

Probabilités – Fiche de cours

1. Expérience aléatoire

Une expérience aléatoire est une expérience dont on ne peut pas prévoir le résultat. Elle dépend uniquement du hasard.

L'univers Ω est constitué par toutes les issues possibles.

2. Événement

On appelle événement une condition qui peut être réalisée lors d'une expérience aléatoire

- événement élémentaire : événement réalisé par une seule issue possible
- événement impossible : événement qui ne peut pas être réalisé
- événement certain : événement qui est toujours réalisé
- événements incompatibles : 2 événements qui ne peuvent pas se produire en même temps
- événement contraire : événement qui se réalise lorsqu'un premier événement n'est pas réalisé

3. Notion de probabilité

La probabilité p est interprétée comme la fréquence f obtenue pour un grand nombre de réalisations d'un événement.

$$p(A) = \frac{\text{nombre de cas favorables à A}}{\text{nombre total de cas}}$$

- équiprobabilité

Lorsque les n issues d'une expérience aléatoire ont la même probabilité, on parle d'équiprobabilité :

$$p = \frac{1}{n}$$

4. Propriétés

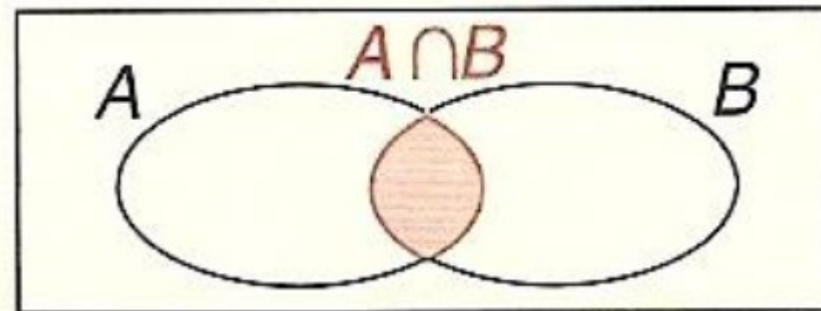
- une probabilité p est un nombre compris entre 0 et 1 : $0 \leq p \leq 1$
- la somme des probabilités de toutes les issues est égale à 1 :

$$p_1 + p_2 + \dots + p_n = 1$$

- $P(\Omega) = 1$; $P(\emptyset) = 0$; $P(\bar{A}) = 1 - P(A)$

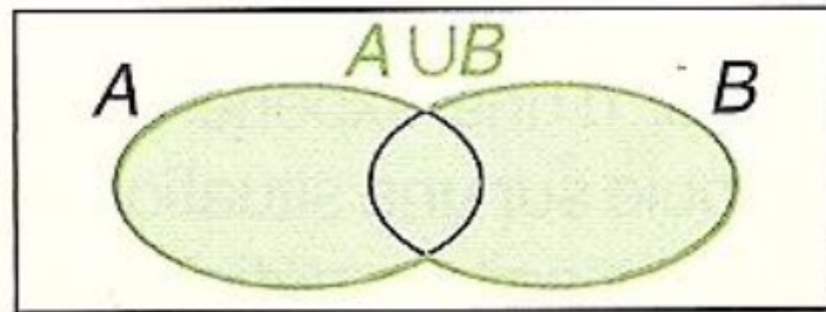
5. Opérations sur les événements

a. Intersection d'événements



La probabilité de l'intersection de 2 événements A et B est notée :
 $P(A \cap B)$

b. Union d'événements



La probabilité de l'union de 2 événements A et B est notée :
 $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

6. Conditionnement

a. Probabilités conditionnelles

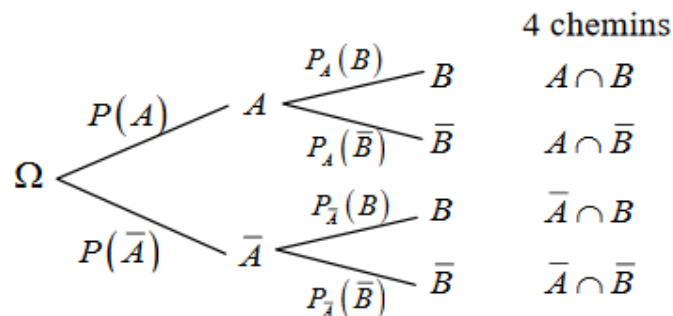
Définition

Soient A et B deux événements d'un univers Ω

On appelle probabilité conditionnelle de l'événement A sachant B le réel :

