Classes préparatoires aux grandes écoles

Filière scientifique • Mathématiques

Espaces probabilisés

Mouad El Assadi

Professeur agrégé de mathématiques

Ingénieur d'État en génie civil — École Hassania des Travaux Publics (EHTP)

elassadimouad1@gmail.com

Table des matières

I Espaces probablisés **2**
I.1 Tribus et espaces probabilisables 2
I.2 Mesures de probabilité, espaces probabilisés 4
I.3 Espaces probabilisés discrets 5
I.4 Continuité monotone 6
I.5 Événements négligeables, presque sûrs 7

II Conditionnement et indépendance **8**
II.1 Probabilités conditionnelles 8
II.2 Formules liées aux probabilités conditionnelles 9
II.3 Indépendance 10

III Bilan des méthodes calculatoires en probabilités **12**

I. Espaces probablisés

Le cours de probabilité de première année se construit entièrement dans le cadre d'un univers Ω fini. Dans ce cadre, une expérience d'univers des possibles Ω est un processus aléatoire amenant un résultat (une issue) ω appartenant à Ω. Un événement A est alors un sous-ensemble de Ω. Dire que A lors d'une expérience signifie que l'expérience débouche sur une issue ω appartenant à A. Nous voyons dans ce chapitre comment élargir ce cadre, en acceptant notamment des univers infinis. Ce cadre formel un peu abstrait est indispensable pour pouvoir définir de façon rigoureuse une mesure de probabilité.

I.1 Tribus et espaces probabilisables

Dans cette section, Ω désigne un ensemble quelconque, fini ou infini.
Souvent, la définition de la mesure de probabilité sur tous les sous-ensembles de Ω, n'est pas pertinente, ou peut poser des problèmes techniques. Pour cette raison, il peut être intéressant de pouvoir ne définir la mesure de probabilité que sur une classe particulière de sous-ensembles de Ω. L'objet de ce paragraphe et du suivant est de définir un type satisfaisant de classe de sous-ensembles à laquelle se restreindre. Voici nos exigences pour la construction d'une telle classe :
- L'événement certain Ω et l'événement impossible ont une probabilité, donc appartiennent à la classe ; - Pour tout événement A dont on sait calculer la probabilité, on voudrait pouvoir calculer la probabilité de l'événement contraire $\overline{A}$; - Si on dispose d'une probabilité sur A et B, on voudrait disposer d'une probabilité de $A \cup B$. - Plus généralement, on souhaite pouvoir calculer la probabilité d'une union infinie dénombrable d'événements.

Exemple I.1. *Le dernier point est motivé par l'étude d'un tirage infini à Pile ou Face : quelle est la probabilité qu'un tirage donné (par exemple n'obtenir que des Piles) ait lieu ? Pour ce calcul, on peut encore se dispenser des unions (ou intersections par passage au complémentaire) infinies, en effectuant des majorations, si on définit correctement la probabilité ; mais en revanche, si on se demande maintenant quelle est la probabilité qu'un résultat soit stationnaire, on a besoin de regrouper une infinité de cas possibles (stationnaires à partir d'un certain rang).*

Cela nous amène à la définition suivante :

Définition I.2 (σ-algèbre, ou tribu)**.** *Soit Ω un univers (fini ou non). Une σ-algèbre $\mathcal{T}$ d'événements sur Ω (ou tribu) est un sous-ensemble $\mathcal{T}$ de $\mathcal{P}(\Omega)$ telle que :*

1. $\Omega \in \mathcal{T}$*;*
2. $\forall A \in \mathcal{P}(\Omega), A \in \mathcal{T} \implies \overline{A} \in \mathcal{T}$ *(stabilité par complémentation);*
3. *Pour tout famille dénombrable $(A_i)_{i\in I}$ d'éléments de $\mathcal{T}$, $\bigcup_{i\in I} A_i$ est dans $\mathcal{T}$ (stabilité par union dénombrable).*

Nous complétons les propriétés imposées par la définition par les suivantes :

Proposition I.3 (Propriétés des tribus)**.** *Soit $\mathcal{T}$ une tribu d'événements sur Ω.*

1. $\emptyset \in \mathcal{T}$*;*
2. $\forall (A, B) \in \mathcal{P}(\Omega), (A, B) \in \mathcal{T}^2 \implies A \cup B \in \mathcal{T}$ *(stabilité par union finie);*
3. *Pour tout famille dénombrable $(A_i)_{i\in I}$ d'éléments de $\mathcal{T}$, $\bigcap_{i\in I} A_i \in \mathcal{T}$;*
4. $\forall (A, B) \in \mathcal{P}(\Omega), (A, B) \in \mathcal{T}^2 \implies A \cap B \in \mathcal{T}$ *(stabilité par intersection finie).*

Éléments de preuve. Le point 2 se montre en complétant la famille (A, B) en une famille dénombrable, en ajoutant l'ensemble vide, ce qui ne change pas l'union. Les autres s'obtiennent par passage au complémentaire.
Par une récurrence immédiate, les points 2 et 4 se généralisent facilement aux unions ou intersections d'un nombre fini quelconque d'événements. □

Définition I.4 (Espace probabilisable et événements). • *Un espace probabilisable est un couple $(\Omega, \mathcal{T})$, où Ω est un ensemble quelconque, appelé univers, et $\mathcal{T}$ une σ-algèbre d'événements sur Ω.*

• *Les éléments de $\mathcal{T}$ sont appelés événements (de l'espace probabilisable $(\Omega, \mathcal{T})$).*

Remarque I.5. *1. Une tribu est aussi fréquemment notée $\mathcal{A}$ au lieu de $\mathcal{T}$ (en référence à la terminologie σ-algèbre). C'est la notation utilisée dans le programme officiel.*

2. « La manipulation des tribus n'est pas un objectif du programme », dixit le programme officiel. La notion est surtout donnée pour avoir le bon cadre de travail. Ce cadre sera souvent fourni dans les exercices ou problème, sans que vous ayez à vous demander précisément à quoi il correspond.

