Chapitre 8

Probabilités sur un univers fini ou dénombrable

Table des matières

Introduction

8.1 Ensembles dénombrables

DÉFINITION 8.1 ★★★ **Ensemble dénombrable**
On dit qu'un ensemble X est :
- *fini* s'il existe un entier $n \in \mathbb{N}$ tel que X est en bijection avec $[\![1, n]\!]$.
- *dénombrable* si X est en bijection avec $\mathbb{N}$.
- *au plus dénombrable* s'il est fini ou dénombrable.

Exemple 8.1
- L'ensemble $\mathbb{N}$ est dénombrable... En effet, l'identité définit une bijection de $\mathbb{N}$ dans lui-même.
- L'ensemble des nombre pairs $\mathscr{P}$ est dénombrable. En effet, si on définit $\varphi : \mathbb{N} \to \mathscr{P}$ par, pour tout $n \in \mathbb{N}$, $\varphi(n) = 2n$ alors φ est bijective.

Remarque 8.1 Un ensemble au plus dénombrable X peut toujours être décrit en extension. En effet, si on note φ la bijection (à valeurs dans $\mathcal{N} = [\![1, n]\!]$ si X est fini ou à valeurs dans $\mathcal{N} = \mathbb{N}$ si X est dénombrable) et qu'on pose $x_k = \varphi(k)$ pour tout $k \in \mathcal{N}$ alors

$$X = \left\{x_k \mid k \in \mathcal{N}\right\}.$$

Réciproquement, si un ensemble X peut être décrit en extension sous la forme précédente, alors il est au plus dénombrable.

PROPOSITION 8.1 ★ Une partie infinie de $\mathbb{N}$ est dénombrable
Soit $X \subset \mathbb{N}$ une partie infinie de $\mathbb{N}$ alors il existe un bijection $\varphi : \mathbb{N} \to X$ strictement croissante.

Démonstration *Comme $X \subset \mathbb{N}$, X admet un plus petit élément qu'on note x_0. On pose alors $\varphi(0) = x_0$. Alors $X \setminus \{x_0\}$ admet aussi un plus petit élément qu'on note x_1 et on pose $\varphi(1) = x_1$. On construit ainsi par récurrence une suite $(x_n) \subset X$ en prenant pour x_{n+1} le plus petit élément de $X \setminus \left\{x_k \mid k \in [\![0, n]\!]\right\}$ et pour tout $n \in \mathbb{N}$, on pose $\varphi(n) = x_n$. L'application φ est bien définie et strictement croissante par construction donc injective.*
Montrons qu'elle est aussi surjective. Soit $x \in X$, alors il existe $n \in \mathbb{N}$ tel que $\varphi(n) > x$. Sinon, φ est une injection de $\mathbb{N}$ dans $[\![0, x]\!]$ ce qui n'est pas possible. Alors $x \notin X \setminus \left\{\varphi(0), \dots, \varphi(n-1)\right\}$ et forcément $x \in \left\{\varphi(0), \dots, \varphi(n-1)\right\}$. Donc il existe $k \in [\![0, n-1]\!]$ tel que $\varphi(k) = x$ et φ est bien surjective.

Exemple 8.2 L'ensemble des nombres premiers est une partie infinie de $\mathbb{N}$ et donc dénombrable.

PROPOSITION 8.2 ★ La réunion disjointe de deux ensembles dénombrables est dénombrable
Soient X, Y deux ensembles disjoints dénombrables. Alors $X \cup Y$ est dénombrable.

Démonstration *Soient $\varphi : X \to \mathbb{N}$ et $\psi : Y \to \mathbb{N}$ deux bijections de X et Y sur $\mathbb{N}$. On pose*

$$\theta : \begin{cases} X \cup Y & \longrightarrow & \mathbb{N} \\ z & \longmapsto & \begin{cases} 2\varphi(k) & \text{si } z \in X \\ 2\psi(z) + 1 & \text{si } z \in Y \end{cases} \end{cases}.$$

Comme X et Y sont disjoints, θ est bien définie.
Si pour $z, z' \in X \cup Y$, $p = \theta(z) = \theta(z')$ alors :
- *si p est pair alors $z, z' \in X$ et l'égalité s'écrit $\varphi(z) = \varphi(z')$. Comme φ est injective, $z = z'$*
- *si p est impair alors $z, z' \in Y$ et l'égalité s'écrit $\psi(z) = \psi(z')$. Comme ψ est injective, $z = z'$*

Alors θ est injective. Montrons qu'elle est aussi surjective. Soit $p \in \mathbb{N}$.
- *si p est pair, alors il existe $k \in \mathbb{N}$ tel que $p = 2k$. Comme $\varphi : X \to \mathbb{N}$ est surjective, il existe $x \in X$ tel que $\varphi(x) = k$. Alors $\theta(x) = p$.*
- *On fait de même si p est impair en considérant $\psi : Y \to \mathbb{N}$.*

Donc θ est surjective. Alors θ réalise bien une bijection de $X \cup Y$ vers $\mathbb{N}$.

Un corollaire immédiat est que :

PROPOSITION 8.3 ★ $\mathbb{Z}$ est dénombrable
$\mathbb{Z}$ est dénombrable.

Démonstration *En effet, $Z = \mathbb{N} \cup \left(-\mathbb{N}^*\right)$ est la réunion de deux ensembles dénombrables.*

PROPOSITION 8.4 ★ $\mathbb{N}^2$ est dénombrable
L'ensemble $\mathbb{N}^2$ est dénombrable.

Démonstration *Considérons*

$$\varphi : \begin{cases} \mathbb{N}^2 & \longrightarrow & \mathbb{N} \\ (u, v) & \longmapsto & 2^u(2v+1) \end{cases}$$

et montrons que φ est bijective.
Soient $(u, v), (u', v') \in \mathbb{N}^2$ tels que $\varphi(u, v) = \varphi(u', v')$ alors

$$2^u(2v+1) = 2^{u'}(2v'+1)$$

et supposons que $u \neq u'$. On peut par exemple considérer que $u' > u$. Alors

$$2^{u'-u}(2v'+1) = 2v+1$$

ce qui est absurde car le nombre de gauche dans cette égalité est pair alors que celui de droite est impair. Donc $u = u'$. Mais alors $2v + 1 = 2v' + 1$ et $v = v'$ et φ est injective.
Soit $n \in \mathbb{N}$ et considérons $A = \left\{k \in \mathbb{N}^* \mid 2^k \text{ divise } n\right\}$. L'ensemble A est non vide car $1 \in A$, il est aussi majoré par n donc il possède un maximum qu'on note u. Alors $n = 2^u m$ avec m impair (sinon u n'est pas le maximum de A) et on a bien $\varphi(u, v) = n$ avec $2v + 1 = m$. Donc φ est surjective.
En conclusion φ est bijective.

COROLLAIRE 8.5 ★ Un produit cartésien d'ensembles dénombrables est dénombrable
Soient X, Y deux ensembles dénombrables. Alors $X \times Y$ est dénombrable.

Démonstration *Il existe une bijection de X dans $\mathbb{N}$ et de Y dans $\mathbb{N}$. En utilisant une bijection de $\mathbb{N}^2$ dans $\mathbb{N}$, on construit alors une bijection de $X \times Y$ dans $\mathbb{N}$.*

COROLLAIRE 8.6 ★ $\mathbb{Q}$ est dénombrable
L'ensemble $\mathbb{Q}$ est dénombrable.

Démonstration *En effet, $\mathbb{Q}$ est en bijection avec $\left\{(p,q) \in \mathbb{Z} \times \mathbb{N}^* \mid p \wedge q = 1\right\}$ qui est une partie infinie de l'ensemble dénombrable $\mathbb{Z} \times \mathbb{N}^*$ et qui est du coup aussi dénombrable.*

THÉORÈME 8.7 ★ $\mathbb{R}$ n'est pas dénombrable
L'ensemble $\mathbb{R}$ n'est pas dénombrable.

Démonstration *Supposons que $\mathbb{R}$ soit dénombrable. Alors il en est de même de $[0,1[$ qui est une partie infinie de $\mathbb{R}$. On peut donc trouver une suite (x_n) qui parcourt tous les éléments de $[0,1[$. Construisons alors un élément α de $[0,1[$ qui n'est pas un point de (x_n) ce qui constituera une contradiction.*
On construit α via son développement décimal. Si le $n$ième terme dans le développement décimal de x_n est :
- *non nul alors on suppose que le $n$ième terme du développement décimal de α est nul ;*
- *nul alors on suppose que le $n$ième terme du développement décimal de α vaut 1.*

Supposons qu'il existe $N \in \mathbb{N}$ tel que $x_N = \alpha$. Alors si le Nième terme du développement décimal de X_N est non nul, celui de α devrait égal à 1, contradiction ! S'il est nul, alors celui de α devrait être égal à 1, nouvelle contradiction. Le nombre α n'est donc pas un élément de (x_n).

