Chapitre 2

Probabilités sur un univers fini

2.1 Notion de probabilité

2.1.1 Vocabulaire probabiliste

On appelle ***expérience aléatoire*** une expérience renouvelable dont les résultats possibles sont connus sans que l'on puisse prévoir lequel sera obtenu.

On appelle ***univers*** (ou encore ***ensemble fondamental***) d'une expérience aléatoire l'ensemble de tous les résultats possibles de cette expérience. Il est souvent noté Ω. Cet ensemble peut être fini, infini dénombrable, ou infini non dénombrable mais on ne s'intéressera ici qu'au cas où il est fini.
Un élément $\omega \in \Omega$ est appelé ***éventualité*** ou ***issue***. C'est donc un résultat possible de l'expérience aléatoire.

Exemple. Un lancer de dé (en observant le chiffre visible) est une expérience aléatoire d'univers $\Omega = \{1, 2, 3, 4, 5, 6\}$.

Parfois, au lieu de s'intéresser à toute l'expérience, on ne s'intéresse qu'à certains résultats. On appelle ***évènement*** toute partie $A \in \mathscr{P}(\Omega)$ de l'univers. C'est donc une combinaison de résultats possibles de l'expérience aléatoire.
Un évènement A est dit ***réalisé*** lorsque le résultat de l'expérience appartient à A.

Exemple. On lance deux dés distincts, l'un est rouge et l'autre est bleu et on observe les chiffres obtenus. Un résultat de l'expérience peut être vu comme un couple (i, j), avec $i \in \{1, \cdots, 6\}$ et $j \in \{1, \cdots, 6\}$, où i représente la valeur du dé rouge et j représente la valeur du dé bleu. L'univers est alors
$\Omega = \{(1, 1), (1, 2), (1, 3), \cdots, (2, 1), (2, 2), \cdots, (6, 5), (6, 6)\} = \{1, \cdots, 6\}^2$.
On peut considérer les évènements suivants :

- A : « la somme des points est égale à 6 ». On a alors $A = \{(1, 5), (5, 1), (2, 4), (4, 2), (3, 3)\}$,
- B : « la somme des points est un multiple de 3 ». On a alors

$$B = \{(1, 2), (2, 1), (1, 5), (5, 1), (2, 4), (4, 2), (3, 3), (3, 6), (6, 3), (4, 5), (5, 4), (6, 6)\}.$$

On appelle ***évènement impossible*** l'ensemble vide $\emptyset$, et ***évènement certain*** l'univers Ω.
On appelle ***évènement élémentaire*** un évènement qui ne contient qu'un seul élément. (En terminologie ensembliste on parle de singleton.)

2.1.2 Opérations sur les évènements

Soit A et B deux évènements d'une expérience aléatoire d'univers Ω.

- **Complémentaire**. Le complémentaire $\overline{A}$ de A dans Ω (aussi noté A^c) est l'***évènement contraire*** de A : $\overline{A}$ contient tous les éléments de Ω qui ne sont pas dans A, $\overline{A} = \Omega \backslash A$. $\overline{A}$ est l'évènement qui se réalise lorsque A ne se réalise pas.
- **Intersection** : L'évènement $C = A \cap B$ est réalisé lorsque A ET B sont réalisés simultanément. On parle donc de l'évènement ***« A ET B »***.
- ***Union*** : L'évènement $D = A \cup B$ est réalisé lorsque l'un des évènements A OU B est réalisé. On parle donc de l'évènement ***« A OU B »***
- Deux évènements A et B sont dits ***incompatibles*** ou disjoints si $A \cap B = \emptyset$.
- La relation « l'évènement A implique l'évènement B » signifie que si A se réalise alors B se réalise aussi. La relation ensembliste associée est donc celle de l'inclusion : $A \subset B$.

Exemple. On revient à l'exemple précédent et on considère les évènements A : « la somme des points vaut 6 », B : « la somme des points est un multiple de 3 », C : « on a au moins un 6 ». On a alors :

$$\begin{array}{ll} A & = \{(1,5),(5,1),(2,4),(4,2),(3,3)\} \\ B & = \{(1,2),(2,1),(1,5),(5,1),(2,4),(4,2),(3,3),(3,6),(6,3),(4,5),(5,4),(6,6)\} \\ C & = \{(1,6),(6,1),(2,6),(6,2),(3,6),(6,3),(4,6),(6,4),(5,6),(6,5),(6,6)\} \end{array}$$

On a $A \subset B$, $A \cap C = \emptyset$, et $\overline{C}$ est l'évènement « n'avoir aucun 6 ».

Définition 2.1. *Soit $A_1, \cdots, A_p$ des évènements. On dit que la famille $(A_i)_{1 \leqslant i \leqslant p}$ est un **système complet d'évènements** si ces évènements sont deux à deux incompatibles (c'est-à-dire $\forall i \neq j$, $A_i \cap A_j = \emptyset$) et si $\bigcup_{i=1}^{p} A_i = \Omega$.*

Exemples.

- Pour tout évènement A, la famille $(A, \overline{A})$ est un système complet d'évènements.
- Si $\Omega = \{\omega_1, \cdots, \omega_n\}$ alors la famille $(\{\omega_i\})_{1 \leqslant i \leqslant n}$ est un système complet d'évènements.

2.1.3 Bases axiomatiques des probabilités

Considérons une expérience aléatoire dont l'univers est un ensemble fini Ω. Intuitivement, la probabilité d'un évènement est la fréquence d'apparition de cet évènement lorsqu'on réalise une infinité de fois dans les mêmes conditions l'expérience.

Définition 2.2. *On appelle **probabilité**, ou mesure de probabilité, sur l'univers fini Ω toute application $\mathbb{P} : \mathscr{P}(\Omega) \longrightarrow \mathbb{R}$ vérifiant les trois axiomes :*

(1) $\forall A \in \mathscr{P}(\Omega) \qquad \mathbb{P}(A) \geqslant 0$ *(positivité)*

(2) $\mathbb{P}(\Omega) = 1$ *(totalité)*

(3) $\forall (A,B) \in \mathscr{P}(\Omega)^2 \qquad (A \cap B = \emptyset \Longrightarrow \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B))$. *($\sigma$-additivité)*

Lorsque Ω est muni d'une probabilité $\mathbb{P}$, on dit que $(\Omega, \mathbb{P})$ est un espace probabilisé (fini).

Remarque. Une probabilité sur Ω est une application définie sur $\mathscr{P}(\Omega)$: on calcule des probabilités d'évènements.

Proposition 2.3. *Soit $\mathbb{P}$ une probabilité sur Ω.*

1. $\mathbb{P}(\emptyset) = 0$
2. *Pour tout évènement A,* $\mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A)$
3. *Si $A \subset B$, alors $\mathbb{P}(A) \leqslant \mathbb{P}(B)$. (On dit qu'une probabilité est croissante.)*

4. *Pour tout évènement A, $\mathbb{P}(A) \in [0,1]$.*
5. *Si $A_1, A_2, \cdots, A_p$ est une suite finie d'évènements deux à deux disjoints alors on a $\mathbb{P}\left(\bigcup_{i=1}^{p} A_i\right) = \sum_{i=1}^{p} \mathbb{P}(A_i)$*

Démonstration : 1. Il suffit de prendre $A = \Omega$ et $B = \emptyset$. 2. On remarque que $\Omega = A \cup \overline{A}$ (union disjointe). 3. On remarque que si $A \subset B$, on a $B = A \cup (B \backslash A)$ (union disjointe), où $B \backslash A = B \cap \overline{A}$. On a donc $\mathbb{P}(B) = \mathbb{P}(A) + \mathbb{P}(B \backslash A) \geqslant \mathbb{P}(A)$. 4. Conséquence de 3. et des axiomes de positivité et de totalité. 5. On procède par récurrence sur $p \geqslant 2$. □

Exercice. Montrer que si A et B sont deux évènements alors $\mathbb{P}(A \backslash B) = \mathbb{P}(A) - \mathbb{P}(A \cap B)$.

