Chapitre 26

Probabilités sur un univers fini

I Vocabulaire probabiliste

I.1 Expérience aléatoire

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.

Expérience aléatoire contrairement à une expérience déterministe : par exemple, avec une tension et une résistance connues, vous devriez pouvoir prédire l'intensité grâce à la loi d'Ohm (modulo les imprécisions dues aux mesures).

Exemple : Lancer d'un dé, lancer (avec un r ! pas un lancé) d'une pièce, tirage d'une boule dans une urne ou d'une carte dans un jeu.

Remarque : La théorie des probabilités est apparue historiquement pour étudier les jeux de hasard. Elle est désormais utilisée pour elle-même mais également 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.

I.2 Univers des résultats observables, éventualités

Étant donnée une expérience aléatoire, on note Ω l'ensemble des résultats possibles de l'expérience. L'ensemble Ω est appelé univers des possibles ou des résultats observables.

Il va sans dire que $\Omega \neq \varnothing$, et c'est ce qu'on supposera dans la suite. De plus, on se restreint en première année au cas où Ω est un ensemble fini. Vous étudierez le cas où Ω est infini (dénombrable) l'année prochaine.

Exemples :

- On lance un dé. On peut obtenir les entiers $1, 2, 3, 4, 5, 6$ donc $\Omega = [\![1\,;\,6]\!]$.
- On lance trois fois une pièce. On peut obtenir $(F, F, F), (F, F, P), (F, P, F), \dots$ donc $\Omega = \{F, P\}^3$.
- On a une urne contenant 5 boules numérotées de 1 à 5, on tire successivement et sans remise 3 boules. On peut obtenir $(1, 3, 5), (1, 3, 2), (5, 4, 1), \dots$ c'est-à-dire qu'on peut prendre Ω l'ensemble des arrangements à 3 éléments de $[\![1\,;\,5]\!]$.

Plus généralement :

Exemples :

- Si on lance n dés successivement (avec $n \in \mathbb{N}^*$), on peut coder le résultat de l'expérience par un n-uplet $(x_1, \dots, x_n)$ d'éléments de $[\![1\,;\,6]\!]$ de telle sorte que, pour tout $i \in [\![1\,;\,n]\!]$, x_i est le résultat du $i^{\text{ième}}$ lancer. Dans ce cas $\Omega = [\![1\,;\,6]\!]^n$ et il y a 6^n issues à l'expérience.
- 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 $0 \leqslant p \leqslant n$), alors on peut coder le résultat de l'expérience par une partie $\{x_1; \dots; x_p\}$ (ou une combinaison) de $[\![1\,;\,n]\!]$ où $x_1, \dots, x_n$ sont les numéros des boules tirées (l'ordre ne compte pas et ils sont distincts). Dans ce cas Ω est l'ensemble des parties de $[\![1\,;\,n]\!]$ à p éléments. Il y a $\dbinom{n}{p}$ issues à l'expérience.
- Si cette fois on tire les boules successivement et sans remise, on peut coder le résultat de l'expérience par un p-uplet $(x_1, \dots, x_n)$ d'éléments distincts de $[\![1\,;\,n]\!]$ de telle sorte que, pour tout $i \in [\![1\,;\,n]\!]$, x_i est le numéro de la $i^{\text{ième}}$ boule tirée. Dans ce cas il y a $\dfrac{n!}{(n-p)!}$ issues à l'expérience.

Comme dit plus haut : on réalise **une** expérience (même si celle-ci comporte plusieurs étapes, par exemple ici on tire plusieurs boules), et on se demande la nature de l'objet, du résultat qu'on a entre les mains.

Comme dans le chapitre 17, on se demande si l'ordre compte et si on peut avoir plusieurs fois le même élément.

Définition. Les éléments de Ω (notés ω) sont appelés éventualités.

Remarque : Ω est donc parfois appelé univers des éventualités.

Exemple : Si on lance 3 fois une pièce, $\omega = (P, P, F)$ est une éventualité.

Remarques :

- Une éventualité ne doit pas être confondue avec un événement (cf. paragraphe I.3). Une éventualité, c'est un résultat possible pour une expérience aléatoire. En clair : on effectue **une** expérience, on se demande la nature de l'objet qu'on a entre les mains. Dans l'exemple ci-dessus : on lance trois fois une pièce (mais ce n'est qu'une répétition de l'expérience, même si on la lance trois fois), le résultat de l'expérience est un triplet (x, y, z) avec x, y, z égaux à P ou F donc $\Omega = \{P; F\}^3$.
- Le choix de Ω dépend de l'expérience qu'on cherche à modéliser. Par exemple, si on lance deux dés indiscernables, on peut prendre $\Omega = [\![1\,;\,6]\!]^2$ l'ensemble des couples (a, b) avec a et b dans $[\![1\,;\,6]\!]$ mais on peut également prendre

$$\Omega = \left\{ \{a; b\} \;\middle|\; (a, b) \in [\![1\,;\,6]\!]^2 \right\}$$

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

$$\binom{6}{1} + \binom{6}{2} = 21$$

En pratique, on prendra toujours le premier, on expliquera pourquoi dans la suite.

On se donne dans la suite un ensemble fini non vide Ω qu'on appelle donc univers des possibles, des éventualités, des résultats observables etc.

À part dans les situations dites « d'équiprobabilité » (cf. paragraphe II.3), 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.

I.3 Événements

I.3.a Introduction et définitions

On peut s'intéresser à des résultats plus généraux que des éventualités. Par exemple, si on lance un dé, on peut s'intéresser aux événements

- A_1 : « on tire un nombre pair. »
- A_2 : « on tire un 3 ou un 5. »
- A_3 : « on tire un 2 ou un 4. »

Intuitivement, un événement est une propriété de l'expérience qui peut être vérifiée ou non. On cherche à en donner une définition plus maniable.

Soit $\Omega = [\![1\,;\,6]\!]$ et soit $\omega \in \Omega$ une éventualité. On dit que l'événement A_i est réalisé par l'éventualité ω si ω vérifie la propriété de l'événement.

En d'autres termes, A_1 est réalisé par ω si et seulement si $\omega = 2, 4$ ou 6 c'est-à-dire si et seulement si $\omega \in \{2; 4; 6\}$. On convient donc de représenter A_1 par $\{2; 4; 6\}$. On notera $A_1 = \{2; 4; 6\}$. De même on note $A_2 = \{3; 5\}$ et $A_3 = \{2; 4\}$.

Prenons un autre exemple et lançons deux dés. On s'intéresse aux événements

- B_1 : « la somme vaut 6. »
- B_2 : « on tire au moins un 6 »
- B_3 : « les deux tirages sont impairs. »

On prend $\Omega = [\![1\,;\,6]\!]^2$ et on notera de même

- $B_1 = \{(1, 5); (2, 4); (3, 3); (4, 2); (5, 1)\}$

- $B_2 = \{(1,6);(2,6);(3,6);(4,6);(5,6);(6,6);(6,1);(6,2);(6,3);(6,4);(6,5);(6,5)\}$
- $B_3 = \{(1,1);(1,3);(1,5);(3,1);(3,3);(3,5);(5,1);(5,3);(5,5)\}$

Cela justifie la définition suivante.

Définition.

- Un élément $A \in \mathscr{P}(\Omega)$ est appelé événement. On dit qu'il est réalisé si le résultat de l'expérience est un élément de A. En d'autres termes, un événement A est réalisé par une éventualité ω si $\omega \in A$.
- Si $\omega \in \Omega$, le singleton $\{\omega\}$ est appelé événement élémentaire.
- L'événement $\varnothing$ n'est jamais réalisé : c'est l'événement impossible. L'événement Ω est toujours réalisé : c'est l'événement certain.

Exemple : Avec les notations ci-dessus, le couple $(1,5)$ est une éventualité qui réalise les événements B_1 et B_3.

Remarques :

- Comme dit ci-dessus, à part (et encore) quand on travaillera dans des situations d'équiprobabilité, on n'explicitera jamais les événements comme on l'a fait ci-dessus pour B_1, B_2, B_3. Cependant, définir des événements comme des parties est tout de même très pratique car on peut utiliser tout ce qu'on sait en théorie des ensembles (union, intersection etc.).
- Quand on fait un exercice de probas, la première chose à faire est d'introduire des événements. En général, on appelle A l'événement dont on cherche à calculer la probabilité, mais calculer $P(A)$ n'est pas toujours simple. Par conséquent, on commence en général une liste d'événements « simples ».

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

Exemples :

- Si on lance un dé, on peut poser, pour tout $k \in [\![1\,;6]\!]$, 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 $k \in [\![1\,;n]\!]$, A_k : « obtenir Pile au $k^{\text{ième}}$ lancer ».
- Dans l'exemple précédent, définir simplement A : « obtenir Pile » n'est pas assez précis. À quel lancer ? Au premier ? À tous ? À un lancer quelconque ? Il ne faut pas oublier de préciser à quel lancer on fait allusion !

Réflexe en début d'exercice de probabilités : définir des événements (tous ceux auxquels on peut penser, mieux vaut trop que pas assez !) : cf. paragraphe V pour une marche à suivre dans les exercices de probabilités.

I.3.b Dictionnaire langage ordinaire/langage ensembliste

Définition. Un espace probabilisable fini est un couple $(\Omega, \mathscr{P}(\Omega))$ où Ω est un ensemble fini appelé univers.

En clair : le but d'un espace probabilisable est de se faire probabiliser, i.e. de se faire adjoindre une probabilité : cf. paragraphe II.1.b.

