

Document 1 : cotes

La **cote** d'un événement (odds en anglais) est le ratio de la probabilité que l'événement se produise par la probabilité qu'il ne se produise pas. On l'exprime en général comme une paire de nombres (le numérateur et le dénominateur).

Par exemple, si un événement a une probabilité de 5% d'advenir, il a donc aussi une probabilité de 95% de ne pas se produire et sa cote est alors de 5 contre 95 (ou 1 contre 19 qu'on peut aussi noter 1:19).

Document 2 : théorème de Bayes exprimé en terme de cotes

La cote de H sachant I vaut la cote a priori de H multipliée par le **facteur de Bayes** $\left(\frac{P(I|H)}{P(I|\bar{H})}\right)$.

Document 3 : expérience

Le psychologue Gerd Gigerenzer présente le problème suivant dans un séminaire de statistique à des gynécologues en activité :

« Une femme de 50 ans sans symptôme passe une mammographie de routine. L'examen se révèle positif. Alarmée, elle veut savoir avec quelle certitude cela implique qu'elle a un cancer du sein.

À part le résultat du test, vous ne savez rien sur cette femme.

La prévalence des cancers du sein est de 1% chez les femmes de cet âge.

La sensibilité du test est de 90%.

Et sa spécificité est de 91%.

Parmi les femmes dont le test est positif, combien sont atteintes d'un cancer du sein ?

A : 9 sur 10 ; B : 8 sur 10 ; C : 1 sur 10 ; D : 1 sur 100 »

Document 4 : vocabulaire

- La **sensibilité** d'un test mesure sa capacité à donner un résultat positif lorsqu'une hypothèse est vérifiée = capacité à détecter un maximum de malades (avoir le moins possible de faux négatifs).
- La **spécificité** d'un test mesure sa capacité à donner un résultat négatif lorsque l'hypothèse n'est pas vérifiée = capacité à ne détecter que les malades (avoir le moins possible de faux positifs).

En notant VP et FP les vrais et les faux positifs, et VN et FN les vrais et faux négatifs, on a :

	Malade	Non malade
Test positif	VP	FP
Test négatif	FN	VN

$$\text{sensibilité} = \frac{VP}{VP + FN}$$

$$\text{spécificité} = \frac{VN}{VN + FP}$$

En bon bayésien, il ne faut pas considérer qu'un test détermine si on a une maladie, ni même qu'il détermine les chances d'avoir une maladie. Tout ce qu'il fait, c'est **mettre à jour** les chances d'avoir une maladie !

Que vaut le **facteur de Bayes** dans le cas du Doc. 3 ?

$$\frac{P(+ | \text{Cancer})}{P(+ | \overline{\text{Cancer}})} = \frac{\text{probabilité de vrais positifs}}{\text{probabilité de faux positifs}} = \frac{\text{sensibilité}}{1-\text{spécificité}}$$

Le théorème de Bayes version cote devient donc :

$$cote(\text{cancer} | +) = \frac{\text{sensibilité}}{1-\text{spécificité}} \times cote(\text{cancer})$$

1. Supposons qu'un labo pharmaceutique vous demande un test spécifique à 100% sans vous parler de sensibilité. Comment pourriez-vous mettre au point un tel test très facilement ?
2. Et un test sensible à 100% sans contrainte sur la spécificité ?
3. Si un pari consiste à obtenir un 5 ou un 6 au dé, que vaut alors sa cote ? À quelle cote correspond une probabilité de 50% ?
4. Comment passer de la prévalence évoquée dans le doc. 3 à $cote(\text{cancer})$?
5. Quelle est alors la réponse à la question du Doc. 3 ?

Plus de la moitié des docteurs présents ont choisi la réponse A, ce qui est très à côté de la plaque, et seulement 1 sur 5 ont choisi la bonne réponse...

6. Comment est mise à jour la cote et par suite la probabilité d'avoir un cancer si on repasse le test et qu'il est à nouveau positif ?

Augmentons la spécificité à 99%.

7. Que devient la probabilité d'avoir un cancer en cas de test positif ?

Et si le test est négatif ?

8. Reprendre les données de départ et trouvez la probabilité de ne pas avoir de cancer si on a été testé négatif.
9. Et si on passe un second test négatif ?