Proposition 2.4. *Si A et B sont deux évènements quelconques, alors*

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

Démonstration : On décompose les évènements en unions disjointes (comme dans le théorème 1.8 du chapitre 1) à l'aide du système complet d'évènements induit par B (par exemple) : $A \cup B = B \cup (A \cap \overline{B})$ et $A = (A \cap B) \cup (A \cap \overline{B})$. La σ-additivité de $\mathbb{P}$ permet alors d'écrire $\mathbb{P}(A \cup B) = \mathbb{P}(B) + \mathbb{P}(A \cap \overline{B})$ et $\mathbb{P}(A) = \mathbb{P}(A \cap B) + \mathbb{P}(A \cap \overline{B})$. Le résultat annoncé s'en déduit. □

Exercice. Montrer que si $A_1, \cdots, A_n$ sont n évènements alors $\mathbb{P}(\bigcup_{i=1}^{n} A_i) \leqslant \sum_{i=1}^{n} P(A_i)$.

Proposition 2.5 (Formule des probabilités totales)**.** *Si $A_1, \cdots A_p$ forment un système complet d'évènements alors, pour tout évènement B, on a :*

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

Démonstration : On a $B = B \cap \Omega = B \cap \bigcup_{i=1}^{p} A_i$. Par distributivité de l'intersection par rapport à la réunion, on en déduit $B = \bigcup_{i=1}^{p} (B \cap A_i)$ et donc le résultat annoncé par σ-additivité. □

Proposition 2.6 (Construction d'une probabilité)**.** *Si $\Omega = \{\omega_1, \cdots, \omega_n\}$ et si $p_1, \ldots, p_n$ sont des nombres positifs tels que $p_1 + \cdots + p_n = 1$, alors il existe une unique probabilité $\mathbb{P}$ sur Ω telle que :* $\forall i \in [[1,n]], \quad \mathbb{P}(\{\omega_i\}) = p_i$.
Cette probabilité est définie par : $\forall A \in \mathscr{P}(\Omega), \quad \mathbb{P}(A) = \sum_{\omega \in A} \mathbb{P}(\{\omega\})$.

Démonstration :

- Unicité. Si $p_1, \ldots, p_n$ sont des nombres positifs tels que $p_1 + \cdots + p_n = 1$ et si $\mathbb{P}$ est une probabilité sur Ω telle que $\forall i \in [[1,n]], \quad \mathbb{P}(\{\omega_i\}) = p_i$, alors, pour $A \subset \Omega$, $A = \bigcup_{\omega \in A} \{\omega\}$ et la σ-additivité donne bien $\mathbb{P}(A) = \sum_{\omega \in A} \mathbb{P}(\{\omega\})$.
- Existence. Il s'agit de vérifier que l'application $\mathbb{P}$ ainsi définie est bien une probabilité. On a déjà, $\forall A \subset \Omega, \mathbb{P}(A) = \sum_{\omega \in A} \mathbb{P}(\{\omega\}) \geqslant 0$ (car les p_i sont positifs) mais aussi $\mathbb{P}(\Omega) = \sum_{\omega \in \Omega} \mathbb{P}(\{\omega\}) = \sum_{i=1}^{n} p_i = 1$. Soit à présent A et B deux évènements incompatibles.

$\mathbb{P}(A\cup B) = \sum_{\omega\in A\cup B} \mathbb{P}(\{\omega\}) = \sum_{\omega\in A} \mathbb{P}(\{\omega\}) + \sum_{\omega\in B} \mathbb{P}(\{\omega\})$ (car A et B sont disjoints) et donc $\mathbb{P}(A\cup B) = \mathbb{P}(A) + \mathbb{P}(B)$.

□

Remarques.

- On dit parfois que la famille $\{p_1, \cdots, p_n\}$ est une ***distribution de probabilité*** sur Ω.
- Se donner une probabilité sur $\Omega = \{\omega_1, \cdots, \omega_n\}$ revient donc à se donner n réels positifs $p_1, \ldots, p_n$ de somme 1. La partie réciproque de cette affirmation est l'objet de la proposition précédente et, pour la partie directe, si $\mathbb{P}$ est une probabilité sur Ω, il suffit de poser $p_i = \mathbb{P}(\{\omega_i\})$ (les axiomes de positivité et de totalité permettant alors de conclure).

Exemples. 1. On jette une pièce bien équilibrée (même chance de tomber sur pile que sur face) et on regarde la face apparente. On a $\Omega = \{P, F\}$. Comme la pièce est bien équilibrée, la probabilité liée à cette expérience vérifie $\mathbb{P}(\{P\}) = \mathbb{P}(\{F\})$. Comme $\mathbb{P}$ est une probabilité, on a forcément $\mathbb{P}(\{P\}) + \mathbb{P}(\{F\}) = 1$. On en déduit que,

$$\mathbb{P}(\{P\}) = \mathbb{P}(\{F\}) = \frac{1}{2}.$$

2. Jetons une pièce biaisée de telle sorte qu'on ait deux fois plus de chance d'avoir pile que face, i.e $\mathbb{P}(\{P\}) = 2\mathbb{P}(\{F\})$. Comme $\mathbb{P}(\{P\}) + \mathbb{P}(\{F\}) = 1$, on a alors

$$\mathbb{P}(\{P\}) = \frac{2}{3} \text{ et } \mathbb{P}(\{F\}) = \frac{1}{3}$$

3. On jette un dé et on regarde le chiffre de la face apparente. On a $\Omega = \{1, 2, \cdots, 6\}$. On suppose que le dé est bien équilibré (non truqué), c'est à dire que chaque face à la même chance d'apparaître. On a donc $\mathbb{P}(\{i\}) = \frac{1}{6}$ pour $i \in \{1, 2, \cdots, 6\}$.
Soit A l'évènement « obtenir un chiffre pair », on a alors

$$\mathbb{P}(A) = \mathbb{P}(\{2, 4, 6\}) = \mathbb{P}(\{2\} \cup \{4\} \cup \{6\}) = \mathbb{P}(\{2\}) + \mathbb{P}(\{4\}) + \mathbb{P}(\{6\}) = \frac{1}{2}$$

Soit B l'évènement « obtenir au moins deux ». On remarque que $\overline{B} =$« tomber sur 1 », d'où

$$\mathbb{P}(B) = 1 - \mathbb{P}(\overline{B}) = 1 - \frac{1}{6} = \frac{5}{6}.$$

2.1.4 Cas particulier de l'équiprobabilité

Considérons une expérience d'univers fini Ω, disons $\Omega = \{\omega_1, \cdots, \omega_n\}$ dont chaque évènement élémentaire a la même chance d'apparaître. On dit que l'expérience est équiprobable. Comme $\sum_{i=1}^{n} \mathbb{P}(\{\omega_i\}) = 1$, on en déduit que

$$\mathbb{P}(\{\omega_1\}) = \mathbb{P}(\{\omega_2\}) = \cdots = \mathbb{P}(\{\omega_n\}) = \frac{1}{n}$$

Cette probabilité est appelée ***probabilité uniforme*** sur l'ensemble fini $\Omega = \{\omega_1, \cdots, \omega_n\}$. Considérons un évènement A, on a alors

$$\mathbb{P}(A) = \sum_{\omega\in A} \frac{1}{n} = \frac{\text{card}(A)}{n} = \frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}}$$

Exemples. 1. On jette deux dés équilibrés distinguables (l'un rouge, l'autre bleu) et on note les chiffres qui apparaissent. On peut modéliser l'expérience en introduisant l'univers $\Omega = \{1, 2, \cdots, 6\}^2$ constitué des couples (i, j) où i désigne le chiffre du dé rouge et j celui du bleu. Les dés étant équilibrés, l'expérience est équiprobable et on a donc

$$\mathbb{P}(\{(i, j)\}) = \frac{1}{36} \text{ pour tout } (i, j) \in \Omega$$

2. On jette maintenant deux dés équilibrés et parfaitement identiques et on note les chiffres qui apparaissent. On peut modéliser l'expérience en introduisant l'univers

$$\Omega' = \Big\{(i, j) \in \{1, \cdots, 6\}^2 \text{ tels que } i \leqslant j\Big\} = \{(1, 1), \cdots, (1, 6), (2, 2), \cdots, (5, 5), (5, 6), (6, 6)\}$$