On se donne dans la suite du chapitre un espace probabilisable $(\Omega, \mathscr{P}(\Omega))$.

Définition. Soient A et B des événements.

- L'événement $\overline{A}$ est réalisé si A ne l'est pas. On dit qu'il s'agit de l'événement contraire ou complémentaire de A.
- La relation $A \subset B$ signifie que la réalisation de l'événement A implique 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.

Puisque les événements sont des parties de Ω, toutes les opérations vues dans le chapitre 4 sont valables pour les événements. Seul le vocabulaire change.

- L'événement $A \backslash B$ est réalisé si l'événement A est réalisé mais pas l'événement B.
- Si $A \cap B = \varnothing$, les événements A et B sont dits incompatibles ou disjoints.

En d'autres termes, les événements A et B sont incompatibles lorsqu'ils ne peuvent pas être réalisés simultanément, puisqu'il n'existe aucune éventualité $\omega \in A \cap B$.

Exemple : A_3 implique A_1.

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.

Langage ordinaire	**Langage ensembliste**
Expérience aléatoire	Univers aléatoire Ω
Résultat possible, éventualité	$\omega \in \Omega$
Événement	$A \subset \Omega, A \in \mathscr{P}(\Omega)$
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	$\omega \in \overline{A}$
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	$\omega \in A \cap \overline{B} = A \backslash B$
A et B sont incompatibles	$A \cap B = \varnothing$
Au moins un des $A_1, \ldots, A_n$ est réalisé	$\omega \in \displaystyle\bigcup_{i=1}^{n} A_i$
Tous les A_k sont réalisés	$\omega \in \displaystyle\bigcap_{k=1}^{n} A_k$
Aucun A_k n'est réalisé	$\omega \in \overline{\displaystyle\bigcup_{k=1}^{n} A_k} = \displaystyle\bigcap_{k=1}^{n} \overline{A_k}$
Les A_k sont deux à deux incompatibles	$\forall i \neq j, A_i \cap A_j = \varnothing$

Exemple : On dispose d'un jeu de 52 cartes. On tire trois cartes successivement et avec remise. On peut travailler avec l'univers $\Omega = E^3$ où

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

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 $\overline{A_1}$.
- « obtenir $\spadesuit$ aux deux premiers tirages mais pas au troisième » est $A_1 \cap A_2 \cap \overline{A_3}$.
- « obtenir $\spadesuit$ aux deux premiers tirages » est $A_1 \cap A_2$.
- « obtenir au moins un $\spadesuit$ » est $A_1 \cup A_2 \cup A_3$.
- « ne pas obtenir de $\spadesuit$ » est $\overline{A_1} \cap \overline{A_2} \cap \overline{A_3}$.
- « obtenir exactement deux $\spadesuit$ » est $(A_1 \cap A_2 \cap \overline{A_3}) \cup (A_1 \cap \overline{A_2} \cap A_3) \cup (\overline{A_1} \cap A_2 \cap A_3)$.

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.

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 $k \in [\![1 ; n]\!]$, notons A_k l'événement « obtenir Pile au $k^{\text{ième}}$ lancer ».

Poser A : « obtenir Pile » n'est pas assez précis, voir ci-dessus.

- L'événement « n'obtenir que des faces » est $\displaystyle\bigcap_{k=1}^{n} \overline{A_k}$.
- Pour tout $k \in \llbracket 1\,;n \rrbracket$, posons B_k : « obtenir Pile pour la première fois au $k^{\text{ième}}$ lancer ». On a
$$B_k = \left(\bigcap_{j=1}^{k-1} \overline{A_j}\right) \cap A_k.$$

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

- $\Omega = \displaystyle\bigcup_{i\in I} A_i$,
- Les événements de la famille sont incompatibles deux à deux, c'est-à-dire :
$$\forall (i,j) \in I^2, \qquad i \neq j \qquad \Longrightarrow \qquad A_i \cap A_j = \varnothing.$$

En d'autres termes, un système complet d'événements est une famille d'événements deux à deux incompatibles dont l'union est égale à Ω. En pratique, il faut 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.

Exemples :

- Si $\Omega = \llbracket 1\,;6 \rrbracket^2$, alors A_1 : « tirer deux nombres pairs », A_2 : « tirer deux nombres impairs » et A_3 : « tirer un seul nombre pair », forment 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 $A_1 \cap A_2 \neq \varnothing$.
  - ⋆ 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 $n = \text{card}(\Omega) \in \mathbb{N}^*$ (et donc les ω_i deux à deux distincts), alors $(\{\omega_i\})_{1\leqslant i\leqslant n}$ est un système complet d'événements.

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

Exemple : Si $\Omega = \llbracket 1\,;6 \rrbracket$,

- A : « tirer un nombre pair. »
- $\overline{A}$: « tirer un nombre impair. »

Remarque : On parle parfois de « partition » à la place de « système complet d'événements », mais il faut alors que les différents événements soient tous non vides (rappelons que, par définition, une partition est composée d'ensembles tous non vides). En général, on ne le demande pas, donc on ne confondra pas une partition (vocabulaire ensembliste, avec des ensembles non vides) avec un système complet d'événements (vocabulaire probabiliste, avec des ensembles qui peuvent être vides).

Si j'osais, je dirais que parler de « partition » fait mauvais genre, est un manque de classe : chaque domaine des mathématiques a un vocabulaire qui lui est propre, on utilise le vocabulaire probabiliste quand on fait des probas. Cela permet d'ailleurs de mieux comprendre et mieux deviner certains résultats.

II Probabilité sur un univers fini

II.1 Définitions

II.1.a Introduction

Lors de N répétitions indépendantes les unes des autres d'une expérience aléatoire, la fréquence d'apparition d'un événement A est le nombre $\dfrac{N(A)}{N}$ 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 $N \to +\infty$. D'un point de vue intuitif, on veut poser :

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

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 :

$$\frac{N(A \cup B)}{N} \xrightarrow[N \to +\infty]{} P(A \cup B)$$

et d'autre part :

$$\frac{N(A \cup B)}{N} = \frac{N(A)}{N} + \frac{N(B)}{N} \xrightarrow[N \to +\infty]{} P(A) + P(B)$$

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

$$\begin{aligned} P(\Omega) &= \lim_{N \to +\infty} \frac{N(\Omega)}{N} \\ &= \lim_{N \to +\infty} \frac{N}{N} \\ &= 1 \end{aligned}$$

La définition du paragraphe suivant est donc naturelle si on veut être conforme à l'intuition.

II.1.b Définition proprement dite et un exemple un peu surprenant

Définition. Une probabilité est une application $P : \mathscr{P}(\Omega) \to [\,0\,;1\,]$ vérifiant les deux conditions suivantes :

- $P(\Omega) = 1$.
- Pour tout couple (A, B) d'événements **incompatibles**, $P(A \cup B) = P(A) + P(B)$.

Cette deuxième propriété est appelée additivité. Par conséquent, une probabilité est une application additive de $\mathscr{P}(\Omega)$ dans $[\,0\,;1\,]$ vérifiant $P(\Omega) = 1$, mais quand on demande la définition, c'est la définition ci-contre qui doit sortir. En particulier, le mot « incompatibles » doit apparaître !

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

Exemple : La probabilité de Dirac.
Soit $x_0 \in \Omega$. On définit l'application $P_{x_0} : \mathscr{P}(\Omega) \to [\,0\,;1\,]$ par :

$$P_{x_0}(A) = \begin{cases} 1 & \text{si } x_0 \in A \\ 0 & \text{si } x_0 \notin A \end{cases}$$

Montrons que P_{x_0} est une probabilité.

- P_{x_0} est à valeurs dans $[\,0\,;1\,]$ par définition (car elle ne prend que les valeurs 0 et 1). Ne pas oublier de vérifier cette condition !

- $x_0 \in \Omega$ donc $P_{x_0}(\Omega) = 1$.
- Soient A et B deux événements incompatibles. Il y a deux cas de figure : soit x_0 n'appartient ni à A ni à B, soit x_0 appartient à un seul des ensembles A ou B (il ne peut pas appartenir à A et à B car A et B sont incompatibles i.e. disjoints). Dans le premier cas, $P_{x_0}(A) = P_{x_0}(B) = 0$ et $x_0 \notin A \cup B$ donc $P_{x_0}(A \cup B) = 0$. Dans le second cas, $P_{x_0}(A) = 1$ et $P_{x_0}(B) = 0$ ou le contraire, et $x_0 \in A \cup B$ donc $P_{x_0}(A \cup B) = 1$. Dans tous les cas, $P_{x_0}(A \cup B) = P_{x_0}(A) + P_{x_0}(B)$.

P_{x_0} est donc une probabilité, appelée probabilité de Dirac en x_0. C'est par exemple la probabilité naturelle dans le cadre d'un lancer de dé pipé (par exemple, si le 4 sort à chaque lancer, on munira l'espace probabilisable $(\Omega, \mathscr{P}(\Omega))$ de la probabilité P_4).

Définition. Un espace probabilisé fini est un triplet $(\Omega, \mathscr{P}(\Omega), P)$ où

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

Comme dit plus haut, le but d'un espace probabilisable est de se faire probabiliser, i.e. de se faire adjoindre une probabilité, et quand c'est fait, on a un espace probabilisé. La différence entre un espace probabilisable et un espace probabilisé est que l'un est muni d'une probabilité, et que l'autre attend qu'on lui en donne une.

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

Théorème. Soient A et B des événements.