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

b. Arbre pondéré



Les branches sont pondérées par des probabilités conditionnelles

La somme des probabilités issues d'un même sommet (ou partition) est égale à 1

c. Inférence bayésienne

L'inférence bayésienne est un inversement du conditionnement

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

7. Indépendance

Définition :

Soit A et B deux événements de probabilité non nulle

A et B sont indépendants lorsque la réalisation de l'un ne change pas la réalisation de l'autre :

Deux événements A et B sont indépendants ssi :

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

Lorsque A et B sont deux événements indépendants, alors :

- $\bar{A}$ et B sont indépendants
- A et $\bar{B}$ sont indépendants
- $\bar{A}$ et $\bar{B}$ sont indépendants

8. Dénombrement

a. Factorielle d'un entier naturel (nombre de permutations distinctes)

Soit n un entier naturel non nul ; on appelle factorielle n :

$$n! = 1 \times 2 \times \dots \times n \quad \text{avec } 0! = 1$$

Une permutation d'un ensemble E à n éléments est un n -uplet distinct

Le nombre de permutations de E est $n! = 1 \times 2 \times \dots \times n$

b. Combinaisons distinctes avec ordre

Le nombre d'arrangements ou k -uplets distincts est :

$$\forall n \in \mathbb{N}^* \quad \forall k \in \mathbb{N} \quad 0 \leq k \leq n$$

$$A_n^k = n \times (n-1) \times \dots \times (n-(k-1)) = \frac{n!}{(n-k)!}$$

Le nombre de combinaisons distinctes avec ordre de k éléments parmi n s'appelle nombre d'arrangements de k parmi n

c. Combinaisons distinctes sans ordre

Le nombre de combinaisons à k éléments de E :

$$\forall n \in \mathbb{N}^* \quad \forall k \in \mathbb{N} \quad 0 \leq k \leq n$$

$$\binom{n}{k} = C_n^k = \frac{A_n^k}{k!} = \frac{n!}{k!(n-k)!}$$

Le nombre de combinaisons distinctes sans ordre de k éléments parmi n s'appelle coefficient binomial de k parmi n

9. Les tests diagnostiques

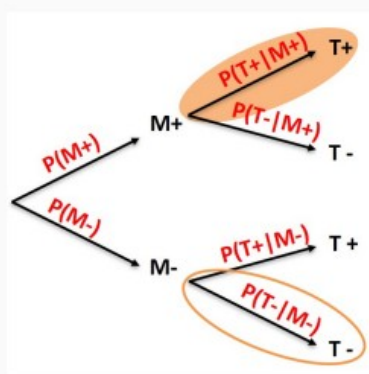
On cherche à savoir si un patient est atteint d'une maladie M, pour cela on va utiliser un test T.

a. Les types de tests

- test qualitatif (généralement binaire)
- test quantitatif (implique le choix d'un seuil)

b. Test qualitatif

	M+	M-	
T+	VP	FP	Vrais Positifs (VP) = $T+ \cap M+$ Faux Positifs (FP) = $T+ \cap M-$
T-	FN	VN	Vrais Négatifs (VN) = $T- \cap M-$ Faux Négatifs (FN) = $T- \cap M+$



Les performances du test sont évaluées :

- ultérieurement
- par un test de référence

c. test quantitatif

Le seuil est défini comme la valeur du résultat (variable biologique) au-delà de laquelle le résultat du test quantitatif est considéré positif

10. Valeurs informatives des tests

a. Prévalence

La prévalence (ou taux de prévalence) est le nombre de cas qui a contracté une maladie (ou pourcentage).

b. Sensibilité

La sensibilité est la capacité du test à donner un résultat positif, lorsque la présence de la maladie est vérifiée

$$Se = P_{M^+}(T^+) = \frac{VP}{VP + FN}$$

c. Spécificité

La spécificité est la capacité du test à donner un résultat négatif, lorsque la présence de la maladie n'est pas vérifiée

$$Spe = P_{M^-}(T^-) = \frac{VN}{VN + FP}$$

d. Valeur prédictive positive

La valeur prédictive positive est définie comme la probabilité que la maladie soit présente lorsque le test est positif

$$VPP = P_{T^+}(M^+) = \frac{VP}{VP + FP}$$

e. Valeur prédictive négative

La valeur prédictive négative est définie comme la probabilité que la maladie ne soit pas présente lorsque le test est négatif

