

Dans tout ce qui suit, on note Ω l'univers (fini) d'une expérience aléatoire, muni d'une probabilité \mathbf{P} .

1 Probabilités conditionnelles

Lorsque on cherche à calculer une probabilité, il est assez courant qu'on sache qu'un certain évènement est déjà réalisé.

Un exemple type est : « **Sachant** qu'un patient est positif à un test de dépistage médical, quelle est la probabilité qu'il soit malade ? »

On veut donc calculer une probabilité sur un univers restreint : celui des patients positifs. Notre façon de calculer les probabilités doit rendre compte de cette information.

On propose la définition suivante.

Définition 1

Soit A, B des évènements, tel que $\mathbf{P}(A) \neq 0$. On appelle la **probabilité de B sachant A** le nombre

$$\mathbf{P}_A(B) := \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(A)}.$$

Remarques

- Pouvez-vous montrer que $\mathbf{P}_A(B)$ est bien dans l'intervalle $[0; 1]$? C'est-à-dire que ce nombre mérite bien le nom de "probabilité".
- On est sensé "savoir" que A est réalisé. Montrez que $\mathbf{P}_A(A) = 1$.
- L'intersection avec A au numérateur traduit l'idée de restriction de l'univers.
- On a supposé $\mathbf{P}(A) \neq 0$. Serait-il intéressant de chercher à définir $\mathbf{P}_A(B)$ dans le cas $\mathbf{P}(A) = 0$?

Exemple Je tire une carte au hasard dans un jeu de 52 cartes. Je vous assure (et on me suppose honnête) que la carte tirée est une carte rouge.

Quelle est la probabilité que ce soit une dame? Essayez de trouver la réponse intuitivement d'abord. Appliquons ensuite la définition. On note

— R l'évènement « la carte tirée est rouge ». Donc $\mathbf{P}(R) = \frac{13}{52} = \frac{1}{4}$.

— D l'évènement « la carte tirée est une dame ». Et donc $\mathbf{P}(D \cap R) = \frac{2}{52}$.

Alors

$$\mathbf{P}_R(D) = \frac{\mathbf{P}(D \cap R)}{\mathbf{P}(R)} = \frac{\frac{2}{52}}{\frac{13}{52}} = \frac{2}{13} \approx 0,15.$$

On a ainsi créé une nouvelle façon de calculer des probabilités qui rend compte de notre "savoir". \mathbf{P}_A vérifie les mêmes formules que d'habitude :

Théorème 1

Soit A, B, C des évènements, tel que $\mathbf{P}(A) \neq 0$.

- $\mathbf{P}_A(B \cup C) = \mathbf{P}_A(B) + \mathbf{P}_A(C) - \mathbf{P}_A(B \cap C)$.
- $\mathbf{P}_A(\emptyset) = 0$, $\mathbf{P}_A(\Omega) = 1$.
- $\mathbf{P}_A(\overline{B}) = 1 - \mathbf{P}_A(B)$.

Démonstration en annexe.

2 Arbres pondérés

On peut représenter les expériences aléatoires par

- des tableaux dans les cas simples, comme vu en introduction du chapitre,
- ou par des arbres.

Ces derniers sont assez intuitifs pour décrire des situations avec une certaine « chronologie » Décrivons un exemple type : un tirage sans remise.

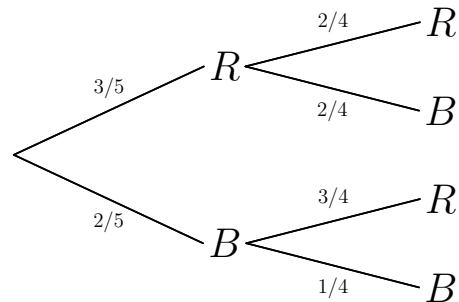
Problème : Un sac contient 3 boules rouges et 2 boules bleues. On effectue deux tirages sans remise. Quelle est, par exemple, la probabilité d'obtenir deux boules de la même couleur ?

Notons

R l'évènement « la boule tirée est rouge »

B l'évènement « la boule tirée est bleue »

On représente le tirage par l'arbre suivant :



Définition 2

On appelle

branche le chemin menant d'un nœud à un autre.

feuille l'extrémité d'une suite de branches.