1. $P(\varnothing) = 0$.
2. $P(\overline{A}) = 1 - P(A)$.
3. Si $A \subset B$, alors $P(A) \leqslant P(B)$.
4. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.
5. $P(B \backslash A) = P(B) - P(A \cap B)$. En particulier, si $A \subset B$, alors $P(B \backslash A) = P(B) - P(A)$.
6. Si $A_1, \dots, A_n$ sont des événements, alors : $P\left(\displaystyle\bigcup_{k=1}^n A_k\right) \leqslant \displaystyle\sum_{k=1}^n P(A_k)$.

   De plus, si les A_i sont deux à deux incompatibles, alors : $P\left(\displaystyle\bigcup_{k=1}^n A_k\right) = \displaystyle\sum_{k=1}^n P(A_k)$.

Démonstration.
1. On a $P(\varnothing) = P(\varnothing \cup \varnothing) = P(\varnothing) + P(\varnothing)$ puisque $\varnothing \cap \varnothing = \varnothing$. Ainsi $\mathscr{P}(\varnothing) = 0$.
2. A et $\overline{A}$ sont incompatibles donc $P(A \cup \overline{A}) = P(A) + P(\overline{A})$. Or, $A \cup \overline{A} = \Omega$ donc $P(\overline{A}) = P(\Omega) - P(A) = 1 - P(A)$.
3. Si $A \subset B$, alors $B = A \cup (B \backslash A)$ et les deux événements A et $B \backslash A$ sont incompatibles donc $P(B) = P(A) + \underbrace{P(B \backslash A)}_{\geqslant 0}$ si bien que $P(B) \geqslant P(A)$.
4. $A \cup B = A \cup (B \backslash A)$ et ces deux événements (A et $B \backslash A$) sont incompatibles donc $P(A \cup B) = P(A) + P(B \backslash A)$. Or, $B = (B \backslash A) \cup (A \cap B)$ et ces deux événements sont incompatibles donc $P(B) = P(B \backslash A) + P(A \cap B)$ ce qui permet de conclure.
5. Découle de la démonstration du point précédent.
6. Par récurrence sur n : $\rightsquigarrow$ Exercice.

Corollaire. Si $(A_1, \dots, A_n)$ est un système complet d'événements, $\displaystyle\sum_{k=1}^n P(A_k) = 1$.

Attention, la réciproque est fausse !

Démonstration. $\Omega = \displaystyle\bigcup_{i=1}^n A_i$ et cette union est disjointe, si bien que $P(\Omega) = 1 = \displaystyle\sum_{i=1}^n P(A_i)$.

Corollaire (Formule du crible ou de Poincaré pour trois événements). Soient A, B, C trois événements.

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

HP : à savoir redémontrer !

DÉMONSTRATION. S'inspirer du chapitre 17 : ⇝ EXERCICE.

Activité - Inégalité de Bonferroni : Il y a une formule générale pour n événements mais elle est HP. En pratique, la majoration du 6 suffit. Puisqu'on a une majoration, cherchons une minoration simple.

cf. exercice 38 du chapitre 27.

D'après ce qui précède,

$$P(A_1\cup A_2\cup A_3)\geqslant P(A_1)+P(A_2)+P(A_3)-P(A_1\cap A_2)-P(A_2\cap A_3)-P(A_1\cap A_2)$$

Plus généralement, montrons par récurrence que si on a une suite d'événements $(A_n)_{n\geqslant 1}$, alors :

$$\forall n\geqslant 3, P\left(\bigcup_{i=1}^n A_i\right)\geqslant \sum_{i=1}^n P(A_i)-\sum_{1\leqslant i<j\leqslant n}P(A_i\cap A_j)$$

Le résultat est aussi vrai pour $n=2$ et c'est même une égalité.

Le résultat est vrai au rang 3 d'après ce qui précède. Soit $n\geqslant 3$. Supposons le résultat vrai au rang n et prouvons qu'il est vrai au rang $n+1$.

$$\begin{aligned} P\left(\bigcup_{i=1}^{n+1} A_i\right) &= P\left(\left(\bigcup_{i=1}^{n} A_i\right)\cup A_{n+1}\right)\\ &= P\left(\bigcup_{i=1}^{n} A_i\right)+P(A_{n+1})-P\left(\left(\bigcup_{i=1}^{n} A_i\right)\cap A_{n+1}\right)\\ &= P\left(\bigcup_{i=1}^{n} A_i\right)+P(A_{n+1})-P\left(\bigcup_{i=1}^{n}(A_i\cap A_{n+1})\right)\\ &\geqslant \sum_{i=1}^n P(A_i)-\sum_{1\leqslant i<j\leqslant n}P(A_i\cap A_j)+P(A_{n+1})-P\left(\bigcup_{i=1}^{n}(A_i\cap A_{n+1})\right)\\ &\geqslant \sum_{i=1}^{n+1} P(A_i)-\sum_{1\leqslant i<j\leqslant n}P(A_i\cap A_j)-P\left(\bigcup_{i=1}^{n}(A_i\cap A_{n+1})\right)\end{aligned}$$

Par distributivité de l'intersection sur l'union.

Par hypothèse de récurrence.

On regroupe le premier et le troisième terme.

Or,

$$P\left(\bigcup_{i=1}^{n}(A_i\cap A_{n+1})\right)\leqslant \sum_{i=1}^n P(A_i\cap A_{n+1})$$

Dès lors :

$$\begin{aligned} P\left(\bigcup_{i=1}^{n+1} A_i\right) &\geqslant \sum_{i=1}^{n+1} P(A_i)-\underbrace{\sum_{1\leqslant i<j\leqslant n}P(A_i\cap A_j)-\sum_{i=1}^n P(A_i\cap A_{n+1})}_{\downarrow}\\ &\geqslant \sum_{i=1}^{n+1} P(A_i)-\overbrace{\sum_{1\leqslant i<j\leqslant n+1}P(A_i\cap A_j)}\end{aligned}$$

c'est-à-dire que le résultat est vrai au rang $n+1$, ce qui clôt la récurrence.

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

Définition. Un événement A est dit :

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

II.2 Construction de probabilités

Si $\Omega = \{\omega_1, \dots, \omega_n\}$ alors $\Omega = \bigcup_{i=1}^{n} \{\omega_i\}$ et cette union est disjointe donc

Un ensemble est toujours l'union disjointe des singletons formés par ses éléments.

$$P(\Omega) = \sum_{i=1}^{n} P(\{\omega_i\})$$

En d'autres termes, $\displaystyle\sum_{i=1}^{n} P(\{\omega_i\}) = 1$ ce qu'on résume souvent par : « la somme des probabilités vaut 1 ».

On se pose à présent la question inverse : si on a des réels positifs dont la somme vaut 1, est-ce que ces réels sont des probabilités ?

J'avoue... plus vague, tu meurs !

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 $(p_\omega)_{\omega \in \Omega}$, il existe une unique probabilité $P : \mathscr{P}(\Omega) \to [\,0\,;1\,]$ telle que, pour tout $\omega \in \Omega$, $P(\{\omega\}) = p_\omega$.

Remarque : En d'autres termes :

- Une probabilité P sur Ω est entièrement déterminée par la distribution de probabilités $(P(\{\omega\}))_{\omega \in \Omega}$.
- Une probabilité $\mathscr{P}$ 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 $P(\{\omega\})$, pour tout $\omega \in \Omega$. D'une part, caractérise entièrement la probabilité, c'est-à-dire qu'il y a unicité : si deux probabilités coïncident en tous les $\{\omega\}$, alors elles sont égales, et il y a aussi existence : pour toute famille $(p_\omega)_{\omega \in \Omega}$ finie de somme 1, il existe une probabilité donc ces valeurs sont les différentes $P(\{\omega\})$.

Exemples :

- Il n'est pas possible qu'un dé truqué donne 1 avec probabilité $1/22$, 2 avec probabilité $2/22$, $\dots$, 6 avec probabilité $6/22$ puisque $1/22 + \cdots + 6/22 = 21/22 \neq 1$.
- Cependant, il est possible de truquer un dé pour que chaque face ait une probabilité de tomber qui soit proportionnelle au numéro de la face. On prend $\Omega = [\![\,1\,;\,6\,]\!]$. On veut définir P telle que, pour tout $k \in [\![\,1\,;\,6\,]\!]$, $P(\{k\}) = kP(\{1\})$. Pour que P soit

une probabilité sur $(\Omega, \mathscr{P}(\Omega))$, il faut et il suffit que

$$\begin{aligned} 1 &= \sum_{k=1}^{6} P(\{k\}) \\ &= P(\{1\})\sum_{k=1}^{6} k \\ &= 21P(\{1\}) \end{aligned}$$

c'est-à-dire $P(\{1\}) = \dfrac{1}{21}$. Ainsi on prend P telle que, pour tout tout $k \in [\![1\,;6]\!]$, $P(\{k\}) = \dfrac{k}{21}$, et cela définit une unique probabilité sur $[\![1\,;6]\!]$.

Remarque : Ce résultat repose sur le fait que toute partie est l'union disjointe des singletons formés par ses éléments. En d'autres termes, les singletons forment les « briques élémentaires » qui peuvent former toute partie de Ω i.e. tout événement. En clair, si on connaît la probabilité de chaque singleton, on connaît la probabilité de chaque partie puisque celle-ci est l'union disjointe des singletons de ses éléments. Par exemple, avec le dé pipé ci-dessus, rien qu'en connaissant la probabilité de chaque numéro, on peut donner la probabilité de chaque événement. Par exemple, si on note A l'événement : « tirer un nombre pair », alors $A = \{2\} \cup \{4\} \cup \{6\}$ et cette union est disjointe si bien que

$$\begin{aligned} P(A) &= P(\{2\}) + P(\{4\}) + P(\{6\}) \\ &= \frac{2+4+6}{21} \\ &= \frac{12}{21} \end{aligned}$$

De même (ou simplement avec l'événement complémentaire), la probabilité d'obtenir un nombre impair vaut $11/21$: il est plus probable d'obtenir un nombre pair qu'un nombre impair, ce qui est intuitif puisque les numéros pairs $(2, 4, 6)$ ont une plus grande probabilité de sortir/sont plus chargés que les nombres impairs $(1, 3, 5)$.