PROPOSITION 8.8 ★ $\{0,1\}^{\mathbb{N}}$ n'est pas dénombrable
enombrable
L'ensemble des suites à valeurs dans $\{0,1\}$ n'est pas dénombrable.

Démonstration *Voir l'exercice* **??** *page* **??**.

8.2 Brève introduction aux familles sommables

8.2.1 Cas des familles de nombres positifs

La notion de familles sommables vise à étendre les calculs de somme à des familles avec un nombre infini de termes sans qu'un ordre préalable ne soit fixé sur les éléments de la famille.
Par exemple, la série harmonique alternée $\sum_{n\geqslant 1} x_n$ avec $x_n = \frac{(-1)^n}{n}$ est convergente de somme $-\ln 2$. Mais pour $x \in \mathbb{R}$ quelconque, on montre [1] qu'on peut trouver une bijection $\varphi : \mathbb{N} \to \mathbb{N}$ telle que $\sum_{n\geqslant 0} x_{\varphi(n)}$ converge de somme x. Autrement dit, en réordonnant les termes de la série $\sum_{n\geqslant 1} \frac{(-1)^n}{n}$ on peut la faire converger vers n'importe quel réel x !
On comprend ainsi l'importance de l'ordre des termes dans la série et la nécessité d'une théorie pour manipuler les sommes de familles infinies. L'idée ici n'est pas de rentrer dans les détails de cette théorie mais seulement de se donner quelques outils fondamentaux. Les preuves ne sont pas au programme.
La lettre I désigne dans la suite un ensemble au plus dénombrable.

DÉFINITION 8.2 ★ Famille sommable
Une famille de **réels positifs** $(x_i)_{i\in I}$ est dite *sommable* si l'ensemble

$$\left\{ \sum_{j\in J} x_j : J \subset I, J \text{ fini} \right\}$$

est bornée. La borne supérieure de cette ensemble est alors appelée *somme de la famille* $(x_i)_{i\in I}$. Si la famille n'est pas sommable, on convient que sa somme est $+\infty$. Dans tous les cas, la somme est notée $\sum_{i\in I} x_i$.

1. Voir par exemple le sujet de Centrale Maths 1 PSI 2009 Partie I

PROPOSITION 8.9 ★ **Une sous-famille d'une famille sommable reste sommable**
Soit $(x_i)_{i\in\mathrm{I}}$ une famille sommable de **réels positifs** et soit $\mathrm{J}\subset\mathrm{I}$. Alors $(x_i)_{i\in\mathrm{J}}$ est aussi sommable.

Démonstration *Exercice !*

THÉORÈME 8.10 ★★★ **Théorème de sommation par paquets**
Soit $(x_i)_{i\in\mathrm{I}}$ une famille au plus dénombrable de **réels positifs**. On a équivalence entre :
- La famille (x_i) est sommable ;
- Pour toute partition $(\mathrm{I}_n)_{n\in\mathbb{N}}$ de I (découpage en paquets de I), on a :
  1. pour tout $n\in\mathbb{N}$, la famille $(a_i)_{i\in\mathrm{I}_n}$ est sommable de somme notée σ_n ;
  2. la série $\sum\sigma_n$ est convergente.

On a alors :

$$\sum_{i\in\mathrm{I}} x_i = \sum_{n=0}^{+\infty}\left(\sum_{i\in\mathrm{I}_n} x_i\right)$$

THÉORÈME 8.11 ★★★ **Séries commutativement convergentes**
Soit $(x_i)_{i\in\mathrm{I}}$ une famille au plus dénombrable de **nombres réels positifs** et φ une bijection de J sur I.
La famille $(x_i)_{i\in\mathrm{I}}$ est sommable si et seulement la famille $(x_{\varphi(j)})_{j\in\mathrm{J}}$ est sommable et on a :

$$\sum_{i\in\mathrm{I}} a_i = \sum_{j\in\mathrm{J}} a_{\varphi(j)}$$

Remarque 8.2 En particulier, avec $\mathrm{I}=\mathrm{J}=\mathbb{N}$, le théorème précédent s'écrit : La famille $(x_i)_{i\in\mathbb{N}}$ est sommable si et seulement si pour un certain ordre des termes de la famille, c'est-à-dire pour le choix d'une bijection $\varphi:\mathbb{N}\to\mathbb{N}$, la série $\sum_{n\in\mathbb{N}} x_{\varphi(n)}$ converge.

8.2.2 Cas général

□

DÉFINITION 8.3 ★ **Famille sommable de nombres complexes**
Une famille $(x_i)_{i\in\mathrm{I}}$ au plus dénombrable de nombres complexes est dite sommable si $(|x_i|)_{i\in\mathrm{I}}$ l'est, c'est-à-dire si $(x_i)_{i\in\mathrm{I}}$ est *absolument convergente*.

THÉORÈME 8.12 ★★★ **Un critère simple de sommabilité**
Pour $\mathrm{I}=\mathbb{N}$, la sommabilité d'une famille au plus dénombrable $(x_i)_{i\in\mathrm{I}}$ est équivalente à la convergence absolue de la série $\sum_{i\in\mathrm{I}} x_i$.

Dans le cas d'une série absolument convergente, on peut permuter les termes à volonté :

THÉORÈME 8.13 ★★★ **Série commutativement convergente**
Soit $(x_i)_{i\in\mathbb{N}}$ une suite complexe. On suppose que la série $\sum_{i\in\mathbb{N}} x_i$ est absolument convergente. Alors pour toute bijection $\varphi:\mathbb{N}\to\mathbb{N}$, la série $\sum_{i\in\mathbb{N}} x_{\varphi(i)}$ est absolument convergente et on a :

$$\sum_{i=0}^{+\infty} x_{\varphi(i)} = \sum_{i=0}^{+\infty} x_i .$$

PROPOSITION 8.14 ★★★ Critère de majoration
Soient $(x_i)_{i\in I}$ et $(y_i)_{iu\in I}$ deux familles au plus dénombrables telles que :

(H1) $\forall i \in I, \quad |x_i| \leqslant y_i$;

(H2) $(y_i)_{i\in I}$ est sommable

alors $(x_i)_{i\in I}$ est aussi sommable.

PROPOSITION 8.15 ★★★ Opérations sur les sommes
Soient $(x_i)_{i\in I}$ et $(y_i)_{i\in I}$ deux familles au plus dénombrables sommables à valeurs dans $\mathbb{K}$ alors :

1. Linéarité Pour $\alpha, \beta \in \mathbb{K}$, la famille $(\alpha x_i + \beta y_i)_{i\in I}$ est aussi sommable. De plus :

$$\boxed{\sum_{i\in I}(\alpha x_i + \beta y_i) = \alpha \sum_{i\in I} x_i + \beta \sum_{i\in I} y_i} ;$$

2. Positivité

$$\boxed{\forall i \in I, \quad x_i \geqslant 0 \implies \sum_{i\in I} x_i \geqslant 0} ;$$

3. Croissance

$$\boxed{\forall i \in I, \quad x_i \leqslant y_i \implies \sum_{i\in I} x_i \leqslant \sum_{i\in I} y_i} .$$

8.2.3 Application aux sommes doubles

THÉORÈME 8.16 ★★★ Théorème de Fubini
Soit $(u_{m,n})_{(n,m)\in\mathbb{N}^2}$ une suite « à double entrée » de réels positifs.
La famille $(u_{m,n})_{(m,n)\in\mathbb{N}^2}$ est sommable si et seulement si

1. pour chaque $n \in \mathbb{N}$, la série $\sum_{m\in\mathbb{N}} u_{m,n}$ converge. On note $v_n = \sum_{m=0}^{\infty} u_{m,n}$ sa somme ;
2. la série $\sum_{n\in\mathbb{N}} v_n$ converge.

De plus, dans ce cas, pour chaque $m \in \mathbb{N}$, la série de terme général $u_{m,n}$, $n \in \mathbb{N}$, converge et la série de terme général $w_m = \sum_{n=0}^{+\infty} u_{m,n}$, $m \in \mathbb{N}$, converge et

$$\boxed{\sum_{(m,n)\in\mathbb{N}^2} u_{m,n} = \sum_{m=0}^{+\infty}\left(\sum_{n=0}^{+\infty} u_{m,n}\right) = \sum_{n=0}^{+\infty}\left(\sum_{m=0}^{+\infty} u_{m,n}\right)} .$$

8.2.4 Produit de familles sommables

THÉORÈME 8.17 ★★★ Produit de deux sommes
Soient $(a_i)_{i\in I}$ et $(b_j)_{j\in J}$ deux familles sommables à valeurs dans $\mathbb{K}$. Alors $(a_i b_j)_{(i,j)\in I\times J}$ est sommable et l'on a :

$$\boxed{\sum_{(i,j)\in I\times J} a_i b_j = \sum_{i\in I} a_i \times \sum_{j\in J} b_j}$$

8.3 Espaces probabilisés

8.3.1 Espaces probabilisables

Une expérience dont l'issu dépend du hasard est dite aléatoire. L'ensemble des issus d'une telle expérience est appelé l'univers.
Différents univers finis ont été étudiés en première année, donnons quelques exemples fondamentaux.

