Mathématiques complémentaires
Chapitre 03 - Inférence Bayésienne

Classe de terminale

I. Rappels

- Une **expérience aléatoire** est une expérience renouvelable dont le résultat ne peut être prévu et qui , renouvelée dans des conditions identiques, ne donne pas forcément le même résultat.
- L'ensemble de toutes les issues possibles d'une expérience aléatoire s'appelle l'**univers** noté Ω.
- Une **issue** est un résultat de l'expérience aléatoire.
- Un **événement** est un sous-ensemble de l'univers.
- Un **événement élémentaire** est un événement constitué d'une seule issue.
- L'**événement certain** est noté Ω.
- L'**événement impossible** est noté $\emptyset$.
- Soit $\mathbb{P}$ une **probabilité** alors on a :
  - $\forall A \in \Omega, 0 \leq \mathbb{P}(A) \leq 1$
  - $\mathbb{P}(\Omega) = 1$
  - $\mathbb{P}(\emptyset) = 0$
- Deux événements A et B sont **incompatibles** si et seulement si $\mathbb{P}(A \cap B) = 0$
- L'événement **contraire** d'un événement A est noté $\overline{A}$ et $\mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A)$
- $\forall A \in \Omega, \forall B \in \Omega, \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$

II. Probabilités conditionnelles

Définition

Soit Ω l'univers d'une expérience aléatoire. Soit A un événement tels que $\mathbb{P}(A) \neq 0$. Soit B un événement. La *probabilité conditionnelle que l'événement B se réalise sachant que l'événement A est réalisé* se note $\mathbb{P}_A(B)$ (« **Probabilité de B sachant A** ») et est définie par :

$$\mathbb{P}_A(B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)}$$

Exemple

Dans une classe, 60% des élèves sont des filles et 40% des élèves sont des filles internes. On choisit un élève au hasard. Quelle est la probabilité qu'il soit interne sachant que c'est une fille ?

Réponse

Posons les événements F : « L'élève est une fille » et I : « l'élève est interne ». On a, d'après l'énoncé :

- $\mathbb{P}(F) = \frac{60}{100} = 0,6$
- $\mathbb{P}(F \cap I) = \frac{40}{100} = 0,4$

Donc $\mathbb{P}_F(I) = \frac{\mathbb{P}(F\cap I)}{\mathbb{P}(F)} = \frac{0{,}4}{0{,}6} = \frac{2}{3}$

Exemple

On lance deux dés parfaitement équilibrés à la suite et on ajoute les résultats. Soit les événements suivants :

- A : « la somme des deux dés est supérieure à 9 ».
- B : « Le premier dé donne un 5 ».

1. Calculer la probabilité de l'événement A.
2. Calculer $\mathbb{P}(B)$.
3. Quelle est la probabilité d'obtenir une somme supérieure à 9, sachant qu'on a obtenu un 5 au premier dé ?

Réponse

1. On peut modéliser cette expérience dans un tableau :

	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

   Ainsi, $\mathbb{P}(A) = \frac{10}{36} = \frac{5}{18}$

2. $\mathbb{P}(B) = \frac{1}{6}$
3. $\mathbb{P}_B(A) = \frac{\mathbb{P}(B\cap A)}{\mathbb{P}(B)} = \frac{\frac{3}{36}}{\frac{1}{6}} = \frac{3}{36} \times \frac{6}{1} = \frac{3}{6} = \frac{1}{2}$

Propriété

Soient A et B deux événements de l'univers Ω tels que $\mathbb{P}(A) \neq 0$. On a :

$$0 \leq \mathbb{P}_A(B) \leq 1$$

$$\mathbb{P}_A(B) + \mathbb{P}_A(\overline{B}) = 1$$

Exemple

Soit A et B deux événements tels que $\mathbb{P}_A(B) = 0,6$.
Calculer $\mathbb{P}_A\left(\overline{B}\right)$

Réponse

$\mathbb{P}_A\left(\overline{B}\right) = 1 - \mathbb{P}_A(B) = 1 - 0,6 = 0,4$

Propriété

Soient A et B deux événements de l'univers Ω tels que $\mathbb{P}(A) \neq 0$ et $\mathbb{P}(B) \neq 0$.
On a :

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \times \mathbb{P}_A(B) = \mathbb{P}(B) \times \mathbb{P}_B(A)$$

Exemple

Soit A et B deux événements tels que $\mathbb{P}_A(B) = 0,6$, $\mathbb{P}(A) = 0,3$ et $\mathbb{P}(B) = 0,8$.