Le cardinal de Ω' est 21. On souhaite munir l'espace Ω' d'une probabilité $\mathbb{P}'$, mais il n'y a cette fois-ci aucune raison de supposer l'expérience équiprobable.
Cependant, on peut faire le lien avec l'expérience précédente en mettant un point de couleur sur l'un des dés. Cela ne change pas le déroulement de l'expérience, mais les dés étant maintenant distinguables on se retrouve dans le cas précédent où l'on connaît la probabilité de chaque événement élémentaire. On introduit alors l'univers $\Omega = \{1, \cdots, 6\}^2$ muni de la probabilité uniforme $\mathbb{P}$. Le lien entre les espaces $(\Omega', \mathbb{P}')$ et $(\Omega, \mathbb{P})$ est le suivant : pour $i < j$, on a

$$\mathbb{P}'(\{(i, j)\}) = \mathbb{P}(\{(i, j\}) + \mathbb{P}(\{(j, i)\}) = \frac{1}{18}$$

et $\mathbb{P}'(\{(i, i)\}) = \mathbb{P}(\{(i, i)\}) = \frac{1}{36}$.
De manière générale, dans le cas d'un univers fini et lorsque cela est possible, on essaye de se ramener à des expériences équibrobables afin de simplifier la modélisation et les calculs.
Le calcul des probabilités est alors ramené à un problème de dénombrement...

Exemple. On jette deux dés équilibrés et on observe la somme des chiffres affichés. On cherche la probabilité de l'évènement A : « La somme égale à 7 ». L'univers « naturel » est $\Omega = [\![2, 12]\!]$ (ensemble des résultats possibles de l'expérience aléatoire) mais il n'est alors pas raisonnable de munir Ω de la probabilité uniforme pour décrire cette expérience : il est statistiquement clair que la somme 2 sera obtenue beaucoup moins souvent que la somme 7. Même si les dés ne sont pas supposés distincts, on a vu dans l'exemple précédent que l'on pouvait se ramener au cas de dés distincts. On considère donc $\Omega = \{1, \cdots, 6\}^2$ que l'on munit de la probabilité uniforme. On a alors $A = \{(1, 6), (2, 5), (3, 4), (4, 3), (5, 2), (6, 1)\}$ et $\mathbb{P}(A) = \dfrac{6}{36} = \dfrac{1}{6}$.

2.2 Probabilités conditionnelles

2.2.1 Introduction

- *Le problème de Monty Hall*

  Un jeu oppose un présentateur à un candidat (le joueur).
  Le joueur est placé devant trois portes fermées. Derrière l'une d'entre elles, se trouve une voiture et derrière chacune des deux autres se trouve une chèvre. Le joueur doit tout d'abord désigner une porte. Puis le présentateur ouvre une porte qui n'est ni celle choisie par le candidat, ni celle cachant la voiture (le présentateur sait quelle est la bonne porte dès le début). Le candidat a alors le droit d'ouvrir la porte qu'il a choisie initialement, ou d'ouvrir la troisième porte. S'il ouvre la porte derrière laquelle se trouve la voiture, il l'emporte.
  Que doit faire le joueur pour maximiser ses chances de gagner la voiture ?

- Prenons un autre exemple simple : on lance successivement deux dés équilibrés et on observe la somme des chiffres apparents. On modélise l'expérience, comme déjà vu, par le couple $(\Omega, \mathbb{P})$ où $\Omega = \{1, \cdots, 6\}^2$ et $\mathbb{P}$ est la probabilté uniforme. Si A désigne l'évènement « La somme vaut 6 », on a $A = \{(1,5),(2,4),(3,3),(4,2),(5,1)\}$ et donc $\mathbb{P}(A) = \dfrac{5}{36}$. Supposons maintenant que l'on ait une information supplémentaire : on sait que le premier dé est tombé sur la face 2. Si B désigne l'évènement « Le premier lancer donne 2 » (on a donc $B = \{(2,1),(2,2),(2,3),(2,4),(2,5),(2,6)\}$), les cas favorables ne sont plus qu'au nombre de un : $(2,4)$ (c'est le seul élément de $A \cap B$) tandis que les cas possibles sont ceux de B. La probabilité d'obtenir une somme de 6 sachant que le premier dé est tombé sur 2 vaut donc $\dfrac{1}{6} = \dfrac{\text{Card}(A \cap B)}{\text{Card}(B)}$. C'est la probabilité de A sachant B.
  Généralisons à présent cet exemple.

2.2.2 Définition et premières propriétés

Définition 2.7. *Soit A et B deux évènements avec $P(B) \neq 0$. La* ***probabilité conditionnelle*** *de A sachant B, notée $P_B(A)$ ou $P(A|B)$, est définie par :*

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

Exemple. On jette deux fois une pièce de monnaie équilibrée et on cherche la probabilité que les deux jets donnent face sachant que le premier a donné face. On modélise l'expérience en introduisant l'univers $\Omega = \{(F,F),(F,P),(P,F),(P,P)\}$ que l'on munit de la probabilité uniforme (la pièce étant équilibrée, les résultats sont équiprobables). On introduit les évènements suivants A : « les deux jets donnent face », B : « le premier jet donne face ».
En utilisant, la formule de la définition, on a $\mathbb{P}_B(A) = \dfrac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \dfrac{\mathbb{P}(A)}{\mathbb{P}(B)} = \dfrac{\frac{1}{4}}{\frac{2}{4}} = \dfrac{1}{2}$.

Cependant on remarque que si l'on sait que le premier jet donne face, tout l'aléa de l'expérience est dans le second lancer. Sachant que le premier dé donne face, on peut considérer l'univers $\Omega' = \{(F,F),(F,P)\}$. Ayant autant de chance que le second lancer donne face que pile, on a donc $\mathbb{P}_B(A) = 1/2$. On retrouve le même résultat.
On s'intéresse maintenant à la probabilité que les deux jets donnent face sachant qu'au moins un des jets donne face. On note C l'évènement « au moins un des jets donne face ». D'après la formule, on a

$$\mathbb{P}_C(A) = \frac{\mathbb{P}(A \cap C)}{\mathbb{P}(C)} = \frac{\mathbb{P}(A)}{\mathbb{P}(C)} = \frac{\frac{1}{4}}{\frac{3}{4}} = \frac{1}{3}$$

Sachant qu'au moins un des jets donne face, l'aléa repose sur l'autre jet et on peut considérer l'univers $\Omega' = \{(F,F),(F,P),(P,F)\}$ et donc $\mathbb{P}_C(A) = \frac{1}{3}$.

Proposition 2.8. *Soit $(\Omega, \mathbb{P})$ un espace probabilisé fini et B un évènement tel que $\mathbb{P}(B) \neq 0$. Alors l'application*

$$\begin{array}{rcl} \mathbb{P}_B : \mathscr{P}(\Omega) & \longrightarrow & \mathbb{R} \\ A & \longmapsto & \mathbb{P}_B(A) \end{array}$$

est une probabilité sur Ω.

Démonstration :

- Il est clair que $\forall A \subset \Omega,\ \mathbb{P}_B(A) \geqslant 0$,
- On a bien $\mathbb{P}_B(\Omega) = \dfrac{\mathbb{P}(\Omega \cap B)}{\mathbb{P}(B)} = \dfrac{\mathbb{P}(B)}{\mathbb{P}(B)} = 1$,

- Soit alors $(A, A') \in \mathscr{P}(\Omega)^2$ avec $A \cap A' = \emptyset$.

$$\mathbb{P}_B(A \cup A') = \frac{\mathbb{P}((A \cup A') \cap B)}{\mathbb{P}(B)} = \frac{\mathbb{P}((A \cap B) \cup (A' \cap B))}{\mathbb{P}(B)}.$$

Comme $A \cap B$ et $A' \cap B$ sont disjoints, on en déduit

$$\mathbb{P}_B(A \cup A') = \frac{\mathbb{P}(A \cap B) + \mathbb{P}(A' \cap B)}{\mathbb{P}(B)} = \mathbb{P}_B(A) + \mathbb{P}_B(A').$$

□

Remarque. On a $\mathbb{P}_B(B) = 1$: tout le poids de la probabilité est porté par B. $\mathbb{P}_B$ est donc aussi une probabilité sur B.