1. Les jetés successif d'un dé non pipé. L'univers Ω est formé des 3-listes de $[\![1,6]\!]$ et est donc égale à $\Omega = [\![1,6]\!]^3$. On a alors $\mathrm{Card}(\Omega) = 6^3$.
2. Le tiercé dans l'ordre. On considère par exemple une course de 10 chevaux. On suppose que tous les chevaux ont la même chance d'arriver à une place donnée. On s'intéresse aux trois premiers arrivés. L'univers Ω est formé de tous les tiercés possibles, c'est-à-dire des 3-arrangements de $[\![1,10]\!]$. On a ici $\mathrm{Card}(\Omega) = \mathrm{A}_{10}^3 = 10 \times 9 \times 8$.
3. Le tiercé dans le désordre. Dans la même course de chevaux, si l'ordre d'arrivé n'intervient pas alors Ω est formé des 3-combinaisons de l'ensemble $[\![1,10]\!]$. Alors $\mathrm{Card}(\Omega) = \binom{10}{3} = \frac{\mathrm{A}_{10}^3}{3!}$.

Le tableau suivant est bien utile pour déterminer l'univers de travail dans le cas fini :

Ω	Ordre important	Répétitions possibles	$\mathrm{Card}(\Omega)$	Exemple	Tirage
p-liste	oui	oui	n^p	Jets de p dés distincts	Tirages successifs avec remise
p-arrangement	oui	non	A_n^p	Tiercé dans l'ordre	Tirages successifs sans remise
p-combinaison	non	non	$\binom{n}{p} = \mathrm{C}_n^p$	Tiercé dans le désordre	Tirages simultanés

Remarquons qu'il manque une entrée. On ne considère pas le cas d'une expérience aléatoire comme celle du jet de trois dés non pipés indiscernables. Dans une telle expérience, l'ordre n'a pas d'importance mais les répétitions sont possibles. L'univers Ω est alors celui des *k-combinaisons avec répétitions* d'un ensemble à n éléments et est hors programme. Indiquons seulement que $\mathrm{Card}(\Omega) = \binom{n+k-1}{k}$. Le lecteur désireux d'approfondir le sujet est renvoyé par exemple à `http://fr.wikipedia.org/wiki/Combinaison_avec_répétition`.
Nous allons étudier cette année des univers infini. Nous allons néanmoins nous limiter aux univers dénombrables et à des cas très particuliers d'univers non dénombrables.
Dans un univers infini Ω, il est loisible de considérer une infinité d'événements $(\mathrm{A}_i)_{i \in \mathrm{I}}$ deux à deux disjoints où I est un ensemble quelconque. Si P est une probabilité sur Ω, il faut alors pouvoir écrire

$$\mathrm{P}\left(\bigcup_{i \in \mathrm{I}} \mathrm{A}_i\right) = \sum_{i \in \mathrm{I}} \mathrm{P}(\mathrm{A}_i)$$

et il se pose alors naturellement le problème de la définition de la somme infini.
Dans le cas d'un univers dénombrable, cette somme s'écrit comme une série et elle sera définie si et seulement si cette série est convergente. Dans le cas d'un univers non dénombrable, il faut invoquer des intégrales et on sort immédiatement du cadre de ce cours.
Mais une expérience aléatoire simple comme le jeu de pile ou face infini (où on lance indéfiniment une pièce de monnaie) se déroule dans un univers non dénombrable. En effet, $\Omega = \{0,1\}^{\mathbb{N}}$. On parvient néanmoins à équiper de tels espaces d'une probabilité en se focalisant sur certains événements bien précis, les événements cylindriques, voir plus bas.
On ne pourra alors pas, comme on le faisait en première année, travailler avec l'ensemble de tous les événements possibles $\mathscr{P}(\Omega)$ et il faudra s'intéresser à une sous-famille $\mathscr{T}$ d'événements de l'ensemble. Cette famille doit être stable par les opérations ensemblistes élémentaires. En effet :

1. Si A est un événement, $\overline{\mathrm{A}}$ doit aussi être un événement.
2. Si $(\mathrm{A}_n)_{n \in \mathbb{N}}$ est une famille d'événements alors $\bigcup_{n \in \mathbb{N}} \mathrm{A}_n$ doit aussi être un événement.
3. On doit avoir la même propriété avec l'intersection.
4. On veut que Ω et $\varnothing$ soient des événements.

L'idée du mathématicien Russe Kolmogorov (qui vécut au début du siècle dernier) pour définir une telle famille $\mathscr{T}$ est que cet ensemble d'événements associé à une expérience aléatoire doit posséder une *structure de tribu* :

DÉFINITION 8.4 ★ Tribu, événement, espace probabilisable
Soit Ω un ensemble et soit $\mathscr{T}$ une partie de $\mathscr{P}(\Omega)$. On dit que $\mathscr{T}$ est une *tribu* sur Ω si $\mathscr{T}$ satisfait aux trois axiomes :

1. Si $\mathrm{A} \in \mathscr{T}$, alors son complémentaire $\overline{\mathrm{A}} = \Omega \setminus \mathrm{A}$ est aussi dans $\mathscr{T}$.
2. Si on a une suite dénombrable $(\mathrm{A}_n)_{n \in \mathbb{N}}$ d'éléments de $\mathscr{T}$, alors leur réunion $\bigcup_{n=0}^{+\infty} \mathrm{A}_n$ est aussi dans $\mathscr{T}$.
3. L'ensemble Ω est dans $\mathscr{T}$.

Un élément de $\mathscr{T}$ est appelé un *événement* et le couple $(\Omega, \mathscr{T})$ est appelé un *espace probabilisable*.

Exemple 8.3

- Un premier exemple de tribu élémentaire est évidemment donné par $\mathscr{P}(\Omega)$. Toute tribu sur Ω est contenue dans cette tribu.
- Un autre exemple élémentaire est donné par la tribu $(\varnothing,\Omega)$. Toute tribu sur Ω la contient.
- Si $A\subset\Omega$ alors $\left(\varnothing,A,\overline{A},\Omega\right)$ est une tribu sur Ω. C'est la plus petite contenant A.
- Notons qu'une intersection de tribus sur Ω est encore une tribu. Pour une partie $\mathscr{T}$ de $\mathscr{P}(\Omega)$, on peut alors introduire la plus petite tribu de Ω contenant $\mathscr{T}$. C'est l'intersection de toutes les tribus sur Ω contenant $\mathscr{T}$. On l'appelle *tribu engendrée par* $\mathscr{T}$ et on la note $\sigma(\mathscr{T})$.
- Si B est une partie de Ω et si $\mathscr{T}$ est une tribu sur Ω alors $\mathscr{T}'=\{B\cap T \mid T\in\mathscr{T}\}$ est une tribu sur B. Voir l'exercice **??** page **??** pour une démonstration.

Revenons comme promis à l'exemple consistant à lancer indéfiniment, ici un dé mais on a une construction analogue pour le jeu de Pile ou Face.

PLAN 8.1 : Lancers infinis, épisode 1/3

_infini_de

On s'intéresse donc à l'expérience aléatoire consistant à jeter une infinité de fois un dé à six faces. L'univers Ω *consiste en l'ensemble non dénombrable des suites dont les termes sont pris dans* $[\![1,6]\!]$: $\Omega=[\![1,6]\!]^{\mathbb{N}}$.

L'événement « Obtenir 2*, puis* 3*, puis que* 1 *jusqu'au dixième lancé » peut être modélisé par l'événement suivant* $A=\{(u_n)\in\Omega \mid u_0=2, u_1=3, \forall i\in[\![2,9]\!], u_i=1\}$.

