec k (par exemple A_k et non A).

3) Opérations sur les événements

L'étape suivante est de décrire les événements qui nous intéressent en fonction des événements « simples » que l'on vient d'introduire, à l'aide d'opérations ensemblistes :

Définition. Soient A et B des événements.

- L'événement $\bar{A} = \Omega \setminus A$ est réalisé si A ne l'est pas. On dit qu'il s'agit de l'événement contraire de A .
- La relation $A \subset B$ signifie que la réalisation de l'événement A entraîne celle de l'événement B .
- L'événement $A \cup B$ est réalisé si l'un au moins des événements A et B est réalisé.
- L'événement $A \cap B$ est réalisé si les événements A et B sont tous les deux réalisés.
- L'événement $A \setminus B$ est réalisé si l'événement A est réalisé mais pas l'événement B .
- Si $A \cap B = \emptyset$, les événements A et B sont dits incompatibles ou disjoints.



Puisque les événements sont des parties de Ω , toutes les opérations vues dans le chapitre 15 sont valables pour les événements. Seul le vocabulaire change.

Exemple : On dispose d'un jeu de 52 cartes. On tire trois cartes successivement et avec remise. On peut travailler avec l'univers où

$$E = \{A\clubsuit; K\clubsuit; Q\clubsuit; J\clubsuit; \dots; A\spadesuit; K\spadesuit; \dots; A\diamondsuit; K\diamondsuit; \dots; A\heartsuit; K\heartsuit; \dots; 3\heartsuit; 2\heartsuit\}.$$

Pour $i \in \{1; 2; 3\}$, notons A_i l'événement « obtenir \spadesuit au $i^{\text{ième}}$ tirage ». L'événement

- « obtenir un \clubsuit , \diamondsuit ou \heartsuit au premier tirage » est
- « obtenir \spadesuit aux deux premiers tirages mais pas au troisième » est
- « obtenir \spadesuit aux deux premiers tirages » est
- « obtenir au moins un \spadesuit » est
- « ne pas obtenir de \spadesuit » est
- « obtenir exactement deux \spadesuit » est

Notons B_1 l'événement « obtenir une carte de couleur noire au premier tirage ». Nous avons $A_1 \subset B_1$: si l'événement A_1 est réalisé, alors l'événement B_1 aussi. Les événements « obtenir une paire » et « les trois cartes sont des \heartsuit » sont incompatibles.

Remarque : De façon générale, les assertions du langage ordinaire sur les événements se traduisent dans le langage ensembliste en terme d'inclusion, de réunion, d'intersection et de passage au complémentaire. Ci-dessous le dictionnaire.



Tout simplement, deux événements sont incompatibles s'ils ne peuvent se produire en même temps.

Langage ordinaire	Langage ensembliste
Expérience aléatoire	Univers aléatoire Ω
Résultat possible, éventualité	$\omega \in \Omega$
Événement	$A \subset \Omega, A \in \mathcal{P}(\Omega)$
L'événement A est réalisé (par l'éventualité ω)	$\omega \in A \subset \Omega$
A n'est pas réalisé / L'événement contraire à A est réalisé	$\omega \in \bar{A}$
A et B sont réalisés	$\omega \in A \cap B$
A ou B est réalisé (« ou » non exclusif)	$\omega \in A \cup B$
La réalisation de A entraîne celle de B	$A \subset B$
A est réalisé mais pas B	$\omega \in A \cap \bar{B} = A \setminus B$
A et B sont incompatibles	$A \cap B = \emptyset$
Au moins un des A_1, \dots, A_n est réalisé	$\omega \in \bigcup_{i=1}^n A_i$
Tous les A_k sont réalisés	$\omega \in \bigcap_{k=1}^n A_k$
Aucun A_k n'est réalisé	$\omega \in \bigcup_{k=1}^n \bar{A}_k = \bigcap_{k=1}^n \bar{A}_k$
Les A_k sont deux à deux incompatibles	$\forall i \neq j, A_i \cap A_j = \emptyset$

Exemple : Soit $n \in \mathbb{N}^*$. On lance n fois successivement une pièce de monnaie. Un univers associé à cette expérience est $\Omega = \{P; F\}^n$. Pour tout $k \in \llbracket 1; n \rrbracket$, notons A_k l'événement « obtenir Pile au $k^{\text{ième}}$ lancer ».

- L'événement « n'obtenir que des faces » est

- Pour tout $k \in \llbracket 1; n \rrbracket$, l'événement B_k : « obtenir Pile pour la première fois au $k^{\text{ième}}$ lancer » est

4) Système complet d'événements


Certaines familles d'événements ont la particularité de « partitionner » l'ensemble Ω , c'est-à-dire de découper Ω en plusieurs sous ensembles incompatibles. Dans ce cas, toute issue ω de l'expérience appartient nécessairement à l'un des événement de la famille et un seul. Ce type de famille est appelé « système complet d'événements ».

Définition (système complet d'événements). Soit I une partie finie de \mathbb{N} et soit Ω un univers fini. Une famille $(A_i)_{i \in I}$ d'événements de Ω est un système complet (fini) d'événements si elle vérifie :

- $\Omega = \bigcup_{i \in I} A_i$,
- Les événements de la famille sont incompatibles deux à deux, c'est-à-dire

$$\forall (i, j) \in I^2, \quad i \neq j \quad \implies \quad A_i \cap A_j = \emptyset.$$

La différence avec la notion de partition vue dans le chapitre 15 est que, dans un système complet d'événements, les ensembles (événements en langage probabiliste) peuvent être vides.

 On parle encore d'un système complet d'événements pour un ensemble $\{A_1; \dots; A_n\}$ d'événements deux à deux incompatibles et dont l'union est égale à Ω (ce qui est plus intuitif vu le terme « système »).

Exemples :

- Si $\Omega = \llbracket 1; 6 \rrbracket$, alors $(\{1; 3; 5\}, \{2\}, \{4; 6\})$ est un système complet d'événements.
- On lance deux dés et on somme les numéros des deux faces obtenues. On obtient alors un entier compris entre 2 et 12. Pour tout $k \in \llbracket 2; 12 \rrbracket$, posons A_k : « obtenir une somme égale à k ». Alors (A_2, \dots, A_{12}) est un système complet d'événements.
- On lance successivement n pièces de monnaie et on pose, pour tout $k \in \llbracket 1; n \rrbracket$, A_k : « obtenir Pile au $k^{\text{ième}}$ lancer » et B_k : « on obtient le premier Pile au $k^{\text{ième}}$ lancer »
 - ★ La famille (A_1, \dots, A_n) n'est pas un système complet d'événements. En effet, on peut tout à fait obtenir Pile aux deux premiers lancers si bien que $A_1 \cap A_2 \neq \emptyset$.
 - ★ La famille (B_1, \dots, B_n) n'est pas un système complet d'événements. Ici les événements sont bien deux à deux incompatibles mais leur union est « obtenir au moins un Pile », ce qui n'est pas égal à Ω .
 - ★ Si on pose B_0 : « n'obtenir aucun Pile », alors (B_0, B_1, \dots, B_n) est un système complet d'événements.
- Si $\Omega = \{\omega_1; \dots; \omega_n\}$ avec $n = \text{card}(\Omega) \in \mathbb{N}^*$, alors $(\{\omega_i\})_{1 \leq i \leq n}$ est un système complet d'événements.

Cas particulier important : Si A est un événement, A et \bar{A} forment un système complet d'événements.

Par exemple, si $\Omega = \llbracket 1; 6 \rrbracket$, alors A : « tirer un nombre pair » et \bar{A} : « tirer un nombre impair » forment un système complet d'événements.