C'est exactement la même chose dans le cas général (et c'est d'ailleurs le principe de la démonstration). Pour cette raison, quand on donnera ou quand on voudra caractériser une probabilité, on se contentera parfois de donner la probabilités des événements élémentaires (c'est-à-dire des singletons).

Démonstration. Notons $\Omega = \{\omega_1; \dots; \omega_n\}$ où $n = \text{card}(\Omega)$. Notons $p_1 = p_{\omega_1}, \dots, p_n = p_{\omega_n}$ la distribution de probabilité. Il s'agit donc de prouver qu'il existe une unique probabilité P sur $(\Omega, \mathscr{P}(\Omega))$ telle que, pour tout $k \in [\![1\,;n]\!]$, $P(\{\omega_k\}) = p_k$.

Existence : Définissons $P : A \in \mathscr{P}(\Omega) \longmapsto \displaystyle\sum_{\substack{1 \leqslant k \leqslant n \\ \omega_k \in A}} p_k$.

- Il s'agit bien d'une application définie sur $\mathscr{P}(\Omega)$.
- On a $P(\Omega) = \displaystyle\sum_{k=1}^{n} p_k = 1$.
- Soit A un événement. Comme $p_1, \dots, p_n$ sont positifs, $0 \leqslant P(A) \leqslant \displaystyle\sum_{k=1}^{n} p_k = 1$.
- Donnons-nous A et B deux événements incompatibles, alors

$$\begin{aligned} P(A \cup B) &= \sum_{\substack{1 \leqslant k \leqslant n \\ \omega_k \in A \cup B}} p_k \\ &= \sum_{\substack{1 \leqslant k \leqslant n \\ \omega_k \in A}} p_k + \sum_{\substack{1 \leqslant k \leqslant n \\ \omega_k \in B}} p_k \\ &= P(A) + P(B) \end{aligned}$$

Ici on a utilisé la propriété de sommation par paquets puisque $A \cap B = \varnothing$.

Ainsi P est une probabilité sur $(\Omega, \mathscr{P}(\Omega))$. D'où l'existence.

Unicité : Soient P_1 et P_2 deux probabilités qui conviennent. En particulier, P_1 et P_2 coïncident en tous les événements élémentaires i.e. en tous les $\{\omega_k\}$. Soit $A \in \mathscr{P}(\Omega)$. Alors $A = \bigcup_{\omega \in A} \{\omega\}$ et cette union est (finie et) disjointe, si bien que

$$\begin{aligned} P_1(A) &= \sum_{\omega \in A} P_1(\{\omega\}) \\ &= \sum_{\omega \in A} P_1(\{\omega\}) \\ &= P_2(A) \end{aligned}$$

□

donc $P_1 = P_2$: d'où l'unicité.

II.3 L'exemple fondamental : l'équiprobabilité

Définition. La probabilité uniforme (ou équiprobabilité) est la probabilité définie par

$$P : \left\{ \begin{array}{ccc} \mathscr{P}(\Omega) & \to & [\,0\,;1\,] \\ A & \mapsto & \dfrac{\operatorname{card}(A)}{\operatorname{card}(\Omega)} \end{array} \right.$$

Tous les énoncés parlant de jeux de cartes *bien battus* ou *bien mélangés*, de boules *indiscernables au toucher*, de dés ou de pièces *équilibrés*, indiquent qu'il faut utiliser le modèle de l'équiprobabilité. L'usage veut que l'expression *au hasard*, sans autre indication, indique aussi que l'on se place dans ce contexte.

Remarques :

- Cette définition a bien un sens car Ω est un ensemble fini. Vous ne pourrez plus le faire en deuxième année quand Ω sera infini !
- C'est bien une probabilité d'après les propriétés vérifiées par les cardinaux démontrées dans le chapitre 17.
- L'équiprobabilité est résumée par l'expression traditionnelle « la probabilité de A est le nombre de cas favorables à A sur le nombre de cas possibles ».
- L'équiprobabilité est l'unique (d'après le paragraphe précédent) probabilité définie par $P(\{\omega\}) = 1/\operatorname{card}(\Omega)$ pour tout $\omega \in \Omega$. On la définit donc parfois de cette manière. Par exemple, l'équiprobabilité sur $[\![\,1\,;\,6\,]\!]$ est l'unique probabilité vérifiant $P(\{1\}) = \cdots = P(\{6\}) = 1/6$.

Exemples :

- On lance deux fois un dé équilibré à six faces. On se place sur $\Omega = [\![\,1\,;\,6\,]\!]^2$. Puisque le dé est équilibré, nous prenons P la probabilité uniforme sur Ω, qui est un ensemble de cardinal 36. Notons A l'événement « la somme des deux chiffres obtenus est 8 ». Nous avons $A = \big\{(2,6);(3,5);(4,4);(5,3);(6,2)\big\}$ donc $P(A) = \dfrac{\operatorname{card}(A)}{\operatorname{card}(\Omega)} = \dfrac{5}{36}$.
- On lance n fois une pièce équilibrée. Quelle est la probabilité d'obtenir au moins deux Pile ou deux Face consécutifs ?

  On a $\Omega = \{P;F\}^n$ (l'ensemble des n-uplets d'éléments de $\{P;F\}$). On munit Ω de $\mathscr{P}(\Omega)$ et de l'équiprobabilité (ce qui est légitime puisque la pièce est équilibrée). Notons A : « obtenir au moins deux Pile ou deux Face consécutifs ». Alors $\overline{A}$ est l'événement : « tout Pile est suivi d'un Face, et tout Face est suivi d'un Pile ». Ainsi,

  $$\overline{A} = \{(P,F,P,F,\dots);(F,P,F,P,\dots)\}$$

  On ne précise pas les derniers termes car ils dépendent de la parité de n. En d'autres termes, $\operatorname{card}\left(\overline{A}\right) = 2$. Dès lors :

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

$$\begin{aligned} P\left(\overline{A}\right) &= \frac{\text{card}(\overline{A})}{\text{card}(\Omega)} \\ &= \frac{2}{2^n} \\ &= \frac{1}{2^{n-1}} \end{aligned}$$

En conclusion, $P(A) = 1 - \dfrac{1}{2^{n-1}}$.

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

  Pour simplifier, on suppose qu'aucun élève n'est né un 29 février. On numérote les jours de l'année de 1 à 365 et les élèves de 1 à n. On peut coder la situation par une n-liste $(x_1, \dots, x_n)$ d'éléments de $[\![1\,;365]\!]$ de telle sorte que, pour tout $i \in [\![1\,;n]\!]$, x_i est le jour de naissance du $i^{\text{ième}}$ élève. Ainsi on prend $\Omega = [\![1\,;365]\!]^n$.

  On suppose que les jours de naissance sont répartis de façons équiprobables dans l'année afin de munir $(\Omega, \mathscr{P}(\Omega))$ de P l'équiprobabilité. On introduit alors A l'événement « au moins deux élèves ont la même date de naissance ».

  Au fait il est plus facile de décrire l'événement $\overline{A}$: « les élèves ont tous des dates d'anniversaires différentes ». Il s'agit de l'ensemble des n-uplets d'éléments distincts de $[\![1\,;365]\!]$. On a donc

$$\begin{aligned} P(\overline{A}) &= \frac{\text{card}(\overline{A})}{\text{card}(\Omega)} \\ &= \frac{365 \cdot 364 \cdots (365 - n + 1)}{365^n} \\ &= \prod_{k=0}^{n-1} \frac{365 - k}{365} \end{aligned}$$

  et donc $P(A) = 1 - \mathscr{P}(\overline{A}) = 1 - \displaystyle\prod_{k=0}^{n-1} \left(1 - \frac{k}{365}\right)$.

  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 44 élèves, cette probabilité dépasse 90%.

> Si $n > 365$, il est sûr qu'au moins deux élèves partagent la même date d'anniversaire : principe des tiroirs.

> D'accord ce n'est pas très réaliste, mais c'est pour simplifier les calculs.

II.4 Équiprobabilité or not équiprobabilité ?

On peut se dire au premier abord que l'exemple de Dirac est un exemple exotique, voire pathologique alors que l'équiprobabilité semble la plus naturelle quand on travaille sur un univers fini. Pas toujours !

En effet, 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. Soit 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 $P(A) = \dfrac{3}{4}$.

- **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 $P(A) = \dfrac{2}{3}$ mais obtenir Pile au premier coup est de proba $\dfrac{1}{2}$ tandis que la probabilité des autres éventualités est $\dfrac{1}{4}$. En d'autres termes, il est naturel de munir $\left(\tilde{\Omega}, \mathscr{P}\left(\tilde{\Omega}\right), P\right)$ où P est la probabilité définie par

$$P(\{P\}) = \frac{1}{2} \qquad \text{et} \qquad P(\{(F,P)\}) = P(\{(F,F)\}) = \frac{1}{4}$$

  et on retrouve bien $P(A) = \dfrac{3}{4}$.