1. Calculer $\mathbb{P}(A \cap B)$
2. En déduire $\mathbb{P}_B(A)$

Réponse

1. $\mathbb{P}(A \cap B) = \mathbb{P}(A) \times \mathbb{P}_A(B) = 0,3 \times 0,6 = 0,18$
2. $\mathbb{P}_B(A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{0,18}{0,8} = \frac{9}{40} = 0,225$

Formule de Bayes

Soient A et B deux événements tels que $\mathbb{P}(B) \neq 0$.
On a :

$$\mathbb{P}_B(A) = \frac{\mathbb{P}_A(B) \times \mathbb{P}(A)}{\mathbb{P}(B)}$$

Exemple

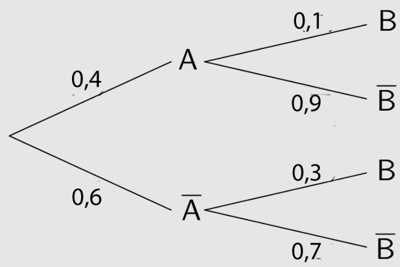
Soit deux événements A et B tels que :

- La probabilité de l'événement A est $0,4$.
- La probabilité de l'événement B sachant l'événement A est $0,1$.
- La probabilité de l'événement B sachant l'événement $\overline{A}$ est $0,3$.

1. Déssiner un arbre de probabilités.
2. Calculer la probabilité de l'événement B.
3. Calculer la probabilité de l'événement A sachant B. Vous arrondirez le résultat à 10^{-3}.

Réponse

1. arbre de probabilité :



2. $\mathbb{P}(B) = \mathbb{P}(A) \times \mathbb{P}_A(B) + \mathbb{P}(\overline{A}) \times \mathbb{P}_{\overline{A}}(B) = 0,4 \times 0,1 + 0,6 \times 0,7 = 0,04 + 0,42 = 0,46$
3. $\mathbb{P}_B(A) = \frac{\mathbb{P}_A(B) \times \mathbb{P}(A)}{\mathbb{P}(B)} = \frac{0,1 \times 0,4}{0,46} = \frac{0,04}{0,46} \approx 0,087$

Les tableaux

On peut représenter les probabilités de certaines expériences aléatoires grâce à un tableau à double entrée.

	B	$\overline{B}$	Total
A	$\mathbb{P}(A \cap B)$	$\mathbb{P}(A \cap \overline{B})$	$\mathbb{P}(A)$
$\overline{A}$	$\mathbb{P}(\overline{A} \cap B)$	$\mathbb{P}(\overline{A} \cap \overline{B})$	$\mathbb{P}(\overline{A})$
Total	$\mathbb{P}(B)$	$\mathbb{P}(\overline{B})$	1

On peut alors calculer facilement les probabilités conditionnelles.

Exemple

Dans un club multisport, on a :

- 60% des adhérents sont des garçons.
- 65% des adhérents pratiquent le football.
- 20% des adhérents sont des garçons qui ne pratiquent pas le football.

Soit les événements suivants :

- G : « l'adhérent est un garçon ».
- F : « l'adhérent pratique du football ».

On choisit un adhérent au hasard.

1. Modéliser l'expérience par un tableau à double entrée.
2. Calculer $\mathbb{P}_G(F)$.
3. Calculer la probabilité que l'adhérent soit soit un garçon sachant qu'il ne joue pas au football.

Réponse

1. Le tableau à double entrée suivant modélise l'expérience.

	F	$\overline{F}$	Total
G	0,4	0,2	0,6
$\overline{G}$	0,25	0,15	0,4
Total	0,65	0,35	1

2. $\mathbb{P}_G(F) = \frac{0{,}4}{0{,}6} = \frac{2}{3}$
3. $\mathbb{P}_{\overline{F}}(G) = \frac{0{,}2}{0{,}35} = \frac{4}{7}$

Définition

Deux événements A et B d'un univers Ω sont **indépendants** si et seulement si $\mathbb{P}(A \cap B) = \mathbb{P}(A) \times \mathbb{P}(B)$.

Exemple

Dans l'exemple précédent, les événements F et G sont-ils indépendants ?

Réponse

$\mathbb{P}(F \cap G) = 0,4$ $\mathbb{P}(F) \times \mathbb{P}(G) = 0,65 \times 0,6 = 0,39$ Or $0,39 \neq 0,4$ donc les événements F et G ne sont pas indépendants.

Propriété

Si deux événements A et B d'un univers Ω tels que $\mathbb{P}(A) \neq 0$ sont indépendants alors $\mathbb{P}_A(B) = \mathbb{P}(B)$

Propriété

Si A et B sont deux événements indépendants alors $\overline{A}$ et B le sont aussi.