5) Probabilités sur un espace probabilisable fini

a) Définition

Intuitivement, une probabilité est une fonction qui à tout événement associe sa probabilité d'occurrence, c'est-à-dire un réel de $[0; 1]$ qui représente la chance qu'il a de se réaliser. Mais que peut-on attendre d'une telle fonction ? Lors de N répétitions indépendantes les unes des autres d'une expérience aléatoire, la fréquence d'apparition d'un événement A est le nombre $\frac{N(A)}{N}$ où $N(A)$ est le nombre de fois où l'événement A a été réalisé, et on suppose que cette fréquence admet bien une limite quand N tend vers $+\infty$. D'un point de vue intuitif, on veut poser :

$$\mathbb{P}(A) = \lim_{N \rightarrow +\infty} \frac{N(A)}{N}$$

Si A et B sont incompatibles, $N(A \cup B) = N(A) + N(B)$ puisqu'ils ne peuvent pas être réalisés simultanément. Dès lors, on a d'une part :

$$\frac{N(A \cup B)}{N} \xrightarrow{N \rightarrow +\infty} \mathbb{P}(A \cup B)$$

et d'autre part :

$$\frac{N(A \cup B)}{N} = \frac{N(A)}{N} + \frac{N(B)}{N} \xrightarrow{N \rightarrow +\infty} \mathbb{P}(A) + \mathbb{P}(B)$$

On s'attend donc, pour être conforme à l'intuition, à ce qu'une probabilité vérifie $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$ dès que les événements A et B sont incompatibles. De plus, une probabilité doit vérifier

$$\mathbb{P}(\Omega) = \lim_{N \rightarrow +\infty} \frac{N(\Omega)}{N} = \lim_{N \rightarrow +\infty} \frac{N}{N} = 1$$


La définition suivante est donc naturelle si on veut être conforme à l'intuition.

La notion de système complet d'événements est ensembliste et non probabiliste. C'est pourquoi nous la définissons dans ce paragraphe. Elle peut sembler abstraite au premier abord mais elle est essentielle puisqu'elle intervient dans la formule la plus importante du chapitre : la formule des probabilités totales.

Il suffit de se poser les deux questions très simples suivantes :


- Est-on dans au moins un des cas de figure ?
- Peut-on être dans plusieurs cas de figure différents en même temps ?

Si on est au moins dans un des cas des figure, et si deux cas de figure ne peuvent pas se produire en même temps, alors les différents cas de figure forment un système complet d'événements. Voir les exemples ci-contre.

 J'utilise personnellement la notation \mathbb{P} pour désigner une probabilité. Le programme utilise la notation P et d'autres ouvrages la notation \mathbf{P} . Cela n'a aucune importance : utilisez les notations de l'énoncé !

Définition. On appelle probabilité sur Ω toute application $\mathbb{P} : \mathcal{P}(\Omega) \rightarrow [0; 1]$ vérifiant :

1. $\mathbb{P}(\Omega) = 1$,
2. Pour tous événements A et B **incompatibles**, $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$.

 Une probabilité est définie sur $\mathcal{P}(\Omega)$ et non pas sur Ω ! Notamment, si $\omega \in \Omega$, alors $\mathbb{P}(\omega)$ n'a pas de sens. En toute rigueur il s'agirait plutôt de $\mathbb{P}(\{\omega\})$.

Avant d'aller plus loin, donnons un exemple de probabilité.

Exemple : Soit $a \in \Omega$. On définit l'application $\mathbb{P}_a : \mathcal{P}(\Omega) \rightarrow [0; 1]$ par :

$$\forall A \in \mathcal{P}(\Omega), \quad \mathbb{P}_a(A) = \begin{cases} 1 & \text{si } a \in A \\ 0 & \text{si } a \notin A \end{cases}$$

Montrons que \mathbb{P}_a est une probabilité.

La deuxième propriété est appelée additivité. Le terme « incompatibles » doit absolument apparaître quand on utilise cette propriété !

Il est pratique d'avoir en tête que :

$$\mathbb{P}(A) \neq 0 \Leftrightarrow \mathbb{P}(A) > 0$$

et

$$\mathbb{P}(A) \neq 1 \Leftrightarrow \mathbb{P}(A) < 1.$$

On l'appelle la probabilité de Dirac (en a). C'est par exemple la probabilité naturelle dans le cadre d'un lancer de dé pipé. Par exemple, si on tombe sur la face 4 à chaque lancer, on munira l'espace probabilisable $(\Omega, \mathcal{P}(\Omega))$ de la probabilité \mathbb{P}_4 .

Définition. On appelle espace probabilisé fini tout triplet $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$ où

- Ω est un ensemble fini (non vide) appelé univers.
- $\mathcal{P}(\Omega)$ est son ensemble des parties, appelé ensemble des événements.
- \mathbb{P} est une probabilité sur Ω (définie sur $\mathcal{P}(\Omega)$ donc).

On se donne dans la suite du chapitre un espace probabilisé fini $(\Omega, \mathcal{P}(\Omega), \mathbb{P})$.

b) Propriétés

Proposition. On a $\mathbb{P}(\emptyset) = 0$.

DÉMONSTRATION.

□

Un couple $(\Omega, \mathcal{P}(\Omega))$ est appelé un espace probabilisable... il n'attend qu'à être « probabilisé » par l'ajout d'une probabilité sur Ω . Il peut sembler surprenant de mettre $\mathcal{P}(\Omega)$ dans l'écriture. Vous comprendrez cela l'an prochain où un espace probabilisé sera un triplet $(\Omega, \mathcal{A}, \mathbb{P})$ où \mathcal{A} est un ensemble d'événements de Ω (vérifiant certaines propriétés) appelé tribu sur Ω .

Par récurrence, on obtient :

Proposition (additivité finie). Si A_1, \dots, A_n sont des événements **deux à deux incompatibles**, alors

$$\mathbb{P}\left(\bigcup_{k=1}^n A_k\right) = \sum_{k=1}^n \mathbb{P}(A_k).$$

↪ DÉMONSTRATION LAISSÉE EN EXERCICE.

Exemple : On lance un dé truqué de telle sorte que les probabilités de tomber sur 1, 2, 3, 4, 5, 6 sont respectivement $1/12, 1/6, 1/4, 1/12, 1/3, 1/12$.

Pour tout $k \in \llbracket 1; 6 \rrbracket$, posons A_k : « tomber sur la $k^{\text{ième}}$ face ». Ces six événements sont deux à deux incompatibles. Ainsi la probabilité de tomber sur un chiffre impair est

On verra dans le paragraphe 1.5.c que, puisque la somme de ces six probabilités est égale à 1, on peut bien construire un espace probabilisé qui modélise cette expérience

Proposition. Soit $n \in \mathbb{N} \setminus \{0; 1\}$. Soit (A_1, \dots, A_n) un système complet d'événements.

$$\text{Alors } \sum_{k=1}^n \mathbb{P}(A_k) = 1.$$

DÉMONSTRATION.

Attention, la réciproque est fausse !

□

Proposition. Soient A , B et C des événements. Nous avons :

1. $\mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A \cap B)$.
2. $\mathbb{P}(\bar{A}) = 1 - \mathbb{P}(A)$.
3. Si $A \subset B$, alors $\mathbb{P}(A) \leq \mathbb{P}(B)$.
4. **Formule de Poincaré pour deux événements.**

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

5. **Formule de Poincaré pour trois événements.**

$$\begin{aligned} \mathbb{P}(A \cup B \cup C) &= \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C) \\ &\quad - \mathbb{P}(A \cap B) - \mathbb{P}(A \cap C) - \mathbb{P}(B \cap C) + \mathbb{P}(A \cap B \cap C). \end{aligned}$$

En fait il suffit de vérifier qu'une probabilité est à valeurs dans \mathbb{R}_+ (au lieu de $[0; 1]$) dans la définition d'une probabilité puisque, si A est un événement, $1 - \mathbb{P}(A) = \mathbb{P}(\bar{A}) \geq 0$ et donc $\mathbb{P}(A) \leq 1$.

DÉMONSTRATION.

La formule de Poincaré peut se généraliser, de proche en proche, à l'aide des opérations usuelles sur les ensembles. La formule générale (celle qui « développe » $\mathbb{P}(A_1 \cup \dots \cup A_n)$ pour tout $n \in \mathbb{N}^*$) est appelé formule du crible (elle est hors programme mais se démontre comme la formule du crible vue dans la feuille d'exercices n° 33).

□

Exemple : On tire une carte au hasard dans un jeu de 52 cartes. Quelle est la probabilité que la carte ne soit ni un roi ni un coeur ?

Notons C l'événement « obtenir un coeur » et R l'événement « obtenir un roi ».