Exemple : On reprend l'exemple du I.2 c'est-à-dire qu'on lance deux dés indiscernables. On peut prendre $\Omega = [\![1\,;\,6]\!]^2$ muni de $\mathscr{P}(\Omega)$ et de l'équiprobabilité, ou

$$\tilde{\Omega} = \left\{ \{a;b\} \;\middle|\; (a,b) \in [\![1\,;\,6]\!]^2 \right\}$$

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

La probabilité la plus naturelle consiste à poser $P(\{a\}) = 1/36$ pour tout $a \in [\![1\,;\,6]\!]$ et $P(\{a;b\}) = 2/36$ pour tous $a \neq b$, c'est-à-dire munir les singletons d'un poids $1/36$ et les paires d'un poids $2/36$.

Remarque : C'est la raison pour laquelle en pratique on « trichera » en discernant les dés, c'est-à-dire qu'on prendra $\Omega = [\![1\,;\,6]\!]^2$ muni de l'équiprobabilité. Intuitivement, cela ne change pas la probabilité (celle-ci ne change pas qu'on sache ou non différencier les dés), la seule conséquence du fait que les dés sont indiscernables est qu'on ne peut considérer que des événements symétriques du type « il y a un nombre pair et un nombre impair » ou « la somme vaut 6 » mais pas « le premier dé donne 1 et le deuxième donne 3 ».

Conclusion : l'équiprobabilité, c'est pas automatique !

III Probabilités conditionnelles.

III.1 Définitions

Question : si on considère deux événements A et B, comment la connaissance de l'information « A est réalisé » modifie-t-elle la probabilité ? On pourrait se dire que cette question est inutile car il est évident que la connaissance de A ne donne aucune information, mais un exemple simple permet de voir que ce n'est pas si simple : par exemple, si B est « obtenir 6 » (quand on lance un dé équilibré), alors $P(B) = 1/6$, mais si on sait que A : « on obtient un nombre pair » alors la probabilité devient égale à $1/3$ car seuls les numéros $2, 4$ et 6 sont alors possibles.

Prenons un autre exemple : l'expérience qui consiste à lancer deux fois une pièce équilibrée. On note A : « Face est sorti » et B : « Pile est sorti ». On prend $\Omega = \{P,F\}^2$ muni de $\mathscr{P}(\Omega)$ et de l'équiprobabilité. On a

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

donc on a $P(A) = P(B) = \dfrac{3}{4}$. 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 ligne 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

$$\frac{\text{Nombre de colonnes} \begin{pmatrix} 1 \\ 1 \end{pmatrix}}{\text{Nombre de colonnes} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \text{ ou } \begin{pmatrix} 1 \\ 0 \end{pmatrix}} = \frac{N(A \cap B)}{N(A)} = \frac{\frac{N(A \cap B)}{N}}{\frac{N(A)}{N}} \xrightarrow[N \to +\infty]{} \frac{P(A \cap B)}{P(A)}$$

On calcule la fréquence (et plus tard la probabilité) en réduisant l'univers au « sous-univers » A. En simplifiant un peu, 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 Q, et de multiplier les probabilités des événements de A par un même facteur pour que leur somme reste égale à 1.

Cela justifie la définition suivante.

Définition. Soient A et B deux événements avec $P(A) \neq 0$. On appelle probabilité conditionnelle de B sachant A le réel

$$P_A(B) = \frac{P(A \cap B)}{P(A)}$$

Remarques :

- On trouve aussi la notation $P(B|A)$.
- En particulier, $P(A \cap B) = P(A) \times P_A(B)$.
- Dans notre exemple, $P_A(B) = \dfrac{2}{3}$ c'est-à-dire que, sachant que Face est sorti, Pile a une proba égale à $\dfrac{2}{3}$ de sortir. En d'autres termes : si vous lancez deux fois une pièce, et si vous savez que vous avez obtenu Face au moins une fois, alors vous avez une probabilité égale à $2/3$ que l'autre lancer soit un Pile.
- Le nombre $P_A(B)$ 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 $P_A(B)$ et pas $P(A \cap B)$. Il ne faut surtout pas les confondre : $P(A \cap B)$ est la probabilité que A et B soient réalisés en même temps, tandis que $P_A(B)$ est la probabilité de B quand on **sait** que A est réalisé.

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.

Exemple : On a un meuble avec trois tiroirs. Un objet a une chance sur deux de se trouver dans le meuble (dans n'importe quel tiroir). On a ouvert deux tiroirs : il n'y est pas. Quelle est la probabilité qu'il soit dans le troisième tiroir ?

On prend $\Omega = \{\text{extérieur}; T_1; T_2; T_3\}$ qu'on munit de $\mathscr{P}(\Omega)$ et de la probabilité P définie par

$$P(\{\text{extérieur}\}) = \frac{1}{2} \qquad \text{et} \qquad P(\{T_1\}) = P(\{T_2\}) = P(\{T_3\}) = \frac{1}{6}$$

En effet, on a vu dans le paragraphe II.2 que, pour définir une probabilité, il suffisait de la définir sur les événements élémentaires (i.e. les singletons). Soient A : « l'objet n'est pas dans les tiroirs 1 et 2 » et B : « l'objet est dans le troisième tiroir ». On cherche $P_A(B)$.

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

$$\begin{aligned} P_A(B) &= \frac{P(A\cap B)}{P(A)} \\ &= \frac{P(B)}{P(A)} \\ &= \frac{1/6}{2/3} \\ &= \frac{1}{4} \end{aligned}$$

Car $B \subset A$.

Exemple : Je sonne chez un ami qui a deux enfants. Un de ses enfants m'ouvre : c'est une fille. Quelle est la probabilité que l'autre enfant soit une fille ?
Notons $\Omega = \{F;G\}^2$ muni de $\mathscr{P}(\Omega)$ et de l'équiprobabilité. Notons A : « mon ami a deux filles » et B : « mon ami a au moins une fille » et on cherche $P_B(A)$. Alors $A = \{(F,F)\}$, $B = \{(F,G);(G,F);(F,F)\}$ et $A\cap B = A$ si bien que $P(A) = P(A\cap B) = 1/4$ (car $\text{card}(\Omega) = 4$) et $P(B) = 3/4$. Finalement,

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

$$\begin{aligned} P_B(A) &= \frac{P(A\cap B)}{P(B)} \\ &= \frac{1/4}{3/4} \\ &= \frac{1}{3} \end{aligned}$$

On trouve le même résultat que pour des Pile et des Face (et ça n'avait choqué personne) : parmi les couples avec deux enfants, il y a une proportion (environ) égale à $2/3$ de couples avec un garçon et une fille, et une proportion de $1/3$ de couples avec deux filles.

Théorème. Soit A un événement avec $P(A) \neq 0$. L'application

$$P_A : \left\{ \begin{array}{ccc} \mathscr{P}(\Omega) & \to & [\,0\,;1\,] \\ B & \mapsto & P_A(B) \end{array} \right.$$

est une probabilité et donc $(\Omega, \mathscr{P}(\Omega), P_A)$ est un espace probabilisé.

Démonstration.

- Tout d'abord,

$$\begin{aligned} P_A(\Omega) &= \frac{P(A\cap \Omega)}{P(A)} \\ &= \frac{P(A)}{P(A)} \\ &= 1 \end{aligned}$$

- Soit $B \in \mathscr{P}(\Omega)$. D'une part,

$$P_A(B) = \frac{P(A\cap B)}{P(A)} \geqslant 0$$

D'autre part, $A\cap B \subset A$ donc $P(A\cap B) \leqslant P(A)$ donc

$$P_A(B) = \frac{P(A\cap B)}{P(A)} \leqslant \frac{P(A)}{P(A)} = 1$$

- Soient B et C appartenant à $\mathscr{P}(\Omega)$ incompatibles.

$$\begin{aligned} P_A(B\cup C) &= \frac{P(A\cap(B\cup C))}{P(A)} \\ &= \frac{P((B\cap A)\cup(C\cap A))}{P(A)} \end{aligned}$$

Or, B et C sont incompatibles donc $B\cap A$ et $C\cap A$ sont incompatibles, et P est une probabilité donc

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

□

En effet, s'il existe $\omega \in (B\cap A)\cap(C\cap A)$, alors en particulier $\omega\in B$ et $\omega\in C$ ce qui est absurde car B et C sont incompatibles.

Remarque : Dans l'exemple avec le meuble, on munit $(\Omega,\mathscr{P}(\Omega))$ de la proba définie par :

$$P_A(\{\text{extérieur}\})=\frac{3}{4} \qquad P_A(\{T_1\})=P_A(\{T_2\})=0 \qquad \text{et} \qquad P_A(\{T_3\})=\frac{1}{4}$$

Remarque : Attention, la notion de probabilité conditionnelle utilise l'expression « proba de B sachant A ». Cela 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, pour les deux pièces, on ne peut se provoquer **qu'après** avoir lancé les deux pièces.

Autre 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 calculer $P_A(B)$.

III.2 Trois théorèmes capitaux

III.2.a 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 $P(A)\neq 0$, alors

$$P(A\cap B)=P(A)\times P_A(B).$$

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

Théorème (formule des probabilités composées). Soit $n\geqslant 2$. Soient $A_1,\dots,A_n$ des événements tels que $P(A_1\cap\dots\cap A_n)\neq 0$. Alors

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

Démonstration. par récurrence : ⇝ Exercice.

