PROBABILITÉS CONDITIONNELLES ET ARBRES PONDÉRÉS

Cours

« Quand je suis curieux d'une chose, mathématique ou autre, je l'interroge. Je l'interroge, sans me soucier si ma question est peut-être stupide ou si elle va paraître telle, sans qu'elle soit à tout prix mûrement pesée. »

Alexandre Grothendieck

I. DÉFINITION

Définition. Soient A et B deux évènements tels que $p(\mathrm{A}) \neq 0$.

La **probabilité conditionnelle de l'évènement B sachant que l'évènement A est réalisé** est notée $p_{\mathrm{A}}(\mathrm{B})$ et est donnée par :

$$p_{\mathrm{A}}(\mathrm{B}) = \frac{p(\mathrm{A} \cap \mathrm{B})}{p(\mathrm{A})}$$

On dira en abrégé « p de B sachant A » pour $p_{\mathrm{A}}(\mathrm{B})$.

On en déduit immédiatement que :

Théorème. Soient A et B deux évènements.

$$p(\mathrm{A} \cap \mathrm{B}) = p_{\mathrm{A}}(\mathrm{B}) \times p(\mathrm{A})$$

Démonstration. On multiplie chaque membre de l'égalité de la définition par $p(\mathrm{A})$, ce qui est légitime car $p(\mathrm{A}) \neq 0$.

J'ajoute une remarque essentielle : $p_{\mathrm{A}}(\mathrm{B})$ est une probabilité donc $0 \leqslant p_{\mathrm{A}}(\mathrm{B}) \leqslant 1$ (si vous obtenez un jour une probabilité conditionnelle qui n'est pas dans cet intervalle, c'est forcément une erreur).

Pour la suite, je vous propose de travailler avec un énoncé unique pour tous les exemples de ce cours.

Si vous avez le temps et un peu de curiosité, vous pouvez regarder une vidéo sur les tests de dépistage par Manu Houdart.

L'énoncé commun aux exemples de ce cours

Dans un pays, il y a 2 % de la population contaminée par un virus. On dispose d'un test de dépistage de ce virus. La probabilité qu'une personne ait un test positif sachant qu'elle est contaminée est de 0,99 (pour les futurs soignants, on appelle cette probabilité la sensibilité du test). La probabilité qu'une personne qui n'est pas contaminée ait un test négatif est de 0,97 (cette probabilité étant appelée spécificité du test). On fait passer un test à une personne choisie au hasard dans cette population. On note C l'évènement « la personne est contaminée par le virus » et P l'évènement « le test est positif ».

Exemple 1. Quelle est la valeur de $p_{\mathrm{C}}(\mathrm{P})$?

Solution. D'après l'énoncé, la probabilité que la personne ait un test positif **sachant** qu'elle est contaminée est de 0,99 : $p_{\mathrm{C}}(\mathrm{P}) = 0{,}99$.

Noter que les expressions « *une personne a un test positif **sachant** qu'elle est contaminée* » et « *une personne contaminée a un test positif* » sont équivalentes.

Exemple 2. Calculer la probabilité de l'évènement $\mathrm{C} \cap \mathrm{P}$ et interpréter celle-ci.

Solution. D'après un théorème, $p(\mathrm{C} \cap \mathrm{P}) = p_{\mathrm{C}}(\mathrm{P}) \times p(\mathrm{C}) = 0{,}99 \times 2\,\% = 0{,}0198$.

La probabilité qu'une personne choisie au hasard soit contaminée **et** ait un test positif est 0,0198.

