

Espaces probabilisés

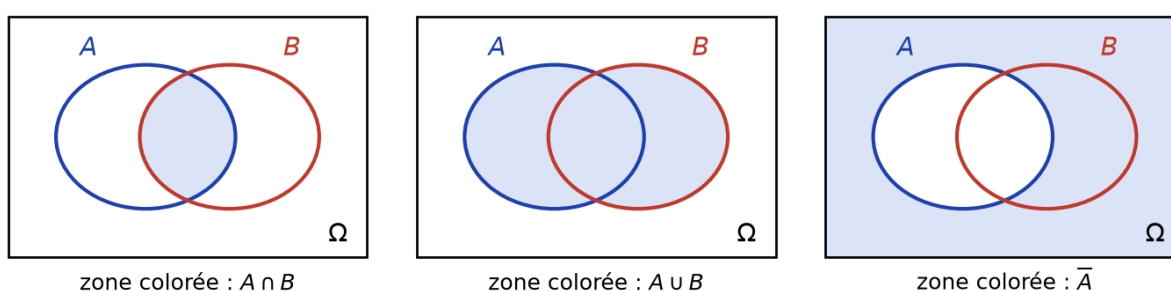
I. Univers, événements et tribus

Une expérience aléatoire est une expérience dont on connaît les issues possibles, mais pas le résultat à l'avance. On lance un dé, on tire une carte, on attend le premier pile d'une pièce. Pour raisonner, on regroupe d'abord toutes les issues dans un ensemble. Ensuite, on décide quelles parties de cet ensemble on veut mesurer.

Définition :

L'**univers** Ω est l'ensemble des issues de l'expérience. Un **événement** est une partie de Ω . L'événement Ω est l'événement certain, $\emptyset$ est l'événement impossible, et un singleton $\{\omega\}$ est un événement élémentaire.

Le langage des ensembles traduit le langage courant. Ainsi, « A ou B » correspond à $A \cup B$, « A et B » à $A \cap B$, et « non A » au complémentaire $\bar{A}$. De plus, deux événements sont incompatibles lorsque $A \cap B = \emptyset$. Enfin, « A implique B » s'écrit $A \subset B$. La figure ci-dessous représente ces opérations dans un diagramme de Venn.



1. Pourquoi une tribu ?

Lorsque Ω est fini ou dénombrable, on peut attribuer une probabilité à toutes les parties de Ω . En revanche, pour des univers plus gros, comme l'ensemble des suites infinies de pile ou face, ce n'est plus possible. On se limite alors à une famille de parties, stable par les opérations usuelles. C'est la notion de tribu.

Définition :

Une **tribu** sur Ω est une partie $\mathcal{A}$ de $\mathcal{P}(\Omega)$ telle que :

- $\Omega \in \mathcal{A}$;
- pour tout $A \in \mathcal{A}$, on a $\bar{A} \in \mathcal{A}$ (stabilité par passage au complémentaire) ;
- pour toute suite $(A_n)_{n \in \mathbb{N}}$ d'éléments de $\mathcal{A}$, on a $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{A}$ (stabilité par réunion dénombrable).

Le couple $(\Omega, \mathcal{A})$ est alors un **espace probabilisable**, et les éléments de $\mathcal{A}$ sont les événements.

Exemple :

Les familles $\{\emptyset, \Omega\}$ et $\mathcal{P}(\Omega)$ sont des tribus : ce sont la plus petite et la plus grande. Par exemple, si A est une partie de Ω , la famille $\{\emptyset, A, \bar{A}, \Omega\}$ est aussi une tribu. C'est la plus petite tribu contenant A .

Proposition :

Une tribu $\mathcal{A}$ contient $\emptyset$. Elle est stable par réunion finie, par intersection finie ou dénombrable et par différence $A \setminus B = A \cap \overline{B}$.

Démonstration :

D'abord, $\emptyset = \overline{\Omega}$ appartient à $\mathcal{A}$. Ensuite, pour une réunion finie $A_0 \cup \dots \cup A_p$, on complète la suite par $A_n = \emptyset$ pour $n > p$. Enfin, pour une intersection dénombrable, on utilise les lois de De Morgan :

$$\bigcap_{n \in \mathbb{N}} A_n = \overline{\bigcup_{n \in \mathbb{N}} \overline{A_n}}.$$

Chaque $\overline{A_n}$ est dans $\mathcal{A}$, donc leur réunion aussi, donc son complémentaire aussi.

Remarque :

Dans ce chapitre, lorsque Ω est fini ou dénombrable, on prend toujours $\mathcal{A} = \mathcal{P}(\Omega)$. La tribu ne joue un vrai rôle que pour les univers non dénombrables, étudiés plus tard avec la théorie de la mesure.