Remarque : Puisque $(A_1\cap\dots\cap A_{n-1})\subset(A_1\cap\dots\cap A_{n-2})\subset\dots\subset A_1\cap A_2\subset A_1$, alors

$$0<P(A_1\cap\dots\cap A_{n-1})\leqslant P(A_1\cap\dots\cap A_{n-2})\leqslant\dots\leqslant P(A_1\cap A_2)\subset P(A_1)$$

En d'autres termes, tous les événements par lesquels on conditionne sont de probabilité non nulle : toutes les probabilités conditionnelles sont bien définies.

⚠ **Cadre :** Cette formule est utile pour calculer la proba d'une intersection d'événements qui ne sont pas indépendants (voir partie suivante), par exemple quand on répète une expérience et quand les conditions changent à chaque fois (typiquement lors de tirages sans remise).

Exemple : Soient $(a,b)\in(\mathbb{N}^*)^2$ et $N=a+b$. On considère une urne contenant a boules blanches et b boules noires. On effectue des tirages successifs de la façon suivante :

- Si on tire une boule blanche, on la remet dans l'urne.
- Si on tire une boule noire, on ne la remet pas dans l'urne et on la remplace par une boule blanche.

Si $k \in [\![1 \,;\, b+1]\!]$, on note A_k : « On tire la première boule blanche au k-ième tirage ». Montrer que $P(A_{b+1}) = \dfrac{b!}{N^b}$ et que, pour tout $k \in [\![1 \,;\, b]\!]$,

$$P(A_k) = \frac{b!}{(b-(k-1))!N^{k-1}} - \frac{b!}{(b-k)!N^k}$$

Soit $k \in [\![1 \,;\, b+1]\!]$. Notons B_k : « la boule tirée lors du k-ième tirage est blanche » et N_k : « la boule tirée lors du k-ième tirage est noire ». Par conséquent,

$$A_k = N_1 \cap \cdots \cap N_{k-1} \cap B_k$$

D'après la formule des probabilités composées :

$$P(A_{b+1}) = P(N_1) \times P_{N_1}(N_2) \times P_{N_1 \cap N_2}(N_3) \times \cdots \times P_{N_1 \cap \cdots \cap N_{b-1}}(N_b) \times P_{N_1 \cap \cdots \cap N_b}(B_{b+1})$$

De plus :

- $P(N_1) = \dfrac{b}{N}$ car, au début de l'expérience, il y a b boules noires et N boules dans l'urne.
- Supposons N_1 réalisé. Alors on a tiré une boule noire lors du premier tirage, qu'on a remplacée par une boule blanche, si bien que l'urne contient ensuite $a-1$ boules blanches et $b+1$ boules noires (donc toujours N boules). Par conséquent, $P_{N_1}(N_2) = \dfrac{b-1}{N}$.
- Supposons $N_1 \cap N_2$ réalisé. Alors on a tiré deux boules noires lors des deux premiers tirages, qu'on a remplacées par deux boules blanches : l'urne contient alors $a+2$ boules blanches et $b-2$ boules noires (et toujours N boules) donc $P_{N_1 \cap N_2}(N_3) = \dfrac{b-2}{N}$.
- Si Plus généralement, soit $k \in [\![1 \,;\, b-1]\!]$, et supposons $N_1 \cap \cdots \cap N_{k-1}$ réalisé. Alors on a tiré $k-1$ boules noires qu'on a remplacées par $k-1$ boules blanches : l'urne contient alors $a+k-1 = a+k-1$ boules blanches et $b-(k-1) = b-k+1$ boules noires (et toujours N boules) donc

$$P_{N_1 \cap \cdots \cap N_{k-1}}(N_k) = \frac{b-k+1}{N}$$

- Enfin, si $k \in [\![1 \,;\, b]\!]$, on trouve de même que

$$P_{N_1 \cap \cdots \cap N_k}(A_{k+1}) = \frac{a+k}{N} = \frac{N-b+k}{N}$$

  En particulier, si $k = b$, cette proba vaut 1.

Dès lors :

$$\begin{aligned} P(A_{b+1}) &= \frac{b}{N} \times \frac{b-1}{N} \times \cdots \times \frac{1}{N} \times 1 \\ &= \frac{b!}{N^b} \end{aligned}$$

Soit à présent $k \in [\![1 \,;\, b]\!]$. Alors $A_k = N_1 \cap \cdots \cap N_{k-1} \cap B_k$ donc, toujours d'après la formule des probabilités composées, et d'après ce qui précède :

Première chose à faire : introduire des noms d'événements, cf. paragraphes I.3.a et V.

Inutile de chercher de l'indépendance (cf. paragraphe IV) ici : la composition de l'urne change à chaque tirage !

Rédaction à retenir pour calculer une probabilité conditionnelle !

On peut se demander la raison de cette attirance des probabilistes pour les urnes... La réponse est simple : beaucoup de problèmes probabilistes peuvent être modélisés par des urnes, en particulier des problèmes avec des possibilités non équiprobables... Par exemple, on peut simuler le dé non équilibré vu plus haut (celui pour lequel $P(\{k\}) = k/21$) à l'aide d'une urne contenant une boule numéro 1, deux boules numérotées 2, etc. jusque 6 boules numérotées 6.

$$
\begin{aligned}
P(A_k) &= P(N_1) \times P_{N_1}(N_2) \times \cdots \times P_{N_1 \cap \cdots \cap N_{k-2}}(N_{k-1}) \times P_{N_1 \cap \cdots \cap N_{k-1}}(B_k) \\
&= \frac{b}{N} \times \frac{b-1}{N} \times \cdots \times \frac{b-(k-2)}{N} \times \frac{N-b+k-1}{N} \\
&= \frac{b}{N} \times \frac{b-1}{N} \times \cdots \times \frac{b-k+2}{N} \times \left(1 - \frac{b-k+1}{N}\right) \\
&= \frac{b!}{N^{k-1}(b-k+1)!} \times \left(1 - \frac{b-k+1}{N}\right) \\
&= \frac{b!}{N^{k-1}(b-k+1)!} - \frac{b!}{N^k(b-k)!}
\end{aligned}
$$

III.2.b Diviser pour régner : la formule des probabilités totales.

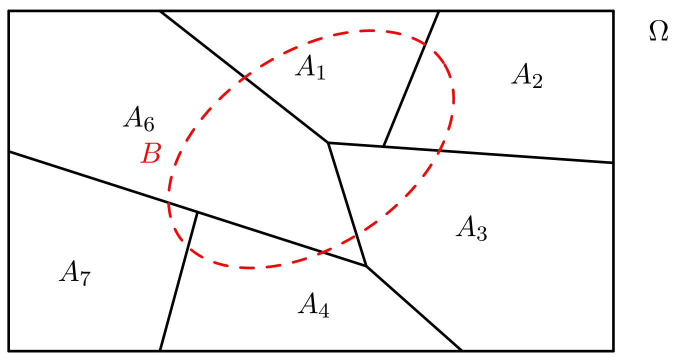
Théorème (Formule des probabilités totales). Soient $n \geqslant 2$, $(A_1, \dots, A_n)$ un système complet d'événements et B un événement. Alors :

$$P(B) = \sum_{i=1}^{n} P(B \cap A_i)$$

De plus, si toutes les $P(A_i)$ sont non nulles :

$$P(B) = \sum_{i=1}^{n} P(A_i) P_{A_i}(B)$$

Faisons simple : la première expression (avec les intersections) n'est utile en pratique que pour calculer des lois marginales, cf. chapitre 27. Dans les autres cas de figure, on utilise la seconde (avec les probas conditionnelles).



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.

Cadre : Quand il y a plusieurs cas de figure et quand, dans chaque cas de figure, la situation est très simple.

Démonstration.

$$
\begin{aligned}
P(B) &= P(B \cup \Omega) \\
&= P\left(B \cap \left(\bigcup_{i=1}^{n} A_i\right)\right) \\
&= P\left(\bigcup_{i=1}^{n} (B \cap A_i)\right)
\end{aligned}
$$

Car $\Omega = \bigcup_{i=1}^{n} A_i$.

Car l'intersection est distributive par rapport à l'union.

Or, les A_i sont deux à deux incompatibles donc les $B \cap A_i$ le sont également. D'où :

$$P(B) = \sum_{i=1}^{n} P(B \cap A_i)$$

□

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

Théorème (formule des probabilités totales). Soient A et B deux événements. Alors :

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

De plus, si $P(A)$ et $P(\overline{A})$ sont non nuls (i.e. si $P(A) \in \,]\, 0\,;1\,[$) :

$$P(B) = P(A) \times P_A(B) + P(\overline{A}) \times P_{\overline{A}}(B)$$

Remarques :

- 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. $P_{A_1}(B)$), la probabilité de B dans le cas où on a obtenu A_2 (i.e. $P_{A_2}(B)$), etc. et enfin la probabilité de B dans le cas où on a obtenu A_n (i.e. $P_{A_n}(B)$). Si les n cas sont deux à deux incompatibles 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.
- Il est parfois compliqué de déterminer si, pour tout $k \in [\![1\,;n]\!]$, $P(A_k) \neq 0$ (hypothèse indispensable si on veut que le terme $P_{A_k}(B)$ soit défini). Mais heureusement, par convention, lorsque $k \in [\![1\,;n]\!]$ est tel que $P(A_k) = 0$, $P_{A_k}(B)P(A_k) = 0$ (cette convention a un sens car cette quantitée est issue du terme $P(B \cap A_k)$, qui est lui-même nul puisque $B \cap A_k \subset A_k$). En conclusion, il n'est pas nécessaire de vérifier que les événements A_k ont une probabilité non nulle pour appliquer la deuxième version de la formule des probas totales (qui est celle qu'on appliquera toujours en pratique).