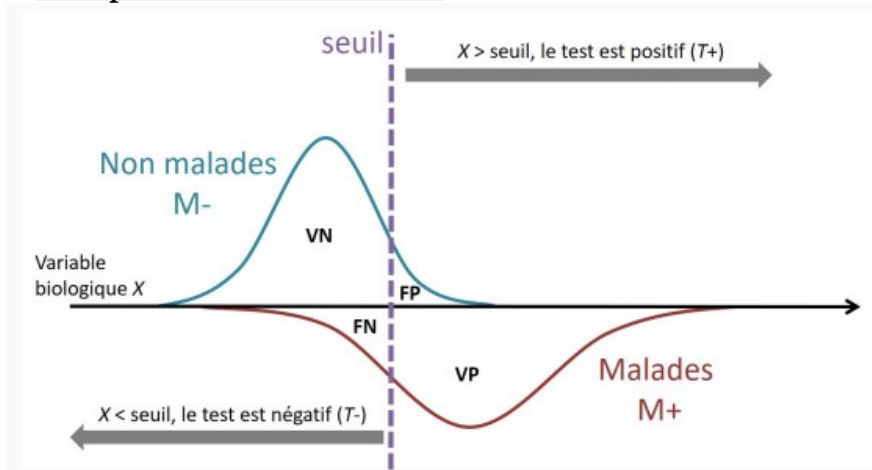
$$VPN = P_{T^-}(M^-) = \frac{VN}{VN + FN}$$

f. Lien de la prévalence et VPP /VPN

Si $p \nearrow$ alors $VPP \nearrow$ et $VPN \searrow$

11. Seuil d'un test et courbe ROC

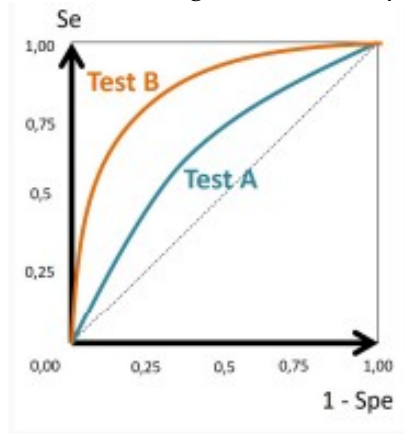
a. Exemple de définition de seuil



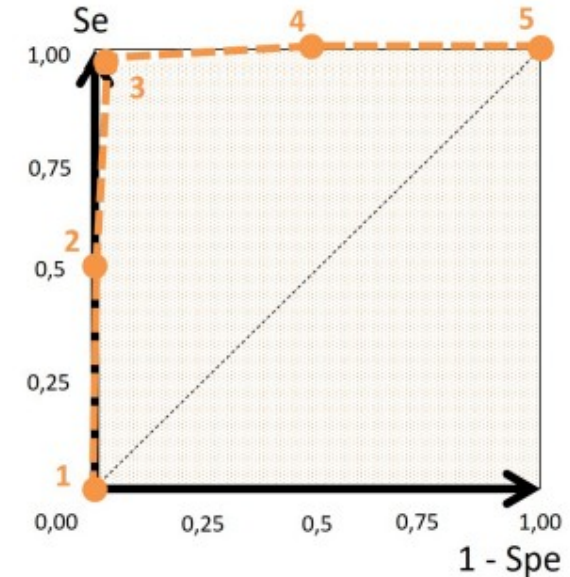
b. Courbe ROC

Une courbe ROC (Receiver Operating Curve) est obtenue de la manière suivante :

- on définit une valeur du seuil
- on calcule les valeurs du couple $(Se; Spe)$
- on place les résultats sur le diagramme $Se = f(1 - Spe)$



c. Résultats d'une courbe ROC et choix du seuil



- Lorsque la courbe ROC est représentée par la diagonale, le test n'a pas d'intérêt
- Plus la courbe ROC se rapproche du coin supérieur gauche, plus le test est discriminant
- Lorsque l'on veut limiter les faux positifs, on privilégie une bonne spécificité
- Lorsque l'on veut limiter les faux négatifs, on privilégie une bonne sensibilité