Le théorème de Bayes

Le théorème de Bayes est une conséquence immédiate des probabilités conditionnelles et des probabilités totales.

Probabilités conditionnelles

- **Exemple**

Dans une bibliothèque comportant 100 ouvrages, il y en a 40 qui sont écrits en anglais dont 8 portent sur la biologie. Considérons les événements suivants:

A = "le livre est écrit en anglais"; $P(A) = \frac{40}{100}$;

B = "le livre porte sur la biologie";

$A \cap B$ = "le livre est écrit en anglais **et** porte sur la biologie"; $P(A \cap B) = \frac{8}{100}$.

- **Probabilité conditionnelle**

$B \mid A$ = "le livre porte sur la biologie **sachant qu'** il est écrit en anglais"

$$P(B \mid A) = \frac{8}{40}$$

(Il s'agit de la fréquence des livres de biologie parmi les livres en langue anglaise.) On a les relations

$$P(B \mid A) = \frac{8}{40} = \frac{\left(\frac{8}{100}\right)}{\left(\frac{40}{100}\right)} = \frac{P(A \cap B)}{P(A)}$$

Retenons

$$\boxed{P(B \mid A) = \frac{P(A \cap B)}{P(A)}}$$

ou encore

$$\boxed{P(A \cap B) = P(A) \cdot P(B \mid A)}$$

Probabilités totales

Considérons une partition $A_1, A_2, \ldots, A_n$ de l'ensemble des événements E, c'est-à-dire P(E) = 1 et

$A_1 \cup A_2 \cup \ldots \cup A_n = E$, $A_i \cap A_j = \emptyset$ pour $i \neq j$. Alors

$$\mathbf{P(B) = P(A_1) \cdot P(B \mid A_1) + P(A_2) \cdot P(B \mid A_2) + \ldots + P(A_n) \cdot P(B \mid A_n)}$$

Démonstration

$$P(B) = P(B \cap E) = P(B \cap A_1) + P(B \cap A_2) + \ldots + P(B \cap A_n) =$$
$$P(A_1) \cdot P(B \mid A_1) + P(A_2) \cdot P(B \mid A_2) + \ldots + P(A_n) \cdot P(B \mid A_n) \quad \blacksquare$$

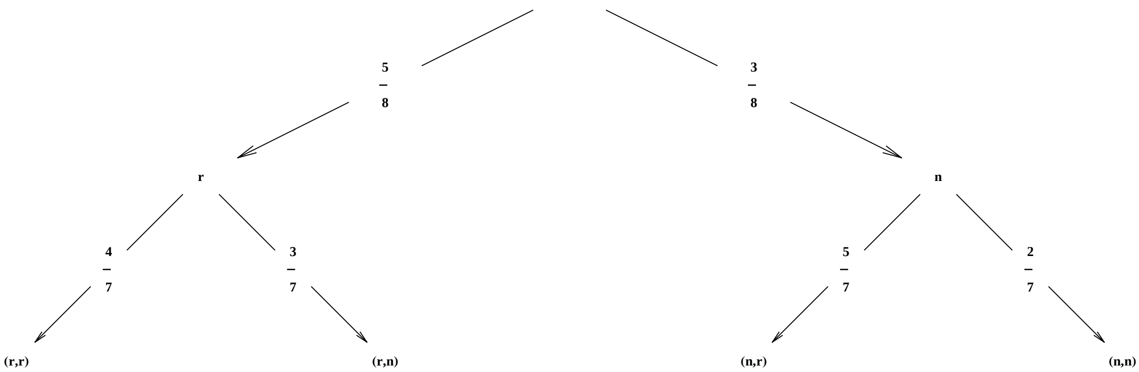
- **Problème**

Une urne contient 5 boules rouges identiques et 3 boules noires identiques. On effectue des tirages de deux boules sans remise.

a) Quelle est la probabilité que la première soit rouge et la deuxième noire ?

b) Quelle est la probabilité que l'une des deux boules au moins soit rouge ?

c) Sachant que l'une des deux boules au moins est rouge, quelle est la probabilité que l'autre soit noire ?



a) $P(r, n) = \frac{5}{8} \cdot \frac{3}{7} = \frac{15}{56}$

avec A = “la première boule est rouge”; B = “la deuxième est noire”; A $\cap$ B = “la première est rouge et la deuxième est noire”; B | A = “la deuxième est noire sachant que la première est rouge”. On retiendra que, dans un arbre, les branches portent des probabilités conditionnelles. La probabilité d’un chemin est est égale au produit des probabilités portées par les branches.