3. En particulier, toute la très riche théorie des tribus, indispensable à une théorie un peu plus poussée des probabilités, ne sera pas abordée.

Exemple I.6. *$\mathcal{P}(\Omega)$ est une tribu sur Ω. Dans le cas où Ω est fini ou dénombrable, on choisira en général cette tribu. Elle s'avère en général trop riche lorsque Ω n'est pas dénombrable, et on devra dans ces cas définir des tribus plus petites pour éviter certains écueils techniques.*

On rappelle alors le vocabulaire probabiliste usuel, qui est le même que dans le cas fini :

Définition I.7 (Vocabulaire probabiliste). *Soit $(\Omega, \mathcal{T})$ un espace probabilisé. On définit les terminologies ci-dessous.*

- *Événement : un objet $A \in \mathcal{T}$;*
- *Événement élémentaire, ou épreuve : un événement de la forme $\{\omega\}$ (attention, tout singleton n'est pas nécessairement un événement);*
- *Événement certain : l'événement Ω;*
- *Événement impossible : l'événement $\emptyset$;*
- *Événement contraire de l'événement A : l'événement $\overline{A} = C_\Omega A$;*
- *A entraîne B si $A \subseteq B$;*
- *Événement A et B : $A \cap B$ (réalisation simultanée);*
- *Événement A ou B : $A \cup B$;*
- *A et B sont dits incompatibles si $A \cap B = \emptyset$;*
- *une famille $\{A_i, i \in I\}$ est dite constituée d'événements deux à deux incompatibles si pour tout $(i, j) \in I^2$ tel que $i \neq j$, A_i et A_j sont incompatibles.*

Définition I.8 (Système complet d'événements, SCE). *Un système complet d'événements est une famille $\{A_i, i \in I\}$ formant une partition de Ω. Autrement dit :*

- *Les événements A_i sont non vides;*
- *La famille est constituée d'événements deux à deux incompatibles;*
- *$\bigcup_{i \in I} A_i = \Omega$.*

I.2 Mesures de probabilité, espaces probabilisés

Intuitivement, une probabilité mesure la possibilité qu'un événement donné se produise. Il s'agit donc d'un réel p indiquant la fréquence probable de réalisation de l'événement lorsqu'on répète un grand nombre de fois l'expérience dans des situations similaires, c'est-à-dire le rapport moyen entre le nombre de réalisations et le nombre d'expériences. Ainsi, une probabilité est un réel positif et inférieur à 1. De plus, toutes les expériences amenant une issue de Ω, la probabilité de l'événement Ω doit être 1. Enfin, si A et B sont deux événements ne pouvant pas être réalisés simultanément, le nombre d'expériences amenant A ou B est la somme du nombre d'expériences amenant A et du nombre d'expériences amenant B (puisqu'une expérience ne peut amener à la fois A et B). De l'interprétation ci-dessus, il doit alors venir que la probabilité de l'union disjointe $A \sqcup B$ doit être la somme des probabilités de A et de B. Cela motive la définition suivante :

Définition I.9 (Mesure de probabilité)**.** *Soit $(\Omega, \mathcal{T})$ un espace probabilisable. Une mesure de probabilité (ou simplement une probabilité) sur $(\Omega, \mathcal{T})$ est une application $\mathbb{P} : \mathcal{T} \to \mathbb{R}$ telle que :*

1. $\forall A \in \mathcal{T}, 0 \leq \mathbb{P}(A) \leq 1$*;*
2. $\mathbb{P}(\Omega) = 1$*;*
3. *(propriété de σ-additivité) Pour tout famille dénombrable $(A_i)_{i\in I}$ d'événements deux à deux incompatibles, $(\mathbb{P}(A_i))_{i\in I}$ est sommable, et on a :*

$$\mathbb{P}\left(\bigcup_{i\in I} A_i\right) = \sum_{i\in I} \mathbb{P}(A_i).$$

Remarque I.10. *Très souvent $I = \mathbb{N}$. Dans ce contexte, la famille $(\mathbb{P}(A_n))_{n\in\mathbb{N}}$ étant positive, sa sommabilité équivaut à la convergence de la série $\sum \mathbb{P}(A_n)$.*

Définition I.11 (Espace probabilisé, ou modèle probabiliste de Kolmogorov)**.** *Un espace probabilisé est un triplet $(\Omega, \mathcal{T}, \mathbb{P})$ où $(\Omega, \mathcal{T})$ est un espace probabilisable, et $\mathbb{P}$ une mesure de probabilité sur $(\Omega, \mathcal{T})$.*

Proposition I.12 (Propriétés d'une mesure de probabilité)**.** *Soit $(\Omega, \mathcal{T}, \mathbb{P})$ un espace probabilisé. Alors :*

1. $\mathbb{P}(\emptyset) = 0$*;*
2. *(Additivité) Pour tout couple (A, B) d'événements incompatibles, $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$;*
3. $\forall A \in \mathcal{T}, \mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A)$*;*
4. $\forall (A, B) \in \mathcal{T}, A \subset B \implies \mathbb{P}(B \setminus A) = \mathbb{P}(B) - \mathbb{P}(A)$*;*
5. $\forall (A, B) \in \mathcal{T}, A \subset B \implies \mathbb{P}(A) \leq \mathbb{P}(B)$*;*
6. $\forall (A, B) \in \mathcal{T}, \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.

