

Inférence bayésienne

Probabilités conditionnelles, arbres, formule de Bayes, indépendance

TERMINALE · MATHS COMPLÉMENTAIRES

1 Probabilité conditionnelle et arbre

Probabilité conditionnelle

Soient A et B deux événements avec $P(A) \neq 0$. La probabilité de B sachant A est

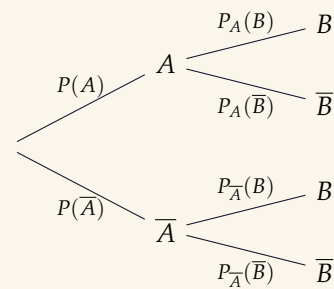
$$P_A(B) = \frac{P(A \cap B)}{P(A)} \quad \text{d'où} \quad P(A \cap B) = P(A) \times P_A(B).$$

$P_A(B)$ est la probabilité que B se réalise lorsqu'on sait que A est réalisé.

Règles de l'arbre pondéré

1. La somme des probabilités des branches issues d'un même nœud vaut 1.
2. La probabilité d'un chemin est le **produit** des probabilités de ses branches : $P(A \cap B) = P(A) \times P_A(B)$.
3. **Probabilités totales** : si A et $\bar{A}$ partitionnent l'univers,

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



Méthode — exploiter un arbre

Exemple : Dans une entreprise, 60 % des salariés travaillent à l'usine (U), les autres au siège. 30 % des salariés de l'usine et 50 % de ceux du siège viennent en transports en commun (T).

$$P(U \cap T) = 0,6 \times 0,3 = 0,18 \quad P(\bar{U} \cap T) = 0,4 \times 0,5 = 0,20 \quad P(T) = 0,18 + 0,20 = 0,38$$

2 Inversement du conditionnement : formule de Bayes

Formule de Bayes

Si $P(A) \neq 0$ et $P(B) \neq 0$:

$$P_B(A) = \frac{P(A \cap B)}{P(B)} = \frac{P(A) \times P_A(B)}{P(A) \times P_A(B) + P(\bar{A}) \times P_{\bar{A}}(B)}$$

L'arbre donne naturellement $P_A(B)$ (cause $\rightarrow$ effet) ; la formule de Bayes permet de remonter de l'effet observé B à la cause A : c'est l'**inférence bayésienne**.

Méthode — l'exemple des transports

Un salarié vient en transports en commun. Probabilité qu'il travaille à l'usine ?

$$P_T(U) = \frac{P(U \cap T)}{P(T)} = \frac{0,18}{0,38} \approx 0,474$$

Avant l'information, $P(U) = 0,6$ (probabilité **a priori**) ; après, $P_T(U) \approx 0,47$ (probabilité **a posteriori**) : savoir que le salarié prend les transports a fait baisser la probabilité qu'il soit à l'usine.

Vocabulaire des tests diagnostiques

Pour une maladie M et un test T (positif) :

- **prévalence** : $P(M)$, proportion de malades dans la population ;
- **sensibilité** : $P_M(T)$, probabilité qu'un malade ait un test positif ;
- **spécificité** : $P_{\bar{M}}(\bar{T})$, probabilité qu'un non-malade ait un test négatif ;
- **valeur prédictive positive (VPP)** : $P_T(M)$, probabilité d'être malade quand le test est positif.

Méthode — le paradoxe du test positif

Exemple : Prévalence 2 %, sensibilité 95 %, spécificité 90 %. Un patient est testé positif : est-il probablement malade ?

On représente la situation par un **tableau d'effectifs** sur 10 000 personnes :

	T (positif)	$\bar{T}$ (négatif)	Total
M (malade)	190	10	200
$\bar{M}$ (sain)	980	8 820	9 800
Total	1 170	8 830	10 000

$$P_T(M) = \frac{190}{1170} = \frac{0,02 \times 0,95}{0,02 \times 0,95 + 0,98 \times 0,10} \approx 0,162$$

Malgré un test « fiable », seulement 16 % des positifs sont malades ! Les faux positifs (980) sont bien plus nombreux que les vrais positifs (190), car la maladie est **rare**.

3 Rappels sur l'indépendance

Événements indépendants

Deux événements A et B sont **indépendants** si

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

ce qui équivaut, si $P(A) \neq 0$, à $P_A(B) = P(B)$: savoir que A est réalisé ne change pas la probabilité de B .

Propriété

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

Méthode — tester l'indépendance

Exemple : Dans l'exemple des transports, $P(U) \times P(T) = 0,6 \times 0,38 = 0,228 \neq 0,18 = P(U \cap T)$: les événements U et T ne sont pas indépendants.

Attention : « indépendants » ne veut pas dire « incompatibles ». Deux événements incompatibles ($A \cap B = \emptyset$) de probabilités non nulles ne sont *jamais* indépendants.

À retenir

- $P_A(B) = \frac{P(A \cap B)}{P(A)}$; dans un arbre, on **multiplie** le long d'un chemin.
- Probabilités totales : $P(B) = P(A)P_A(B) + P(\bar{A})P_{\bar{A}}(B)$.
- Bayes : $P_B(A) = \frac{P(A) \times P_A(B)}{P(B)}$ — on remonte de l'effet à la cause.
- Indépendance : $P(A \cap B) = P(A) \times P(B)$.
- **Piège :** $P_A(B) \neq P_B(A)$! La sensibilité d'un test n'est pas la probabilité d'être malade quand on est positif.