II. Espaces probabilisés : axiomes et propriétés

1. Définition d'une probabilité

Définition :

Soit $(\Omega, \mathcal{A})$ un espace probabilisable. Une **probabilité** est une application $P : \mathcal{A} \rightarrow [0, 1]$ telle que $P(\Omega) = 1$ et, pour toute suite (A_n) d'événements deux à deux incompatibles,

$$P\left(\bigcup_{n=0}^{+\infty} A_n\right) = \sum_{n=0}^{+\infty} P(A_n) \quad (\sigma\text{-additivit}[?][?]).$$

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

La série du membre de droite converge toujours. En effet, ses termes sont positifs et ses sommes partielles sont majorées par 1, comme on le verra. La σ -additivité est l'axiome qui permet de passer du fini au dénombrable.

Propriété :

Pour tous événements A et B :

- $P(\emptyset) = 0$, et $P(A \cup B) = P(A) + P(B)$ si $A \cap B = \emptyset$;
- $P(\overline{A}) = 1 - P(A)$;
- si $A \subset B$, alors $P(B \setminus A) = P(B) - P(A)$ et donc $P(A) \leq P(B)$;
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

Démonstration :

D'abord, on applique la σ -additivité à la suite constante égale à $\emptyset$. On obtient

$P(\emptyset) = \sum P(\emptyset)$, ce qui impose $P(\emptyset) = 0$. Ensuite, en complétant A, B par des

ensembles vides, on obtient l'additivité finie. Par conséquent, $1 = P(\Omega) = P(A) + P(\bar{A})$. Si $A \subset B$, la réunion disjointe $B = A \cup (B \setminus A)$ donne la troisième propriété. Enfin, $A \cup B = A \cup (B \setminus (A \cap B))$ est une réunion disjointe, d'où la dernière formule.

La dernière formule se généralise à n événements. C'est la formule du crible, aussi appelée formule de Poincaré.

Proposition :

Pour des événements $A_1, \dots, A_n$,

$$P\left(\bigcup_{i=1}^n A_i\right) = \sum_{k=1}^n (-1)^{k+1} \sum_{1 \leq i_1 < \dots < i_k \leq n} P(A_{i_1} \cap \dots \cap A_{i_k}).$$

Elle se démontre par récurrence sur n , en appliquant la formule à deux événements à $A_1 \cup \dots \cup A_n$ et A_{n+1} .

2. Univers fini ou dénombrable

Sur un univers fini, la situation de référence est l'équiprobabilité. Chaque issue a alors la probabilité $1/\text{card}(\Omega)$, et on a :

$$P(A) = \frac{\text{card}(A)}{\text{card}(\Omega)}.$$

Sur un univers dénombrable, l'équiprobabilité est impossible. En effet, si tous les singletons de $\mathbb{N}$ avaient la même probabilité c , la série $\sum c$ vaudrait 0 ou $+\infty$, jamais 1 . On décrit donc une probabilité par ses valeurs sur les singletons.

Théorème :

Soit $\Omega = \{\omega_n, n \in \mathbb{N}\}$ un ensemble dénombrable, et (p_n) une suite de réels positifs telle que $\sum_n p_n = 1$. Il existe une unique probabilité P sur $(\Omega, \mathcal{P}(\Omega))$ telle que $P(\{\omega_n\}) = p_n$ pour tout n . Elle est donnée par

$$P(A) = \sum_{n : \omega_n \in A} p_n.$$

Démonstration :

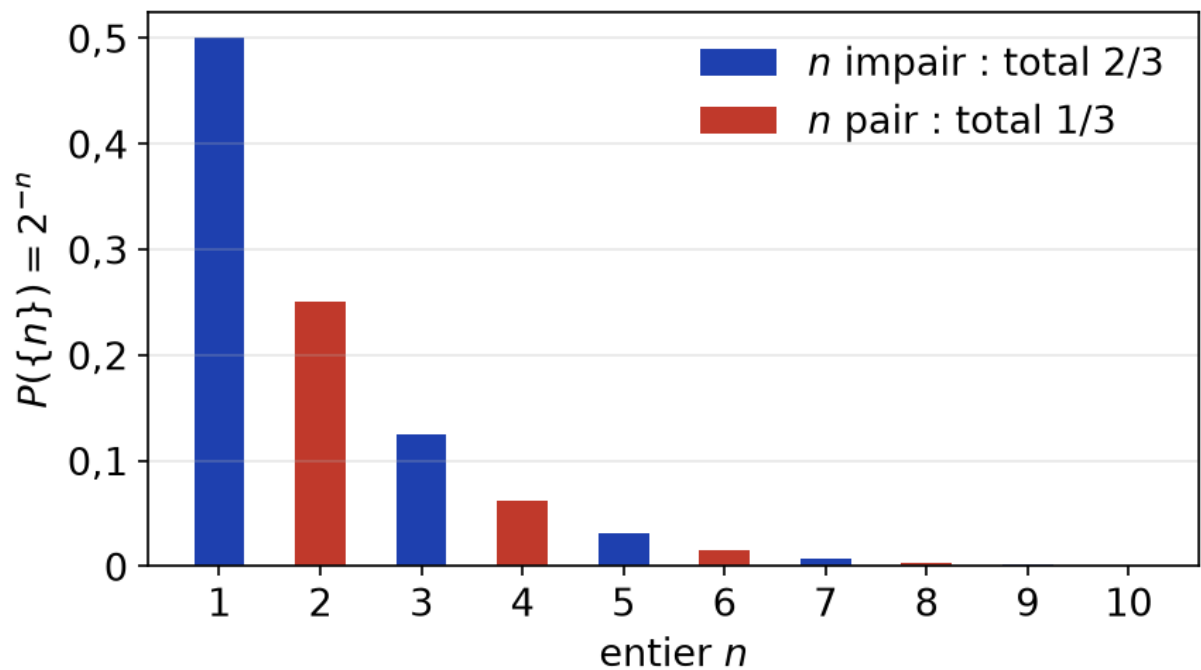
D'abord, l'unicité : toute partie A est la réunion disjointe dénombrable des singletons qu'elle contient. La σ -additivité impose donc la formule. Ensuite, l'existence : la formule définit bien un nombre de $[0,1]$, car (p_n) est une famille sommable de réels positifs. De plus, $P(\Omega) = 1$. Enfin, si les A_k sont deux à deux disjoints, le théorème de sommation par paquets pour les familles positives donne $P(\bigcup_k A_k) = \sum_k P(A_k)$.