Éléments de preuve. $\mathbb{P}(\emptyset)$ s'obtient en étudiant la convergence de $\sum \mathbb{P}(\emptyset)$, convergence assurée par la σ-additivité. L'additivité en découle en complétant la famille finie (A, B) en une famille dénombrable par ajout de $\emptyset$. Les points 3,4,5,6 sont évidents, en décomposant le plus grand ensemble en union disjointe d'ensembles plus petits. □

Proposition I.13 (Sous-additivité finie de la mesure de probabilité)**.** *Soit $(A_i)_{i\in I}$ une famille finie d'événements. Alors*

$$\mathbb{P}\left(\bigcup_{i\in I} A_i\right) \leq \sum_{i\in I} \mathbb{P}(A_i).$$

Éléments de preuve. Sans perte de généralité, considérer $I = [1, n]$, et faire une récurrence sur n à partir de l'expression de $\mathbb{P}(A \cup B)$ qui nous donne le cas $n = 2$. □

On verra un peu plus loin que cette propriété se généralise au cas dénombrable.

Proposition I.14 (Probabilités associées à un SCE). *Soit $(A_i)_{i\in I}$ un SCE au plus dénombrable. Alors $\sum_{i\in I} \mathbb{P}(A_i) = 1$.*

I.3 Espaces probabilisés discrets

Il est fréquent que, même si Ω est de grand cardinal (infini non dénombrable), il existe un sous-ensemble A de Ω, au plus dénombrable, tel que presque sûrement, toutes les issues soient dans A (autrement dit, si A est lui-même un élément de la tribu, $\mathbb{P}(A) = 1$).

Définition I.15 (Espace probabilisé discret). *Soit $(\Omega, \mathcal{T}, \mathbb{P})$ un espace probabilisé tel que $\mathcal{T}$ contienne tous les singletons. On dit que $(\Omega, \mathcal{T}, \mathbb{P})$ est discret si la famille $(\mathbb{P}(\{\omega\}))_{\omega\in\Omega}$ est sommable de somme égale à 1.*

Remarque I.16. *Un résultat sur les familles sommables (démontré dans le chapitre sur la dénombrabilité) montre alors que $\{\omega \mid \mathbb{P}(\{\omega\}) \neq 0\}$ est au plus dénombrable, ce qui justifie la terminologie.*

Définition I.17 (Distribution discrète sur Ω). *Soit Ω un ensemble. Une distribution de probabilités sur Ω est une famille $(p_\omega)_{\omega\in\Omega}$ à valeurs dans $\mathbb{R}_+$ telle que*

$$\sum_{\omega\in\Omega} p_\omega = 1.$$

L'ensemble $\{\omega \in \Omega,\ p_\omega \neq 0\}$ est appelé support de la distribution de probabilités $(p_\omega)_{\omega\in\Omega}$, et est au plus dénombrable, en vertu de la remarque précédente.
Ainsi, $(\Omega, \mathcal{T}, \mathbb{P})$ est un espace probabilisé discret si et seulement si $\mathcal{T}$ contient les singletons, et si $(\mathbb{P}(\{\omega\}))_{\omega\in\Omega}$ est une distribution de probabilités discrète sur Ω.

Proposition I.18 (Probabilité définie par une distribution discrète). *Soit Ω un ensemble quelconque, et $(p_\omega)_{\omega\in\Omega}$ une distribution de probabilités discrète. On définit, pour $A \in \mathcal{P}(\Omega)$,*

$$\mathbb{P}(A) = \sum_{\omega\in A} p_\omega,$$

bien défini dans $\mathbb{R}_+$.
L'application $\mathbb{P}$ est à valeurs dans $[0, 1]$, et définit une mesure de probabilités sur l'espace probabilisable $(\Omega, \mathcal{P}(\Omega))$.

Proposition I.19 (Détermination d'une probabilité sur un univers dénombrable). *Soit Ω un ensemble fini ou dénombrable. Une mesure de probabilités sur l'espace probabilisable $(\Omega, \mathcal{P}(\Omega))$ est entièrement déterminée par la distribution de probabilités $(\mathbb{P}(\{\omega\}))_{\omega\in\Omega}$.*

Éléments de preuve. Unicité car les singletons engendrent $\mathcal{P}(\Omega)$. Expliciter p ainsi obtenue en fonction des probabilités des singletons, et vérifier qu'il s'agit bien d'une mesure de probabilité. Ainsi, pour définir une mesure de probabilité sur Ω fini ou dénombrable, il suffit de se donner une distribution de probabilités $(p_\omega)_{\omega\in\Omega}$. □

Exemple I.20 (Probabilité uniforme sur un univers fini). *Si Ω est fini, on peut définir la mesure de probabilité uniforme en posant $\mathbb{P}(\{\omega\}) = \frac{1}{|\Omega|}$. On obtient alors, pour tout événement $A \in \mathcal{P}(\Omega)$,*

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|} \quad \textit{(formule de Laplace)}.$$

I.4 Continuité monotone

Le théorème suivant découle directement de la σ-additivité.