Cet événement peut être décrit comme un produit cartésien. On note :

- $A_0=\{2\}$;
- $A_1=\{3\}$;
- *Pour tout* $i\in[\![2,9]\!]$, $A_i=\{1\}$;
- *Pour tout* $i\geqslant 10$, $A_i=[\![1,6]\!]$;

alors $A=\prod_{k=0}^{+\infty}A_k$. *Une tel événement qui s'écrit comme un produit cartésien dont tous les termes sont égaux à* $[\![1,6]\!]$ *à partir d'un certain rang, est dit* cylindrique. *On peut alors s'intéresser à la plus petite tribu (au sens de l'inclusion) contenant les éléments cylindriques.*

Le langage de la théorie des ensembles permet des calculs systématiques sur les événements. Toutefois, il faut savoir que le langage courant, que nous utilisons dans une première étape pour décrire des événements a sa traduction ensembliste. Voici un dictionnaire :

Langage probabiliste	Notations	Langage des ensembles
Univers	Ω	Ensemble Ω
Tribu, ensemble de tous les événements	$\mathscr{T}$	Ensemble de parties de Ω
Épreuve	$\omega\in\Omega$	Élément de Ω
Événement élémentaire	$\{\omega\}$	Singleton de Ω
Événement	$A\in\mathscr{T}$	Partie de Ω
A implique B	$A\subset B$	B est inclus dans A
A ou B	$A\cup B$	Union de A et B
A et B	$A\cap B$	Intersection de A et B
Événement contraire de A : $\overline{A}$	A^c ou $\Omega\setminus A$	Complémentaire de A dans Ω
A mais pas B	$A\setminus B := A\cap B^c$	Différence symétrique
Événement impossible	$\varnothing$	Partie vide
Événement certain	Ω	Ensemble Ω
Événements incompatibles	$A\cap B=\varnothing$	Parties disjointes

Nous rappelons aussi les règles de calcul suivantes :

PROPOSITION 8.18 ★ **Règles de calcul**

Soient $A,B,C\in\mathscr{P}(\Omega)$. Alors :

1. $A\cap(B\cup C)=(A\cap B)\cup(A\cap C)$,
2. $A\cup(B\cap C)=(A\cup B)\cap(A\cup C)$,
3. $A\cup\overline{A}=\Omega$,
4. $A\cap\overline{A}=\varnothing$,
5. $\overline{\left(\overline{A}\right)}=A$,
6. $\overline{\Omega}=\varnothing$,
7. $\boxed{\overline{A\cup B}=\overline{A}\cap\overline{B}}$,
8. $\boxed{\overline{A\cap B}=\overline{A}\cup\overline{B}}$.

Les deux dernières égalités sont connues sous le nom de *lois de Morgan*.

Tirons maintenant quelques conséquences des axiomes définissant une tribu.

PROPOSITION 8.19 ★ **Une intersection finie ou dénombrable d'événements est un événement**
Soit $\mathscr{T}$ une tribu de parties de l'ensemble Ω. Alors :

1. $\varnothing \in \mathscr{T}$;
2. Une union finie d'éléments de $\mathscr{T}$ est encore dans $\mathscr{T}$.
3. Une intersection finie d'éléments de $\mathscr{T}$ est encore dans $\mathscr{T}$.
4. Considérons une suite dénombrable $(A_n)_{n\in\mathbb{N}}$ d'éléments de $\mathscr{T}$, alors leur intersection $\displaystyle\bigcap_{n=0}^{+\infty} A_n$ est aussi dans $\mathscr{T}$.

Démonstration

— *En appliquant les axiomes 1 et 3, on a le premier résultat.*

— *Pour le second, on considère $A_0,\dots,A_n$ des éléments de $\mathscr{T}$. On pose, pour tout $k > n$, $A_k = \varnothing$. Alors $(A_n)_{n\in\mathbb{N}}$ est une famille dénombrable d'éléments de $\mathscr{T}$ et d'après l'axiome 2, on a $\bigcup_{k=0}^{+\infty} A_k = \bigcup_{k=0}^{n} A_k \in \mathscr{T}$.*

— *Pour le troisième point, si $A_1,\dots,A_n \in \mathscr{T}$ alors $\overline{A_1},\dots,\overline{A_n} \in \mathscr{T}$ et d'après le point précédent, $A = \displaystyle\bigcup_{k=1}^{n} \overline{A_k} \in \mathscr{T}$. Donc $\overline{A} \in \mathscr{T}$, c'est-à-dire, grâce aux lois de Morgan :*

$$\overline{A} = \overline{\bigcup_{k=1}^{n} \overline{A_k}} = \bigcap_{k=1}^{n} \overline{\overline{A_k}} = \bigcap_{k=1}^{n} A_k \in \mathscr{T}.$$

— *On procède de même dans le quatrième point mais il faut ici considérer une famille infinie (A_n) d'éléments de $\mathscr{T}$. Par l'axiome 3, on sait que les éléments de $(\overline{A_n})$ sont aussi dans $\mathscr{T}$ et d'après l'axiome 2, $\displaystyle\bigcup_{n=0}^{+\infty} \overline{A_n} \in \mathscr{T}$. Mais alors, à nouveau grâce à l'axiome 3, par passage au complémentaire et grâce aux lois de Morgan, il vient*

$$\overline{\bigcup_{n=0}^{\infty} \overline{A_n}} = \bigcap_{n=0}^{+\infty} \overline{\overline{A_n}} = \bigcap_{n=0}^{+\infty} A_n \in \mathscr{T}$$

Remarque 8.3 Soient $(\Omega,\mathscr{T})$ un espace probabilisable et $(A_n)_{n\in\mathbb{N}}$ une suite d'événements de $\mathscr{T}$. Dire que :

1. l'événement $\bigcup_{n\in\mathbb{N}} A_n$ est réalisé revient à dire que l'issue $\omega \in \Omega$ de l'expérience aléatoire est élément de $\bigcup_{n\in\mathbb{N}} A_n$, c'est-à-dire $\omega \in \bigcup_{n\in\mathbb{N}} A_n$ ce qui se traduit avec des quantificateurs par :

$$\exists n \in \mathbb{N} : \quad \omega \in A_n.$$

   Cela révient donc à dire que l'un au moins des événements A_n est réalisé :

$$\bigcup_{n\in\mathbb{N}} A_n \text{ réalisé} \iff \text{l'un au moins des événements } A_n \text{ est réalisé .}$$

2. l'événement $\bigcap_{n\in\mathbb{N}} A_n$ est réalisé revient à dire que $\omega \in \bigcap_{n\in\mathbb{N}} A_n$ c'est-à-dire $\forall n \in \mathbb{N}, \quad \omega \in A_n$ ce qui revient à dire que tous les événements A_n sont réalisés :

$$\bigcap_{n\in\mathbb{N}} A_n \text{ réalisé} \iff \text{tous les événements } A_n \text{ sont réalisés .}$$

3. Intéressons nous maintenant à l'événement $\omega \in \bigcup_{n\in\mathbb{N}} \bigcap_{m\geqslant n} A_m$. Il se traduit avec des quantificateurs par

$$\exists n \in \mathbb{N}, \quad \forall m \geqslant n, \quad \omega \in A_m$$

   ce qui revient à dire que il existe un rang à partir duquel tous les A_n sont réalisés. Cet événement est noté $\underline{\lim} A_n$
4. L'événement $\omega \in \bigcap_{n\in\mathbb{N}} \bigcup_{m\geqslant n} A_m$ se traduit avec des quantificateurs par

$$\forall n \in \mathbb{N}, \quad \exists m \geqslant n : \quad \omega \in A_m$$

   ce qui revient à dire l'événement A_n est réalisé une infinité de fois. Cet événement est noté $\overline{\lim} A_n$

8.3.2 Espaces probabilisés

□

DÉFINITION 8.5 ★★★ **Probabilité, espace probabilisé**
Étant donnés un ensemble Ω et une tribu d'événements $\mathcal{T}$ sur Ω, une *probabilité* P est une application de $\mathcal{T}$ dans $[0,1]$, donc une application qui associe à tout évènement $A \in \mathcal{T}$ un nombre $P(A)$ compris entre 0 et 1 appelé *probabilité de* A, et qui satisfait aux axiomes suivants :

me1_proba

me2_proba

1. L'évènement certain est de probabilité 1 : $P(\Omega) = 1$.
2. **Axiome d'additivité dénombrable** : pour toute suite $(A_n)_{n\in\mathbb{N}}$ d'événements de $\mathcal{T}$ qui sont deux à deux incompatibles (c'est-à-dire tels que pour tout $i, j \in \mathbb{N}$, $i \neq j \implies A_i \cap A_j = \varnothing$) alors la série $\sum_{k\geqslant 0} P(A_k)$ converge et a pour somme $P\left(\bigcup_{k=0}^{+\infty} A_k\right)$:
$$\sum_{k=0}^{+\infty} P(A_k) = P\left(\bigcup_{k=0}^{+\infty} A_k\right).$$

Le triplet $(\Omega, \mathcal{T}, P)$ est alors appelé un *espace probabilisé*.

DÉFINITION 8.6 ★★★ **Evénement presque sûr, événement négligeable**
Soit $(\Omega, \mathcal{T}, P)$ un espace probabilisé et soit $A \in \mathcal{T}$. On dit que
— A est un *évément presque sûr* si $P(A) = 1$;
— A est un *évément presque impossible* si $P(A) = 0$.