Exemple :

Sur $\Omega = \mathbb{N}^*$, posons $p_n = \frac{1}{2^n}$. Ces réels sont positifs et $\sum_{n \geq 1} 2^{-n} = 1$. Ils définissent donc une probabilité. Par exemple, la probabilité d'obtenir un entier pair vaut

$$\sum_{k \geq 1} \frac{1}{4^k} = \frac{1/4}{1 - 1/4} = \frac{1}{3}.$$

C'est la loi du rang du premier pile avec une pièce équilibrée, représentée ci-dessous.



III. Continuité monotone d'une probabilité

La σ -additivité a une conséquence très utile : une probabilité passe à la limite le long des suites monotones d'événements. C'est l'outil central pour les expériences qui se répètent indéfiniment.

Théorème :

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé.

- **Continuité croissante.** Si (A_n) est croissante ($A_n \subset A_{n+1}$), alors

$$P\left(\bigcup_n A_n\right) = \lim_{n \rightarrow +\infty} P(A_n).$$

- **Continuité décroissante.** Si (A_n) est décroissante ($A_{n+1} \subset A_n$), alors

$$P\left(\bigcap_n A_n\right) = \lim_{n \rightarrow +\infty} P(A_n).$$

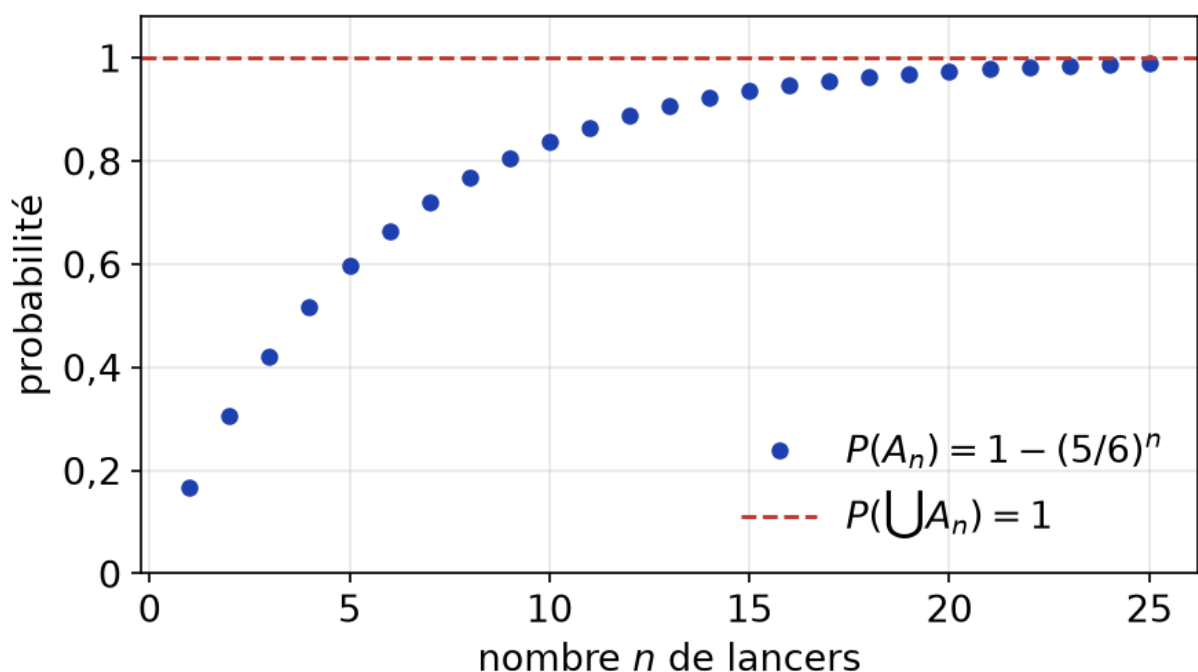
Démonstration :