Théorème I.21 (Propriété de limite monotone, ou continuité monotone de $\mathbb{P}$)**.** *Soit $(\Omega, \mathcal{T}, \mathbb{P})$ un espace probabilisé.*

1. *Soit $(A_n)_{n\in\mathbb{N}}$ une suite croissante d'événements (pour l'inclusion). Alors :*
$$\mathbb{P}\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \lim_{n\to+\infty} \mathbb{P}(A_n).$$
2. *Soit $(A_n)_{n\in\mathbb{N}}$ une suite décroissante d'événements. Alors :*
$$\mathbb{P}\left(\bigcap_{n\in\mathbb{N}} A_n\right) = \lim_{n\to+\infty} \mathbb{P}(A_n).$$

Éléments de preuve. Écrire $\bigcup_{n\in\mathbb{N}} A_n = \bigcup_{n\in\mathbb{N}}(A_n \setminus A_{n-1})$, avec $A_{-1} = \emptyset$, puis σ-additivité. Passer au complémentaire pour l'intersection. □

Exemple I.22. *Lors d'une succession infinie de tirages à Pile ou Face, calculer la probabilité de n'obtenir que des Piles.*

Corollaire I.23. *Soit $(A_n)_{n\in\mathbb{N}}$ une famille quelconque d'événements. Alors*
$$\mathbb{P}\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \lim_{N\to+\infty} \mathbb{P}\left(\bigcup_{n=0}^{N} A_n\right) \quad \textit{et} \quad \mathbb{P}\left(\bigcap_{n\in\mathbb{N}} A_n\right) = \lim_{N\to+\infty} \mathbb{P}\left(\bigcap_{n=0}^{N} A_n\right).$$

Corollaire I.24 (Sous-additivité dénombrable de $\mathbb{P}$, ou inégalité de Boole)**.** *Soit $(A_i)_{i\in I}$ une famille au plus dénombrable d'événements. Alors, dans $\overline{\mathbb{R}}_+$,*
$$\mathbb{P}\left(\bigcup_{i\in I} A_i\right) \leq \sum_{i\in I} \mathbb{P}(A_i).$$

En définissant pour une suite croissante (A_n), $\bigcup_{n\in\mathbb{N}} A_n = \lim \uparrow A_n$ et de même pour une suite décroissante, $\bigcap_{n\in\mathbb{N}} A_n = \lim \downarrow A_n$, la propriété ci-dessus se réécrit :
$$\mathbb{P}(\lim \uparrow A_n) = \lim \mathbb{P}(A_n) \quad \text{et} \quad \mathbb{P}(\lim \downarrow A_n) = \lim \mathbb{P}(A_n).$$

Il s'agit donc d'une propriété de continuité séquentielle, d'où le nom donné au théorème.

Exemple I.25 (Limite supérieure et limite inférieure, HP)**.** *Soit (A_n) une suite d'événements. On définit*

- *la limite supérieure $\limsup A_n = \bigcap_{n\in\mathbb{N}} \bigcup_{k=n}^{\infty} A_k$: réalisation d'une infinité de A_k;*
- *la limite inférieure $\liminf A_n = \bigcup_{n\in\mathbb{N}} \bigcap_{k=n}^{\infty} A_k$: réalisation de presque tous les A_k (tous sauf un nombre fini, c'est-à-dire tous à partir d'un certain rang).*

Ainsi,

- *La probabilité qu'un nombre infini de A_k se réalisent est*
$$\mathbb{P}(\limsup A_n) = \lim \mathbb{P}\left(\bigcup_{k=n}^{\infty} A_k\right).$$

- *La probabilité que presque tous les A_k se réalisent est*

$$\mathbb{P}(\liminf A_n) = \lim \mathbb{P}\left(\bigcap_{k=n}^{\infty} A_k\right).$$

Par exemple, calculer la probabilité que lors d'une suite infinie de tirages à P/F, la suite obtenue soit stationnaire.

Corollaire I.26. *Soit (A_n) une suite d'événements.*

- *La probabilité qu'un nombre infini de A_k se réalisent est*

$$\mathbb{P}(\limsup A_n) = \lim \mathbb{P}\left(\bigcup_{k=n}^{\infty} A_k\right).$$

- *La probabilité que presque tous les A_k se réalisent est*

$$\mathbb{P}(\liminf A_n) = \lim \mathbb{P}\left(\bigcap_{k=n}^{\infty} A_k\right).$$

Exemple I.27. *On tire une infinité de fois à P ou F. Probabilité que la suite de P/F soit stationnaire.*

Remarque I.28. *Si (A_n) est décroissante, $\bigcup_{k=n}^{+\infty} A_k = A_n$ et $\limsup A_n = \lim \downarrow A_n$; de même si (A_n) est croissante, $\liminf A_n = \lim \uparrow A_n$. Le corollaire précédent englobe donc le théorème de continuité monotone.*

I.5 Événements négligeables, presque sûrs

Définition I.29 (Ensembles négligeables)**.** *Soit $(\Omega, \mathcal{T}, \mathbb{P})$ un espace probabilisé, et soit $A \in \mathcal{T}$.*

1. *On dit que A est* négligeable*, ou* presque-impossible*, ou* quasi-impossible *si $\mathbb{P}(A) = 0$.*
2. *On dit que A est* presque sûr*, ou* presque-certain *ou* quasi-certain *si $\mathbb{P}(A) = 1$.*
3. *Une propriété est vraie* presque sûrement *si l'événement « la propriété est satisfaite » est presque-certain.*

Exemple I.30. *Les événements « n'obtenir que des P » lors d'une succession infinie de tirages P/F, ou « obtenir une suite stationnaire » sont quasi-impossibles, sans être impossibles. Leur complémentaire est donc presque sûr, sans être certain.*

Remarque I.31. *La terminologie recommandée par le programme officiel est « négligeable » et « presque sûr », mais les autres terminologies sont aussi très répandues dans les ouvrages.*

Proposition I.32 (Intersection et union d'événements presque impossibles/sûrs)**.**
1. *Toute union d'un nombre au plus dénombrable d'événements négligeables est négligeable.*
2. *Toute intersection d'un nombre au plus dénombrable d'événements presque sûrs est presque sûre.*

Éléments de preuve. Par sous-additivité dénombrable, puis complémentation. □

Remarque I.33.
1. *Dans $(\Omega, \mathcal{T}, \mathbb{P})$ par convention, si I est vide,*

$$\bigcup_{i\in I} A_i = \emptyset \quad \text{et} \quad \bigcap_{i\in I} A_i = \Omega.$$

Ainsi, les propriétés sont aussi vraies pour des familles vides.