2.2.3 Différentes formules

Proposition 2.9 (Formule des probabilités totales - Version 2)**.** *Si $B_1, \cdots B_p$ forment un système complet d'évènements et sont tous de probabilité non nulle alors, pour tout évènement A, on a :*

$$\mathbb{P}(A) = \sum_{i=1}^{p} \mathbb{P}(A \cap B_i) = \sum_{i=1}^{p} \mathbb{P}_{B_i}(A).\mathbb{P}(B_i).$$

Démonstration : Il suffit d'appliquer la première version de cette formule et la définition de la probabilité conditionnelle. □

Exemple. Une compagnie d'assurance répartit ses clients en deux classes : ceux qui sont enclins aux risques cardio-vasculaires (c'est-à-dire à haut risque) et ceux qui ne le sont pas (c'est-à-dire à risque modéré). Elle constate que ceux à haut risque ont une probabilité de 0.4 d'avoir un accident dans l'année, tandis que pour ceux à risque modéré cette probabilité est de 0.05. On suppose que 30% de la population est à haut risque. Quelle est alors la probabilité qu'un nouvel assuré ait un accident cardio-vasculaire durant l'année qui suit la signature de son contrat ?

On note A l'évènement « l'assuré a un accident pendant l'année » et B l'évènement « l'individu est à haut risque ». Alors d'après la formule des probabilités totales, on a

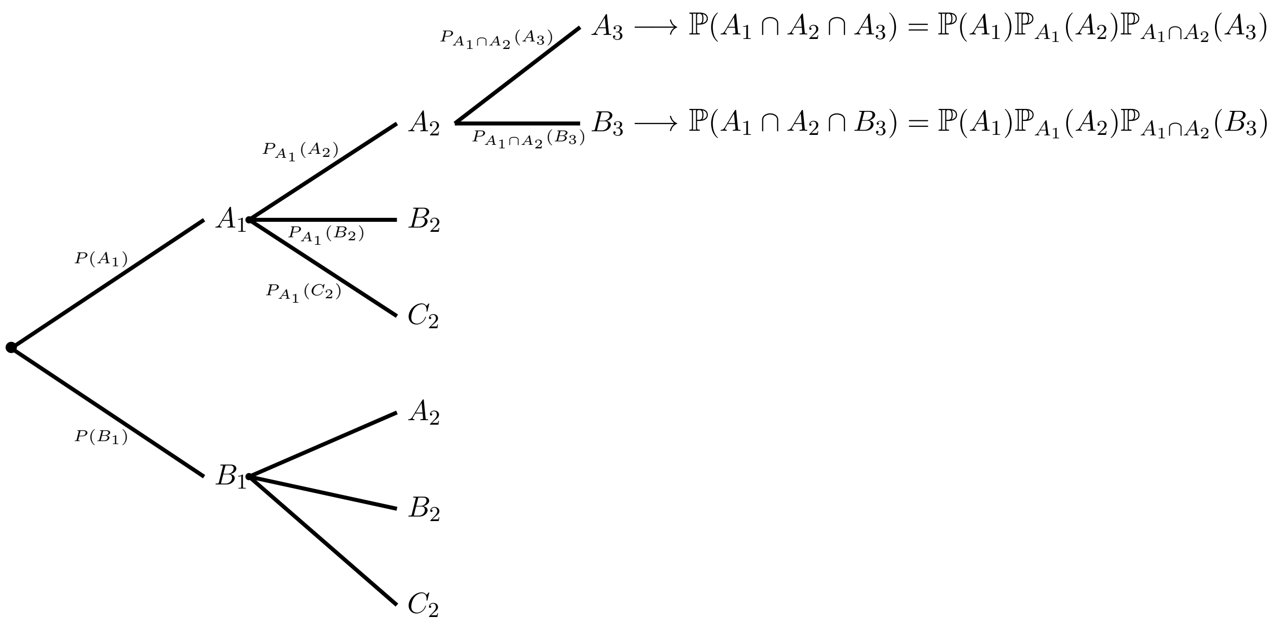
$$\mathbb{P}(A) = \mathbb{P}_B(A).\mathbb{P}(B) + \mathbb{P}_{\overline{B}}(A).\mathbb{P}(\overline{B}) = 0.4 \times 0.3 + 0.05 \times 0.7 = 0.155$$

Proposition 2.10 (Probabilités composées)**.** *Soit $A_1, \cdots, A_{p-1}$ des évènements dont l'intersection est de probabilité non nulle et soit A_p un évènement. Alors :*

$$\mathbb{P}(A_1 \cap \cdots \cap A_p) = \mathbb{P}(A_1) \times \mathbb{P}_{A_1}(A_2) \times \mathbb{P}_{A_1 \cap A_2}(A_3) \cdots \mathbb{P}_{A_1 \cap \cdots \cap A_{p-1}}(A_p).$$

Démonstration : On procède par récurrence en partant de $\mathbb{P}(A_1 \cap A_2) = \mathbb{P}(A_1).\mathbb{P}_{A_1}(A_2)$ (qui résulte de la définition d'une probabilité conditionnelle). □

Remarque. Ce résultat justifie le calcul des probabilités à l'aide d'un arbre de probabilité : dans un arbre pondéré, la probabilité de l'évènement à l'extrémité d'un chemin (qui est l'intersection des évènements rencontrés sur ce chemin) est égale au produit des probabilités rencontrées le long du chemin correspondant.



Exemple. Une urne contient quatre boules blanches et trois noires. On tire successivement (sans remise) trois boules. La probabilité d'obtenir une boule noire puis deux blanches est de $\frac{3}{7} \times \frac{4}{6} \times \frac{3}{5} = \frac{6}{35}$. La probabilité d'obtenir une boule noire et deux blanches est de $\frac{3}{7} \times \frac{4}{6} \times \frac{3}{5} + \frac{4}{7} \times \frac{3}{6} \times \frac{3}{5} + \frac{4}{7} \times \frac{3}{6} \times \frac{3}{5} = \frac{4}{7}$.

Proposition 2.11 (Formule de Bayes)**.** *Si A et B sont deux évènements de probabilité non nulle alors on a :* $\mathbb{P}_A(B) = \dfrac{\mathbb{P}_B(A)\mathbb{P}(B)}{\mathbb{P}_B(A)\mathbb{P}(B) + \mathbb{P}_{\overline{B}}(A)\mathbb{P}(\overline{B})}$.
Plus généralement, si $B_1, \cdots B_p$ forment un système complet d'évènements et sont tous de probabilité non nulle et si A est un évènement de probabilité non nulle alors :

$$\forall i \in [\![1, p]\!] \quad \mathbb{P}_A(B_i) = \frac{\mathbb{P}_{B_i}(A)\mathbb{P}(B_i)}{\displaystyle\sum_{k=1}^{p} \mathbb{P}_{B_k}(A)\mathbb{P}(B_k)}$$

Démonstration : Il suffit d'écrire $\mathbb{P}_A(B_i) = \dfrac{\mathbb{P}(A \cap B_i)}{\mathbb{P}(A)} = \dfrac{\mathbb{P}_{B_i}(A)\mathbb{P}(B_i)}{\mathbb{P}(A)}$ et d'utiliser la formule des probabilités totales. $\square$

Remarque. La formule de Bayes permet de calculer les probabilités a posteriori d'un évènement en fonction des probabilités a priori de cet évènement c'est-à-dire de connaître $\mathbb{P}_A(B)$ quand on connaît $\mathbb{P}_B(A)$ et $\mathbb{P}_{\overline{B}}(A)$. La formule de Bayes est donc souvent utilisée pour calculer des probabilités de causes dans des diagnostics (maladies, pannes, etc.).