Supposons (A_n) croissante. Posons $B_0 = A_0$ et $B_n = A_n \setminus A_{n-1}$ pour $n \geq 1$. Les B_n sont deux à deux disjoints. De plus, $A_n = B_0 \cup \dots \cup B_n$ et $\bigcup A_n = \bigcup B_n$. Par σ -additivité,

$$P\left(\bigcup_n A_n\right) = \sum_{k=0}^{+\infty} P(B_k) = \lim_{n \rightarrow +\infty} \sum_{k=0}^n P(B_k) = \lim_{n \rightarrow +\infty} P(A_n).$$

Si maintenant (A_n) est décroissante, la suite $(\overline{A_n})$ est croissante. Par conséquent, on applique le premier point aux complémentaires et on utilise $\overline{\bigcap A_n} = \bigcup \overline{A_n}$.

La figure suivante illustre la continuité croissante. On lance un dé indéfiniment et A_n est l'événement « un six apparaît parmi les n premiers lancers ». Ainsi, $P(A_n) = 1 - (5/6)^n$ croît vers 1.



Corollaire :

Pour toute suite (A_n) d'événements (non nécessairement disjoints), on a l'inégalité de Boole :

$$P\left(\bigcup_n A_n\right) \leq \sum_{n=0}^{+\infty} P(A_n).$$

En particulier, une réunion dénombrable d'événements de probabilité nulle est de probabilité nulle.

Définition :

Un événement de probabilité 0 est dit **négligeable**. Un événement de probabilité 1 est dit **presque sûr**. Un événement négligeable n'est pas forcément vide.

Méthode :

Pour calculer la probabilité d'un événement décrit par « à un moment » ou « toujours » :

- écrivez-le comme une réunion (« à un moment ») ou une intersection (« toujours ») d'événements A_n ne portant que sur les n premières étapes ;
- rendez la suite monotone en posant $A'_n = A_0 \cup \dots \cup A_n$ ou $A_0 \cap \dots \cap A_n$;
- calculez $P(A'_n)$ par un argument fini, puis passez à la limite par continuité monotone.

Exemple :

On lance indéfiniment une pièce équilibrée. Soit F_n l'événement « face aux n premiers lancers ».

La suite (F_n) est décroissante et $P(F_n) = 2^{-n}$. Donc l'événement « on n'obtient jamais pile », égal à $\bigcap F_n$, a pour probabilité $\lim 2^{-n} = 0$. Il est négligeable, mais il n'est pas vide.

IV. Dénombrement et modèles d'urnes

Dans le cas équiprobable, calculer une probabilité revient à compter. On rappelle les trois résultats de base, pour un ensemble E à n éléments.

Propriété :

- Le nombre de p -listes d'éléments de E est n^p .
- Le nombre de p -listes d'éléments distincts (arrangements) est $\frac{n!}{(n-p)!}$ pour $p \leq n$.
- Le nombre de parties à p éléments (combinaisons) est $\binom{n}{p} = \frac{n!}{p!(n-p)!}$.

Ces trois dénombrements correspondent aux trois façons de tirer p boules dans une urne de n boules numérotées. Le choix de l'univers dépend du protocole, et il doit rendre les issues équiprobables.

Méthode :

- **Tirages successifs avec remise** : Ω est l'ensemble des p -listes, de cardinal n^p .
- **Tirages successifs sans remise** : Ω est l'ensemble des arrangements, de cardinal $n!/(n-p)!$.
- **Tirage simultané** : Ω est l'ensemble des parties à p éléments, de cardinal $\binom{n}{p}$.

Numérotez toujours les boules, même si elles ont la même couleur : c'est ce qui garantit l'équiprobabilité.

Proposition :

Une urne contient N boules, dont b blanches. On en tire n . La probabilité d'obtenir exactement k boules blanches vaut :