Exemple 8.4 L'événement sûr Ω est presque sûr. La réciproque est fausse. On verra par exemple que dans le lancer infini d'une pièce de monaie, l'événement « Tirer au moins une fois pile » est un événement presque sûr et pas sûr.
Comme on va le montrer ci-dessous, l'événement impossible $\varnothing$ est presque impossible. La réciproque est à nouveau fausse : un événement presque impossible n'est pas forcément l'événement impossible $\varnothing$.

Remarque 8.4 Si A est un évenement presque sûr alors $\overline{A}$ est presque impossible. De la même façon, si A est presque impossible alors $\overline{A}$ est presque sûr.

Voici quelques conséquences immédiates des axiomes.

THÉORÈME 8.20 ★★★ **Règles de calcul avec une probabilité**

ne_proba_1

Soit $(\Omega, \mathcal{T}, P)$ un espace probabilisé. Alors
1. $P(\varnothing) = 0$.
2. Si $A_1, A_2, \dots, A_n$ dans $\mathcal{T}$ sont deux à deux disjoints, alors
$$\boxed{P(A_1 \cup \dots \cup A_n) = P(A_1) + \dots + P(A_n)};$$
3. En particulier $\boxed{P(\overline{A}) = 1 - P(A)}$.
4. Si A et B sont dans $\mathcal{T}$ alors $\boxed{A \subset B \implies P(A) \leqslant P(B)}$.
5. Si A et B sont dans $\mathcal{T}$ et si B est un événement presque sûr alors $\boxed{P(A\cap B) = P(A)}$.

Démonstration

1. *L'axiome d'additivité dénombrable est applicable à la suite constante définie par $A_n = \varnothing$, qui est effectivement formée d'événements deux à deux disjoints. La série dont le terme général $P(\varnothing)$ est constant ne peut converger que si ce terme général est 0.*
2. *La première partie du second point se démontre en appliquant l'axiome d'additivité dénombrable à $A_1, A_2, \dots, A_n$ continuée par $\varnothing = A_{n+1} = A_{n+2} = \cdots$, et en utilisant le premier point.*
3. *On applique maintenant le point 2 à $n = 2$, $A_1 = A$ et $A_2 = A'$ ce qui fournit $1 = P(\Omega) = P(A) + P(\overline{A})$ d'après le premier axiome d'une probabilité.*
4. *On écrit $B = A \cup (B \setminus A)$ comme réunion de deux ensembles disjoints (notez que $B \setminus A = B \cap \overline{A}$ est bien dans $\mathcal{T}$), et on applique le 2. : $P(B) = P(A) + P(B \setminus A) \geqslant P(A)$.*
5. *On a $P(A) = P\left(A \cap \left(B \cup \overline{B}\right)\right) = P(A\cap B) + P(A\cap \overline{B})$ par 2. De plus, par 4, $P(A\cap\overline{B}) \leqslant P(\overline{B}) = 0$ donc $P(A\cap\overline{B}) = 0$ d'où le résultat.*

THÉORÈME 8.21 ★★★ **Suite des règles de calcul**
Soit $(\Omega,\mathcal{T},P)$ un espace probabilisé. Alors

1. Si A et B sont des événements de $\mathcal{T}$, mais ne sont pas nécessairement disjoints, alors
$$\boxed{P(A\cup B)=P(A)+P(B)-P(A\cap B)}.$$
2. Si les $A_1,A_2,\dots,A_n$ sont des événements de $\mathcal{T}$ pas nécessairement deux à deux disjoints, alors
$$\boxed{P(A_1\cup\dots\cup A_n)\leqslant P(A_1)+\dots+P(A_n)}.$$
3. **Continuités croissante et décroissante** : Soit une suite $(B_n)_{n\in\mathbb{N}}$ d'événements de $\mathcal{T}$ qui est ou bien croissante (c'est-à-dire que pour tout $n\geqslant 0$ on a $B_n\subset B_{n+1}$) ou bien décroissante (c'est-à-dire que pour tout $n\geqslant 0$ on a $B_n\supset B_{n+1}$). Alors, la suite $P(B_n)$ admet une limite donnée, dans le cas croissant, par :
$$\boxed{\lim_{n\to+\infty}P(B_n)=P(\bigcup_{n=0}^{+\infty}B_n)};$$
et, dans le cas décroissant, par :
$$\boxed{\lim_{n\to+\infty}P(B_n)=P(\bigcap_{n=0}^{+\infty}B_n)}.$$
4. **Sous additivité dénombrable** : Soit une suite $(B_n)_{n\in\mathbb{N}}$ d'événements de $\mathcal{T}$. Alors
$$\boxed{P\left(\bigcup_{n=0}^{+\infty}B_n\right)\leqslant\sum_{k=0}^{+\infty}P(B_k)}$$
où $\sum_{k=0}^{+\infty}P(B_k)\in\overline{\mathbb{R}}_+$.

Démonstration

1. *On écrit comme dans la dernière démonstration du théorème précédent 8.20 :*
$$P(B)=P(A\cap B)+P(B\setminus A),P(A)=P(A\cap B)+P(A\setminus B),$$
puis on écrit $A\cup B=(A\cap B)\cup(B\setminus A)\cup(A\setminus B)$ *comme réunion de trois ensembles deux à deux disjoints et on applique le 2. du théorème 8.20 :*
$$\begin{aligned}P(A\cup B)&=P(A\cap B)+P(B\setminus A)+P(A\setminus B)\\&=P(A\cap B)+(P(B)-P(A\cap B))+(P(A)-P(A\cap B))\\&=P(A)+P(B)-P(A\cap B).\end{aligned}$$
2. *On démontre le résultat par récurrence sur n. C'est trivial pour $n=1$. Si c'est démontré pour $n\in\mathbb{N}$, appliquons la règle du 1. à $A=A_1\cup\dots\cup A_n$ et à $B=A_{n+1}$. On obtient, à l'aide de l'hypothèse de récurrence*
$$P(A\cup B)=P(A)+P(B)-P(A\cap B)\leqslant P(A)+P(B)\leqslant(\sum_{k=1}^{n}P(A_k))+P(A_{n+1}).$$
3. *Dans le cas croissant, posons $A_1=B_1$ et, pour $n\geqslant 2$, $A_n=B_n\setminus B_{n-1}$. Les $A_1,A_2,\dots,A_n,\dots$ sont alors deux à deux disjoints et la série $\sum_{k=1}^{\infty}P(A_k)$ est donc convergente de somme $\displaystyle\sum_{k=1}^{\infty}P(A_k)=P\left(\bigcup_{k=0}^{+\infty}A_k\right)$ d'après l'axiome d'additivité dénombrable. D'après la partie 2. du théorème précédent, on a*
$$P(B_n)=P(A_1\cup\dots\cup A_n)=\sum_{k=1}^{n}P(A_k)$$
et donc la nieme somme partielle de $\sum P(A_k)$ est donnée par $P(B_n)$. Comme cette série converge, il en est de même de $(P(B_n)$ par définition de la convergence d'une série numérique et on a
$$\lim_{n\to+\infty}P(B_n)=\sum_{k=1}^{\infty}P(A_k)=P\left(\bigcup_{k=0}^{+\infty}A_k\right).$$

Dans le cas où la famille (C_n) est décroissante, on se ramène au cas précédent par passage au complémentaire. En effet, à l'aide de la loi de Morgan, le complémentaire d'une union est l'intersection des complémentaires. On sait que

$$P(B_n) = 1 - P(\overline{B}_n) = 1 - P(\bigcup_{k=0}^{n} \overline{A}_n)$$

mais $(P(\overline{B}_n))$ *admet, par propriété de limite croissante, une limite donnée par*

$$P\left(\bigcup_{n=0}^{+\infty} \overline{C}_n\right) = P\left(\overline{\bigcap_{n=0}^{+\infty} C_n}\right) = 1 - P\left(\bigcap_{n=0}^{+\infty} C_n\right).$$

Alors $\lim P(B_n) = P\left(\bigcap_{n=0}^{+\infty} C_n\right)$.

4. *Posons pour tout* $n \in \mathbb{N}$, $C_n = B_1 \cup \dots \cup B_n$. *En utilisant le second point, on a pour tout* $n \in \mathbb{N}$

$$P(C_n) = P\left(\bigcup_{k=0}^{n} B_k\right) \leqslant \sum_{k=0}^{n} P(B_k).$$

La famille (C_n) *est croissante et par propriété de limite croissante, la suite* $P(C_n)$ *converge et a pour limite*

$$\lim P(C_n) = P\left(\bigcup_{n=0}^{+\infty} C_n\right) = P\left(\bigcup_{n=0}^{+\infty} B_n\right).$$

Mais la suite $\left(\sum_{k=0}^{n} P(B_k)\right)_{n\in\mathbb{N}}$ *des sommes partielles de la série* $\sum P(B_k)$ *est croissante et donc d'après le théorème de la limite monotone, elle admet une limite dans* $\overline{\mathbb{R}}_+$. *Par passage à la limite dans une inégalité, on obtient donc*

$$P(\bigcup_{n=1}^{+\infty} B_n) = \lim_{n\to+\infty} P(C_n) \leqslant \lim_{n\to+\infty} (P(B_1) + \dots + P(B_n)) = \sum_{k=1}^{\infty} P(B_k).$$

Il n'y a pas de formule simple pour décrire la probabilité d'une réunion. Dans le cas d'une réunion d'ensemble finis, il existe néanmoins une formule pour calculer le cardinal de la réunion.