III. Partition de l'univers

Définition

Une famille d'événements $(A_i)_{1 \leq i \leq n}$ est une **partition** (ou un **système complet d'événements**) de l'univers Ω si elle vérifie :

$$\forall i \in [\![1;n]\!], A_i \neq \emptyset$$

$$\bigcup_{i=1}^{n} A_i = \Omega$$

$$\forall i \in [\![1;n]\!], \forall j \in [\![1;n]\!], i \neq j \Rightarrow A_i \cap A_j = \emptyset$$

Formule des probabilités totales

Soit Ω l'univers d'une expérience aléatoire. Soit B un événement de Ω. Soit $(A_i)_{1 \leq i \leq n}$ une partition de Ω. On a :

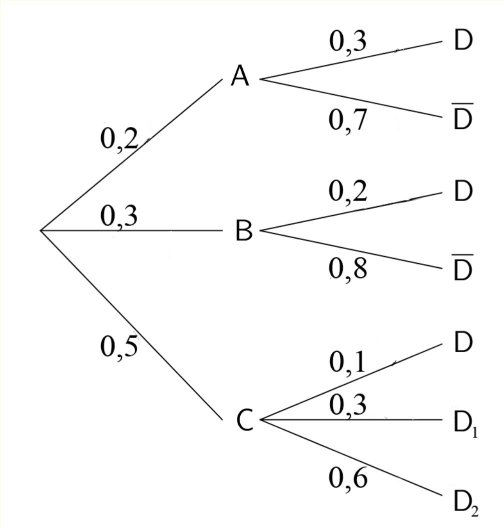
$$\mathbb{P}(B) = \sum_{i=1}^{n} \mathbb{P}(A_i \cap B) = \mathbb{P}(A_1 \cap B) + \mathbb{P}(A_2 \cap B) + ... + \mathbb{P}(A_n \cap B)$$

ou encore :

$$\mathbb{P}(B) = \sum_{i=1}^{n} \mathbb{P}(A_i) \times \mathbb{P}_{A_i}(B) = \mathbb{P}(A_1) \times \mathbb{P}_{A_1}(B) + \mathbb{P}(A_2) \times \mathbb{P}_{A_2}(B) + ... + \mathbb{P}(A_n) \times \mathbb{P}_{A_n}(B)$$

Exemple

On donne l'arbre pondéré ci-dessous :



Calculer $\mathbb{P}(D)$.

Réponse

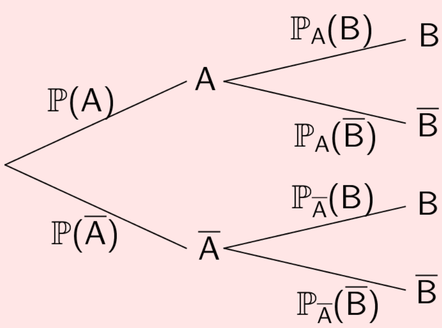
$\mathbb{P}(D) = \mathbb{P}(A) \times \mathbb{P}_A(D) + \mathbb{P}(B) \times \mathbb{P}_B(D) + \mathbb{P}(C) \times \mathbb{P}_C(D) = 0,2 \times 0,3 + 0,3 \times 0,2 + 0,5 \times 0,1 = 0,06 + 0,06 + 0,05 = 0,17$

Remarque

Dans le cas d'une famille de deux événements, A_1 et A_2, la propriété précédente devient :

$$\mathbb{P}(B) = \mathbb{P}(A_1) \times \mathbb{P}_{A_1}(B) + \mathbb{P}(A_2) \times \mathbb{P}_{A_2}(B)$$

et l'arbre :



IV. Exemples de problèmes sur l'inférence Bayésienne

Le raisonnement bayésien est à la base de nombreux algorithmes de décision et se retrouve dans de nombreux domaines pratiques : sport, médecine, justice, etc. où l'on doit raisonner à partir de probabilités et d'informations incomplètes.

IV.1. Tests binaires pour le diagnostic médical

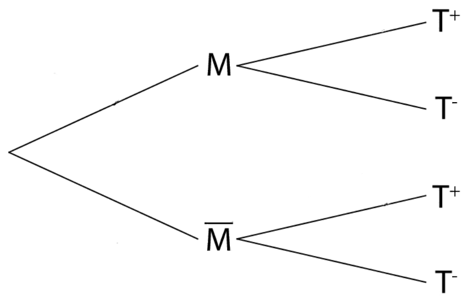
On se propose de mettre en place et d'illustrer le vocabulaire afférent aux tests binaires de diagnostic médical. Dans un second temps, on analysera l'utilité du dépistage systématique d'une maladie rare.
Partie A : le vocabulaire des tests médicaux On étudie un test biologique utilisé pour le diagnostic du cancer du côlon .

- Dans la population française $0,5\%$ des personnes sont atteintes par cette pathologie (prévalence).
- Le fabricant indique que :
  - le test est positif dans 60% des cas chez les sujets atteints du cancer du côlon.
  - le test est positif dans 2% des cas chez les sujets indemnes.