Il faut absolument appliquer cette formule à une famille $(A_k)_{1 \leqslant k \leqslant n}$ qui est un système complet d'événements. Tous les indices $k \in [\![1\,;n]\!]$ doivent figurer dans la somme, même si $\mathscr{P}_{A_k}(B) = 0$. C'est seulement dans un second temps que l'on peut enlever les termes nuls de la somme.

Exemple : On se donne une urne avec quatre jetons numérotés $1, 2, 3, 4$. On retire au hasard deux jetons de l'urne et on les met dans une nouvelle urne, en y ajoutant un jeton numéroté 5. On tire un jeton dans cette nouvelle urne. Quelle est la probabilité que le jeton tiré soit impair ?

Il est naturel de séparer les cas selon le nombre de jetons impairs dans la nouvelle urne. Notons :

- A_1 : « la deuxième urne contient les jetons $2, 4, 5$ ».
- A_2 : « la deuxième urne contient les jetons $1, 2, 5$, $1, 4, 5$, $2, 3, 5$ ou $3, 4, 5$ ».
- A_3 : « la deuxième urne contient les jetons $1, 3, 5$ ».
- B : « on tire un jeton impair dans la nouvelle urne ».

Il y a plusieurs cas particuliers (selon le nombre de jetons impairs dans l'urne) et dans chaque cas particulier on sait répondre (si on connaît le nombre de jetons impairs dans l'urne, le résultat est trivial) : probas totales !

On a $P(A_1) = P(A_3) = 1/6$ et $P(A_2) = 4/6 = 2/3$. Or, A_1, A_2 et A_3 forment un système complet d'événements. D'après la formule des probas totales :

$$\begin{aligned} P(B) &= P_{A_1}(B) \times P(A_1) + P_{A_2}(B) \times P(A_2) + P_{A_3}(B) \times P(A_3) \\ &= \frac{1}{3} \times \frac{1}{6} + \frac{2}{3} \times \frac{2}{3} + 1 \times \frac{1}{6} \\ &= \frac{12}{18} \\ &= \frac{2}{3} \end{aligned}$$

Attention, cela ne veut pas dire qu'il y a deux jetons impairs dans la nouvelle urne !

Au lycée, vous auriez fait un arbre. Maintenant, c'est fini, un arbre n'est pas une démonstration !

III.2.c Formule de Bayes

Théorème (Formule de Bayes). Soient A et B deux événements avec $P(A) \neq 0$ et $P(B) \neq 0$. Alors :

$$P_B(A) = \frac{P_A(B) \times P(A)}{P(B)}$$

DÉMONSTRATION. $P(A \cap B) = P_B(A) \times P(B) = P_A(B) \times P(A)$.

⚠ **Cadre :** Quand on cherche une probabilité conditionnelle, et que la probabilité conditionnelle « miroir » est très simple (i.e. lorsqu'on cherche $P_B(A)$ et qu'on connaît $P_A(B)$).

Remarque : Cette formule est souvent utilisée avec la formule des probabilités totales pour le calcul de $P(B)$: cf TD.

Dans certains livres, c'est cette formule (avec la formule des probabilités totales au dénominateur) qu'on appelle formule de Bayes.

Remarque : ⚠ La formule de Bayes montre que, en général, $\mathscr{P}_A(B) \neq \mathscr{P}_B(A)$. 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.

Exemple : On réalise un test pour détecter une maladie qui touche $0,5\%$ de la population. Sur la population malade, le test est positif dans 95% des cas. Cependant ce test déclare malades (à tort) 5% des personnes saines. Le test est positif : quelle est la probabilité d'être malade ?

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

Notons M : « être malade », P : « le test est positif ». Dès lors :

$$P(M) = \frac{5}{1000} = \frac{1}{200}, \qquad P_M(P) = \frac{95}{100} \qquad \text{et} \qquad P_{\overline{M}}(P) = \frac{5}{100}$$

On cherche $P_P(M)$. D'après la formule de Bayes,

$$P_P(M) = \frac{P_M(P) \times P(M)}{P(P)}$$

Or, M et $\overline{M}$ forment un système complet d'événements donc, d'après la formule des probabilités totales :

$$\begin{aligned} P(P) &= P(M) \times P_M(P) + P\left(\overline{M}\right) \times P_{\overline{M}}(P) \\ &= \frac{1}{200} \times \frac{95}{100} + \frac{199}{200} \times \frac{5}{100} \end{aligned}$$

Après calculs, on trouve que $P_P(M) = \dfrac{19}{218}$: ce test est-il fiable ?

IV Indépendance

IV.1 Indépendance de deux événements

Intuitivement, deux événements sont indépendants quand le fait que l'un est réalisé n'apporte aucune information sur le fait que l'autre est ou non réalisé. Du point de vue probabiliste, cela donne :

$$P_A(B) = \frac{P(A \cap B)}{P(A)} = P(B) \qquad \text{et} \qquad P_B(A) = \frac{P(A \cap B)}{P(B)} = P(A)$$

Inconvénient : on doit avoir $P(A) \neq 0$ et $P(B) \neq 0$. On contourne cet inconvénient avec la définition suivante :

Définition. Deux événements A et B sont indépendants si $P(A \cap B) = P(A)P(B)$.

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

Démonstration. Immédiat.

Exemple : On lance deux fois un dé équilibré. On définit les événements suivants :

- A : « obtenir 6 au premier lancer »
- B : « obtenir 4 au deuxième lancer »
- C : « obtenir au moins un 6 »
- D : « obtenir au moins un 4 »

On se place sur $\Omega = [\![1\,; 6]\!]^2$ muni de $\mathscr{P}(\Omega)$ et de l'équiprobabilité.

- $P(A) = P(B) = 1/6, P(A \cap B) = 1/36 = P(A) \times P(B)$. Les événements A et B sont donc indépendants. C'est logique ! Le résultat du premier lancer ne donne aucune information sur le résultat du second.
- $P(C) = P(D) = 11/36$ mais $P(C \cap D) = 2/36 \neq P(C) \times P(D)$. C et D ne sont pas indépendants, ce qui est également logique : si on a au moins un 4 qui sort, il « reste moins de place » pour avoir au moins un 6.

Remarques :

- 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.
- Il ne faut pas confondre indépendants et incompatibles ! 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. Avec la définition mathématique, c'est presque le cas : deux événements incompatibles sont indépendants si et seulement si l'un des deux est de probabilité nulle (exo).

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.

Remarque : Plus généralement, tout événement est indépendant d'un événement de probabilité nulle. En effet, si B est un événement avec $P(B) = 0$ alors, pour tout événement A, $A \cap B \subset B$ donc $0 \leqslant P(A \cap B) \leqslant P(B) = 0$ d'où $P(A \cap B) = 0 = P(A) \times P(B)$. C'est logique ! Si B ne se produit jamais, le fait qu'il ne se produise pas ne donne aucune information sur A.

Proposition. Soient A et B deux événements.

$$\begin{array}{rl} & A \text{ et } B \text{ sont indépendants} \\ \iff & A \text{ et } \overline{B} \text{ sont indépendants} \\ \iff & \overline{A} \text{ et } B \text{ sont indépendants} \\ \iff & \overline{A} \text{ et } \overline{B} \text{ sont indépendants} \end{array}$$

Nous généraliserons ce résultat dans le paragraphe suivant.

Remarque : Encore une fois, c'est intuitif ! Si le fait que A se produise ne donne aucune information sur B, le fait que qu'il ne se produise pas n'en donne pas non plus !

Démonstration. Supposons A et B indépendants. B et $\overline{B}$ forment un système complet d'événements donc, d'après la formule des probabilités totales :

$$\begin{aligned} P(A) &= P(A \cap B) + P(A \cap \overline{B}) \\ &= P(A) \times P(B) + P(A \cap \overline{B}) \end{aligned}$$

Car A et B sont indépendants.

Dès lors :

$$
\begin{aligned}
P(A\cap\overline{B}) &= P(A)-P(A)\times P(B)\\
&= P(A)\times(1-P(B))\\
&= P(A)\times P\left(\overline{B}\right) \qquad \square
\end{aligned}
$$

c'est-à-dire que A et $\overline{B}$ sont indépendants.

Réciproquement, si A et $\overline{B}$ sont indépendants, d'après ce qui précède, A et $\overline{\overline{B}}=B$ sont indépendants. D'où la première équivalence. On a les autres par symétrie des rôles.

Remarque : Si $P(B)=1$ alors $P\left(\overline{B}\right)=0$ donc tout événement est indépendant de $\overline{B}$ donc de B. Encore une fois, c'est intuitif ! Si B se produit tout le temps, le fait qu'il se produise ne nous apprend rien sur A !

Remarque : Attention, quand on dit que deux événements sont indépendants, il est sous-entendu : « pour la probabilité P ». L'indépendance dépend en effet de la probabilité choisie. Autrement dit, deux événements peuvent être indépendants pour une probabilité et ne pas l'être pour une autre.