Nous verrons dans le prochain paragraphe (consacré aux situations d'équiprobabilité) que l'on peut construire un espace probabilisé fini modélisant cette expérience tel que la probabilité d'obtenir un pique est $\frac{13}{52} = \frac{1}{4}$, la probabilité d'obtenir un roi est $\frac{4}{52} = \frac{1}{13}$ et la probabilité de tirer le roi de pique est $\frac{1}{52}$.

Proposition. Soit $n \in \mathbb{N}^*$. Si A_1, \dots, A_n sont des événements, alors

$$\mathbb{P}\left(\bigcup_{k=1}^n A_k\right) \leq \sum_{k=1}^n \mathbb{P}(A_k).$$

DÉMONSTRATION. Raisonnons par récurrence. Si $n = 1$, il est immédiat que $\mathbb{P}(A_1) \leq \mathbb{P}(A_1)$ pour tout $A_1 \in \mathcal{P}(\Omega)$. Soit $n \in \mathbb{N}^*$. Supposons la formule vraie pour n événements. Donnons-nous A_1, \dots, A_{n+1} événements. La formule de Poincaré entraîne que

La différence avec l'additivité finie est qu'on ne suppose plus que les événements sont deux à deux incompatibles.

$$\begin{aligned} \mathbb{P}\left(\bigcup_{k=1}^{n+1} A_k\right) &= \mathbb{P}\left(\left(\bigcup_{k=1}^n A_k\right) \cup A_{n+1}\right) \\ &= \mathbb{P}\left(\bigcup_{k=1}^n A_k\right) + \mathbb{P}(A_{n+1}) - \mathbb{P}\left(\left(\bigcup_{k=1}^n A_k\right) \cap A_{n+1}\right) \\ &\leq \mathbb{P}\left(\bigcup_{k=1}^n A_k\right) + \mathbb{P}(A_{n+1}) \\ &\leq \mathbb{P}\left(\bigcup_{k=1}^{n+1} A_k\right) \leq \sum_{k=1}^n \mathbb{P}(A_k) + \mathbb{P}(A_{n+1}). \end{aligned}$$

Par hypothèse de récurrence.

D'où la formule pour n événements. On conclut par récurrence. □

Remarque : On a $\mathbb{P}(\emptyset) = 0$ mais on a vu avec l'exemple de la probabilité de Dirac que la réciproque est fautive : on peut avoir $\mathbb{P}(A) = 0$ et $A \neq \emptyset$. De plus, toujours avec cet exemple, on voit qu'on peut avoir $\mathbb{P}(A) = 1$ et $A \neq \Omega$. D'où la définition suivante :

Définition. Un événement A est dit :

- négligeable ou presque impossible si $\mathbb{P}(A) = 0$.
- presque certain ou presque sûr si $\mathbb{P}(A) = 1$.

Lorsque Ω est fini, on s'arrange cependant en général pour que seul l'événement \emptyset soit de probabilité nulle et que seul Ω soit de probabilité égale à 1. Ce ne sera plus du tout le cas en deuxième année.

c) Caractérisation et construction de probabilités

Proposition. Supposons que $\Omega = \{\omega_1; \dots; \omega_n\}$ avec $n = \text{card}(\Omega) \in \mathbb{N}^*$. Alors

$$\forall A \in \mathcal{P}(\Omega), \quad \mathbb{P}(A) = \sum_{k=1}^n \mathbb{P}(A \cap \{\omega_k\}) = \sum_{\substack{1 \leq k \leq n \\ \omega_k \in A}} \mathbb{P}(\{\omega_k\}) = \sum_{\omega \in A} \mathbb{P}(\{\omega\}).$$

En particulier $\sum_{k=1}^n \mathbb{P}(\{\omega_k\}) = 1$.

Ce que l'on résume grossièrement par : la somme des probabilités vaut 1.

DÉMONSTRATION.

□

On se pose à présent la question inverse : si on dispose de réels positifs dont la somme vaut 1, peut-on toujours construire un univers et une probabilité sur cet univers de sorte que ces réels soient les probabilités des événements élémentaires ?

Définition. Soit E un ensemble fini. Une distribution de probabilités sur E est une famille d'éléments de \mathbb{R}_+ indexée par E et de somme 1.

Théorème. Pour toute distribution de probabilités $(p_\omega)_{\omega \in \Omega}$ sur Ω , il existe une unique probabilité $\mathbb{P} : \mathcal{P}(\Omega) \rightarrow [0; 1]$ telle que, pour tout $\omega \in \Omega$, $\mathbb{P}(\{\omega\}) = p_\omega$.

Remarque : En d'autres termes, une probabilité \mathbb{P} sur Ω est entièrement déterminée par la distribution de probabilités $(\mathbb{P}(\{\omega\}))_{\omega \in \Omega}$. Dit autrement, une probabilité \mathbb{P} définie sur un espace probabilisable fini est entièrement déterminée par la donnée des probabilités de chaque événement élémentaire de Ω .

Encore en d'autres termes : pour définir totalement une probabilité, il suffit de connaître $\mathbb{P}(\{\omega\})$, pour tout $\omega \in \Omega$.

DÉMONSTRATION. Soit $(p_\omega)_{\omega \in \Omega}$ une distribution de probabilité sur Ω . Notons $\Omega = \{\omega_1, \dots, \omega_n\}$ et posons, pour tout $k \in \llbracket 1; n \rrbracket$, $p_k = p_{\omega_k}$.

• **Existence.** Définissons sur $\mathcal{P}(\Omega)$ la fonction \mathbb{P} par

$$\forall A \in \mathcal{P}(\Omega), \quad \mathbb{P}(A) = \sum_{\substack{1 \leq k \leq n \\ \omega_k \in A}} p_k.$$

★ Il s'agit bien d'une application sur $\mathcal{P}(\Omega)$.

★ On a $\mathbb{P}(\Omega) = \sum_{k=1}^n p_k = 1$.

★ Soit A un événement. Comme p_1, \dots, p_n sont positifs, $0 \leq \mathbb{P}(A) \leq \sum_{k=1}^n p_k = 1$.

★ Donnons-nous A et B deux événements incompatibles, alors

$$\mathbb{P}(A \cup B) = \sum_{\substack{1 \leq k \leq n \\ \omega_k \in A \cup B}} p_k = \sum_{\substack{1 \leq k \leq n \\ \omega_k \in A}} p_k + \sum_{\substack{1 \leq k \leq n \\ \omega_k \in B}} p_k = \mathbb{P}(A) + \mathbb{P}(B).$$

Ici on a utilisé la propriété de sommation par paquet puisque

$$A \cap B = \emptyset.$$

Ainsi \mathbb{P} est une probabilité sur $(\Omega, \mathcal{P}(\Omega))$.

- **Unicité.** Soient \mathbb{P}_1 et \mathbb{P}_2 deux probabilités sur $(\Omega, \mathcal{P}(\Omega))$ telles que, pour tout $\omega \in \Omega$, $\mathbb{P}_1(\{\omega\}) = \mathbb{P}_2(\{\omega\})$. La proposition précédente assure que, pour tout $A \in \mathcal{P}(\Omega)$,

$$\mathbb{P}_1(A) = \sum_{\omega \in A} \mathbb{P}_1(\{\omega\}) = \sum_{\omega \in A} \mathbb{P}_2(\{\omega\}) = \mathbb{P}_2(A).$$

D'où l'unicité. □

Exemple : On lance un dé truqué tel que chaque face a une probabilité de tomber qui est proportionnelle au numéro de la face. Quelle est la probabilité de tomber sur un nombre impair ?

d) Équiprobabilité

Proposition. Sur tout espace probabilisable fini $(\Omega, \mathcal{P}(\Omega))$, il existe une unique probabilité \mathbb{P} prenant la même valeur sur les événements élémentaires. On appelle cette probabilité la probabilité uniforme (ou équiprobabilité) sur Ω . Nous avons

$$\forall A \in \mathcal{P}(\Omega), \quad \mathbb{P}(A) = \frac{\text{card}(A)}{\text{card}(\Omega)}.$$

En particulier, pour tout $\omega \in \Omega$,

$$\mathbb{P}(\{\omega\}) = \frac{1}{\text{card}(\Omega)}.$$

DÉMONSTRATION. Si une telle probabilité \mathbb{P} existe, notons p la valeur commune de la probabilité de chaque événement élémentaire. On a alors

$$1 = \sum_{\omega \in \Omega} \mathbb{P}(\{\omega\}) = \sum_{\omega \in \Omega} p = p \text{ card}(\Omega).$$

Ainsi $p = \frac{1}{\text{card}(\Omega)}$. Les réels de la famille $(p)_{\omega \in \Omega}$ étant positifs et de somme 1, le théorème précédent nous assure l'existence d'une unique probabilité sur $(\Omega, \mathcal{P}(\Omega))$ telle que, pour tout $\omega \in \Omega$, $\mathbb{P}(\{\omega\}) = \frac{1}{\text{card}(\Omega)}$.