Remarque. La formule de Bayes est à la base de toute une branche de la statistique appelée statistique bayesienne.

Exemple. Un étudiant répond à une question à choix multiples, m réponses sont proposées. Il connait la réponse avec une probabilité p. Dans le cas contraire, il choisit au hasard la réponse (avec une probabilité $1 - p$). Quelle est la probabilité qu'un étudiant connaisse la réponse s'il y a répondu correctement ?
En notant B l'évènement « l'étudiant donne la bonne réponse » et C l'évènement « l'étudiant

connaît la réponse », on a :

$$\mathbb{P}_B(C) = \frac{\mathbb{P}_C(B)\mathbb{P}(C)}{\mathbb{P}_C(B)\mathbb{P}(C) + \mathbb{P}_{\overline{C}}(B)(1 - \mathbb{P}(C))} = \frac{1 \times p}{p + (1-p)/m}$$

Si $p = 1/2$ et $m = 4$, alors $\mathbb{P}_B(C) = 4/5$.

2.3 Indépendance d'évènements

Dans le langage courant, on dit de deux évènements qui ne sont pas liés entre eux qu'ils sont indépendants. Par conséquent, on a tendance à dire que si A et B sont indépendants, la connaissance de B ne donne aucune information utile pour la connaissance de l'évènement A et donc de manière naturelle $\mathbb{P}_B(A)$ doit être égale à $\mathbb{P}(A)$, i.e. $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$. On va donner à partir de cette constatation la définition mathématique de l'indépendance.

Définition 2.12. *Deux évènements A et B sont dits **indépendants** si $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.*

Attention ! Les notions d'évènements indépendants et d'évènements disjoints n'ont aucun rapport.
Par exemple, si on jette deux dés et on considère les évènements A =« le premier dé vaut 3 » et B =« le deuxième dé vaut 1 ». Les évènements sont indépendants car on jette les dés de façon indépendante, mais les évènements ne sont pas disjoints car l'évènement $A \cap B =$ « le premier dé vaut 3 et le deuxième vaut 1 » est possible.

Exercice. À quelle condition deux évènements incompatibles sont-ils indépendants ?

Proposition 2.13. *Si A et B sont deux évènements indépendants, alors A et $\overline{B}$ le sont aussi.*

Démonstration : On écrit A comme l'union de deux évènements disjoints : $A = (A \cap B) \cup (A \cap \overline{B})$. D'où $\mathbb{P}(A) = \mathbb{P}(A \cap B) + \mathbb{P}(A \cap \overline{B}) = \mathbb{P}(A)\mathbb{P}(B) + \mathbb{P}(A \cap \overline{B})$. Ce qui implique que $\mathbb{P}(A \cap \overline{B}) = \mathbb{P}(A) - \mathbb{P}(A)\mathbb{P}(B) = \mathbb{P}(A)\mathbb{P}(\overline{B})$ et donc A et $\overline{B}$ sont indépendants. □

Définition 2.14. *On considère $n \geqslant 2$ évènements $A_1, \cdots, A_n$. Ces évènements sont dits **mutuellement indépendants** si pour tout $2 \leqslant k \leqslant n$, pour toute k-combinaison $\{i_1, \cdots, i_k\}$ de $\{1, \cdots, n\}$ on a $\mathbb{P}(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_k}) = \mathbb{P}(A_{i_1})\mathbb{P}(A_{i_2}) \cdots \mathbb{P}(A_{i_k})$.*
Par exemple, A, B et C sont trois évènements mutuellement indépendants si

$$\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B) \qquad \mathbb{P}(A \cap C) = \mathbb{P}(A)\mathbb{P}(C) \qquad \mathbb{P}(B \cap C) = \mathbb{P}(B)\mathbb{P}(C)$$

$$\text{et} \qquad \mathbb{P}(A \cap B \cap C) = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C).$$

Remarque. Si des évènements sont mutuellements indépendants, alors ils sont indépendants deux à deux, mais la réciproque est fausse. Par exemple, si on jette deux dés, l'un rouge l'autre bleu et si on considère les évènements suivants : $A = \{$le dé rouge donne un chiffre impair$\}$ $B = \{$le dé bleu donne un chiffre impair$\}$ $C = \{$la somme des deux dés est un chiffre impair$\}$ A, B et C sont deux à deux indépendants mais pas mutuellement indépendants.

2.4 Variables aléatoires sur un univers fini

2.4.1 Définition et notations

Après avoir réalisé une expérience, on ne s'intéresse bien souvent qu'à une certaine fonction du résultat et non au résultat en lui-même. Lorsqu'on lance des dés, on peut, par exemple, juste s'intéresser à la somme des chiffres apparus et pas nécessairement au résultat de chacun des dés. Ces grandeurs (ou applications) auxquelles on s'intéresse sont des fonctions réelles définies sur l'univers et sont appelées variables aléatoires.

Définition 2.15. *Soit $(\Omega, \mathbb{P})$ un espace probabilisé fini. On appelle* ***variable aléatoire*** *réelle toute application X de Ω dans $\mathbb{R}$.*

Notations :

- Pour tout ensemble $E \subset \mathbb{R}$, on note $[X \in E]$ l'évènement $\{\omega \in \Omega,\, X(\omega) \in E\} = X^{-1}(E)$.
- Pour tout réel x, on note $[X = x] = \{\omega \in \Omega,\, X(\omega) = x\} = X^{-1}(\{x\})$
- Pour tout réel x, on note $[X < x] = \{\omega \in \Omega,\, X(\omega) < x\} = X^{-1}(\{]-\infty, x[\})$
- On définit de même les évènements $[X > x]$, $[X \geqslant x]$, $[X < x]$ et $[X \leqslant x]$.

Exemples.

1. On jette deux dés distincts et on s'intéresse à la somme des points obtenus. On note X la variable aléatoire associée ; elle est définie par $X : \Omega \longrightarrow \mathbb{R}$, $\omega = (\omega_1, \omega_2) \longmapsto \omega_1 + \omega_2$ avec $\Omega = \{1, \cdots, 6\}^2$. L'ensemble des valeurs prises par X est $X(\Omega) = \{2, 3, \cdots, 12\}$.
   L'évènement $[X = 4]$ est l'ensemble
   $$[X = 4] = \Big\{(\omega_1, \omega_2) \in \{1, \cdots, 6\}^2,\, \omega_1 + \omega_2 = 4\Big\} = \{(1,3), (3,1), (2,2)\}$$
2. On lance toujours deux dés, mais cette fois on s'intéresse au plus grand chiffre Y obtenu. On a alors $Y : \Omega \longrightarrow \mathbb{R}$, $\omega = (\omega_1, \omega_2) \longmapsto \max(\omega_1, \omega_2)$ avec $\Omega = \{1, \cdots, 6\}^2$.
   La variable Y est à valeurs dans $\{1, 2, \cdots, 6\}$. L'évènement $[Y = 3]$ est égal à :
   $$[Y = 3] = \{(\omega_1, \omega_2) \in \{1, \cdots, 6\}^2,\, \max(\omega_1, \omega_2) = 3\} = \{(1,3), (3,1), (2,3), (3,2), (3,3)\}$$
3. Si a est un réel donné, l'application $X : \Omega \longrightarrow \mathbb{R}$, $\omega \longmapsto a$ est une variable aléatoire constante.

Opérations sur les variables aléatoires

En conséquence de la définition, si X et Y sont deux variables aléatoires sur Ω alors $X + Y$, XY et $X + a$ (où $a \in \mathbb{R}$) sont encore des variables aléatoires sur Ω.