Exemple : On lance deux fois une pièce équilibrée. On se place sur $\Omega=\{P,F\}^2$ muni de $\mathscr{P}(\Omega)$ et de l'équiprobabilité. Soient A : « On obtient une seule fois Face », B : « On obtient Face au premier lancer » et C : « On obtient Face au second lancer », c'est-à-dire qu'on a

$$
A=\{(F,P),(P,F)\} \qquad B=\{(F,P),(F,F)\} \qquad \text{et} \qquad C=\{(P,F),(F,F)\}
$$

B et C sont indépendants (intuitivement, et par le calcul) mais ne le sont pas pour la proba P_A. En effet,

$$
P_A(B\cap C)=\frac{P(A\cap B\cap C)}{P(A)}=0 \qquad \text{et} \qquad P_A(B)\times P_A(C)=\frac{P(A\cap B)}{P(A)}\times\frac{P(A\cap C)}{P(A)}=\frac{1}{4}
$$

Encore une fois, c'est logique ! Quand on conditionne par A, les événements B et C « deviennent incompatibles » : la réalisation de B sachant que A est réalisé entraîne la non-réalisation de C. Ils ne sont donc plus indépendants.

IV.2 Indépendance mutuelle de n événements

Définition (indépendance mutuelle). On dit que les événements $A_1,\dots,A_n$ sont mutuellement indépendants si, pour toute partie I de $[\![1\,;\,n]\!]$:

$$
P\left(\bigcap_{i\in I}A_i\right)=\prod_{i\in I}P(A_i)
$$

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

Remarque : Il est inutile de vérifier l'égalité si $I=\varnothing$ ou si I a un seul élément (si $I=\{i_0\}$, l'égalité à vérifier devient $P(A_{i_0})=P(A_{i_0})$ qui est toujours vraie...).

Exemple : Pour montrer que trois événements A_1,A_2,A_3 sont indépendants, il faut donc montrer que :

- $P(A_1\cap A_2)=P(A_1)\times P(A_2)$.
- $P(A_1\cap A_3)=P(A_1)\times P(A_3)$.
- $P(A_2\cap A_3)=P(A_2)\times P(A_3)$.
- $P(A_1\cap A_2\cap A_3)=P(A_1)\times P(A_2)\times P(A_3)$.

Pour quatre événements A_1,A_2,A_3,A_4 : exo, il faut vérifier 11 égalités. Dans le cas général, il faut vérifier $2^n-(n+1)$ égalités (cf. exercice 18).

Deux erreurs courantes (à ne pas commettre) :

- On dit que $A_1, \dots, A_n$ sont indépendants deux à deux si pour tous $i \neq j$, A_i et A_j sont indépendants. Si $A_1, \dots, A_n$ sont mutuellement indépendants, alors ils sont indépendants deux à deux (prendre $I = \{i ; j\}$) mais

  LA RÉCIPROQUE EST FAUSSE ! ! ! ! !

  Exemple : Prenons la même expérience (on lance deux fois une pièce) associée au même espace probabilisé. Notons A : « On obtient Face au premier lancer », B : « On obtient Face au deuxième lancer » et C : « les deux lancers donnent un résultat différent ». Intuitivement, A, B, C ne sont pas mutuellement indépendants car, si on sait que A et B sont réalisés, alors on sait que C ne l'est pas. Prouvons-le par le calcul. En explicitant $A, B, C, A \cap B$ etc. (par exemple, $A = \{(F, F) ; (F ; P)\}$), on trouve que

  - $\star$ $P(A \cap B) = \dfrac{1}{4} = P(A) \times P(B)$.
  - $\star$ $P(A \cap C) = \dfrac{1}{4} = P(A) \times P(C)$.
  - $\star$ $P(B \cap C) = \dfrac{1}{4} = P(B) \times P(C)$.

  si bien que A, B, C sont indépendants deux à deux mais $P(A \cap B \cap C) = 0 \neq P(A)P(B)P(C)$: A, B, C ne sont pas indépendants.

- Autre erreur courante : si $A_1, \dots, A_n$ sont indépendants alors $P(A_1 \cap \cdots \cap A_n) = P(A_1) \times \cdots \times P(A_n)$ (prendre $I = [\![1 ; n]\!]$) mais là aussi

  LA RÉCIPROQUE EST FAUSSE ! ! ! ! !

  c'est-à-dire qu'avoir $P(A_1 \cap \cdots \cap A_n) = P(A_1) \times \cdots \times P(A_n)$ ne suffit pas pour que les événements $A_1, \dots, A_n$ soient mutuellement indépendants.

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

Morale de l'histoire : Pour montrer que n événements sont mutuellement indépendants, il n'y a rien d'autre que la définition.

Théorème. Soient $A_1, \dots, A_n$ des événements mutuellement indépendants. Si pour tout $k \in [\![1 ; n]\!]$, $B_k = A_k$ ou $\overline{A_k}$, alors $B_1, \dots, B_k$ 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.

Démonstration. Prouvons tout d'abord que $\overline{A_1}, A_2, \dots, A_n$ sont mutuellement indépendants. Soit I une partie de $[\![1 ; n]\!]$. Notons $B_1 = \overline{A_1}, B_2 = A_2, \dots, B_n = A_n$, et prouvons que

$$P\left(\bigcap_{i \in I} B_i\right) = \prod_{i \in I} P(B_i)$$

Si $1 \notin I$, le résultat découle de l'indépendance de $A_1, \dots, A_n$. Supposons donc que $1 \in I$, et notons $B = \displaystyle\bigcap_{i \in I, i \neq 1} B_i$, si bien que $B = \displaystyle\bigcap_{i \in I, i \neq 1} A_i$. Les événements $A_1, \dots, A_n$ étant mutuellement indépendants :

$$P\left(\bigcap_{i\in I} A_i\right) = \prod_{i\in I} P(A_i) \qquad \text{et} \qquad P(B) = P\left(\bigcap_{i\in I, i\neq 1} A_i\right) = \prod_{i\in I, i\neq 1} P(A_i)$$

Or :

$$\bigcap_{i\in I} A_i = A_1 \cap \bigcap_{i\in I, i\neq 1} A_i = A_1 \cap B \qquad \text{et} \qquad \prod_{i\in I} P(A_i) = P(A_1) \times \prod_{i\in I, i\neq 1} P(A_i) = P(A_1) \cap P(B)$$

On en déduit que $P(A_1 \cap B) = P(A_1) \times P(B)$ c'est-à-dire que A_1 et B sont indépendants. D'après le paragraphe précédent, $\overline{A_1}$ et B sont indépendants, c'est-à-dire que

$$P(\overline{A_1} \cap B) = P(\overline{A_1}) \times P(B)$$

En utilisant les résultats précédents :

$$\begin{aligned}
P\left(\bigcap_{i\in I} B_i\right) &= P\left(\overline{A_1} \bigcap_{i\in I, i\neq 1} B_i\right) \\
&= P(\overline{A_1} \cap B) \\
&= P(\overline{A_1}) \times P(B) \\
&= P(\overline{A_1}) \times \prod_{i\in I, i\neq 1} P(A_i) \\
&= P(B_1) \times \prod_{i\neq i, i\neq 1} P(B_i) \\
&= \prod_{i\in I} P(B_i)
\end{aligned}$$

□

ce qui est le résultat voulu : $\overline{A_1}, A_2, \ldots, A_n$ sont indépendants. Par symétrie des rôles, si on a n événements indépendants, ils restent indépendants lorsqu'on remplace un événement (pas forcément le premier) par son complémentaire. On peut évidemment itérer le processus : en remplaçant plusieurs A_k par leur complémentaire, les événements restent indépendants, ce qui est le résultat voulu.

V Comment aborder un exercice de probabilités

Nous exposons ici le principe général du calcul (et de la rédaction de ce calcul) d'une probabilité. Évidemment, dans la pratique, les situations sont très variées, et on peut être amené à s'écarter légèrement de la voie tracée ci-dessous.

- Dans une situation d'équiprobabilité, on peut se diriger vers un argument combinatoire (mais c'est assez rare).
- Il faut donner des noms à tous les événements auxquels on peut penser. Il ne faut pas hésiter à en introduire beaucoup, mieux vaut trop que pas assez, ce travail n'est pas forcément fait pour vous dans le sujet.
- Cas particulier important : lorsqu'on répète plusieurs fois une expérience, ne pas oublier d'indiquer dans le nom et dans la description de l'événement à quelle répétition on fait allusion. Par exemple, quand on lance n fois une pièce, poser P : « obtenir Pile » est trop vague, on préférera poser P_k : « obtenir Pile au k-ième lancer ».

- Une fois ces événements simples définis, exprimer l'événement « global » (i.e. celui dont on cherche la probabilité) à l'aide de ces événements simples et d'opérations ensemblistes (intersections, unions, complémentaire, différence etc.).
- Si on a un « au moins » ou « au plus » dans l'énoncé : se demander s'il ne serait pas plus simple de calculer la probabilité du complémentaire.
- Si on a une intersection : si les événements sont indépendants, la probabilité de l'intersection est le produit des probabilités. Sinon (par exemple lors de tirages avec remise), on applique la formule des probabilités composées.
- Si on a une union : on se demande s'ils sont deux à deux incompatibles. Si c'est le cas, la probabilité de l'union est la somme des probabilités. Sinon, on fait autrement.
- Si on a une union d'événements indépendants : l'indépendance « marchant » bien avec l'intersection, on « passe au complémentaire » et, avec les lois de Morgan, on se ramène à une intersection d'événements indépendants. Sinon, on fait autrement.
- Si on a une expérience lors de laquelle plusieurs issues sont possibles, et si on sait répondre dans chaque cas de figure, alors on applique la formule des probabilités totales. De toute façon, **dans chaque exercice de probabilités**, cela ne coûte pas cher de se demander si on peut appliquer cette formule. C'est le résultat le plus important du chapitre !
- Pour calculer une probabilité conditionnelle : la rédaction est toujours la même et est non négociable. Pour calculer $P_A(B)$, on écrit : « Supposons A réalisé. Alors… »