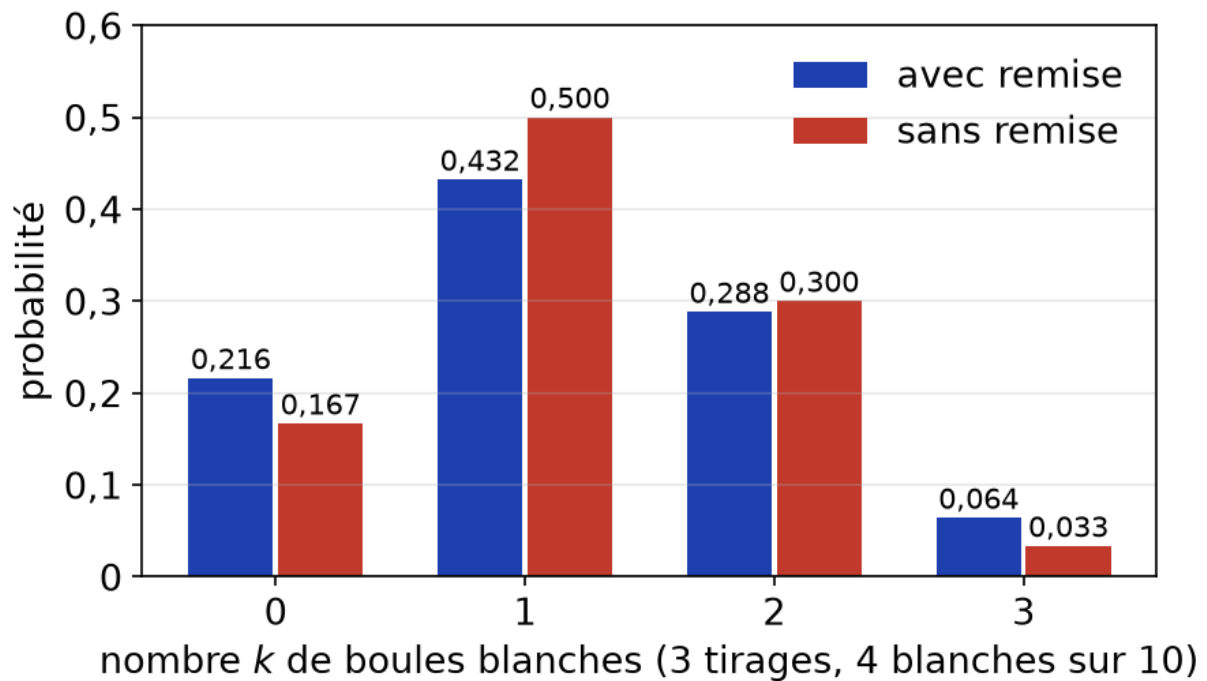
- avec remise : $\binom{n}{k} \left(\frac{b}{N}\right)^k \left(1 - \frac{b}{N}\right)^{n-k}$;
- sans remise, ou simultanément : $\frac{\binom{b}{k} \binom{N-b}{n-k}}{\binom{N}{n}}$.

Démonstration :

Sans remise, avec l'univers des parties à n éléments, on choisit k boules parmi les b blanches et $n-k$ parmi les $N-b$ autres. Le principe multiplicatif donne le numérateur. Avec remise, on choisit d'abord les k rangs des boules blanches, soit $\binom{n}{k}$ choix. Ensuite, il reste $b^k (N-b)^{n-k}$ listes. On divise enfin par N^n .

Exemple :

Une urne contient 10 boules dont 4 blanches, et on en tire 3. Sans remise, la probabilité d'avoir exactement une blanche vaut $\frac{\binom{4}{1}\binom{6}{2}}{\binom{10}{3}} = 60/120 = 1/2$. Avec remise, elle vaut $3 \times 0,4 \times 0,6^2 = 0,432$. La figure ci-dessous compare les deux modèles.



V. Probabilités conditionnelles

1. Définition

Savoir qu'un événement B est réalisé modifie l'information disponible. On restreint alors l'univers à B et on renormalise.

Définition :

Soit B un événement tel que $P(B) > 0$. La **probabilité conditionnelle** de A sachant B est

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

Proposition :

L'application $P_B : \mathcal{A} \rightarrow [0,1]$ est une probabilité sur $(\Omega, \mathcal{A})$.

Démonstration :

D'abord, $0 \leq P(A \cap B) \leq P(B)$, donc P_B est à valeurs dans $[0,1]$. Ensuite, $P_B(\Omega) = P(B)/P(B) = 1$. Enfin, si les A_n sont deux à deux disjoints, les $A_n \cap B$ le sont aussi. Donc $P\left(\bigcup A_n \cap B\right) = \sum P(A_n \cap B)$, et on divise par $P(B)$.

Propriété :

Formule des probabilités composées. Si $P(A_1 \cap \dots \cap A_{n-1}) > 0$, alors

$$P(A_1 \cap \dots \cap A_n) = P(A_1) P(A_2 | A_1) \dots P(A_n | A_1 \cap \dots \cap A_{n-1}).$$

Cette formule se démontre par récurrence, en écrivant chaque facteur comme un quotient : le produit se télescope. Elle traduit la lecture d'une branche d'un arbre de probabilités.

2. Probabilités totales et formule de Bayes

Définition :

Une famille finie ou dénombrable $(B_i)_{i \in I}$ d'événements est un **système complet d'événements** si les B_i sont deux à deux incompatibles et si leur réunion est Ω .