2. *À condition d'enlever le cas des familles vides, la propriété est aussi trivialement vraie pour les intersections quelconques d'ensembles négligeables, et les unions quelconques d'ensembles presque sûrs.*

On définit maintenant une notion voisine des SCE, mais qui permet plus de souplesse dans les hypothèses de certains théorèmes.

Définition I.34 (Système quasi-complet d'événements, SQCE)**.** *Soit $(\Omega, \mathcal{T}, \mathbb{P})$ un espace probabilisable, et $(A_i)_{i\in I}$ une famille au plus dénombrable d'éléments de $\mathcal{T}$. On dit que $(A_i)_{i\in I}$ est un système quasi-complet d'événements si :*

(i) les A_i sont 2 à 2 incompatibles;

(ii) $\mathbb{P}(\bigcup_{i\in I} A_i) = 1$.

Ainsi, la différence avec un système complet réside dans le fait que certaines parts peuvent être vides, et que l'union $\bigcup_{i\in I} A_i$ n'est égale à Ω qu'à un ensemble négligeable près.
Un système complet est bien entendu aussi quasi-complet, la réciproque étant fausse.

Proposition I.35 (Système quasi-complet à probabilités non nulles)**.** *Soit $(A_i)_{i\in I}$ un SQCE, et $J = \{i \in I, \mathbb{P}(A_i) > 0\}$. Alors $(A_j)_{j\in J}$ est encore un SQCE.*

II. Conditionnement et indépendance

II.1 Probabilités conditionnelles

Les définitions relatives aux probabilités conditionnelles s'étendent bien au contexte étendu qu'on vient de définir.

Définition II.1 (Probabilités conditionnelles)**.** *Soit $(\Omega, \mathcal{T}, \mathbb{P})$ un espace probabilisé, A et B deux événements de $\mathcal{T}$ tels que B ne soit pas presque-impossible. Alors, la probabilité conditionnelle de A en B (ou la probabilité de A sachant B) est :*

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

On trouve également souvent la notation $\mathbb{P}(A|B)$.

Cette définition se comprend bien pour la mesure uniforme : il s'agit dans ce cas de la proportion des éléments de B qui satisfont A (c'est-à-dire la proportion des éléments de B qui sont dans $A \cap B$).

Proposition II.2. *Soit B un événement non presque-impossible. Alors $\mathbb{P}_B$ définit une mesure de probabilités sur l'espace probabilisable $(\Omega, \mathcal{T})$.*

Éléments de preuve. Vérifications sans difficulté. □

Remarque II.3. *La définition des probabilités conditionnelles est souvent utilisée sous la forme suivante : $\mathbb{P}(A \cap B) = \mathbb{P}(A|B)\mathbb{P}(B)$. En effet, expérimentalement, on a plus facilement accès aux probabilités conditionnelles qu'aux probabilités d'intersections, et on s'en sert pour le calcul de probabilités d'intersections. C'est d'ailleurs le cas initial de la formule plus générale des probabilités composées.*

II.2 Formules liées aux probabilités conditionnelles

Théorème II.4 (Formule des probabilités totales)**.** *Soit $(A_i)_{i\in I}$ un système quasi-complet fini ou dénombrable, tels que les événements A_i ne sont pas presque-impossibles, et soit $B \in \mathcal{T}$. Alors la famille $(\mathbb{P}(B|A_i)\mathbb{P}(A_i))_{i\in I}$ est sommable, et :*

$$\mathbb{P}(B) = \sum_{i\in I} \mathbb{P}(B|A_i)\mathbb{P}(A_i).$$

Éléments de preuve. Écrire $B = \left(\bigcup_{i\in I} B\cap A_i\right) \cup (B\cap C)$, où $\mathbb{P}(C) = 0$. □

Remarque II.5.
1. *La formule des probabilités totales permet de retrouver une probabilité $\mathbb{P}(B)$ connaissant les probabilités conditionnelles $\mathbb{P}(B|A_i)$. Ainsi, on peut calculer $\mathbb{P}(B)$ en « distinguant suivant un certain nombre de cas » (suivant que A_i est réalisé), si ces cas partagent l'ensemble (ou presque) des cas possibles. Ainsi, cette formule est adaptée au cas où le résultat d'une expérience dépend de résultats antérieurs, donc si on est amené à faire une discussion pour décrire le résultat de l'expérience.*
2. *Si un SCE ou un SQCE contient des événements quasi-impossibles, on a vu ci-dessus qu'en enlevant ces événements, on obtient encore un SQCE avec lequel la formule des probabilités totales est valide. Ainsi, plus généralement, si $(A_i)_{i\in I}$ est un SQCE quelconque,*

$$\mathbb{P}(B) = \sum_{\substack{i\in I\\ \mathbb{P}(A_i)\neq 0}} \mathbb{P}(B|A_i)\mathbb{P}(A_i).$$

Théorème II.6 (Formule des probabilités composées)**.** *Soit $n \geq 2$, et $(A_1,\dots,A_n) \in \mathcal{T}^n$ tel que $\mathbb{P}(A_1\cap\cdots\cap A_{n-1}) \neq 0$ (i.e. $A_1\cap\cdots\cap A_{n-1}$ n'est pas quasi-impossible). Alors :*

$$\mathbb{P}\left(\bigcap_{i=1}^{n} A_i\right) = \mathbb{P}(A_1)\mathbb{P}(A_2|A_1)\mathbb{P}(A_3|A_1\cap A_2)\cdots\mathbb{P}(A_n|A_1\cap\cdots\cap A_{n-1})$$

$$= \mathbb{P}(A_1)\prod_{i=2}^{n}\mathbb{P}\left(A_i\Big|\bigcap_{j=1}^{i-1} A_j\right).$$