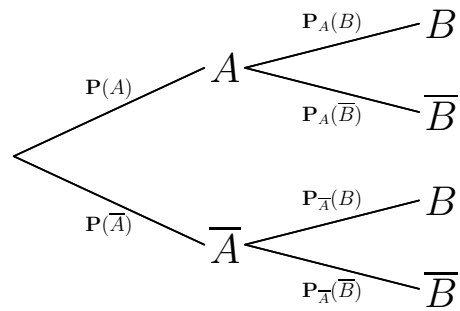
Théorème 2 (Règles d'utilisation d'un arbre)

- La somme des probabilités des branches issues d'un même nœud est égale à 1.
- La probabilité d'une feuille est égale au produit des probabilités des branches aboutissant à la feuille.
- La probabilité d'un évènement réalisé par plusieurs feuilles est la somme des probabilités des feuilles en question.

La probabilité cherchée est donc $\frac{3}{5} \cdot \frac{2}{4} + \frac{2}{5} \cdot \frac{1}{4} = \frac{2}{5}$.

Plus généralement, étant donné deux évènements A, B on a l'arbre suivant :

La troisième des règles ci-dessus peut se formaliser ainsi :

**Théorème 3**

Soit A, B deux évènements, tels que $P(A) \neq 0$.

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

Ce qui peut se réécrire

$$P(B) = P_A(B)P(A) + P_{\bar{A}}(B)P(\bar{A}).$$

3 Indépendance de deux évènements

L'idée est la suivante :

Deux évènements A, B sont indépendants quand savoir que l'un des deux est réalisé n'apporte pas de nouvelle information sur l'autre (en termes de probabilités).

Plus précisément :

Définition 3

Soit A, B deux évènements. Ils sont dits **indépendants** quand

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

Une conséquence très importante de cette définition est le théorème suivant :

Théorème 4

Soit A, B deux évènements, tels que $P(A) \neq 0$ et $P(B) \neq 0$. Au moins une des deux égalités suivantes est vraie

$$P_A(B) = P(B)$$

$$P_B(A) = P(A)$$

si et seulement si A et B sont indépendants.

En effet, ce théorème correspond mieux à l'idée intuitive énoncée ci-dessus. Pourquoi ne pas prendre ce théorème comme définition de l'indépendance? La définition est **une** égalité concise, **symétrique**, ne nécessitant pas d'hypothèse supplémentaire.

Théorème 5

Si A, B sont indépendants alors \bar{A} et B le sont aussi.

Démonstration

$$\begin{aligned}\mathbf{P}(\bar{A})\mathbf{P}(B) &= (1 - \mathbf{P}(A))\mathbf{P}(B) \\ &= \mathbf{P}(B) - \mathbf{P}(A)\mathbf{P}(B).\end{aligned}$$

Or par hypothèse $\mathbf{P}(A \cap B) = \mathbf{P}(A)\mathbf{P}(B)$. Donc

$$\mathbf{P}(\bar{A})\mathbf{P}(B) = \mathbf{P}(B) - \mathbf{P}(A \cap B).$$

Puis, par le théorème 3

$$\mathbf{P}(\bar{A})\mathbf{P}(B) = \mathbf{P}(\bar{A} \cap B).$$

Remarques On en déduit que \bar{A} et \bar{B} sont indépendants, ainsi que A et \bar{B} .

4 Annexe

Démonstration (du théorème 1)

Premier point On part de la définition pour écrire :

$$\mathbf{P}_A(B \cup C) = \frac{\mathbf{P}((B \cup C) \cap A)}{\mathbf{P}(A)}.$$

Un dessin (faites-le!) permet de se rendre compte que $(B \cup C) \cap A = (B \cap A) \cup (C \cap A)$. Et donc

$$\begin{aligned} \mathbf{P}_A(B \cup C) &= \frac{\mathbf{P}((B \cap A) \cup (C \cap A))}{\mathbf{P}(A)} \\ &= \frac{\mathbf{P}(B \cap A) + \mathbf{P}(C \cap A) - \mathbf{P}((B \cap A) \cap (C \cap A))}{\mathbf{P}(A)} \\ &= \mathbf{P}_A(B) + \mathbf{P}_A(C) - \frac{\mathbf{P}((B \cap A) \cap (C \cap A))}{\mathbf{P}(A)}. \end{aligned}$$