Théorème :

Formule des probabilités totales. Soit $(B_i)_{i \in I}$ un système complet d'événements, avec I fini ou dénombrable. Pour tout événement A ,

$$P(A) = \sum_{i \in I} P(A \cap B_i) = \sum_{\substack{i \in I, \\ P(B_i) > 0}} P(A | B_i) P(B_i).$$

Démonstration :

Les événements $A \cap B_i$ sont deux à deux disjoints et leur réunion est $A \cap \Omega = A$. La σ -additivité donne la première égalité. Ensuite, on remplace $P(A \cap B_i)$ par $P(A | B_i) P(B_i)$ lorsque $P(B_i) > 0$. Les autres termes sont nuls, car $P(A \cap B_i) \leq P(B_i) = 0$.

Théorème :

Formule de Bayes. Avec les mêmes hypothèses, si $P(A) > 0$ et $P(B_j) > 0$,

$$P(B_j | A) = \frac{P(A | B_j) P(B_j)}{\sum_{\substack{i \in I, \\ P(B_i) > 0}} P(A | B_i) P(B_i)}.$$

La formule de Bayes « remonte le temps » : elle donne la probabilité d'une cause connaissant l'effet observé. Par exemple, elle mesure la fiabilité réelle d'un test médical.

Méthode :

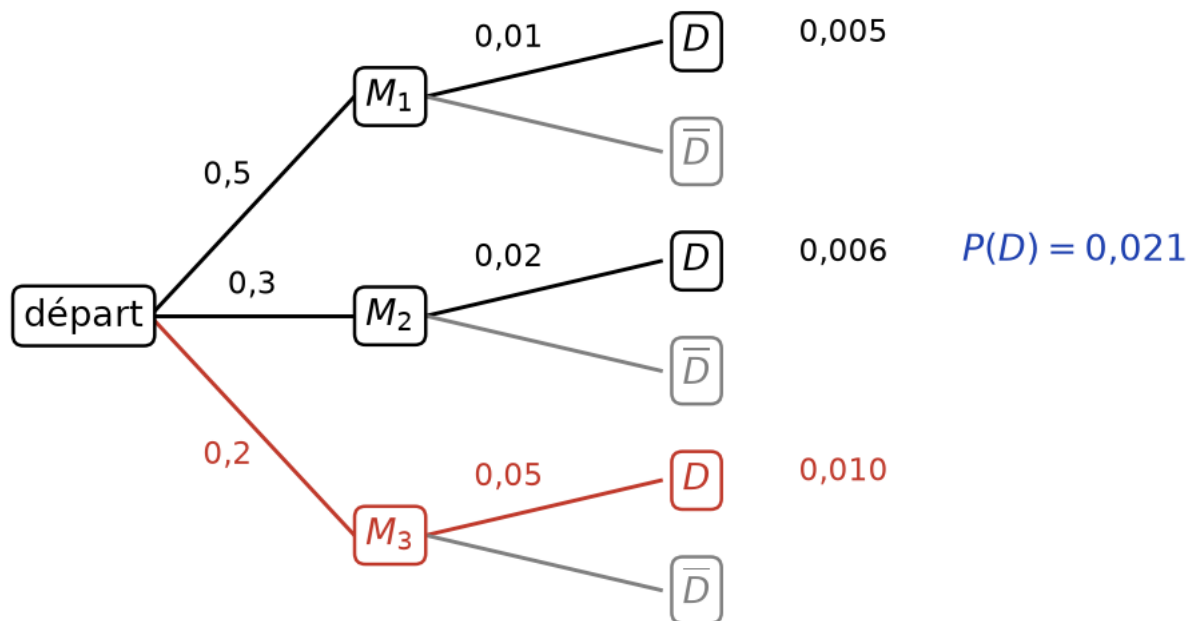
- Identifiez l'étape « cause » : ses issues forment le système complet (B_i) .
- Traduisez chaque donnée de l'énoncé : « parmi les B_i , une proportion... » donne $P(A | B_i)$, pas $P(A \cap B_i)$.
- Calculez $P(A)$ par les probabilités totales, puis $P(B_j | A)$ par Bayes. Un arbre aide à ne rien oublier.

Exemple :

Trois machines produisent respectivement 50 %, 30 % et 20 % des pièces d'une usine. Leurs taux de pièces défectueuses sont 1 %, 2 % et 5 %. Soit D l'événement « la pièce est défectueuse ». Comme le montre l'arbre ci-dessous, les probabilités totales donnent

$$P(D) = 0,5 \times 0,01 + 0,3 \times 0,02 + 0,2 \times 0,05 = 0,021.$$

Ensuite, la formule de Bayes donne $P(M_3 | D) = 0,010 / 0,021 = 10/21 \approx 0,48$. Autrement dit, près d'une pièce défectueuse sur deux vient de la troisième machine.