⚠ Si Ω est infini (ce qui pourra être le cas l'an prochain), il n'y a pas de probabilité uniforme.

Enfin, pour tout $A \in \mathcal{P}(\Omega)$, on a

$$\mathbb{P}(A) = \sum_{\omega \in A} \mathbb{P}(\{\omega\}) = \sum_{\omega \in A} \frac{1}{\text{card}(\Omega)} = \frac{\text{card}(A)}{\text{card}(\Omega)}. \quad \square$$

Remarques :

- Munir un espace probabilisable de la probabilité uniforme permet de modéliser une expérience aléatoire pour laquelle tous les événements élémentaires ont la même chance de réalisation. On dit que les événements élémentaires sont équiprobables. C'est l'hypothèse que l'on fera pour des lancers de dés ou de pièces non truqués (ou équilibrées), des tirages de cartes ou de boules numérotées dans une urne, etc.
- Dans le cas de l'équiprobabilité, on a $\mathbb{P}(A) = \ll \frac{\text{nombre de cas favorables}}{\text{nombre total de cas}} \gg$.

Exemples :

- On lance deux fois un dé équilibré à six faces. Quelle est la probabilité de l'événement A : « la somme des deux chiffres obtenus est 8 ».

- On lance n fois une pièce équilibrée. Quelle est la probabilité d'obtenir au moins deux Pile ou deux Face consécutifs ?

Parfois, calculer $\mathbb{P}(\bar{A})$ est plus simple que calculer $\mathbb{P}(A)$. Y penser quand il y a un « au moins » ou un « au plus ».

- **Un grand classique : le problème des anniversaires.** Quelle est la probabilité que, dans une classe de n élèves (avec $n \in \llbracket 2; 365 \rrbracket$), au moins deux partagent la même date d'anniversaire ?

Si $n > 365$, il est sûr qu'au moins deux élèves partagent la même date d'anniversaire (car sinon 366 élèves auraient une date d'anniversaire différente) d'après le principe des tiroirs.

On calcule facilement ce produit à l'aide de Python. On trouve par exemple que, dans une classe de 23 élèves, il y a plus d'une chance sur deux qu'au moins deux élèves partagent la même date de naissance. Dans une classe de 41 élèves, cette probabilité dépasse 90%.


Remarque : Nous l'avons déjà dit mais en général Ω est bien trop compliqué pour être décrit mathématiquement et la plupart des énoncés se contentent d'admettre qu'il existe un univers décrivant l'expérience étudiée (voire ne disent rien du tout). Si on ne connaît pas précisément Ω , on ne risque pas de savoir construire rigoureusement une probabilité

Si l'énoncé ne demande pas clairement de construire un espace probabilisé, ne vous y risquez pas !

qui modélise l'expérience, y compris en situation d'équiprobabilité : si on ne connaît pas le cardinal de Ω , on ne peut pas utiliser la formule « $\frac{\text{card}(A)}{\text{card}(\Omega)}$ ». Parfois seule une petite partie de l'expérience présente une situation d'équiprobabilité.

De la formule « $\frac{\text{nombre de cas favorables}}{\text{nombre total de cas}}$ » dans un cas d'équiprobabilité, on retient néanmoins le principe suivant :

Soient $N \in \mathbb{N}^*$ et $a \in \llbracket 1; N \rrbracket$. Considérons N objets distincts dont un nombre a d'entre eux vérifient une caractéristique C . Si on choisit au hasard (de façon équiprobable) un objet parmi ces N objets, alors la probabilité que l'objet tiré possède la caractéristique C est $\frac{a}{N}$.

 L'équiprobabilité semble la plus naturelle quand on travaille sur un univers fini mais ce n'est pas toujours le cas ! Par exemple, on a déjà vu que la probabilité de Dirac est la plus naturelle dans le cadre d'un lancer de dé pipé. Donnons deux autres exemples.

Exemple : On tire deux fois à Pile ou Face. On cherche la probabilité d'obtenir au moins un Pile. Notons A cet événement.

- Premier cas. On prend $\Omega = \{(P, P); (P, F); (F, P); (F, F)\}$ muni de l'équiprobabilité. On a $A = \{(P, F); (F, P); (P, P)\}$ donc $\mathbb{P}(A) = \frac{3}{4}$.
- Deuxième cas. On peut ne considérer que 3 éventualités différentes, à savoir :
 - ★ Pile au premier coup, ce qui dispense d'en jouer un deuxième.
 - ★ Face au premier coup, Pile au deuxième.
 - ★ Face aux deux coups.

L'univers est alors $\tilde{\Omega} = \{P; (F, P); (F, F)\}$. Si on met l'équiprobabilité sur $\tilde{\Omega}$, on obtient $\mathbb{P}(A) = \frac{2}{3}$, ce qui n'est pas le résultat attendu. Le problème vient du fait qu'obtenir Pile au premier coup est de probabilité $\frac{1}{2}$ tandis que la probabilité des autres éventualités est $\frac{1}{4}$. Il est plus naturel de munir $(\tilde{\Omega}, \mathcal{P}(\tilde{\Omega}), \mathbb{P})$ où \mathbb{P} est la probabilité définie par $\mathbb{P}(\{P\}) = \frac{1}{2}$ et $\mathbb{P}(\{(F, P)\}) = \mathbb{P}(\{(F, F)\}) = \frac{1}{4}$. On retrouve bien $\mathbb{P}(A) = \frac{3}{4}$.

Exemple : On reprend l'exemple du lancer de deux dés indiscernables du paragraphe 1.2. On peut prendre $\Omega = \llbracket 1; 6 \rrbracket^2$ muni de $\mathcal{P}(\Omega)$ et de l'équiprobabilité ou bien prendre

$$\tilde{\Omega} = \{\{a; b\} \mid (a, b) \in \llbracket 1; 6 \rrbracket^2\}$$

mais alors, si on met l'équiprobabilité sur $\tilde{\Omega}$, on obtient que la probabilité d'obtenir deux 6 est $\frac{1}{21}$ alors qu'il est « évident » qu'elle vaut $\frac{1}{36}$. Là encore, l'équiprobabilité n'est pas la plus naturelle.

La probabilité la plus naturelle consiste à poser $\mathbb{P}(\{a\}) = \frac{1}{36}$ pour tout $a \in \llbracket 1; 6 \rrbracket$ et $\mathbb{P}(\{a; b\}) = \frac{2}{36}$ pour tous $(a, b) \in \llbracket 1; 6 \rrbracket^2$ tel que $a \neq b$.

II Probabilité conditionnelle

1) Définition et premiers exemples

Si on considère deux événements A et B , comment la connaissance de l'information « A est réalisé » modifie-t-elle la probabilité ?

Par exemple, on lance un dé équilibré. Notons B l'événement « obtenir 6 ». On a alors $\mathbb{P}(B) = \frac{1}{6}$, mais si on sait que A : « on obtient un nombre pair » est réalisé, alors la probabilité devient égale à $\frac{1}{3}$ car seuls les numéros 2, 4 et 6 sont alors possibles.

Comment rendre compte de ce phénomène rigoureusement ? Appuyons nous sur un exemple pour motiver la définition de la notion de probabilité conditionnelle.