b) P (l' une des deux boules au moins est rouge) =
P (la première est rouge et la deuxième noire) +
P (la première est noire et la deuxième rouge) +
P (la première est rouge et la deuxième rouge) =
P (la première est rouge) · P (la deuxième est noire | la première est rouge) +
P (la première est noire) · P (la deuxième est rouge | la première est noire) + P (la première est rouge) · P (la deuxième est rouge | la première est rouge) $= \frac{5}{8} \cdot \frac{3}{7} + \frac{3}{8} \cdot \frac{5}{7} + \frac{5}{8} \cdot \frac{4}{7} = \frac{25}{28}$

c) P (l' autre est noire | l' une des deux boules au moins est rouge) =

$$\frac{P\,(\text{une boule est rouge et l' autre est noire})}{P\,(\text{l' une des deux boules au moins est rouge})}$$

$$= \frac{P(r, n) + P(n, r)}{P(r, n) + P(n, r) + P(r, r)} = \frac{\frac{5}{8}\,\frac{3}{7} + \frac{3}{8}\,\frac{5}{7}}{\frac{5}{8}\,\frac{3}{7} + \frac{3}{8}\,\frac{5}{7} + \frac{5}{8}\,\frac{4}{7}} = \frac{3}{5}$$

Préparation au théorème de Bayes

En intervertissant l'événement et la condition

$$\mathbf{P(A \mid B) = \frac{P(A) \cdot P(B \mid A)}{P(B)}}$$

La démonstration découle directement de la définition des probabilités conditionnelles

$$\frac{P(A) \cdot P(B \mid A)}{P(B)} = \frac{P(A) \cdot \frac{P(A \cap B)}{P(A)}}{P(B)} = \frac{P(A \cap B)}{P(B)} = P(A \mid B) \qquad \blacksquare$$

Formule de Bayes

Considérons une partition $A_1, A_2, \ldots, A_n$ de l'ensemble des événements E. Alors

$$\boxed{\begin{aligned} P(B) &= P(A_1)\cdot P(B\mid A_1) + P(A_2)\cdot P(B\mid A_2) + \ldots + P(A_n)\cdot P(B\mid A_n) \\ P(A_1\mid B) &= \frac{P(A_1)\cdot P(B\mid A_1)}{P(B)} \\ P(A_2\mid B) &= \frac{P(A_2)\cdot P(B\mid A_2)}{P(B)} \\ &\ldots \\ P(A_n\mid B) &= \frac{P(A_n)\cdot P(B\mid A_n)}{P(B)} \end{aligned}}$$

La démonstration a été faite au préalable sous "Probabilités totales" et "Préparation au théorème de Bayes".

- **L'apport d'une nouvelle information permet de corriger les probabilités à priori**

Les nombres suivants sont applelés "Probabilité à priori de A_k":

$$P(A_1),\ P(A_2),\ \ldots,\ P(A_n).$$

Les nombres suivants, appelés "fonction de vraisemblance de A_k" expriment des apports d'informations:

$$P(B\mid A_1),\ P(B\mid A_2),\ \ldots,\ P(B\mid A_n).$$

Les nombres suivants, applelés "Probabilité à postériori de A_k", expriment comment les probabilités à priori doivent être adaptées à la sous-population B:

$$P(A_1\mid B),\ P(A_2\mid B),\ \ldots,\ P(A_n\mid B).$$

- **Probabilités des causes**

Si les $A_1,\ A_2,\ \ldots,\ A_n$ expriments les causes possibles de B, on peut maintenant établir la cause la plus probable (éventuellement les causes les plus probables) A_k : c'est celle où $P(A_k\mid B)$ diffère le plus de $P(A_k)$, ce qui indique que les événements A_k et B ne sont pas indépendants.