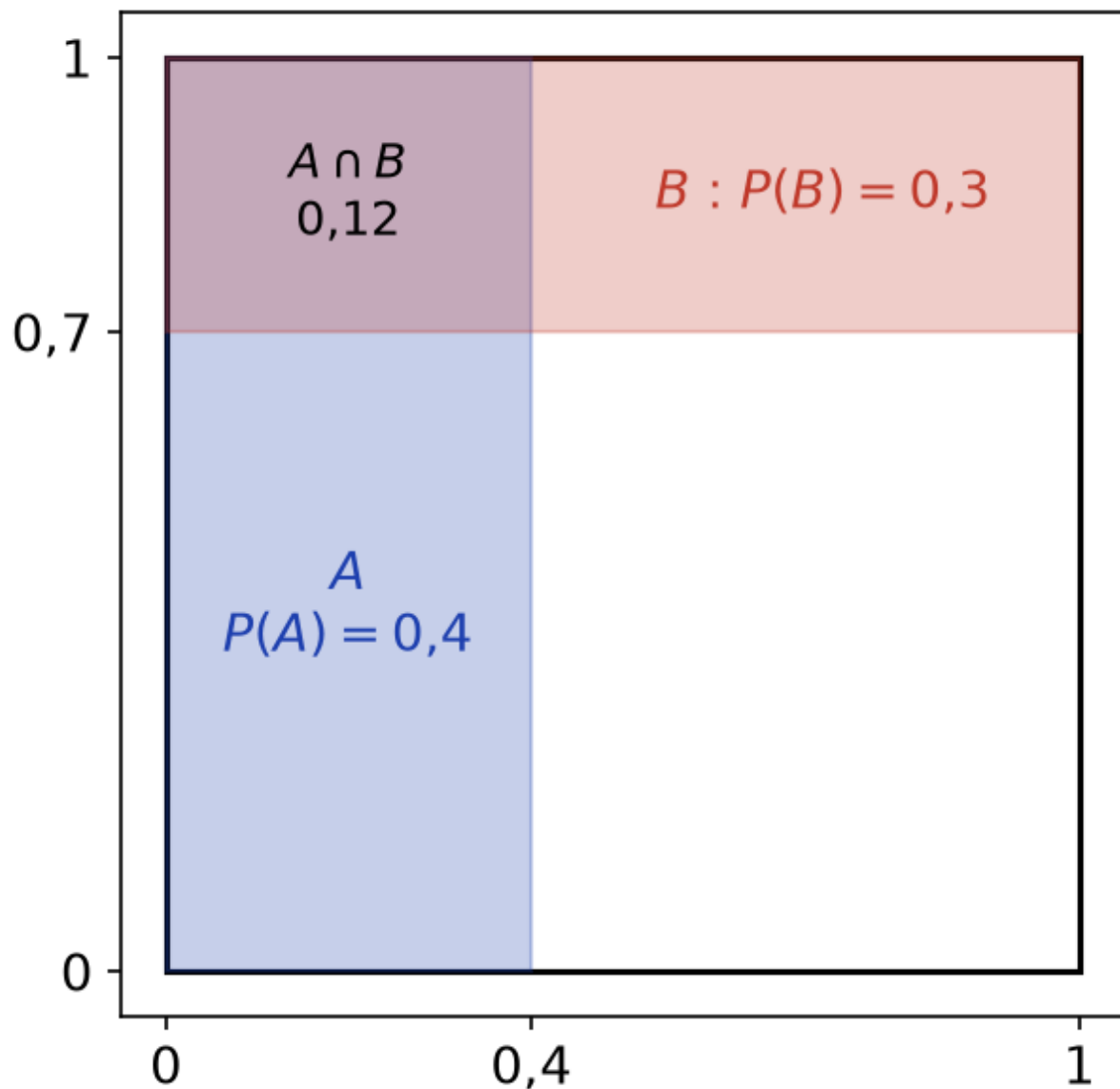
VI. Indépendance d'événements

1. Deux événements

Définition :

Deux événements A et B sont **indépendants** si $P(A \cap B) = P(A) P(B)$. Lorsque $P(B) > 0$, cela équivaut à $P(A | B) = P(A)$.

Sur le carré unité muni de l'aire, une bande verticale et une bande horizontale sont indépendantes. En effet, l'aire de leur intersection est le produit des largeurs, comme le montre la figure.



Propriété :

Si A et B sont indépendants, alors A et $\bar{B}$ le sont, ainsi que $\bar{A}$ et $\bar{B}$.

Démonstration :

On a $P(A \cap \bar{B}) = P(A) - P(A \cap B) = P(A) - P(A)P(B) = P(A)P(\bar{B})$. En appliquant ce résultat au couple $(\bar{B}, A)$, on obtient ensuite l'indépendance de $\bar{B}$ et $\bar{A}$.

Attention :

Il ne faut pas confondre indépendants et incompatibles. Si $A \cap B = \emptyset$ avec $P(A) > 0$ et $P(B) > 0$, alors $P(A \cap B) = 0 \neq P(A)P(B)$: les événements ne sont pas indépendants.

2. Famille d'événements

Définition :

Les événements $A_1, \dots, A_n$ sont **mutuellement indépendants** si, pour toute partie $J \subset \{1, \dots, n\}$,

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

Une suite infinie d'événements est mutuellement indépendante si toute sous-famille finie l'est.