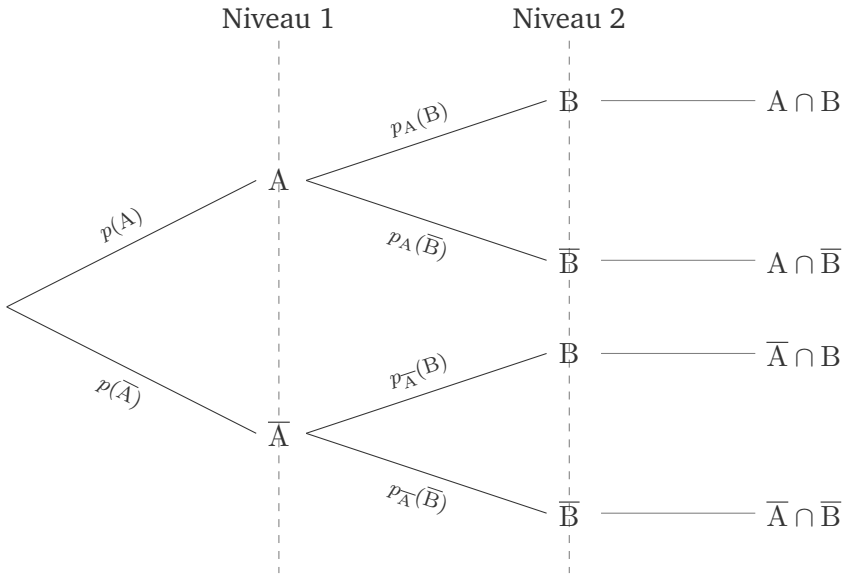
On peut donc déterminer des probabilités par la lecture de l'énoncé (exemple 1) ou par un calcul direct (exemple 2).

Grâce aux arbres pondérés, on va pouvoir représenter de façon condensée et très pratique cette situation : cette représentation va se révéler très efficace dans la résolution de problèmes de probabilités, comme lorsque l'on va chercher à connaître la probabilité $p_{\mathrm{P}}(\mathrm{C})$ qu'une personne soit contaminée sachant que son test est positif (dans l'exemple 1, on a extrait de l'énoncé $p_{\mathrm{C}}(\mathrm{P})$, noter la différence).

II. Arbres pondérés

Un arbre pondéré permet de modéliser certaines expériences aléatoires.

Arbre pondéré à deux niveaux et deux branches à chaque nœud



Règle de la somme : $p_{\mathrm{A}}(\mathrm{B}) + p_{\mathrm{A}}(\overline{\mathrm{B}}) = 1$

Règle du produit : $p(\mathrm{A}) \times p_{\mathrm{A}}(\mathrm{B}) = p(\mathrm{A} \cap \mathrm{B})$

Règle de la somme

La somme des probabilités inscrites sur les branches partant d'un même nœud est égale à 1.

Démonstration. Distinguons les niveaux 1 et 2.

Pour tout évènement A d'un univers Ω et $\overline{\mathrm{A}}$ son évènement contraire, on a, d'après un théorème de seconde :

$$p(\mathrm{A}) + p(\overline{\mathrm{A}}) = 1$$

ce qui prouve la règle au niveau 1.

Maintenant, pour le niveau 2, il s'agit de montrer par exemple que $p_{\mathrm{A}}(\mathrm{B}) + p_{\mathrm{A}}(\overline{\mathrm{B}}) = 1$.

L'idée est de considérer l'évènement A comme étant l'union de $\mathrm{A} \cap \mathrm{B}$ et de $\mathrm{A} \cap \overline{\mathrm{B}}$ (toute issue de A est soit dans B soit dans son contraire $\overline{\mathrm{B}}$). On obtient :

$$\begin{aligned} p(\mathrm{A}) &= p\big((\mathrm{A} \cap \mathrm{B}) \cup (\mathrm{A} \cap \overline{\mathrm{B}})\big) \\ &= p(\mathrm{A} \cap \mathrm{B}) + p(\mathrm{A} \cap \overline{\mathrm{B}}) - p\big((\mathrm{A} \cap \mathrm{B}) \cap (\mathrm{A} \cap \overline{\mathrm{B}})\big) \qquad \text{(d'après un autre théorème de seconde)} \\ &= p(\mathrm{A} \cap \mathrm{B}) + p(\mathrm{A} \cap \overline{\mathrm{B}}) \end{aligned}$$

