

Inférence bayésienne

Exercice 1. Le virus de l’immunodéficience humaine (VIH) est responsable du SIDA. Il existe aujourd’hui des tests rapides appelés « Tests rapides d’orientation diagnostique » (TROD) qui ont l’avantage de pouvoir être réalisés à partir d’un échantillon de salive ou d’une goutte de sang prélevée sur le bout du doigt.

1. On a testé avec les deux types de TROD (salivaire et sanguin) deux populations : une première composée de 10 000 personnes infectées par le VIH et une seconde composée de 100 000 personnes non infectées. On a obtenu les résultats suivants.

	Personnes infectées	Personnes non infectées
Tests salivaires positifs	9 803	260
Tests sanguins positifs	9 968	90

Calculer la sensibilité et la spécificité de chacun des deux tests.

2. **a.** En 2017, la population mondiale exposée au VIH était estimée à 6 milliards et, dans cette population, le nombre de personnes infectées par le VIH était estimé à 37 millions.

Calculer la valeur prédictive positive de chacun des deux tests pour la population mondiale exposée.

- b.** En 2017, la population française exposée était estimée à 50 millions et, dans cette population, le nombre de personnes infectées par le VIH était estimé à 150 000.

Calculer la valeur prédictive positive de chacun des deux tests pour la population française exposée.

- c.** En 2017, la population sud-africaine exposée était estimée à 35 millions et, dans cette population, le nombre de personnes infectées par le VIH était estimé à 7 millions.

Calculer la valeur prédictive positive de chacun des deux tests pour la population sud-africaine exposée.

- d.** Que mettent en évidence les trois exemples précédents ?

Exercice 2. Parmi les femmes de 40 ans ayant effectué une mammographie, 1% a un cancer du sein. À la suite de mammographies sur un échantillon, on a établi que :

- pour 82% des femmes ayant un cancer du sein, la mammographie détecte une anomalie ;
- pour 9% des femmes n’ayant pas de cancer du sein, la mammographie détecte une anomalie.

On suppose que 10 000 femmes de 40 ans ont effectué une mammographie.

1. Déterminer la sensibilité et la spécificité d’une mammographie.
2. Compléter le tableau suivant.

	Anomalie détectée	Pas d’anomalie détectée	Total
Personnes malades			
Personnes non malades			
Total			10 000

3. Une femme de 40 ans a effectué une mammographie qui a permis de détecter une anomalie. Quelle est la probabilité qu'elle soit atteinte d'un cancer du sein ?
4. Calculer les valeurs prédictives positive et négative d'une mammographie chez les femmes de 40 ans.

Exercice 3. On suppose que la somme de la sensibilité et de la spécificité d'un test est égale à 1 (c'est-à-dire, avec les notations du cours, $Se + Sp = 1$).

Expliquer pourquoi le test est inutile dans ce cas.

Les deux exercices suivants sont destinés aux élèves suivant la spécialité mathématiques ou l'option mathématiques complémentaires.

Exercice 4. Trois maladies virales peuvent être transmises par les moustiques : dengue, chikungunya et zika. Elles provoquent des symptômes qui peuvent être assez proches et il est donc parfois difficile de les différencier directement. On s'intéresse à la mise en place d'une aide statistique au diagnostic. Pour cela, on va s'appuyer sur des données obtenues chez des personnes dont le diagnostic a pu être certifié par des examens biologiques. Pour simplifier, on supposera que ces caractères apparaissent indépendamment chez les personnes infectées.

Symptômes	Dengue	Chikungunya	Zika
Fièvre	95%	75%	75%
Courbatures	75%	95%	50%
Douleur oculaire	50%	25%	50%
Déficit en globules blancs	50%	50%	25%
Hémorragies	25%	5%	5%

À partir de ces données, on veut déterminer les probabilités de chaque maladie selon les symptômes présentés et dans des conditions différentes.

1. On suppose qu'une personne malade revient d'un pays dans lequel aucune de ces maladies n'est épidémique. On considère donc a priori que les trois maladies sont équiprobables.
 - a. Quelles sont les probabilités de chaque maladie si cette personne présente à la fois de la fièvre, pas de courbatures et des douleurs oculaires ?
 - b. Quel est le diagnostic le plus probable dans ce cas ?
2. On suppose qu'une personne malade revient d'un pays dans lequel sévit une épidémie de Zika. On suppose qu'il y a 80% de chances qu'elle ait été infectée par Zika et 10% par chacune des deux autres maladies.
 - a. Quelles sont les probabilités de chaque maladie si cette personne présente à la fois de la fièvre, pas de courbatures et des douleurs oculaires ?
 - b. Quel est le diagnostic le plus probable dans ce cas ?

Exercice 5. À l'aide de la formule des probabilités totales, démontrer la formule de Bayes.