Éléments de preuve. Récurrence sur n. □

Remarque II.7. *L'intérêt de cette formule est de donner une formule pour le calcul des probabilités d'intersections d'événements, notamment dans le cas d'une succession d'expériences. Cette formule dévoile toute son utilité lorsque les événements considérés ne sont pas mutuellement indépendants. En cas d'indépendance mutuelle, elle ne dit rien de plus que la proposition ci-dessous.*

Théorème II.8 (Formule de Bayes simple)**.** *Soit A et B deux événements tels que $\mathbb{P}(A) > 0$ et $\mathbb{P}(B) > 0$. Alors*

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

Éléments de preuve. Calculer $\mathbb{P}(A\cap B)$ en conditionnant de deux manières. □

De façon plus générale, la connaissance des probabilités d'un système complet, et des probabilités d'un événement B conditionnées au système complet permet de retourner une à une les probabilités conditionnelles :

Théorème II.9 (Formule de Bayes sur un système complet)**.** *Soit $(A_i)_{i\in I}$ un système quasi-complet fini ou dénombrable, et B un événement non quasi-impossible. On suppose que pour tout $i \in I$, $\mathbb{P}(A_i) \neq 0$. Soit $j \in I$. Alors :*

$$\mathbb{P}(A_j|B) = \frac{\mathbb{P}(B|A_j)\mathbb{P}(A_j)}{\sum_{i\in I}\mathbb{P}(B|A_i)\mathbb{P}(A_i)}.$$

Éléments de preuve. Exprimer $\mathbb{P}(B)$ avec la formule des probabilités totales dans la formule précédente. □

Remarque II.10 (Intérêt de la formule de Bayes)**.** *Le système (A_i) représente souvent une liste de causes pouvant amener l'événement B lors d'une étape suivante de l'expérience. Il est alors généralement facile de déterminer la probabilité qu'une certaine conséquence B ait lieu, sachant que la cause A_i a eu lieu, c'est-à-dire la probabilité conditionnelle $\mathbb{P}(B|A_i)$ (on respecte dans ce cas l'ordre temporel). Ces données permettent, grâce à la formule de Bayes, de remonter le temps, en déterminant la probabilité qu'une certaine cause A_i ait eu lieu sachant la conséquence B. Pour cette raison, cette formule est aussi souvent appelée formule de probabilité des causes.*

Exemple II.11. *Dans une population composée d'autant d'hommes que de femmes, 5 % des hommes et 0,25 % des femmes sont daltoniens. Quelle est la probabilité pour qu'une personne daltonienne choisie au hasard soit une femme ?*

II.3 Indépendance

Les probabilités conditionnelles mesurent l'impact de la réalisation d'un événement sur un autre. Si cet impact est nul (c'est-à-dire si la connaissance de la réalisation d'un événement n'influe pas sur la probabilité de l'autre), on dira que ces deux événements sont indépendants.
Ainsi, intuitivement, deux événements A et B sont indépendants si la connaissance de l'un ne modifie pas la probabilité de l'autre, donc si $\mathbb{P}(A|B) = \mathbb{P}(A)$ et $\mathbb{P}(B|A) = \mathbb{P}(B)$. L'inconvénient de cette définition réside dans le fait qu'elle n'est valable que sous une hypothèse de non quasi-impossibilité, afin de pouvoir considérer les probabilités conditionnelles. On constate cependant que dans le cas de non quasi-impossibilité, les deux égalités sont équivalentes à l'égalité : $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$. Cette égalité pouvant être considérée sans hypothèse de quasi-impossibilité, c'est elle que nous prenons comme définition de l'indépendance. Elle a en outre l'avantage d'être symétrique.

Définition II.12 (Événements indépendants)**.** *Soit $(\Omega, \mathcal{T}, \mathbb{P})$ un espace probabilisé, et $(A, B) \in \mathcal{T}^2$. Les événements A et B sont indépendants si $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.*

En faisant le cheminement inverse de celui précédant la définition, il vient :

Proposition II.13 (Probabilités conditionnelles et indépendance)**.** *Soit A et B deux événements indépendants tel que B ne soit pas quasi-impossible. Alors*

$$\mathbb{P}(A) = \mathbb{P}(A|B).$$

On a évidemment aussi $\mathbb{P}(B) = \mathbb{P}(B|A)$ si A n'est pas quasi-impossible.

On note dans ce qui suit $\mathcal{P}_f(I)$ l'ensemble des parties finies de I.

Définition II.14 (Indépendance d'une famille)**.** *Soit $(\Omega, \mathcal{T}, \mathbb{P})$ un espace probabilisé, et $(A_i)_{i\in I}$ une famille d'événements. On dit que les $A_i, i \in I$, sont indépendants (ou mutuellement indépendants), si et seulement si pour tout $J \in \mathcal{P}_f(I)$,*

$$\mathbb{P}\left(\bigcap_{j\in J} A_j\right) = \prod_{j\in J}\mathbb{P}(A_j).$$

En particulier si I lui-même est fini en prenant $J = I$:

Proposition II.15 (Probabilité d'une intersection d'événements indépendants)**.** *Soit $(A_i)_{i\in I}$ une famille finie d'événements indépendants. Alors :*

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

Remarque II.16. *L'égalité précédente ne suffit pas à avoir l'indépendance. Il faut bien contrôler TOUS les produits des sous-familles (finies).*
On verra dans le paragraphe suivant comment calculer la probabilité d'une intersection lorsque les événements ne sont pas indépendants (formule des probabilités composées).