Proposition 2.16. *Si X est une variable aléatoire sur Ω alors les $[X = x]$ pour $x \in X(\Omega)$ forment un système complet d'évènements. Autrement dit, on a $\Omega = \bigcup_{x \in X(\Omega)} [X = x]$ (réunion disjointe).*

Démonstration : Si x et y sont deux éléments distincts de $X(\Omega)$ alors $[X = x] \cap [X = y] = \emptyset$. En effet, un ω de $[X = x] \cap [X = y]$ serait tel que $X(\omega) = x$ et $X(\omega) = y$ ce qui est impossible pour $x \neq y$.
Montrons à présent, par double inclusion, l'égalité d'ensembles annoncée.

- Soit $x \in X(\Omega)$. Par définition même, $[X = x] \subset \Omega$. On a donc bien $\bigcup_{x \in X(\Omega)} [X = x] \subset \Omega$.
- Soit $\omega \in \Omega$. Posons $x = X(\omega)$. On a alors $\omega \in [X = x]$ et donc $\omega \in \bigcup_{x \in X(\Omega)} [X = x]$.

On a bien montré l'égalité voulue. □

Remarque. On a donc, pour toute variable aléatoire X sur Ω, $\sum_{x \in X(\Omega)} P([X = x_i]) = 1$.

2.4.2 Loi d'une variable aléatoire, fonction de répartition

Théorème 2.17. *Soient* $(\Omega, \mathbb{P})$ *un espace probabilisé fini et* X *une variable aléatoire sur cet espace. L'application*

$$\begin{array}{rcl} \mathbb{P}_X : \mathscr{P}(X(\Omega)) & \longrightarrow & [0,1] \\ E & \longmapsto & \mathbb{P}([X \in E]) \end{array}$$

est une probabilité sur l'espace probabilisable $(X(\Omega), \mathscr{P}(X(\Omega)))$*.*

Démonstration : On a vu plus haut que pour chaque partie E de $X(\Omega)$, $[X \in E]$ est un évènement. $\mathbb{P}_X$ est donc bien une application de $\mathscr{P}(X(\Omega))$ dans $[0,1]$.

- Il est clair que $\forall E \in \mathscr{P}(X(\Omega))$, $\mathbb{P}_X(E) \geqslant 0$.
- $\mathbb{P}_X(X(\Omega)) = \mathbb{P}([X \in X(\Omega)]) = \mathbb{P}(\Omega) = 1$.
- Soit (E, F) un couple de parties de $X(\Omega)$ avec $E \cap F = \emptyset$. Alors, $[X \in E]$ et $[X \in F]$ sont deux évènements disjoints (un élément de l'intersection aurait une image qui serait à la fois dans E et dans F). Donc, ,

$$\begin{aligned} \mathbb{P}_X(E \cup F) &= \mathbb{P}([X \in E \cup F]) \\ &= \mathbb{P}([X \in E] \cup [X \in F]) \\ &= \mathbb{P}([X \in E]) + \mathbb{P}([X \in F]) \quad \text{par } \sigma\text{-additivité de } \mathbb{P} \\ &= \mathbb{P}_X(E) + \mathbb{P}_X(F) \end{aligned}$$

On a montré que $\mathbb{P}_X$ est une probabilité sur l'espace probabilisable $(X(\Omega), \mathscr{P}(X(\Omega))$. □

Définition 2.18. *L'application* $\mathbb{P}_X$ *du théorème s'appelle la* ***loi de probabilité*** *(ou encore* ***probabilité image****) de la variable aléatoire* X*.*

Remarque. La loi $\mathbb{P}_X$ de X est entièrement déterminée par $X(\Omega)$ et les atomes de probabilité $\mathbb{P}_X(\{x\}) = \mathbb{P}(X \in \{x\}) = \mathbb{P}([X = x])$ pour tous les x de $X(\Omega)$.
En effet, $\forall E \subset X(\Omega)$, comme $[X \in E] = \bigcup_{x \in E} [X = x]$ et que cette réunion est disjointe, $\mathbb{P}([X \in E]) = \sum_{x \in E} \mathbb{P}([X = x])$.

Conséquence pratique

Pour répondre à une question du type « déterminer la loi de probabilité de la variable X », on donne $X(\Omega)$ et les $\mathbb{P}([X = x])$ pour tous les x de $X(\Omega)$.

Notation. La loi de X est en général notée $\mathscr{L}(X)$. Deux variables aléatoires X et Y ont la même loi de probabilité si et seulement si $X(\Omega) = Y(\Omega)$ et $\forall a \in X(\Omega)$, $\mathbb{P}([X = a]) = \mathbb{P}([Y = a])$. Quand deux variables aléatoires X et Y ont la même loi de probabilité, on écrit $X \sim Y$. Quand une variable aléatoire suit une loi de probabilité $\mathscr{L}$ donnée, on écrit $X \sim \mathscr{L}$.

Définition 2.19. *Soient* $(\Omega, \mathbb{P})$ *un espace probabilisé fini et* X *une variable aléatoire sur cet espace. On appelle* ***fonction de répartition*** *de la variable* X *la fonction* F_X*, donnée par*

$$F_X : \mathbb{R} \longrightarrow [0,1],\ x \longmapsto \mathbb{P}([X \leqslant x])$$

Proposition 2.20. *La fonction de répartition* F_X *d'une variable aléatoire finie est croissante et constante par morceaux.*

Démonstration : Soit $(x,y) \in \mathbb{R}^2$ tel que $x \leqslant y$, on a

$$[X \leqslant x] = \{\omega \in \Omega,\ X(\omega) \leqslant x\} \subset \{\omega \in \Omega,\ X(\omega) \leqslant y\} = [X \leqslant y]$$

Par conséquent, $F_X(x) = \mathbb{P}(X \leqslant x) \leqslant F_X(y)$ et f_X est bien croissante.
Si on pose $X(\Omega) = \{x_1, \cdots, x_n\}$ avec $x_1 < \cdots < x_n$ alors pour $x \in \mathbb{R}$

$$F_X(x) = \begin{cases} 0 & \text{si } x < x_1 \\ \displaystyle\sum_{i=1}^{k} \mathbb{P}([X = x_i]) & \text{si } x \in [x_k, x_{k+1}[\\ 1 & \text{si } x \geqslant x_n \end{cases}$$

□

Remarques.

- Les sauts de la fonction de répartition F_X ont lieu en les points x_k et la hauteur du saut au point x_k est égale à $\mathbb{P}([X = x_k])$.
- On a $\lim\limits_{x \to -\infty} F_X(x) = 0$ et $\lim\limits_{x \to +\infty} F_X(x) = 1$.

Exemples. Reprenons les deux exemples précédents.

1. (Lancer de deux dés distincts et X la variable aléatoire donnant la somme des points obtenus.) X est à valeurs dans $\{2, \cdots, 12\}$, donc $\mathbb{P}([X = k]) = 0$ pour $k \notin \{2, 3, \cdots, 12\}$.

k	2	3	4	5	6	7	8	9	10	11	12
$\mathbb{P}([X = k])$	1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

2. (Lancer de deux dés distincts et Y la variable aléatoire égale au plus grand des deux chiffres obtenus.)
   Y est à valeurs dans $\{1, \cdots, 6\}$, donc $\mathbb{P}([Y = k]) = 0$ pour $k \notin \{1, 2, \cdots, 6\}$.

k	1	2	3	4	5	6
$\mathbb{P}([Y = k])$	1/36	3/36	5/36	7/36	9/36	11/36
$F_Y(k)$	1/36	4/36	9/36	16/36	25/36	1

2.4.3 Fonctions de variables aléatoires

Soit $(\Omega, \mathbb{P})$ un espace probabilisé fini. Si X est une variable aléatoire sur Ω à valeurs dans un ensemble E et si f est une application de E vers un ensemble F, on peut définir la composée $f \circ X$ que l'on note abusivement $f(X)$. C'est une nouvelle variable aléatoire sur Ω. Ainsi, si X est une variable aléatoire réelle, on peut définir les variables aléatoires $X^2, |X|, e^X$...

On s'intéresse maintenant à la loi de probabilité de $f(X)$.