En effet, $\mathrm{A} \cap \mathrm{B}$ et $\mathrm{A} \cap \overline{\mathrm{B}}$ étant incompatibles (une issue ne peut être à la fois dans A, dans B et dans $\overline{\mathrm{B}}$), $p\big((\mathrm{A} \cap \mathrm{B}) \cap (\mathrm{A} \cap \overline{\mathrm{B}})\big) = 0$. On a ainsi :

$$p(\mathrm{A}) = p(\mathrm{A} \cap \mathrm{B}) + p(\mathrm{A} \cap \overline{\mathrm{B}})$$

ce qui donne, puisque $p(\mathrm{A}) \neq 0$,

$$1 = \frac{p(\mathrm{A} \cap \mathrm{B})}{p(\mathrm{A})} + \frac{p(\mathrm{A} \cap \overline{\mathrm{B}})}{p(\mathrm{A})}$$

et enfin, par définition :

$$p_{\mathrm{A}}(\mathrm{B}) + p_{\mathrm{A}}(\overline{\mathrm{B}}) = 1$$

La démonstration est similaire pour les autres branches du niveau 2.

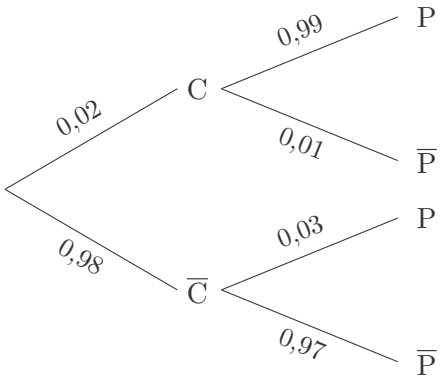
Exemple 3. Traduire la situation à l'aide d'un arbre de probabilités.

Solution. $p(\overline{\mathrm{C}}) = 1 - p(\mathrm{C}) = 1 - 0{,}02 = 0{,}98$

$p_{\mathrm{C}}(\overline{\mathrm{P}}) = 1 - p_{\mathrm{C}}(\mathrm{P}) = 1 - 99\ \% = 0{,}01$

$p_{\overline{\mathrm{C}}}(\mathrm{P}) = 1 - p_{\overline{\mathrm{C}}}(\overline{\mathrm{P}}) = 1 - 97\ \% = 0{,}03$

D'où l'arbre pondéré suivant :



Une succession de plusieurs branches reliées par des nœuds s'appelle un **chemin**. L'arbre ci-dessus possède quatre chemins.

Règle du produit

La probabilité de l'intersection des évènements rencontrés le long d'un chemin est égale au produit des probabilités rencontrées le long du chemin.

Démonstration. En suivant le chemin en haut de l'arbre, on rencontre les évènements A et B et le théorème donne directement :

$$p(\mathrm{A} \cap \mathrm{B}) = p(\mathrm{A}) \times p_{\mathrm{A}}(\mathrm{B})$$

La démonstration est similaire pour chaque chemin.

Exemple 4. On admet que la probabilité que le test soit positif est 0,0492. Déterminer la probabilité $p_{\mathrm{P}}(\mathrm{C})$ qu'une personne soit contaminée sachant que son test est positif. Donner le résultat sous la forme d'une fraction irréductible, puis une valeur approchée à 10^{-2}.

Aide au calcul : $198 = 6 \times 33$; $492 = 6 \times 82$; $\dfrac{33}{82} \approx 0{,}402$

Solution. Par définition, $p_{\mathrm{P}}(\mathrm{C}) = \dfrac{p(\mathrm{C} \cap \mathrm{P})}{p(\mathrm{P})} = \dfrac{0{,}0198}{0{,}0492} = \dfrac{198}{492} = \dfrac{6 \times 33}{6 \times 82} = \dfrac{33}{82} \approx 0{,}40$

La probabilité $p_{\mathrm{P}}(\mathrm{C})$ qu'une personne soit contaminée **sachant** que son test est positif est $\dfrac{33}{82}$, soit environ 0,40 (40 %).

Avec ce dernier exemple, la boucle est bouclée : la probabilité recherchée n'apparaît pas sur l'arbre (elle appartient à l'arbre inverse, celui dont le niveau 1 est constitué de P et de $\overline{\mathrm{P}}$) et il faut revenir à la définition de départ pour l'obtenir.