Exemple II.17. *On fait des tirages à pile ou face*

- *A est réalisé si et seulement si le premier tirage est Pile.*
- *B est réalisé si et seulement si lors des 3 premiers tirages, il y a au plus un Pile.*
- *$C = B$.*

On a $\mathbb{P}(A\cap B\cap C) = \frac{1}{8} = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C)$, mais les événements ne sont clairement pas indépendants !

De la définition même découle :

Proposition II.18 (Stabilité de l'indépendance par restriction)**.** *Soit $(A_i)_{i\in I}$ une famille d'événements indépendants et $J \subset I$. Alors $(A_j)_{j\in J}$ est encore composée d'événements indépendants.*

En particulier, en prenant toutes les sous-familles de 2 événements, l'indépendance implique l'indépendance 2 à 2 des événements.

Remarque II.19. *La réciproque est fausse : l'indépendance 2 à 2 d'une famille n'implique pas l'indépendance (mutuelle).*

Exemple II.20. *On procède à trois tirages à Pile ou Face. Soit A l'événement consistant à obtenir exactement 1 Pile lors des tirages 2 et 3, B de même lors des tirages 1 et 3, C de même lors des tirages 1 et 2. Les événements A, B et C sont 2 à 2 indépendants, mais pas mutuellement.*

Proposition II.21 (Indépendance des intersections de sous-familles disjointes, HP)**.** *Soit $(A_i)_{i\in I}$ une famille d'événements indépendants, et $(I_k)_{k\in K}$ une famille de sous-ensembles finis de I, 2 à 2 disjoints. Alors la famille $\bigcap_{i\in I_k} A_i$ est constituée d'événements indépendants.*

Éléments de preuve. Prendre une intersection finie de tels ensembles, c'est aussi une intersection finie de A_i. Écrire par indépendance la probabilité comme produit de $\mathbb{P}(A_i)$, puis regrouper par paquets de I_k en réutilisant l'indépendance sur ces paquets. □

Proposition II.22 (Indépendance et complémentation)**.** *Si A et B sont indépendants, alors A et $\overline{B}$ sont indépendants.*

Éléments de preuve. Calculer $\mathbb{P}(A\cap\overline{B}) = \mathbb{P}(A\setminus(A\cap B))$. □

Proposition II.23 (Indépendance et complémentation pour une famille finie)**.** *Si $A_1, \dots, A_n$ sont indépendants, il en est de même de $\overline{A_1}, A_2, \dots, A_n$.*

Éléments de preuve. Considérer l'intersection d'une sous-famille. Seul cas à voir : $\overline{A_1}$ est dans la sous-famille. Se ramener au cas précédent en justifiant que A_1 est indépendant de l'intersection des autres termes de la sous-famille. □

Corollaire II.24 (Indépendance et complémentation pour une famille quelconque)**.** *Soit $(A_i)_{i\in I}$ une famille d'événements indépendants et $(B_i)_{i\in I}$ une famille telle que pour tout $i \in I$, $B_i = A_i$ ou $\overline{A_i}$. Alors $(B_i)_{i\in I}$ est une famille d'événements indépendants.*

III. Bilan des méthodes calculatoires en probabilités

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

Méthode III.1 (Comment aborder un calcul de probabilité)**.**

1. *Si on est dans une situation d'équiprobabilité, on peut se diriger vers un argument combinatoire (et utiliser la formule de Laplace). Dans les autres cas, la démarche générale est indiquée ci-dessous.*
2. *Définir des événements simples issus de l'expérience décrite, en n'oubliant pas la numérotation de ces événements en cas d'expérience répétée. Ces événements simples auront pour vocation de décrire ceux qui nous intéressent. Il ne faut pas hésiter à introduire des événements par vous-même, à donner des notations. Ce travail n'est pas nécessairement fait pour vous dans le sujet.*
3. *Décrire à l'aide d'une phrase l'événement dont on recherche la probabilité, sous forme d'une condition nécessaire et suffisante de réalisation :*

   « L'événement A est réalisé si et seulement si ... »

   Cette condition nécessaire et suffisante de réalisation de l'événement A doit s'exprimer en fonction de la réalisation ou non des événements simples définis dans la première étape. Il s'agit donc de « décomposer » l'événement A en événements simples.
4. *Traduire cette description dans le langage ensembliste (c'est-à-dire sous forme d'union, intersection ou complémentaires d'événements élémentaires). Les deux étapes sont nécessaires : la première (description verbale) fournit une explication facilement lisible, et éclaire la formule, la seconde (description ensembliste) donne de la rigueur à l'argument, et permet ensuite de s'orienter vers la bonne méthode de calcul.*
5. *Calcul de la probabilité $\mathbb{P}(A)$, en se servant de la décomposition ensembliste précédente, et des règles de calcul des probabilités d'unions, d'intersections ou de complémentaires. Nous rappelons l'essentiel de ces règles ci-dessous.*

Ainsi, il est important de savoir s'orienter rapidement vers la bonne technique de calcul (probabilité d'une union, d'une intersection), suivant la situation rencontrée. Voici les différents cas rencontrés les plus fréquemment :

Méthode III.2 (Calcul de la probabilité d'un complémentaire)**.** *Une seule formule à connaître, issue de la définition d'une mesure de probabilité :*

$$\forall A \in \mathcal{T},\ \mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A).$$

Méthode III.3 (Calcul de la probabilité d'une intersection)**.** *Suivant les cas :*

- *Intersection dénombrable d'événements décroissants pour l'inclusion : On utilise le théorème de la limite monotone.*
- *Intersection d'un nombre fini d'événements mutuellement indépendants : On utilise la relation issue de la définition de l'indépendance mutuelle :*

$$\mathbb{P}(A_1 \cap A_2 \cap \cdots \cap A_n) = \mathbb{P}(A_1)\mathbb{P}(A_2)\ldots\mathbb{P}(A_n).$$

- *Intersection d'un nombre fini d'événements non nécessairement indépendants : On utilise la formule des probabilités composées. Cette formule est surtout utile lorsqu'il y a un enchaînement*

temporel des événements (A_1 étant le premier à advenir, A_n le dernier), chaque résultat influant sur les suivants. C'est le cas par exemple dans le cadre de tirages successifs dans une urne évolutive, l'évolution se faisant différemment suivant la boule tirée à un tirage donné. Cette formule permet de formaliser le calcul de la probabilité associée à une branche de l'arbre des possibilités.