THÉORÈME 8.22 ★ **Formule du crible, HORS PROGRAMME**
Soient $A_1, A_2, \dots, A_n$, n ensembles finis ayant ou non des intersections. Le cardinal de l'union de ces n ensembles est donné par :

$$\operatorname{Card}\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{k=1}^{n}\left((-1)^{k+1} \sum_{1\leqslant i_1<i_2<\dots<i_k\leqslant n} \operatorname{Card}\left(\bigcap_{j=1}^{k} A_{i_j}\right)\right).$$

Terminons cette section par une condition suffisante d'existence d'une probabilité sur un univers Ω dénombrable.

THÉORÈME 8.23 ★★★ **Condition suffisante d'existence d'une probabilité**
Soit $\Omega = \{x_n \mid n \in \mathbb{N}\}$ un univers dénombrable. Soit $(p_n) \in \mathbb{R}_+^{\mathbb{N}}$ une suite de réels positifs telle que :

H1 la série $\sum p_n$ est convergente de somme 1 :

$$\boxed{\sum_{n=0}^{+\infty} p_n = 1}$$

alors il existe une unique probabilité $P : \mathscr{P}(\Omega) \to [0,1]$ telle que $\boxed{P(\{x_n\}) = p_n}$.

Démonstration

Analyse *Supposons qu'il existe une probabilité* P *sur* $\mathscr{P}(\Omega)$ *telle que pour tout* $x_n \in \Omega$, $P(\{x_n\}) = p_n$. *Alors, comme pour tout* $A \subset \Omega$, *il existe* $I_A \subset \Omega$ *tel que* $A = \{x_i \mid i \in I_A\}$ *et comme les événements* $\{x_i\}$ *pour* $i \in I_A$ *sont deux à deux indépendants, on a :*

$$P(A) = \sum_{i\in I_A} P(\{x_i\}) = \sum_{i\in I_A} p_i.$$

On admet, conformément au programme, que cette somme est bien définie car indépendante de l'ordre d'énumération de ses termes.

Synthèse *Considérons une application* $P : \mathscr{P}(\Omega) \to \mathbb{R}$ *définie de la façon suivante. Pour tout* $A \subset \Omega$, *il existe* $I_A \subset \Omega$ *tel que* $A = \{x_i \mid i \in I_A\}$ *et on pose* $P(A) = \sum_{i \in I_A} P(\{x_i\}) = \sum_{i \in I_A} p_i$. *On pose aussi* $P(\varnothing) = 0$. *Vérifions que* P *est bien une probabilité sur* $(\Omega, \mathscr{P}(\Omega))$.

1. *Comme* $I_A \subset \mathbb{N}$ *et que* $\forall n \in \mathbb{N}$, $p_n \geqslant 0$ *alors* $0 \leqslant P(A) \leqslant \sum_{n=0}^{+\infty} p_n = 1$. *Donc* P *est bien à valeurs dans* $[0,1]$.
2. *On a* $P(\Omega) = \sum_{n=0}^{+\infty} p_n = 1$.
3. *Soit* (A_n) *une famille dénombrable d'événements de* $\mathscr{P}(\Omega)$ *deux à deux incompatibles. Alors*

$$I_{\bigcup_{n=0}^{+\infty} A_n} = \bigcup_{n=0}^{+\infty} I_{A_n}$$

et cette dernière réunion est disjointe donc par définition de P :

$$P\left(\bigcup_{n=0}^{+\infty} A_n\right) = \sum_{n=0}^{+\infty} \sum_{i \in I_{A_n}} p_i = \sum_{n=0}^{+\infty} P(A_n)$$

donc P *vérifie l'axiome d'additivité dénombrable.*

On vérifie ainsi que $(\Omega, \mathscr{P}(\Omega), P)$ *est un espace probabilisé.*

8.4 Indépendance et conditionnement

ionnement

8.4.1 Probabilités conditionnelles

DÉFINITION - PROPOSITION 8.7 **Probabilité conditionnelle**
Soit $(\Omega, \mathscr{T}, P)$ est un espace probabilisé, soit $B \in \mathscr{T}$ un évènement tel que $P(B) > 0$. On définit alors la nouvelle probabilité P_B sur $\mathscr{T}$ par

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

qu'on note aussi $P(A|B)$, et qui se lit *"probabilité de* A *conditionnée par* B", ou *"sachant* B", ou "sachant que B est réalisé".

Démonstration *Le triplet* $(\Omega, \mathscr{T}, P_B)$ *est un authentique espace probabilisé puisque :*

- P_B *est bien à valeurs dans* $[0,1]$.
- $P_B(\Omega) = P(\Omega \cap B)/P(B) = 1$;
- *Si les termes de la suite* $(A_n)_{n \geqslant 0}$ *sont deux à deux incompatibles et dans* $\mathscr{T}$, *on a bien*

$$P_B\left(\bigcup_{n \geqslant 0} A_n\right) = \frac{1}{P(B)} P\left(\left(\bigcup_{n \geqslant 0} A_n\right) \cap B\right) = \frac{1}{P(B)} P\left(\bigcup_{n \geqslant 0} (A_n \cap B)\right)$$

et comme $(A_n \cap B)_{n \in \mathbb{N}}$ *est une suite d'événements deux à deux disjoints de* $\mathscr{T}$ *et que* P *est une probabilité, la série* $\sum P(A_n \cap B)$ *converge et on a bien :*

$$P_B\left(\bigcup_{n \geqslant 0} A_n\right) = \frac{1}{P(B)} \sum_{n=0}^{+\infty} P(A_n \cap B) = \sum_{n=0}^{+\infty} P_B(A_n).$$

Il faut toutefois réaliser que la probabilité P_B est concentrée sur B et ne charge pas $\overline{B}$.

8.4.2 Formule des probabilités composées

composees

La formule suivante apparaît pour la première fois en 1718 dans le fameux « Doctrine of chance » de De Moivre.

THÉORÈME 8.24 ★★ **Formule des probabilités composées**
Soit $(\Omega, \mathscr{T}, P)$ un espace probabilisé fini et soit $(A_i)_{i \in [\![1,n]\!]}$ des événements dont l'intersection est de probabilité non nulle. On a la *formule des probabilités composées* :

$$\boxed{P\left(\bigcap_{k=1}^{n} A_k\right) = P(A_1) P_{A_1}(A_2) P_{A_1 \cap A_2}(A_3) \dots P_{A_1 \cap A_2 \cap \dots \cap A_{n-1}}(A_n)}.$$

Démonstration Par définition des probabilités conditionnelles, on a

$$P(A_1 \cap A_2) = P(A_1)P_{A_1}(A_2).$$

De la même façon :

$$P(A_1 \cap A_2 \cap A_3) = P(A_1 \cap A_2)P_{A_1\cap A_2}(A_2) = P(A_1)P_{A_1}(A_2)P_{A_1\cap A_2}(A_2)$$

et on termine par une récurrence facile.

Exemple 8.5 Un soldat a 1 chance sur 3 de devenir sergent (événement S). Un sergent à 1 chance sur 4 de devenir lieutenant(événement L). Un lieutenant a 1 chance sur 5 de devenir général(événement G). Quelle est la probabilité que le soldat Etienne devienne général ?

$$P(S \cap L \cap G) = P(S)P_S(L)P_{S\cap L}(G) = \frac{1}{3} \times \frac{1}{4} \times \frac{1}{5}.$$

8.4.3 Formules des probabilités totales

Pour énoncer le prochain résultat, il est commode d'introduire les notions suivantes :

DÉFINITION 8.8 ★★★ **Système complet d'événements**
Soit $(\Omega, \mathscr{T})$ un espace probabilisable. Une suite $(B_i)_{i\in I}$ d'événements de $\mathscr{T}$ où I est une partie de $\mathbb{N}$ est un *système complet d'événements* de Ω si les B_i sont deux à deux disjoints et si leur réunion est égale à Ω c'est-à-dire :

$$\begin{cases} \forall i,j \in I, \quad i \neq j \implies B_i \cap B_j = \varnothing \\ \bigcup_{i\in I} B_i = \Omega \end{cases}$$

DÉFINITION 8.9 ★★★ **Système quasi-complet d'événements**
Soit $(\Omega, \mathscr{T})$ un espace probabilisable. Une suite $(B_i)_{i\in I}$ d'événements de $\mathscr{T}$ où I est une partie de $\mathbb{N}$ est un *système complet d'événements* de Ω si les B_i sont deux à deux disjoints et si leur réunion est égale à Ω c'est-à-dire :

$$\begin{cases} \forall i,j \in I, \quad i \neq j \implies B_i \cap B_j = \varnothing \\ \sum_{i\in I} P(B_i) = 1 \end{cases}$$

THÉORÈME 8.25 ★★★ **Formule des probabilités totales**
Soit $(\Omega, \mathscr{T}, P)$ un espace probabilisé, soit $(B_n)_{n\in\mathbb{N}}$ un système complet ou quasi-complet d'événements de Ω. Alors pour tout $A \in \mathscr{T}$, la série $\sum P(A \cap B_n)$ converge et on a la **formule des probabilités totales** :

$$\boxed{P(A) = \sum_{n=0}^{+\infty} P(A \cap B_n) = \sum_{n=0}^{+\infty} P_{B_n}(A)P(B_n)}$$

en adoptant la convention que $P_{B_n}(A)P(B_n) = 0$ dès que $P(B_n) = 0$.