Proposition 2.21. *Soit $(\Omega, \mathbb{P})$ un espace probabilisé fini. Soient X une variable aléatoire sur Ω à valeurs dans un ensemble non vide E et f une application de E vers un ensemble non vide F. On pose $Y = f(X)$. La loi de probabilité $\mathbb{P}_Y$ de Y est donnée par*

$$\forall y \in F \quad \mathbb{P}_Y(\{y\}) = \mathbb{P}([Y = y]) = \sum_{x \in f^{-1}(\{y\})} \mathbb{P}([X = x])$$

Démonstration : Soit $y \in F$. On a $[Y = y] = [f \circ X = y] = \{\omega \in \Omega,\ f[X(\omega)] = y\}$. Or, $f(a) = b \iff a \in f^{-1}(\{b\})$ (on rappelle que $f^{-1}(B) = \{x \in E,\ f(x) \in B\}$). On a donc $[Y = y] = \{\omega \in \Omega,\ X(\omega) \in f^{-1}(\{y\})\} = [X \in f^{-1}(\{y\})]$. $[Y = y]$ est donc la réunion disjointe des $[X = x]$ pour $x \in f^{-1}(\{y\})$ et le résultat découle alors de la σ-additivité de $\mathbb{P}$. □

2.4.4 Lois conditionnelles

Définition 2.22. *Soit X une variable aléatoire réelles sur un espace probabilisé fini $(\Omega, \mathbb{P})$ et $A \subset \Omega$ un évènement tel que $\mathbb{P}(A) \neq 0$. La **loi conditionnelle** de X sachant A est la loi de la variable aléatoire X dans l'espace probabilisé $(\Omega, \mathbb{P}_A)$ c'est-à-dire :*

$$\forall x \in X(\Omega) \quad \mathbb{P}_A([X = x]) = \frac{\mathbb{P}([X = x] \cap A)}{\mathbb{P}(A)}$$

2.4.5 Variables aléatoires indépendantes

Définition 2.23. *Soit X et Y deux variables aléatoires réelles sur un espace probabilisé fini $(\Omega, \mathbb{P})$. On dit que X et Y sont **indépendantes** si pour tout x de $X(\Omega)$ et tout y de $Y(\Omega)$ on a $\mathbb{P}([X = x] \cap [Y = y]) = \mathbb{P}([X = x])\mathbb{P}([Y = y])$.*

Remarque. On définit, de même que pour les évènements, la notion d'indépendance (mutuelle) de n variables aléatoires.

Exemple. Les variables aléatoires X et Y données en exemple plus haut ne sont pas indépendantes.
On a en effet $[X = 4] \cap [Y = 4] = \emptyset$ donc $0 = \mathbb{P}([X = 4] \cap [Y = 4]) \neq \mathbb{P}([X = 4])\mathbb{P}([Y = 4])$.

2.4.6 Épreuves successives indépendantes

Considérons deux expériences aléatoires modélisées respectivement par les espaces probabilisés finis $(\Omega_1, \mathbb{P}_1)$ et $(\Omega_2, \mathbb{P}_2)$. On peut alors modéliser la succession de ces deux expériences indépendantes par l'espace $(\Omega_1 \times \Omega_2, \mathbb{P})$ où

$$\forall (A_1, A_2) \in \mathscr{P}(\Omega_1) \times \mathscr{P}(\Omega_2) \quad \mathbb{P}(A_1 \times A_2) = \mathbb{P}_1(A_1)\mathbb{P}_2(A_2).$$

Proposition 2.24. *Il existe une unique probabilité $\mathbb{P}$ sur $\Omega_1 \times \Omega_2$ vérifiant la propriété précédente.*

Remarque.
Attention à ne pas confondre $\mathscr{P}(\Omega_1 \times \Omega_2)$ et $\mathscr{P}(\Omega_1) \times \mathscr{P}(\Omega_2)$! Si, pour $(A_1, A_2) \in \mathscr{P}(\Omega_1) \times \mathscr{P}(\Omega_2)$ on a bien $A_1 \times A_2 \in \mathscr{P}(\Omega_1 \times \Omega_2)$, la réciproque est en général fausse. Par exemple, pour $\Omega_1 = \Omega_2 = \{0, 1\}$ on a bien $X = \{(0, 0), (1, 1)\} \in \mathscr{P}(\Omega_1 \times \Omega_2)$ mais X ne peut pas s'écire sous la forme $X = A_1 \times A_2$ avec $(A_1, A_2) \in \mathscr{P}(\Omega_1) \times \mathscr{P}(\Omega_2)$.

Démonstration : L'ensemble $\Omega = \Omega_1 \times \Omega_2$ étant fini, on utilise 2.6. Les réels $\mathbb{P}_1(\{\omega_1\})\mathbb{P}_2(\{\omega_2\})$ sont tous positifs et on a

$$\sum_{(\omega_1,\omega_2) \in \Omega_1 \times \Omega_2} \mathbb{P}_1(\{\omega_1\})\mathbb{P}_2(\{\omega_2\}) = \sum_{\omega_1 \in \Omega_1} \sum_{\omega_2 \in \Omega_2} \mathbb{P}_1(\{\omega_1\})\mathbb{P}_2(\{\omega_2\}) = \sum_{\omega_1 \in \Omega_1} \mathbb{P}_1(\{\omega_1\})\mathbb{P}_2(\{\Omega_2\})$$

soit finalement $\displaystyle\sum_{(\omega_1,\omega_2) \in \Omega_1 \times \Omega_2} \mathbb{P}_1(\{\omega_1\})\mathbb{P}_2(\{\omega_2\}) = \mathbb{P}_1(\Omega_1)\mathbb{P}_2(\Omega_2) = 1$.
Il existe donc une unique probabilité $\mathbb{P}$ sur Ω telle que, pour tous les (ω_1, ω_2) de Ω, $\mathbb{P}(\{(\omega_1, \omega_2)\}) = \mathbb{P}_1(\{\omega_1\})\mathbb{P}_2(\{\omega_2\})$. Cette probabilité vérifie la propriété voulue puisqu'en particulier, pour $(A_1, A_2) \in \mathscr{P}(\Omega_1) \times \mathscr{P}(\Omega_2)$, $\mathbb{P}(A_1 \times A_2) = \displaystyle\sum_{(\omega_1,\omega_2) \in A_1 \times A_2} \mathbb{P}(\{(\omega_1, \omega_2)\})$ et donc
$\mathbb{P}(A_1 \times A_2) = \displaystyle\sum_{(\omega_1,\omega_2) \in A_1 \times A_2} \mathbb{P}_1(\{\omega_1\})\mathbb{P}_2(\{\omega_2\}) = \mathbb{P}_1(A_1)\mathbb{P}_2(A_2)$. □

Remarque. $\mathbb{P}$ est appelée ***probabilité produit***.

Généralisation

Plus généralement, une succession de n expériences aléatoires indépendantes modélisées par des espaces probabilisés $(\Omega_i, \mathbb{P}_i)$ peut être modélisée par l'espace $(\Omega, \mathbb{P})$ avec $\Omega = \Omega_1 \times \cdots \times \Omega_n$ et $\mathbb{P}$ l'unique probabilité sur Ω telle que

$$\forall (A_1, \cdots, A_n) \in \mathscr{P}(\Omega_1) \times \cdots \times \mathscr{P}(\Omega_n) \quad \mathbb{P}(A_1 \times \cdots \times A_n) = \mathbb{P}_1(A_1) \cdots \mathbb{P}_n(A_n)$$

En particulier, lorsque l'on répète n fois, de manière indépendante, la même expérience aléatoire modélisée par l'espace $(\Omega_1, \mathbb{P}_1)$ alors la suite de ces n épreuves est modélisée par $(\Omega_1^n, \mathbb{P})$ avec $\forall (A_1, \cdots, A_n) \in (\mathscr{P}(\Omega_1))^n \quad \mathbb{P}(A_1 \times \cdots \times A_n) = \mathbb{P}_1(A_1) \cdots \mathbb{P}_1(A_n)$. On a donc dans ce cas $\forall \omega = (\omega_1, \cdots, \omega_n) \in \Omega$, $\mathbb{P}(\{\omega\}) = \mathbb{P}_1(\{\omega_1\}) \cdots \mathbb{P}_1(\{\omega_n\})$.