Méthode III.4 (Calcul de la probabilité d'une union)**.**

- *Union dénombrable d'événements croissants pour l'inclusion : On utilise ici aussi le théorème de la limite monotone.*
- *Union d'un nombre fini ou dénombrable d'événements 2 à 2 incompatibles : On utilise l'additivité (cas fini) ou la σ-additivité (cas dénombrable) de la mesure de probabilité.*
- *Union d'un nombre fini d'événements mutuellement indépendants : On a intérêt dans ce cas à considérer l'événement complémentaire. On est alors ramené au calcul de la probabilité d'une intersection d'événements mutuellement indépendants.*
- *Union d'un nombre fini d'événements lorsqu'on a des informations sur les intersections : On utilise la formule du crible de Poincaré. Les cas $n = 2$ et $n = 3$ sont les plus utiles :*
  - $n = 2$: $\mathbb{P}(A_1 \cup A_2) = \mathbb{P}(A_1) + \mathbb{P}(A_2) - \mathbb{P}(A_1 \cap A_2)$
  - $n = 3$: $\mathbb{P}(A_1 \cup A_2 \cup A_3) = \mathbb{P}(A_1) + \mathbb{P}(A_2) + \mathbb{P}(A_3) - \mathbb{P}(A_2 \cap A_3) - \mathbb{P}(A_1 \cap A_3) - \mathbb{P}(A_1 \cap A_2) + \mathbb{P}(A_1 \cap A_2 \cap A_3)$

Méthode III.5 (Cas d'une expérience dont l'issue dépend de résultats antérieurs)**.** *Dans certaines situations, par exemple lorsque le mode opératoire d'une expérience (par exemple la composition d'une urne) dépend du résultat d'une expérience précédente, il peut être nécessaire de discuter suivant le résultat obtenu lors de l'expérience antérieure : autrement dit, dans ces situations, il est aisé de calculer des probabilités conditionnellement à chacun des résultats possibles de la première expérience. Il faut ensuite récupérer la probabilité globale à partir de toutes ces probabilités conditionnelles. On utilise pour cela la formule des probabilités totales.*
À retenir : discussion à faire $\rightarrow$ formule des probabilités totales.

Enfin, vu qu'elles interviennent dans de nombreuses formules, nous faisons quelques remarques concernant le calcul des probabilités conditionnelles.

Méthode III.6 (Calcul de probabilités conditionnelles)**.** *Puisque la probabilité conditionnelle $\mathbb{P}_B$ sachant un événement B donné est une mesure de probabilité, tous les résultats, et toutes les règles vues ci-dessus pour le calcul des probabilités s'appliquent au calcul des probabilités conditionnelles. Lorsque ces formules font elles-mêmes intervenir des probabilités conditionnelles, les conditions s'ajoutent à celle déjà présente (sous forme d'une intersection) :*

$$(\mathbb{P}_B)_C = \mathbb{P}_{B\cap C}.$$

À titre d'exemple, voici la formule des probabilités totales pour une probabilité conditionnelle : $(A_i)_{i\in I}$ étant un système complet au plus dénombrable tel que pour tout $i \in I$, $\mathbb{P}_C(A_i) \neq 0$, on a :

$$\mathbb{P}(B|C) = \sum_{i\in I} \mathbb{P}(A_i|C)\mathbb{P}(B|C \cap A_i).$$

Remarque III.7. *Attention, on ne peut pas définir de notion d'événement conditionnel. Ce n'est pas l'événement qui est conditionnel, mais sa réalisation. Ainsi, si vous adoptez la démarche générale de rédaction :*

- *ne décrivez pas l'événement conditionnel (cela n'a pas de sens), mais donnez bien une condition nécessaire et suffisante de réalisation, conditionnellement à un certain événement :*

« Sachant que B est réalisé, A est réalisé si et seulement si ... »

- *Gardez-vous bien de transcrire cette phrase de manière ensembliste ; cela vous amènerait inévitablement à considérer des « événements conditionnels ». Sautez cette étape et passez directement aux probabilités. Vous pouvez transcrire votre phrase de façon ensembliste à l'intérieur des probabilités, la condition étant alors précisée non pas au niveau des ensembles, mais au niveau de la probabilité.*

Remarque III.8. *L'indépendance est une notion dépendant de la mesure de probabilité. Ce n'est pas parce que A et B sont indépendants (lorsqu'on dit cela, on sous-entend l'indépendance pour la mesure de probabilité totale $\mathbb{P}$) qu'ils sont indépendants pour une mesure de probabilité conditionnelle $\mathbb{P}_C$.*

Exemple III.9. *On effectue 3 tirages à Pile ou Face indépendants. L'événement C est réalisé si et seulement si on obtient exactement 1 Pile lors des tirages 1 et 2, A de même avec les tirages 2 et 3, et B de même avec les tirages 1 et 3. Les événements A et B sont indépendants pour $\mathbb{P}$, mais pas pour $\mathbb{P}_C$.*