On lance deux fois une pièce équilibrée. On prend $\Omega = \{P, F\}^2$ muni de $\mathcal{P}(\Omega)$ et de l'équiprobabilité. Notons A : « Face est sorti » et B : « Pile est sorti ». On a

$$A = \{(F, F); (F, P); (P, F)\} \quad \text{et} \quad B = \{(P, P); (P, F); (F, P)\}.$$

On a donc $\mathbb{P}(A) = \mathbb{P}(B) = \frac{3}{4}$. On effectue à présent N répétitions (indépendantes les unes des autres) de notre expérience aléatoire. On va noter les résultats obtenus sous forme de tableau :

A	1	0	0	1	1	...
B	0	1	1	1	0	...

où chaque colonne correspond à une répétition de l'expérience et on met 1 dans la ligne A (respectivement la ligne B) si A est réalisé (respectivement B) et 0 sinon. La fréquence conditionnelle de B sachant que A est réalisé est la fréquence de B calculée uniquement à l'aide des répétitions lors desquelles A est réalisé. Elle est donc donnée par

$$\frac{\text{Nombre de colonnes} \binom{1}{1}}{\text{Nombre de colonnes} \binom{1}{1} \text{ ou } \binom{1}{0}} = \frac{N(A \cap B)}{N(A)} = \frac{N(A \cap B)}{N} \xrightarrow{N \rightarrow +\infty} \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)}$$

Cela justifie la définition suivante :

Définition. Soient A et B deux événements tels que $\mathbb{P}(A) \neq 0$. On appelle probabilité conditionnelle de B sachant A le réel

$$\mathbb{P}_A(B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)}.$$

Remarques :

- En particulier, lorsque $\mathbb{P}(A) \neq 0$, on a $\mathbb{P}(A \cap B) = \mathbb{P}(A) \times \mathbb{P}_A(B)$.
- Le nombre $\mathbb{P}_A(B)$ représente la probabilité de B calculée du point de vue d'un observateur qui arriverait en cours d'expérience au moment où A vient de se réaliser. Il dispose de plus d'informations qu'un observateur qui assiste à l'expérience depuis son début : pour lui « l'univers devient A ».
- Souvent on connaît $\mathbb{P}_A(B)$ et non pas $\mathbb{P}(A \cap B)$. Il ne faut surtout pas les confondre : $\mathbb{P}(A \cap B)$ est la probabilité que A et B soient réalisés en même temps, tandis que $\mathbb{P}_A(B)$ est la probabilité de B quand on **sait** que A est réalisé.

Exemples :

- Dans le premier exemple ci-dessus...

- Dans le deuxième exemple ci-dessus...

- On dispose de trois boîtes fermées. Un objet a une chance sur deux de se trouver dans l'une des boîtes (équiprobablement). On a ouvert les deux premières boîtes sans trouver l'objet : il n'y est pas. Quelle est la probabilité qu'il soit dans la troisième boîte ?

On calcule la fréquence (et plus tard la probabilité) en réduisant l'univers au « sous-univers » A . Pour simplifier : conditionner par A revient à faire une sorte de zoom sur A : il ne s'agit pas de supprimer ce qui est en dehors de A mais d'attribuer une probabilité nulle au complémentaire de A , et de multiplier les probabilités des événements de A par un même facteur pour que leur somme reste égale à 1.

On rencontre aussi la notation $\mathbb{P}(B|A)$.

En d'autres termes, parmi les couples de lancers comportant au moins un Face, on trouve une proportion (environ) égale à $2/3$ de couples comportant un Pile et un Face (dans n'importe quel ordre) et une proportion égale à $1/3$ de couples comportant deux Face.

Encore un exemple où la probabilité naturelle n'est pas l'équiprobabilité.

• **Le « paradoxe » des deux enfants.** David a deux enfants (je n'en sais pas plus). Je sonne chez lui et l'un des deux enfants m'ouvre : c'est une fille. Quelle est la probabilité que l'autre enfant soit une fille ?

Au moment où une fille de David m'ouvre, je sais qu'il ne reste plus que trois configurations équiprobables de familles : Fille/Fille, Fille/-Garçon, Garçon/Fille.

Surprenant ? C'est pourtant le même résultat que pour des Piles et des Faces ci-dessus (et ça n'avait choqué personne) : parmi les couples avec deux enfants dont une fille, il y a une proportion égale à 2/3 de couples avec un garçon et une fille, et une proportion de 1/3 de couples avec deux filles.

Les résultats faisant intervenir des probabilités conditionnelles sont souvent contre-intuitifs !

Proposition. Soit A un événement de probabilité non nulle. L'application

$$P_A : \begin{cases} \mathcal{P}(\Omega) & \longrightarrow & [0, 1] \\ B & \longmapsto & \mathbb{P}_A(B) \end{cases}$$

est une probabilité sur $(\Omega, \mathcal{P}(\Omega))$.

DÉMONSTRATION.

Plus généralement, si B est un événement tel que $A \subset B$, alors $\mathbb{P}_A(B) = 1$. En particulier $\mathbb{P}_A(A) = 1$.

Puisque \mathbb{P}_A est une probabilité, elle vérifie toutes les propriétés vues dans le paragraphe 1.5, notamment :


- Si B est un événement, alors

$$\mathbb{P}_A(\overline{B}) = 1 - \mathbb{P}_A(B).$$
- (formule de Poincaré) Si B et C sont des événements, alors

$$\mathbb{P}_A(B \cup C) = \mathbb{P}_A(B) + \mathbb{P}_A(C) - \mathbb{P}_A(B \cap C).$$

Exemple : Dans l'exemple avec les trois boîtes, on munit $(\Omega, \mathcal{P}(\Omega))$ de la probabilité définie par :

$$\mathbb{P}_A(\{\text{absent}\}) = \frac{3}{4}, \mathbb{P}_A(\{\text{boite 1}\}) = \mathbb{P}_A(\{\text{boite 2}\}) = 0 \text{ et } \mathbb{P}_A(\{\text{boite 3}\}) = \frac{1}{4}$$

 La notion de probabilité conditionnelle utilise l'expression « probabilité de B sachant A » qui donne l'impression qu'il y a un ordre temporel, que la réalisation de l'événement A précède la réalisation de B . Il n'en est rien !

Par exemple, si on regarde deux lancers consécutifs d'une même pièce, on peut prendre B : « Pile sort lors du premier lancer », A : « Pile sort lors du seconde lancer », et on peut tout à fait calculer $\mathbb{P}_A(B)$.

2) Formule des probabilités composées

Généralement il est plus facile de déterminer la probabilité conditionnelle que de déterminer la probabilité d'une intersection. Si A et B sont des événements tels que $\mathbb{P}(A) \neq 0$, alors

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \times \mathbb{P}_A(B).$$

Voici une formule qui généralise cela à l'intersection d'un nombre fini d'événements.

Théorème (formule des probabilités composées). Soit $n \in \mathbb{N} \setminus \{0; 1\}$. Soient A_1, \dots, A_n des événements tels que $\mathbb{P}(A_1 \cap \dots \cap A_n) \neq 0$. Alors

$$\mathbb{P}(A_1 \cap \dots \cap A_n) = \mathbb{P}(A_1) \times \mathbb{P}_{A_1}(A_2) \times \mathbb{P}_{A_1 \cap A_2}(A_3) \times \dots \times \mathbb{P}_{A_1 \cap \dots \cap A_{n-1}}(A_n).$$

DÉMONSTRATION. Pour tout $k \in [1; n-1]$, $A_1 \cap \dots \cap A_n \subset A_1 \cap \dots \cap A_k$ donc

$$0 < \mathbb{P}(A_1 \cap \dots \cap A_n) \leq \mathbb{P}(A_1 \cap \dots \cap A_k).$$

Ainsi toutes les probabilités conditionnelles intervenant dans la formule sont bien définies. Montrons-la par récurrence :

- **Initialisation.** Par définition, si A_1 et A_2 sont deux événements tels que $\mathbb{P}(A_1 \cap A_2) \neq 0$, on a $\mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_1) \times \mathbb{P}_{A_1}(A_2)$ donc la formule est vraie au rang 2.
- **Hérédité.** Soit $n \geq 2$. Supposons la formule vraie au rang n . Soient A_1, \dots, A_n et A_{n+1} des événements tels que $\mathbb{P}(A_1 \cap \dots \cap A_{n+1}) \neq 0$. On a

$$\mathbb{P}(A_1 \cap \dots \cap A_{n+1}) = \mathbb{P}(A_1 \cap \dots \cap A_n) \times \mathbb{P}_{A_1 \cap \dots \cap A_n}(A_{n+1})$$

et l'hypothèse de récurrence entraîne alors que

$$\mathbb{P}(A_1 \cap \dots \cap A_{n+1}) = \mathbb{P}(A_1) \times \mathbb{P}_{A_1}(A_2) \times \dots \times \mathbb{P}_{A_1 \cap \dots \cap A_{n-1}}(A_n) \times \mathbb{P}_{A_1 \cap \dots \cap A_n}(A_{n+1}).$$

Ainsi la formule est vraie au rang $n+1$

D'où la proposition par récurrence. \square

Exemples :

- On considère une urne qui contient 10 boules dont 6 rouges et 4 bleues. On tire successivement 3 boules avec les règles suivantes :

- ★ Si on tire une boule rouge, alors on l'enlève.
- ★ Si on tire une boule bleue, on la remplace par une boule rouge.