2.5 Lois discrètes classiques

Loi Uniforme sur un ensemble fini

La ***loi uniforme*** (équirépartie) sur l'ensemble à n éléments $\{x_1, \cdots, x_n\}$ est telle que tous les éléments sont équiprobables. Elle est donc définie par $\mathbb{P}(\{x_i\}) = 1/n$ pour tout $i \in \{1, \cdots, n\}$ (ce qui, toujours d'après 2.6, caractérise bien $\mathbb{P}$).
Lorsque la variable aléatoire X suit la loi uniforme sur l'ensemble E, on note : $X \sim \mathscr{U}(E)$.
Dire que X suit la loi uniforme revient donc à dire que :

$$\forall x \in X(\Omega) \quad \mathbb{P}([X = x]) = \frac{1}{\mathrm{Card}(X(\Omega))}.$$

Loi de Bernoulli

On réalise une expérience aléatoire qui n'a que deux issues possibles : l'échec (auquel on associe la valeur 0) et le succès (auquel on associe la valeur 1). On note $p \in [0, 1]$ la probabilité de succès de l'expérience. Cette expérience est appelée épreuve de Bernoulli. La variable qui rend compte du résultat de l'expérience est appelée variable aléatoire de Bernoulli de paramètre p. Elle ne prend que deux valeurs et sa loi est donnée par

$$X(\Omega) = \{0, 1\} \quad \text{et} \quad \mathbb{P}([X = 1]) = p \qquad \mathbb{P}([X = 0]) = 1 - p.$$

Lorsque la variable aléatoire X suit la ***loi de Bernoulli de paramètre*** p, on note : $X \sim \mathscr{B}(p)$.

Remarque. Bien des situations se ramènent en fait à une épreuve de Bernoulli : si on s'intéresse à un évènement A particulier, seuls deux résultats sont possibles : A est réalisé (succès) ou A n'est pas réalisé (échec)...

Loi binomiale

On s'intéresse à la loi du nombre de succès lorsque l'on renouvelle n fois ($n \in \mathbb{N}^*$) de manière indépendante une même épreuve de Bernoulli de paramètre p ($p \in [0, 1]$). On note X le nombre de succès obtenus à l'issue des n épreuves. Ce nombre est à valeurs dans $\{0, 1, \cdots, n\}$. Sa loi de probabilité s'appelle la ***loi binomiale de paramètres*** n et p.
Lorsque la variable aléatoire X suit la loi binomiale de paramètres n et p, on note : $X \sim \mathscr{B}(n, p)$.

Proposition 2.25. *Si la variable aléatoire X suit la loi binomiale de paramètres n et p, alors :*

$$X(\Omega) = \{0, 1, \cdots, n\} \quad \text{et} \quad \forall k \in \{0, 1, \cdots, n\} \quad \mathbb{P}([X = k]) = \binom{n}{k} p^k (1 - p)^{n-k}$$

Démonstration : Chaque épreuve comporte deux issues : S (le succès) et E (l'échec). On introduit alors l'univers $\Omega_1 = \{S, E\}$ que l'on munit de la probabilité $\mathbb{P}_1$ caractérisée par $\mathbb{P}_1(\{S\}) = p$ et $\mathbb{P}_1(\{E\}) = 1 - p$. La suite des n épreuves est modélisée par $(\Omega_1^n, \mathbb{P})$ où $\mathbb{P}$ est la probabilité produit.
Il est clair que $X(\Omega) = \{0, 1, \cdots, n\}$. Soit alors $k \in \{0, 1, \cdots, n\}$. Soit $\omega \in \Omega$.

$$\begin{aligned} \omega = (\omega_1, \cdots, \omega_n) \in [X = k] &\iff \mathrm{Card}(\{i \in [\![1, n]\!],\, \omega_i = S\}) = k \\ &\iff \exists A \in \mathscr{C}_k([\![1, n]\!]),\ \{i \in [\![1, n]\!] \quad \omega_i = S\} = A \end{aligned}$$

On en déduit $[X = k] = \displaystyle\bigcup_{A \in \mathscr{C}_k([\![1,n]\!])} \{\omega \in \Omega,\, \{i \in [\![1, n]\!],\, \omega_i = S\} = A\}$ puis, cette réunion étant disjointe, $\mathbb{P}([X = k]) = \displaystyle\sum_{A \in \mathscr{C}_k([\![1,n]\!])} \mathbb{P}(\{\omega \in \Omega,\, \{i \in [\![1, n]\!],\, \omega_i = S\} = A\})$. Or, la probabilité d'obtenir k succès ne dépend pas de la place de ces k-succès dans la suite des n épreuves indépendantes. On a donc, pour tout A de $\mathscr{C}_k([\![1, n]\!])$,

$$\begin{aligned} \mathbb{P}(\{\omega \in \Omega,\, \{i \in [\![1, n]\!],\, \omega_i = S\} = A\}) &= \mathbb{P}(\{\omega \in \Omega,\, \{i \in [\![1, n]\!],\, \omega_i = S\} = [\![1, k]\!]\}) \\ &= \mathbb{P}(\{\omega \in \Omega,\, \omega_1 = S, \cdots, \omega_k = S, \omega_{k+1} = E, \cdots, \omega_n = E\}) \\ &= \mathbb{P}_1(\{S\}) \cdots \mathbb{P}_1(\{S\})\mathbb{P}_1(\{E\}) \cdots \mathbb{P}_1(\{E\}) \\ &= p^k (1 - p)^{n-k} \end{aligned}$$

On a donc finalement : $\mathbb{P}([X = k]) = \displaystyle\sum_{A \in \mathscr{C}_k([\![1,n]\!])} p^k (1 - p)^{n-k} = \binom{n}{k} p^k (1 - p)^{n-k}$. □

Remarque. Dans le même cadre, on aurait aussi pu introduire, pour $i \in [\![1, n]\!]$, les variables aléatoires

$$\begin{array}{cccc} X_i : & \Omega & \longrightarrow & [0, 1] \\ & \omega = (\omega_1, \cdots, \omega_n) & \longmapsto & \begin{cases} 1 \text{ si } \omega_i = S \\ 0 \text{ sinon} \end{cases} \end{array}$$

On constaterait alors que :

- Les variables aléatoires $X_1, \cdots, X_n$ sont indépendantes,
- Les variables aléatoires $X_1, \cdots, X_n$ sont ***identiquement distribuées*** (elles suivent toutes la même loi de Bernoulli de paramètre p,
- $X = X_1 + \cdots + X_n$.

Une variable aléatoire de loi binomiale de paramètres n et p peut donc s'écrire comme somme de n variables aléatoires de Bernoulli de même paramètre p.

Exemple. Un sac contient trois boules numérotées respectivement 0, 1 et 2, indiscernables au toucher. On tire successivement et avec remise deux boules du sac et on note leurs numéros x et y .

À chaque tirage on associe dans le plan, muni d'un repère orthonormal, le point M de coordonnées $(x\ ;\ y)$. On désigne par D le disque de centré en l'origine et de rayon 1,7.

1. Déterminer la loi de probabilité de la variable aléatoire X qui, à chaque tirage associe la somme $x^2 + y^2$.
2. On renouvelle n fois de suite, de façon indépendante, le tirage de deux boules successivement et avec remise. Quelle est la loi de la variable aléatoire Y qui donne le nombre de points appartenant à D.

<u>Solution :</u> On trouve $X(\Omega) = \{0, 1, 2, 4, 5, 8\}$, $\mathbb{P}([X = 0]) = \mathbb{P}([X = 2]) = \mathbb{P}([X = 8]) = \dfrac{1}{9}$ et $\mathbb{P}([X = 1]) = \mathbb{P}([X = 4]) = \mathbb{P}([X = 5]) = \dfrac{2}{9}$. On en déduit $Y \sim \mathscr{B}(n, \frac{4}{9})$.