Un autre dessin donne $(B \cap A) \cap (C \cap A) = (B \cap C) \cap A$. D'où

$$\mathbf{P}_A(B \cup C) = \mathbf{P}_A(B) + \mathbf{P}_A(C) - \mathbf{P}_A(B \cap C).$$

Deuxième point Par définition

$$\mathbf{P}_A(\emptyset) = \frac{\mathbf{P}(\emptyset \cap A)}{\mathbf{P}(A)}$$

Or, $\emptyset \cap A = \emptyset$. Donc

$$\mathbf{P}_A(\emptyset) = \frac{\mathbf{P}(\emptyset)}{\mathbf{P}(A)} = 0.$$

De même, comme $\Omega \cap A = A$,

$$\mathbf{P}_A(\Omega) = \frac{\mathbf{P}(A)}{\mathbf{P}(A)} = 1.$$

Dernier point La dernière formule se déduit des deux précédentes car

$$B \cup \bar{B} = \Omega$$

$$B \cap \bar{B} = \emptyset$$

la première formule donne

$$\mathbf{P}_A(\Omega) = \mathbf{P}_A(B) + \mathbf{P}_A(\bar{B}) - \mathbf{P}_A(\emptyset).$$

Et le deuxième point donne alors

$$1 = \mathbf{P}_A(B) + \mathbf{P}_A(\bar{B})$$

Définition 4

Un évènement est dit **négligeable** si sa probabilité est nulle.

Théorème 6 (Formule des probabilités composées)

Soit $n \geq 1$ un entier. Soit A_1, \dots, A_n des évènements d'intersection non négligeable. Alors

$$\mathbf{P}(A_1 \cap \dots \cap A_n) = \mathbf{P}(A_1) \cdot \mathbf{P}_{A_1}(A_2) \dots \mathbf{P}_{A_1 \cap \dots \cap A_{n-1}}(A_n)$$

Démonstration

Par récurrence sur n .

Initialisation La formule est triviale dans ce cas : $\mathbf{P}(A_1) = \mathbf{P}(A_1)$.

Hérédité Soit $n \geq 1$. Supposons la formule vraie. Montrons qu'elle est vraie pour $n + 1$. Soit donc A_1, \dots, A_{n+1} des évènements d'intersection non négligeable.

$$\begin{aligned} \mathbf{P}(A_1 \cap \dots \cap A_{n+1}) &= \mathbf{P}(A_1 \cap \dots \cap A_n) \times \mathbf{P}_{A_1 \cap \dots \cap A_n}(A_{n+1}) \\ &= \mathbf{P}(A_1) \cdot \mathbf{P}_{A_1}(A_2) \dots \mathbf{P}_{A_1 \cap \dots \cap A_{n-1}}(A_n) \cdot \mathbf{P}_{A_1 \cap \dots \cap A_n}(A_{n+1}) \end{aligned}$$

Théorème 7 (Formule des probabilités totales)

Soit A_1, \dots, A_n une partition de l'univers Ω . Et B un évènement. Alors

$$\mathbf{P}(B) = \mathbf{P}(A_1 \cap B) + \dots + \mathbf{P}(A_n \cap B).$$

Si de plus aucun des A_i n'est négligeable, alors

$$\mathbf{P}(B) = \mathbf{P}_{A_1}(B)\mathbf{P}(A_1) + \dots + \mathbf{P}_{A_n}(B)\mathbf{P}(A_n).$$

L'hypothèse de non-négligeabilité des A_i est superflue si on prend la convention $\mathbf{P}_{A_i}(B)\mathbf{P}(A_i) = 0$ quelque soit B .

Démonstration

Soit A_1, \dots, A_n une partition de l'univers, et B un évènement. Alors

$$B = B \cap \Omega = B \cap (A_1 \cup \dots \cup A_n) = (B \cap A_1) \cup \dots \cup (B \cap A_n)$$

Les unions du membre de droite de la dernière égalité sont disjointes. Donc

$$\mathbf{P}(B) = \mathbf{P}(B \cap A_1) + \dots + \mathbf{P}(B \cap A_n).$$

La deuxième formule découle de la définition des probabilités conditionnelles.