Quelle est la probabilité de tirer une boule bleue, une boule rouge puis une boule bleue (dans cet ordre) ?

Cadre d'utilisation : on utilise cette formule pour calculer la probabilité d'une intersection d'événements qui ne sont pas indépendants (cf. paragraphe III), typiquement quand on procède à des tirages et quand la composition de l'urne change à chaque étape : voir l'exemple ci-dessous.

Il n'y a pas de notion d'indépendance (cf. paragraphe III) ici car la composition de l'urne à une étape dépend des boules tirées lors des étapes précédentes.

- On dispose d'un jeu de 52 cartes. On tire des cartes successivement et sans remise jusqu'à ce que l'on obtienne un trèfle. Soit $k \in \llbracket 1 ; 40 \rrbracket$. Quelle est la probabilité de réaliser k tirages ?

Forcément, vu qu'il y a 13 trèfles, on va finir par en obtenir un en moins de 40 tirages.

3) Formule des probabilités totales

La formule des probabilités totales est la formule la plus importante de ce chapitre. Avant de l'énoncer, adoptons une convention :

Lorsque $\mathbb{P}(A) = 0$, on convient que $\mathbb{P}(A)\mathbb{P}_A(B) = 0$.

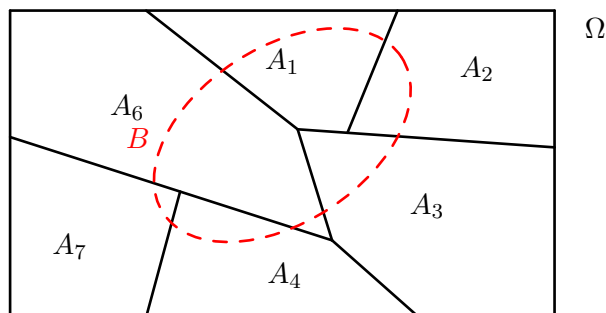
La raison d'être et le bien-fondé de cette convention seront justifiés dans les remarques ci-dessous.

Théorème (formule des probabilités totales). Soit $n \in \mathbb{N} \setminus \{0; 1\}$. Soit (A_1, \dots, A_n) un système complet d'événements tels que, pour tout $k \in \llbracket 1; n \rrbracket$, $\mathbb{P}(A_k) \neq 0$. Alors, pour tout événement B ,

$$\mathbb{P}(B) = \sum_{k=1}^n \mathbb{P}(B \cap A_k) = \sum_{k=1}^n \mathbb{P}(A_k) \mathbb{P}_{A_k}(B).$$

Remarques :

- Il faut absolument appliquer cette formule à une famille $(A_k)_{1 \leq k \leq n}$ qui est un système complet d'événements. Tous les indices $k \in \llbracket 1; n \rrbracket$ doivent figurer dans la somme, même on sait d'avance que $\mathbb{P}_{A_k}(B) = 0$. C'est seulement dans un second temps que l'on peut enlever les termes nuls de la somme.
- La formule des probabilités totales permet en quelque sorte de faire une disjonction de cas dont la réalisation est aléatoire. Supposons que l'on connaisse la probabilité de B dans le cas où on a obtenu A_1 (i.e. $\mathbb{P}_{A_1}(B)$), la probabilité de B dans le cas où on a obtenu A_2 (i.e. $\mathbb{P}_{A_2}(B)$), etc. et enfin la probabilité de B dans le cas où on a obtenu A_n (i.e. $\mathbb{P}_{A_n}(B)$). Si les n cas sont disjoints et couvrent l'ensemble des possibilités (autrement dit si (A_1, \dots, A_n) est un système complet d'événements), alors la probabilité de B est obtenue via la formule des probabilités totales.



Cadre d'utilisation : on utilise cette formule quand il y a plusieurs cas de figure et que, dans chacun de ses cas, on sait calculer la probabilité (conditionnelle au cas).

Cela se comprend très bien avec un dessin : la probabilité de B est la somme des probabilités de « chaque morceau » i.e. la somme des probabilités de chaque intersection de B avec un A_i .

- Un mot sur la convention avant de procéder à la démonstration : supposons que $\mathbb{P}(A) = 0$ et soit $B \in \mathcal{P}(\Omega)$. On a $A \cap B \subset A$ donc $\mathbb{P}(A \cap B) \leq 0$ et donc $\mathbb{P}(A \cap B) = 0$. Dès lors, bien que $\mathbb{P}_A(B)$ ne soit pas défini, on pourrait lui attribuer n'importe quelle valeur réelle arbitraire et on aurait alors $0 = 0 \times \mathbb{P}_A(B)$ donc $\mathbb{P}(A \cap B) = \mathbb{P}(A) \mathbb{P}_A(B)$. Dans la somme de la formule des probabilités totales, les termes faisant intervenir des A_k de probabilité nulle ne sont donc pas réellement présents. La raison d'être de cette convention et que, dans la pratique, il n'est pas toujours facile de s'assurer qu'une probabilité est nulle ou non avant d'appliquer cette formule.

On arrive parfois même à donner un sens à des probabilités conditionnelles sachant un événement qu'on sait de probabilités nulle.

DÉMONSTRATION.

[Empty box for the proof of the theorem]

□

Retenons le cas particulier d'un système complet de deux événements :

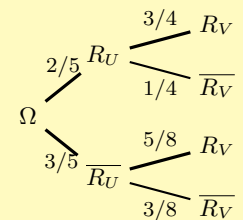
Théorème (formule des probabilités totales). Soit A un événement tel que $\mathbb{P}(A) \in]0; 1[$. Pour tout événement B , nous avons

$$\mathbb{P}(B) = \mathbb{P}(A \cap B) + \mathbb{P}(\bar{A} \cap B) = \mathbb{P}(A)\mathbb{P}_A(B) + \mathbb{P}(\bar{A})\mathbb{P}_{\bar{A}}(B).$$

Exemple : On dispose d'une urne U contenant 4 boules rouges et 6 boules bleues ainsi que d'une urne V contenant 5 boules rouges et 2 boules bleues. On tire une boule dans l'urne U (de façon équiprobable) et, sans la regarder on la met dans l'urne V . On tire ensuite une boule dans l'urne V (de façon équiprobable). Quelle est la probabilité qu'elle soit rouge ?

Si on sait quelle boule on a tirée dans l'urne U alors on connaît la composition de l'urne V et donc c'est trivial : il y a plusieurs cas particuliers, et dans chaque cas particulier on sait répondre. On pense donc à la formule des probabilités totales.

Au lycée, pour traiter cet exemple, on aurait rédigé ainsi :



Pour obtenir la probabilité de R_V , on aurait ajouté les probabilités des chemins menant à R_V , la probabilité d'un chemin étant obtenue en multipliant entre elles les probabilités conditionnelles placées sur les arêtes du chemin. Cette règle de calcul n'est rien d'autre que la formule des probabilités totales en moins rigoureux. Désormais on **ARRÊTE** de raisonner avec des arbres et on rédige avec des formules.

Exemple : Soit $n \in \mathbb{N}^*$. On dispose de n pièces de monnaie truquées de telle sorte que, pour tout $k \in \llbracket 1; n \rrbracket$, la $k^{\text{ième}}$ pièce a une probabilité $\frac{k}{n}$ de tomber sur Pile. On choisit une pièce au hasard et on la lance. Quelle est la probabilité d'obtenir Face ?