Attention :

L'indépendance deux à deux n'entraîne pas l'indépendance mutuelle. Par exemple, lançons deux pièces équilibrées. Soit A : « pile au premier lancer », B : « pile au second », C : « les deux résultats sont égaux ». Chacun a pour probabilité $1/2$, et chaque intersection de deux d'entre eux vaut $\{(P, P)\}$, de probabilité $1/4$. Cependant, $P(A \cap B \cap C) = 1/4 \neq 1/8$.

Proposition :

Si $A_1, \dots, A_n$ sont mutuellement indépendants, on obtient encore une famille mutuellement indépendante en remplaçant certains A_i par leurs complémentaires. En particulier,

$$P\left(\bigcup_{i=1}^n A_i\right) = 1 - \prod_{i=1}^n (1 - P(A_i)).$$

La preuve consiste à remplacer un seul événement, puis à itérer. La formule finale vient du passage au complémentaire : l'événement contraire de « au moins un A_i » est « aucun A_i ».

Remarque :

L'indépendance sert surtout à modéliser. Lorsque des lancers sont réalisés « sans influence mutuelle », on choisit une probabilité qui rend indépendants les événements liés à des lancers différents. Par exemple, pour une infinité de lancers de pièce, on admet l'existence d'un espace probabilisé où les événements « pile au n -ième lancer » sont mutuellement indépendants, de probabilité p .

Exemple :

Un réseau comporte 4 liaisons en parallèle. Chacune tombe en panne avec probabilité 0,1, indépendamment des autres. Le réseau fonctionne si au moins une liaison fonctionne. Sa probabilité de fonctionner vaut donc $1 - 0,1^4 = 0,9999$.

Ce qu'il faut retenir

- Un espace probabilisé est un triplet $(\Omega, \mathcal{A}, P)$: un univers, une tribu d'événements, une probabilité σ -additive.

- Sur un univers fini ou dénombrable, on prend $\mathcal{A} = \mathcal{P}(\Omega)$, et une probabilité est donnée par une suite (p_n) positive de somme 1.
- Il n'existe pas d'équiprobabilité sur un univers dénombrable infini.
- Continuité monotone : la probabilité d'une réunion croissante ou d'une intersection décroissante est la limite des probabilités.
- Inégalité de Boole : une réunion dénombrable d'événements négligeables est négligeable.
- Tirages avec remise : listes ; sans remise : arrangements ; simultanés : combinaisons. Les deux derniers donnent les mêmes probabilités.
- La probabilité conditionnelle P_B est une probabilité ; la formule des probabilités composées décrit une branche d'arbre.
- Probabilités totales sur un système complet fini ou dénombrable, puis formule de Bayes pour remonter à la cause.
- L'indépendance mutuelle exige la formule du produit pour toutes les sous-familles ; elle est stable par passage aux complémentaires.

Questions fréquentes sur espaces probabilisés

Pourquoi a-t-on besoin d'une tribu si l'univers est dénombrable ?

Lorsque l'univers est fini ou dénombrable, on prend simplement la tribu de toutes les parties, et toute partie est un événement. La notion de tribu devient indispensable pour les univers non dénombrables, comme l'ensemble des suites infinies de pile ou face. Sur ces univers, on ne peut plus définir une probabilité raisonnable sur toutes les parties.

Un événement de probabilité nulle est-il impossible ?

Non. Un événement négligeable peut contenir des issues. Par exemple, pour une infinité de lancers d'une pièce équilibrée, l'issue « face à tous les lancers » existe, mais la continuité décroissante montre que sa probabilité vaut $\lim 2^{-n} = 0$.

Quelle différence entre indépendance deux à deux et indépendance mutuelle ?

L'indépendance deux à deux demande seulement $P(A_i \cap A_j) = P(A_i)P(A_j)$ pour $i \neq j$.
L'indépendance mutuelle exige la formule du produit pour toutes les sous-familles. Avec deux pièces, les événements « pile au premier », « pile au second » et « résultats égaux » sont indépendants deux à deux, mais pas mutuellement.

Quand utiliser la formule de Bayes plutôt que les probabilités totales ?

Les probabilités totales calculent la probabilité d'un effet à partir des probabilités des causes. La formule de Bayes fait le chemin inverse : elle donne la probabilité d'une cause sachant l'effet observé. En pratique, on calcule d'abord $P(A)$ par les probabilités totales, puis on l'utilise au dénominateur de Bayes.

Pour aller plus loin en L2

- Les énoncés : [exercices de maths en L2 sur espaces probabilisés](#)
- Chapitre précédent : [Différentiabilité, formule de Taylor et extrema](#)
- Chapitre suivant : [Variables aléatoires discrètes](#)
- Tester vos connaissances : [QCM de maths en L2 par chapitre](#)
- Le sommaire : [tous les chapitres de maths de L2](#) et la [licence de maths de L1 à L3](#)