Démonstration Dans le cas où la famille (B_n) est un système complet d'événements, on observe que les événements de la suite $(A \cap B_n)_{n\in\mathbb{N}}$ sont deux à deux incompatibles et leur réunion est A donc, par axiome d'additivité dénombrable, la série de terme général $P(A \cap B_n)$ converge et on a en utilisant que $A = A \cap \Omega = A \cap \cup_{n\in\mathbb{N}} B_n$:

$$P(A) = P(A \cap \Omega) = P\left(A \cap \left(\bigcup_{n=0}^{+\infty} B_n\right)\right) = P\left(\bigcup_{n=0}^{+\infty} (A \cap B_n)\right) = \sum_{n=0}^{+\infty} P(A \cap B_n) = \sum_{n=0}^{+\infty} P_{B_n}(A)P(A_n).$$

Dans le cas où la famille (B_n) est un système quasi-complet d'événements, les événements de la suite $(A \cap B_n)_{n\in\mathbb{N}}$ restent deux à deux incompatibles et comme avant la série de terme général $P(A \cap B_n)$ converge.
Par les régles de calcul 8.20, comme $\cup_{n\in\mathbb{N}} B_n$ est un événement presque sûr, on sait que $P(A) = P\left(A \cup \left(\cup_{n\in\mathbb{N}} B_n\right)\right)$ et on obtient comme avant :

$$P(A) = P\left(A \cap \left(\bigcup_{n=0}^{+\infty} B_n\right)\right) = P\left(\bigcup_{n=0}^{+\infty} (A \cap B_n)\right) = \sum_{n=0}^{+\infty} P(A \cap B_n) = \sum_{n=0}^{+\infty} P_{B_n}(A)P(A_n).$$

Exemple 8.6 On considère trois urnes U_1, U_2 et U_3 chacune contenant 10 boules. La première contient une seule boule blanche, la deuxième en contient 2 et la troisième 3. On tire une boule au hasard. Quelle est la probabilité qu'elle soit blanche ? On note B l'événement la boule est blanche et U_i l'événement « la boule provient de l'urne U_i ». L'univers est ici formé de l'ensemble des trente boules. Les événements U_1, U_2 et U_3 forment un système complet d'événements. On utilise la formule des probabilités totales :

$$P(B) = P(U_1)P_{U_1}(B) + P(U_2)P_{U_2}(B) + P(U_3)P_{U_3}(B) = \frac{1}{3}\times\frac{1}{10} + \frac{1}{3}\times\frac{2}{10} + \frac{1}{3}\times\frac{3}{10} = \frac{1}{5}.$$

Remarque 8.5 Le tirage des urnes dans cet exemple est équiprobable aussi cette expérience aléatoire est équivalente à celle de mettre toutes les boules dans une seule et même urne et à tirer une boule de cette urne.

8.4.4 Formules de Bayes

THÉORÈME 8.26 ★★★ **Formule de Bayes**
Soit $(\Omega, \mathscr{T}, P)$ un espace probabilisé, soit $(B_n)_{n\in\mathbb{N}}$ un système complet d'événements de Ω. Alors pour tout $A \in \mathscr{T}$, la série $\sum P(A\cap B_k)$ converge et on a la **formule de Bayes** :

$$\forall k \in \mathbb{N}, \quad \boxed{P_A(B_k) = \frac{P_{B_k}(A)P(B_k)}{\sum_{n=0}^{+\infty} P_{B_n}(A)P(B_n)}}$$

en adoptant la convention que $P_{B_n}(A)P(B_n) = 0$ dès que $P(B_n) = 0$.

Démonstration *Pour tout $k \in \mathbb{N}$, on a*

$$P_{B_k}(A)P(B_k) = P(A\cap B_k) = P_A(B_k)P(A) = P_A(B_k)\sum_{n=0}^{+\infty} P_{B_n}(A)P(B_n)$$

d'après la formule des probabilités totales pour $P(A)$.

8.4.5 Evénements indépendants

Parfois A et B sont tels que $P_B(A) = P(A)$: savoir que B est réalisé ne modifie pas la probabilité de A. Ainsi dans le schéma succès échec fini avec $N = 2$, Ω a 4 éléments SS, SE, ES, EE de probabilités respectives $p^2, p(1-p), (1-p)p, (1-p)^2$. Si $B = \{SS, SE\}$ est l'évènement : « le premier essai est un succès » et $A = \{SS, ES\}$ est l'évènement : « le second essai est un succès » alors $A\cap B = \{SS\}$, $P(A) = p^2 + (1-p)p = p$, $P(B) = p^2 + p(1-p) = p$, $P(A\cap B) = p^2$ et donc $P_B(A) = P(A)$. C'est le phénomène essentiel pour les probabilités des événements indépendants (qu'il ne faut pas confondre avec les événements incompatibles) et que nous allons définir.

DÉFINITION 8.10 ★★★ **Événements indépendants**
Deux événements A, B d'un espace probabilisé $(\Omega, \mathscr{T}, P)$ sont dits *indépendants* si et seulement si $P(A\cap B) = P(A)P(B)$.

Remarque 8.6 Si $P(B) > 0$ alors l'indépendance de A et B équivaut à $P_B(A) = P(A)$.

⚠ *Attention 8.7* Ne pas confondre indépendance et incompatibilité. Deux événements incompatibles sont en général dépendants. En effet, la réalisation de l'un entraîne l'impossibilité de la réalisation de l'autre.

PROPOSITION 8.27 ★★★ **Indépendance et événements contraires**
Soient A, B deux événements d'un espace probabilisé $(\Omega, \mathscr{T}, P)$. Si A et B sont indépendants alors $\overline{A}$ et B sont aussi indépendants :

$$\boxed{A \text{ et } B \text{ indépendants} \implies \overline{A} \text{ et } B \text{ indépendants}}.$$

Démonstration *On a par indépendance de* A *et* B.

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

de quoi on tire $(1-P(A))P(B) = P(\overline{A}\cap B)$ *c'est-à-dire* $P(\overline{A})P(B) = P(\overline{A})P(B)$ *et* $\overline{A}$ *et* B *sont bien indépendants.*

Remarque 8.7 On a alors A et $\overline{\mathrm{B}}$, ou $\overline{\mathrm{A}}$ et $\overline{\mathrm{B}}$ qui sont aussi indépendants.

DÉFINITION 8.11 ★★★ **Suite d'événements deux à deux indépendants, suite d'événements mutuellement indépendants**
Soit $(\Omega,\mathcal{T},\mathrm{P})$ un espace probabilisé et soit $(\mathrm{A}_i)_{i[\![1,n]\!]}$ une suite finie d'événements de $\mathcal{T}$. Les événements de cette famille sont dits :

— *deux à deux indépendants* si et seulement si A_i et A_j sont indépendants pour tout $i,j\in[\![1,n]\!]$, $i\neq j$.
— *mutuellement indépendants* si et seulement si pour toute partie $\mathrm{I}\subset[\![1,n]\!]$ non vide,

$$\mathrm{P}\left(\bigcap_{i\in\mathrm{I}}\mathrm{A}_i\right)=\prod_{i\in\mathrm{I}}\mathrm{P}(\mathrm{A}_i).$$