On pense à la formule des probabilités totales car, si on sait quelle pièce on a tirée, alors c'est trivial : il y a plusieurs cas particuliers, et dans chaque cas particulier on sait répondre.



Même si on remarque d'avance que $\mathbb{P}_{A_n}(F) = 0$, on fait tout de même figurer ce terme dans la formule avant de remarquer qu'il est nul. On applique impérativement la formule des probabilités totales à un système complet d'événements (ce que $(A_k)_{1 \leq k \leq n-1}$ n'est pas).

4) Formule de Bayes

Théorème (formule de Bayes). Soient A et B deux événements qui sont tous les deux de probabilité non nulle. Alors

$$\mathbb{P}_A(B) = \frac{\mathbb{P}(B)\mathbb{P}_B(A)}{\mathbb{P}(A)}.$$

DÉMONSTRATION. On a $\mathbb{P}_A(B)\mathbb{P}(A) = \mathbb{P}(A \cap B) = \mathbb{P}(B \cap A) = \mathbb{P}_B(A)\mathbb{P}(B)$. \square

Exemples :

- Reprenons l'exemple du paragraphe précédent avec les urnes U et V . On vient d'obtenir une boule rouge. Quelle est la probabilité que la boule tirée dans l'urne U soit aussi rouge ?


- Reprenons l'exemple du paragraphe précédent avec les n pièces truquées. On vient d'obtenir Face. Quelle est la probabilité d'avoir tiré la première pièce ?

Comme on le voit sur les exemples ci-dessus, on utilise souvent conjointement la formule des probabilités totales et la formule de Bayes : si A et B sont deux événements tels que $\mathbb{P}(B) \neq 0$ et $\mathbb{P}(A) \in]0; 1[$, alors

$$\mathbb{P}_A(B) = \frac{\mathbb{P}(B)\mathbb{P}_B(A)}{\mathbb{P}(A)\mathbb{P}_A(B) + \mathbb{P}(\bar{A})\mathbb{P}_{\bar{A}}(B)}.$$

Plus généralement, si (A_1, \dots, A_n) est un système complet d'événements, alors pour tout événement B de probabilité non nulle,


$$\forall i \in [1; n], \quad \mathbb{P}_B(A_i) = \frac{\mathbb{P}(A_i)\mathbb{P}_{A_i}(B)}{\sum_{k=1}^n \mathbb{P}(A_k)\mathbb{P}_{A_k}(B)}.$$

 La formule de Bayes montre que, en général, $\mathbb{P}_A(B) \neq \mathbb{P}_B(A)$. Penser que ces deux quantités sont égales est non seulement une erreur à éviter dans les épreuves de concours mais surtout une erreur à ne pas faire dans la vie de tous les jours. Examinons cela à travers un exemple supplémentaire :

On réalise un test sanguin pour détecter une maladie dont la probabilité d'occurrence est 0,5%. Si le patient est malade, le test détecte la maladie dans 95% des cas. Cependant ce test déclare malades (à tort) 1% des personnes saines. Quelle est la probabilité qu'un patient soit malade sachant que le test est positif ?

Ces formules générales ne sont pas à connaître par cœur. On les retrouve facilement en appliquant consécutivement la formule de Bayes puis la formule des probabilités totales pour le terme du dénominateur.

Ce n'est vrai que lorsque $\mathbb{P}(A) = \mathbb{P}(B)$.

 La plupart des gens répondraient (à tort) 95%.

Cet exemple montre qu'il faut se méfier de son intuition concernant le calcul de probabilités, surtout concernant les probabilités conditionnelles, et que l'usage des formules que l'on vient de voir s'impose systématiquement.

III Indépendance

1) Indépendance de deux événements

Intuitivement deux événements sont indépendants si la réalisation (ou non) de l'un d'entre eux n'affecte pas la probabilité que l'autre se réalise (ou non). Si A et B sont ces événements, cela se traduit par $\mathbb{P}_A(B) = \mathbb{P}(B)$... bien sûr si A est de probabilité non nulle. Pour éviter ce problème, voici la définition que l'on prendra pour l'indépendance d'événements :

Définition. On dit que deux événements A et B sont indépendants si $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.

Proposition. Soient A et B des événements tels que $\mathbb{P}(A) \neq 0$. Les événements A et B sont indépendants si et seulement si $\mathbb{P}_A(B) = \mathbb{P}(B)$.

Typiquement si on effectue deux lancers successifs d'une pièce ou d'un dé (ou deux tirages successifs avec remise de boules dans une urne), alors on fera l'hypothèse que les deux lancers sont indépendants : tout événement concernant le résultat du premier lancer seul sera indépendant d'un événement concernant le résultat du deuxième lancer seul.


Exemple : On lance deux fois de façon indépendante une pièce de monnaie telle que la probabilité d'obtenir Pile est $p \in]0; 1[$. Notons A l'événement « les deux lancers donnent le même résultat » et B l'événement « le deuxième lancer donne face ». A priori l'événement A semble dépendre de l'événement B ... analysons cela avec des calculs.

Posons F_1 l'événement « obtenir Face au premier lancer » et F_2 l'événement « obtenir Face au deuxième lancer ». On suppose que F_1 et F_2 sont indépendants.

Cadre d'utilisation de la formule de Bayes : quand on cherche une probabilité conditionnelle, et que la probabilité conditionnelle « miroir » est connue ou facile à calculer, c'est-à-dire si on cherche $\mathbb{P}_A(B)$ alors que l'on connaît $\mathbb{P}_B(A)$. Par exemple, ci-contre, on cherche $\mathbb{P}_T(M)$ et on connaît $\mathbb{P}_M(T)$.

donc $\frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)} = \mathbb{P}(B)$
et donc
 $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.

En pratique nous verrons que l'indépendance ne se démontre pas. En général l'indépendance est une hypothèse de modélisation qui permet de simplifier les calculs.

 L'indépendance de deux événements n'a rien à voir avec le fait que ces deux événements soient incompatibles ou non. D'ailleurs aucune des deux hypothèses n'entraîne l'autre a priori.

★ Quand on lance une pièce équilibrée, les événements A : « obtenir Pile » et B : « obtenir Face » sont incompatibles mais pas indépendants puisque $\mathbb{P}(A \cap B) = 0 \neq \frac{1}{4} = \mathbb{P}(A)\mathbb{P}(B)$.

★ Les événements « j'ai faim » et « ma voiture est grise » peuvent tout à fait arriver en même temps mais semblent totalement indépendants.

Plus précisément, l'incompatibilité est une notion ensembliste (elle est définie sur un espace probabilisable) alors que l'indépendance est une notion probabiliste (elle dépend de la probabilité considérée). En particulier, si A , B et C sont trois événements, il se peut que A et B soient indépendants pour \mathbb{P} mais pas pour \mathbb{P}_C (par définition, A et B sont indépendants pour \mathbb{P}_C si $\mathbb{P}_C(A \cap B) = \mathbb{P}_C(A)\mathbb{P}_C(B)$), on dit aussi que A et B sont indépendants conditionnellement à C .

Par exemple, si on dispose d'un dé équilibré et de deux pièces : une équilibrée et une telle que Face tombe deux fois plus souvent que Pile. On lance le dé et

★ Si on obtient 6, alors on lance deux fois la pièce équilibrée.

★ Sinon on lance deux fois la pièce truquée.

Notons A : « obtenir Pile au premier lancer », B : « obtenir Pile au second lancer » et C : « obtenir 6 au lancer de dé ». Nous pouvons raisonnablement supposer que A et B sont indépendants conditionnellement à C (et à \bar{C}). Par contre A et B ne sont pas indépendants. En effet :

Plus fort : si deux événements sont incompatibles, le fait que l'un est réalisé nous dit que l'autre ne l'est pas. Ainsi, intuitivement, on « voit » que deux événements incompatibles ne sont pas indépendants...

Là encore, si on sait quelle pièce on lance, calculer la probabilité que A ou B soit réalisé est enfantin : il y a plusieurs cas particuliers, dans chaque cas particulier on sait répondre. On pense donc à appliquer la formule des probabilités totales.