Pour un individu de la population, on note les événements :

- M : « L'individu a le cancer du côlon ».
- T^+ (resp. T^-) : « L'individu a un test positif (resp. négatif) ».

1. (a) Réaliser l'arbe ci-dessous et le compléter avec les probabilités qui conviennent.



   (b) La sensibilité S_e d'un test est $\mathbb{P}_M(T^+)$ et la spécificité S_p d'un test est $\mathbb{P}_{\overline{M}}(T^-)$. Quelles sont les valeurs de S_e et S_p pour ce test biologique ?

2. (a) La valeur prédictive négative (VPN) d'un test est la probabilité de ne pas avoir la maladie sachant que le test est négatif. Calculer la VPN pour ce test biologique. Arrondir au millième.
   (b) La valeur prédictive positive (VPP) d'un test est la probabilité d'avoir la maladie sachant que le test est positif. Calculer la VPP pour ce test biologique.

Partie B : VPP et maladie rare On dispose d'un test pour dépister une maladie rare. Le fabricant du test fournit les caractéristiques suivantes :

- la probabilité qu'un individu malade ait un test positif est de $0,99$.
- la probabilité qu'un individu non malade ait un test négatif est de $0,99$.

On envisage un dépistage systématique sur une population dans laquelle on estime à f (avec $0 \leq f \leq 1$) la proportion de gens malades.

1. Expliquer pourquoi $VPP = \frac{99f}{98f+1}$.
2. (a) Reproduire et compléter le tableau ci-dessous. Arrondir au millième.

f	0,001	0,01	0,1	0,3	0,5	0,8
VPP						

   (b) La VPP varie donc fortement selon la population cible.
   Quel inconvénient majeur présente, dans une population, le dépistage d'une maladie rare ?

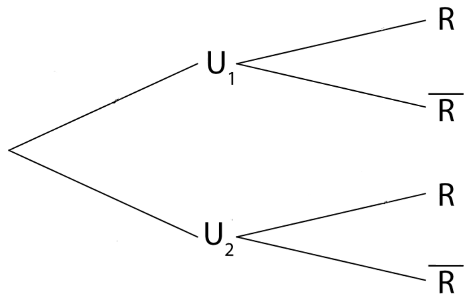
IV.2. De quelle urne vient la boule ?

Partie A : la formule de Bayes

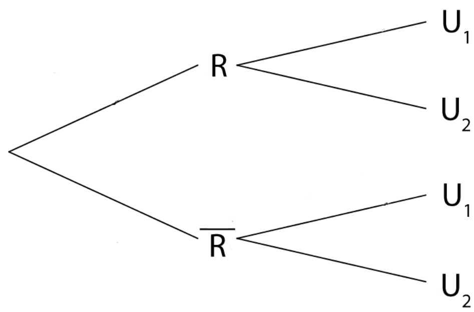
Nous possédons deux urnes : l'urne 1 a 6 boules bleues et 4 boules rouges, et l'urne 2 a 3 boules bleues et 9 boules rouges.

1. **Avec un bandeau**
   Une personne qui a les yeux bandés tire une boule au hasard dans l'une des deux urnes.
   (a) Peut-elle dire de quelle urne provient la boule tirée ?
   (b) On note U_1 : « La boule vient de l'urne 1 » et U_2 : « Labou le vient de l'urne 2 ».
   Calculer $\mathbb{P}(U_1)$ et $\mathbb{P}(U_2)$.

2. **On retire le bandeau**
   On retire le bandeau à cette personne. Elle voit maintenant la couleur de la boule qu'elle a tirée, cette boule est rouge. On note R l'événement : « La boule tirée est rouge ».
   (a) Réaliser l'arbre ci-dessous en complétant par les probabilités qui conviennent.



   (b) Vérifier que $\mathbb{P}(R) = 0,575$.
   (c) Expliquer pourquoi $\mathbb{P}_R(U_1) = \frac{\mathbb{P}(U_1)\times\mathbb{P}_{U_1}(R)}{\mathbb{P}(R)}$.
   Calculer alors $\mathbb{P}_R(U_1)$. Arrondir au centième.
   (d) Quelle influence a eu l'effet d'avoir retiré le bandeau à la personne sur la probabilité que la boule tirée provienne de l'urne 1 ?
   (e) Calculer alors $\mathbb{P}_R(U_2)$. Interpréter cette probabilité.
   (f) Réaliser l'arbre pondéré ci-dessous en le complétant.



Partie B : application de la formule de Bayes

En France, la population de daltoniens est d'environ 8% chez les hommes et de $0,4\%$ chez les femmes. On estime qu 'il y a environ 52% de femmes en France. Quelle est la probabilité pour qu'une personne daltonienne soit une femme ? Arrondir au millième.