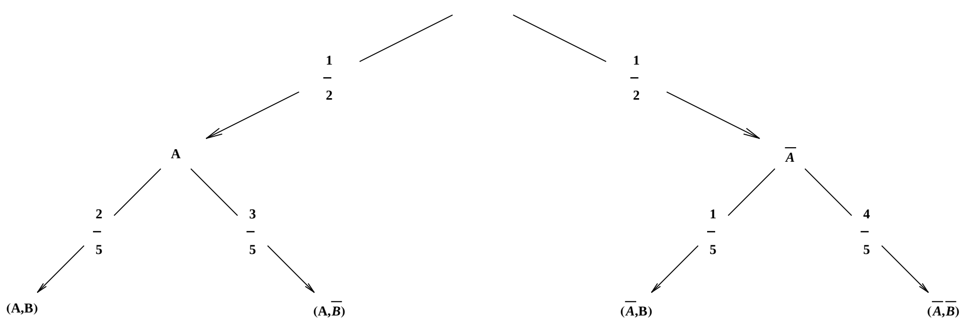
- **Exemple**

Dans un laboratoire, on a fait les constats suivants:

si une souris porte l'anticorps A, alors 2 fois sur 5 elle porte aussi l'anticorps B;

si une souris ne porte pas l'anticorps A, alors 4 fois sur 5 elle ne porte pas l'anticorps B.

La moitié de la population porte l'anticorps A.



- **a)**

Calculez la probabilité que, si une souris porte l'anticorps B, alors elle porte aussi l'anticorps A.

Pour la partition $\left(A,\ \overline{A}\right)$, la formule de Bayes donne

$$P(B) = P(A)\cdot P(B\mid A) + P(\overline{A})\cdot P\left(B\mid \overline{A}\right) = \frac{1}{2}\,\frac{2}{5} + \frac{1}{2}\,\frac{1}{5} = \frac{3}{10}$$

$$P(A \mid B) = \frac{P(A) \cdot P(B \mid A)}{P(B)} = \frac{\frac{1}{2}\,\frac{2}{5}}{\frac{3}{10}} = \frac{2}{3}$$

$$P(\overline{A} \mid B) = \frac{P(\overline{A}) \cdot P(B \mid \overline{A})}{P(B)} = \frac{\frac{1}{2}\,\frac{1}{5}}{\frac{3}{10}} = \frac{1}{3}$$

Interprétation:

probabilités à priori: $P(A) = \frac{1}{2}$, $P(\overline{A}) = \frac{1}{2}$;

apports d'informations: $P(B \mid A) = \frac{2}{5}$, $P(B \mid \overline{A}) = \frac{1}{5}$;

probabilités à postériori pour les porteuses de l'anticorps B: $P(A \mid B) = \frac{2}{3}$, $P(\overline{A} \mid B) = \frac{1}{3}$.

Probabilité des causes :

Les 2/3 des souris porteuses de l'anticorps B portent aussi l'anticorps A, ce qui dénote une nette incidence de A sur B.

- **b)**

Calculez la probabilité que, si une souris ne porte pas l'anticorps B, alors elle ne porte pas l'anticorps A.

Pour la partition $(A, \overline{A})$, la formule de Bayes donne

$$P(\overline{B}) = P(A) \cdot P(\overline{B} \mid A) + P(\overline{A}) \cdot P(\overline{B} \mid \overline{A}) = \frac{1}{2}\,\frac{3}{5} + \frac{1}{2}\,\frac{4}{5} = \frac{7}{10}$$

$$P(A \mid \overline{B}) = \frac{P(A) \cdot P(\overline{B} \mid A)}{P(\overline{B})} = \frac{\frac{1}{2}\,\frac{3}{5}}{\frac{7}{10}} = \frac{3}{7}$$

$$P(\overline{A} \mid \overline{B}) = \frac{P(\overline{A}) \cdot P(\overline{B} \mid \overline{A})}{P(\overline{B})} = \frac{\frac{1}{2}\,\frac{4}{5}}{\frac{7}{10}} = \frac{4}{7}$$

Interprétation:

probabilités à priori: $P(A) = \frac{1}{2}$, $P(\overline{A}) = \frac{1}{2}$;

apports d'informations: $P(\overline{B} \mid A) = \frac{3}{5}$, $P(\overline{B} \mid \overline{A}) = \frac{4}{5}$;

probabilités à postériori pour les non porteuses de l'anticorps B: $P(A \mid \overline{B}) = \frac{3}{7}$, $P(\overline{A} \mid \overline{B}) = \frac{4}{7}$.

Probabilité des causes :

Les 4/7 des souris qui ne portent pas l'anticorps B ne portent pas non plus l'anticorps A, ce qui dénote peut-être une légère incidence de $\overline{A}$ sur $\overline{B}$.

- **Lien hypertexte vers la page mère**

[http : // www.deleze.name/marcel/culture/probabilites/index.html](http://www.deleze.name/marcel/culture/probabilites/index.html)