Proposition. Si A et B sont des événements indépendants, alors A et \bar{B} sont indépendants, \bar{A} et B sont indépendants et \bar{A} et \bar{B} sont indépendants.

DÉMONSTRATION.

C'est intuitif ! Si le fait que A soit réalisé n'apporte aucune information sur B , alors le fait que A ne soit pas réalisé n'apporte aucune information !

⚠ Par contre, si des événements A , B et C sont indépendants deux à deux, alors on ne peut rien dire sur l'indépendance de A et $B \cap C$, ou encore de A et $B \cup C$.

Par exemple : on dispose de deux dés équilibrés de couleurs respectives bleue et rouge. On lance les deux dés et on relève les chiffres obtenus. On prend $\Omega = \llbracket 1; 6 \rrbracket^2$ et \mathbb{P} la probabilité uniforme sur $(\Omega, \mathcal{P}(\Omega))$. Notons A : « La somme des chiffres vaut 7 », B : « Le dé bleu a donné 5 » et C : « Le dé rouge a donné 2 ». Nous avons

★ $\mathbb{P}(A) = \mathbb{P}(\{(1, 6); (2, 5); (3, 4); (4, 3); (5, 2); (6, 1)\}) = \frac{6}{36} = \frac{1}{6}$.

★ De même $\mathbb{P}(B) = \mathbb{P}(C) = \frac{1}{6}$.

L'indépendance deux à deux n'est donc pas la bonne notion d'indépendance lorsque l'on a plusieurs événements. Nous verrons l'indépendance dite mutuelle dans le paragraphe suivant.

$$\star \mathbb{P}(A \cap B) = \mathbb{P}(A \cap C) = \mathbb{P}(B \cap C) = \mathbb{P}(\{(5, 2)\}) = \frac{1}{36}.$$

Par conséquent A et B sont indépendants, A et C sont indépendants et B et C sont indépendants. Ensuite

\star D'après la formule de Poincaré,

$$\mathbb{P}(B \cup C) = \mathbb{P}(B) + \mathbb{P}(C) - \mathbb{P}(B \cap C) = \frac{1}{6} + \frac{1}{6} - \frac{1}{36} = \frac{11}{36}.$$

$$\star \mathbb{P}(A \cap B \cap C) = \mathbb{P}(A \cap (B \cup C)) = \mathbb{P}(\{(5, 2)\}) = \frac{1}{36}.$$

On a donc $\mathbb{P}(A \cap (B \cap C)) = \frac{1}{36} \neq \frac{1}{6} \times \frac{1}{36} = \mathbb{P}(A)\mathbb{P}(B \cap C)$ si bien que A et $B \cap C$ ne sont pas indépendants.


On a enfin $\mathbb{P}(A \cap (B \cup C)) = \frac{1}{36} \neq \frac{1}{6} \times \frac{11}{36} = \mathbb{P}(A)\mathbb{P}(B \cup C)$ si bien que A et $B \cup C$ ne sont pas non plus indépendants.

2) Événements mutuellement indépendants

Définition (indépendance mutuelle). Soit $n \in \mathbb{N} \setminus \{0; 1\}$. On dit que les événements A_1, \dots, A_n sont mutuellement indépendants si :


$$\forall J \subset \llbracket 1; n \rrbracket, \quad \mathbb{P}\left(\bigcap_{i \in J} A_i\right) = \prod_{j \in J} \mathbb{P}(A_j).$$

Si J est l'ensemble vide ou un singleton, l'égalité ci-dessus est vérifiée. On peut donc se contenter de la prouver pour les parties à au moins deux éléments. En particulier, pour montrer que trois événements A , B et C sont mutuellement indépendants, on doit montrer que $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$, $\mathbb{P}(A \cap C) = \mathbb{P}(A)\mathbb{P}(C)$, $\mathbb{P}(B \cap C) = \mathbb{P}(B)\mathbb{P}(C)$ et que $\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C)$.

 Des événements indépendants sont deux à deux indépendants (prendre J n'importe quelle partie de $\llbracket 1; n \rrbracket$ à deux éléments) mais la réciproque est fautive ! Montrer que des événements sont deux à deux indépendants ne permet pas de conclure qu'ils sont mutuellement indépendants.

Dans l'exemple ci-dessus avec les dés rouge et bleu, les événements A , B et C sont deux à deux indépendants mais ils ne sont pas mutuellement indépendants car

$$\mathbb{P}(A \cap B \cap C) = \frac{1}{36} \neq \frac{1}{6^3} = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C).$$

 Autre erreur classique : montrer que $\mathbb{P}(A_1 \cap \dots \cap A_n) = \mathbb{P}(A_1) \dots \mathbb{P}(A_n)$ et conclure que A_1, \dots, A_n sont mutuellement indépendants. Il faut vérifier toutes les autres intersections finies.

Par exemple, si on prend $A = B$ un événement quelconque de probabilité $p \in]0; 1[$ et $C = \emptyset$, alors $A \cap B \cap C = \emptyset$ donc $\mathbb{P}(A \cap B \cap C) = 0 = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C)$ mais $\mathbb{P}(A \cap B) = p \neq p^2 = \mathbb{P}(A)\mathbb{P}(B)$ donc A , B et C ne sont pas mutuellement indépendants.

Théorème. Soit $n \in \mathbb{N} \setminus \{0; 1\}$. Soient A_1, \dots, A_n des événements mutuellement indépendants. Si pour tout $k \in \llbracket 1; n \rrbracket$, $B_k = A_k$ ou \overline{A}_k , alors B_1, \dots, B_n sont mutuellement indépendants.



Il se peut que le mot « mutuellement » ne soit pas précisé. Si on dit que des événements sont indépendants, on sous entend toujours qu'ils sont mutuellement indépendants. Il s'agit de la bonne notion d'indépendance.

DÉMONSTRATION. Prouvons tout d'abord que $\overline{A_1}, A_2, \dots, A_n$ sont mutuellement indépendants.

En d'autres termes, quand on a des événements mutuellement indépendants, toute famille qu'on peut construire à l'aide de ces événements en prenant ou non le complémentaire reste une famille d'événements mutuellement indépendants.

□

Proposition. Si A, B et C sont trois événements mutuellement indépendants, alors A et $B \cap C$ sont indépendants et A et $B \cup C$ sont indépendants.

Rappelons qu'on a vu plus haut que c'était faux en général pour des événements deux à deux indépendants seulement.

DÉMONSTRATION. Nous avons $\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C) = \mathbb{P}(A)\mathbb{P}(B \cap C)$ donc A et $B \cap C$ sont indépendants. Ensuite la formule de Poincaré nous donne

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

On utilise bien ici la mutuelle indépendance de A, B et C .

donc A et $B \cup C$ sont indépendants.

□

Plus généralement, on a :

Théorème (des coalitions). Soit $n \in \mathbb{N} \setminus \{0; 1\}$. Soient A_1, \dots, A_n des événements mutuellement indépendants.

1. Si $J \subset \llbracket 1; n \rrbracket$, alors $(A_i)_{i \in J}$ est une famille d'événements mutuellement indépendants.
2. Si $J \subset \llbracket 1; n \rrbracket$, alors tout événement construit à partir de $(A_i)_{i \in J}$ est indépendant de tout événement construit à partir de $(A_i)_{i \notin J}$.

~> DÉMONSTRATION LAISSÉE EN EXERCICE.

Exemple : Soit $n \in \mathbb{N}^*$. On lance n fois successivement une pièce de monnaie. La pièce est telle que la probabilité de tomber sur Pile est $p \in]0; 1[$. Pour tout $k \in \llbracket 1; n \rrbracket$, notons A_k l'événement « obtenir Pile au $k^{\text{ième}}$ lancer ». On fait donc l'hypothèse que ces n événements sont mutuellement indépendants.

- La probabilité de n'obtenir que des Piles est

Par « construit avec », on comprend avec des unions, intersections, complémentaires, etc.

- Pour tout $k \in \llbracket 1; n \rrbracket$, la probabilité d'obtenir Pile pour la première fois au $k^{\text{ième}}$ lancer est

C'est le cadre d'une loi géométrique, cf. chapitre