Exemple 8.8

— Par exemple si $\mathrm{N}=2$, la famille d'événements $\{\mathrm{A},\mathrm{B}\}$ est indépendante si et seulement si $\mathrm{P}(\mathrm{A}\cap\mathrm{B})=\mathrm{P}(\mathrm{A})\mathrm{P}(\mathrm{B})$; dans le cas où $\mathrm{P}(\mathrm{B})>0$ il serait équivalent de dire $\mathrm{P}_{\mathrm{B}}(\mathrm{A})=\mathrm{P}(\mathrm{A})$. On a coutume de dire par abus de langage que A et B sont indépendants (abus, car l'adjectif qualificatif "indépendant" n'a de sens que s'il s'applique à la paire) ou plus correctement que A est indépendant de B, expression qui ne rend toutefois pas justice à la symétrie de la définition d'indépendance.
— Si $\mathrm{N}=3$ la famille d'événements $\{\mathrm{A},\mathrm{B},\mathrm{C}\}$ est indépendante si et seulement si

$$\mathrm{P}(\mathrm{A}\cap\mathrm{B})=\mathrm{P}(\mathrm{A})\mathrm{P}(\mathrm{B}),\mathrm{P}(\mathrm{B}\cap\mathrm{C})=\mathrm{P}(\mathrm{B})\mathrm{P}(\mathrm{C}),\mathrm{P}(\mathrm{C}\cap\mathrm{A})=\mathrm{P}(\mathrm{C})\mathrm{P}(\mathrm{A}),$$

$$\mathrm{P}(\mathrm{A}\cap\mathrm{B}\cap\mathrm{C})=\mathrm{P}(\mathrm{A})\mathrm{P}(\mathrm{B})\mathrm{P}(\mathrm{C}).$$

Notez que la deuxième ligne n'est pas entraînée par la première. Si Ω a 4 éléments 1,2,3,4 de probabilité 1/4 chacun, les 3 événements $\mathrm{A}=\{1,2\}$, $\mathrm{B}=\{1,3\}$ et $\mathrm{C}=\{1,4\}$ satisfont la première ligne et pas la deuxième : ils sont seulement deux à deux indépendants.
— Si n est quelconque, il n'y a pour montrer l'indépendance que 2^n-1-n égalités à vérifier, puisque l'ensemble vide pour I est exclu et que les n cas où I est un singleton sont triviaux. Notez aussi que l'ensemble vide et l'ensemble Ω sont indépendants de n'importe quoi et qu'une sous famille d'une famille indépendante est encore indépendante.

Exemple 8.9 Comme exemple d'indépendance de n événements, considérons dans le schéma succès échec fini avec n essais un élément particulier $a=(a_1,\dots,a_n)$ de Ω, c'est-à-dire une suite particulière de succès et d'échecs. Notons $k=\mathrm{X}(a)$ le nombre de succès que comprend la suite a. Soit

$$\mathrm{A}_j=\left\{\omega=(\omega_1,\dots,\omega_n)\in\Omega\mid\omega_j=a_j\right\}.$$

Alors $\{\mathrm{A}_1,\dots,\mathrm{A}_n\}$ est une famille indépendante. En effet $\mathrm{P}(\mathrm{A}_j)=p$ si $a_j=\mathrm{S}$ et $1-p$ si $a_j=\mathrm{E}$. De plus, par définition du schéma, $\mathrm{P}(\{a\})=p^k(1-p)^{n-k}$ Comme $\bigcap_{j=1}^n\mathrm{A}_j=\{a\}$ on a bien $\mathrm{P}\left(\bigcap_{j=1}^n\mathrm{A}_j\right)=\prod_{j=1}^n\mathrm{P}(\mathrm{A}_j)$. La démonstration pour n'importe quel sous ensemble I est analogue.

PLAN 8.2 : $\boxed{\text{Lancers infinis, épisode 2/3}}$

_de_proba

Reprenons l'exemple de la tribu permettant de modéliser l'expérience consistant à jeter indéfiniment un dé à six faces et posons nous la question de transformer l'espace probabilisable construit page 7 en un espace probabilisé.

On sait que l'univers modélisant cette expérience aléatoire est $\Omega=\prod_{k=0}^{+\infty}\Omega_n$ *avec* $\Omega_n=[\![1,6]\!]$ *pour tout* $n\in\mathbb{N}$. *On sait aussi que pour tout* $n\in\mathbb{N}$, $(\Omega_n,\mathcal{P}(\Omega_n),\mathrm{P})$ *est un espace probabilisé où* P *est la probabilité uniforme sur* Ω_n.

Le problème est de définir une probabilité P *sur* $(\Omega,\mathcal{T})$ *où* $\mathcal{T}$ *est la tribu engendrée par les événements cylindriques.*

Si un telle probabilité existait, alors comme les tirages sont indépendants, pour un cylindre $\prod_{k=0}^{+\infty}\mathrm{A}_k$ *avec* $\mathrm{A}_k\in\mathcal{P}(\Omega_k)$ *pour tout* $k\in\mathbb{N}$ *et* $\mathrm{A}_k=\Omega_k$ *à partir d'un certain rang* N, *on devrait avoir*

$$\mathrm{P}\left(\prod_{k=0}^{+\infty}\mathrm{A}_k\right)=\prod_{k=0}^{\mathrm{N}}\mathrm{P}(\mathrm{A}_k).$$

Le problème est alors de vérifier que l'application P *s'étend à tous les événements de* $\mathcal{T}$ *et que l'application ainsi étendue fait de* $(\Omega,\mathcal{T},\mathrm{P})$ *un espace probabilisé.*

Le théorème suivant, qui est hors programme, nous dit que c'est le cas.

ylindriques

THÉORÈME 8.28 ★ ***Tribu des parties cylindriques***

On considère une suite dénombrable d'expériences aléatoires mutuellement indépendantes. On note $(\Omega_n, \mathscr{P}(\Omega_n), \mathrm{P}_n)$ les espaces probabilisés associés. On note $\Omega = \prod_{k=0}^{+\infty} \Omega_n$.

Alors il existe une tribu $\mathscr{T}$ sur Ω et une probabilité $\mathrm{P} : \mathscr{T} \to [0,1]$ telles que : les cylindres $\prod_{k=0}^{+\infty} \mathrm{A}_k$ où $\mathrm{A}_k \in \mathscr{P}(\Omega_k)$ pour tout $k \in \mathbb{N}$ et où $\mathrm{A}_k = \Omega_k$ à partir d'un certain rang N (qui dépend de la suite (A_k) considérée) sont éléments de $\mathscr{T}$.

Pour un tel cylindre, on a

$$\boxed{\mathrm{P}\left(\prod_{k=0}^{+\infty} \mathrm{A}_k\right) = \prod_{k=0}^{\mathrm{N}} \mathrm{P}_k(\mathrm{A}_k)}.$$

C'est en ce sens que les expériences aléatoires sont dites mutuellement indépendantes.

Par exemple, si on considère l'événement « on obtient 6 aux N premiers lancers » alors cet événement correspond au cylindre

$$\mathrm{A}_\mathrm{N} = \{6\}^\mathrm{N} \times \left(\prod_{n \geqslant \mathrm{N}} [\![1,6]\!]\right)$$

et $\mathrm{P}(\mathrm{A}_\mathrm{N}) = \prod_{k=0}^{n} \mathrm{P}_k(\{\mathrm{P}\}) = \dfrac{1}{6^\mathrm{N}}$.

On peut maintenant se pencher sur le problème de calculer la probabilité d'événements du type « on obtient pile à chaque lancer » qu'on appelle A et qui s'exprime en fonction des A_k par $\mathrm{A} = \bigcap\limits_{k=0}^{\infty} \mathrm{A}_k$. Remarquons que A est bien un élément de $\mathscr{T}$ car c'est une intersection dénombrable de cylindres de $\mathscr{T}$. Cette intersection est de plus décroissante donc d'après le théorème de continuité décroissante, on a :

$$\mathrm{P}(\mathrm{A}) = \mathrm{P}\left(\bigcap_{k=0}^{\infty} \mathrm{A}_k\right) = \lim_{n \to +\infty} \mathrm{P}(\mathrm{A}_n) = 0.$$

Le résultat obtenu est bien conforme à l'intuition...

On tire directement de la définition la proposition suivante :

PROPOSITION 8.29 ★ **Propriétés des événements mutuellement indépendants**

1. Des événements mutuellement indépendants sont deux à deux indépendants.
2. Si $\mathrm{A}_1, \ldots, \mathrm{A}_n \in \mathscr{T}$ sont mutuellement indépendants alors les événements A_1 et $\bigcap_{k=2}^{n} \mathrm{A}_k$ sont indépendants.
3. Si les événements $\mathrm{A}_1, \ldots, \mathrm{A}_n \in \mathscr{T}$ sont mutuellement indépendants alors pour toute partie $\mathrm{I} \subset [\![1,n]\!]$ telle que $\mathrm{P}\left(\bigcap_{k\in\mathrm{I}} \mathrm{A}_k\right) > 0$, on a :
$$\forall i \in [\![1,n]\!], \quad i \notin \mathrm{I} \implies \mathrm{P}_{\bigcap_{k\in\mathrm{I}} \mathrm{A}_k}(\mathrm{A}_i) = \mathrm{P}(\mathrm{A}_i).$$

Démonstration *Laissée en exercice.*

⚠ *Attention 8.10* L'indépendance deux à deux n'entraîne pas l'indépendance mutuelle.