

Exercices de mathématiques sur la fiabilité de tests d'hypothèses

UNIQUEMENT LES ELEVES AYANT CHOISI DE POURSUIVRE LA SPECIALITE MATHÉMATIQUES EN TERMINALE

Ce document propose, dans des situations concrètes, des exemples d'applications de notions figurant aux programmes de mathématiques des classes de premières et terminales, générale ou technologiques. Le fil conducteur est le test d'hypothèse.

La pandémie Covid-19 a mis en lumière les **tests d'hypothèses** (par exemple les tests sérologiques pour dépister les anticorps). Leurs caractéristiques sont : **sensibilité, spécificité, valeurs prédictives positives et négatives** (VPP et VPN). Les deux dernières, VPP et VPN, sont fonctions homographiques de la **prévalence** (proportion de personnes atteintes par une maladie dans une population). De nombreuses activités abordant ces sujets peuvent être proposées dans le cadre de la classe de mathématiques de lycée général et technologique.

Sont mobilisés dans les 10 exercices proposés :

- *Calculs de probabilités simples et conditionnelles, arbre de probabilité*
- *Proportions, fréquences, tri croisé*
- *Fonctions affines, homographiques*
- *Calcul littéral*
- *Inéquation du premier degré de la forme $(ax+b) / (cx+d) \leq \text{Constante}$*
- *Exploiter un graphique, formuler des conjectures, interprétations,...*

(Exercices guidés, avec prise d'initiatives,...)

Liens utiles , sources :

- Sensibilité et spécificité sur Wikipédia : https://fr.wikipedia.org/wiki/Sensibilité_et_spécificité
- Valeur prédictive sur Wikipédia : https://fr.wikipedia.org/wiki/Valeur_prédictive
- Rapport de la Haute Autorité de Santé du 02/05/20 à télécharger sur <https://www.has-sante.fr>

Exercices préparés par : OBELE MVOMO Julien

Exercices de mathématiques sur la fiabilité de tests d'hypothèses

[Tapez ici]

Introduction : contexte, définitions et hypothèses

1/ Tests sérologiques : (Extraits choisis d'un rapport de la Haute Autorité de Santé du 02/05/20)

Après avoir été infectés par un virus, la plupart des individus développent une réponse immunitaire par la production d'anticorps dirigés contre ce virus. **Des tests sérologiques permettent de déterminer si une personne a produit des anticorps.** Ces tests ont une place dans la surveillance épidémiologique et dans l'identification des personnes étant ou ayant été contact avec ce virus. Les données recueillies permettent d'alimenter des modèles mathématiques dont un des objectifs est d'anticiper la trajectoire de l'épidémie. Les tests sérologiques sont des tests d'hypothèses.

2/ Tests d'hypothèses - Vocabulaire, définitions

En statistiques, un **test** est une démarche consistant à valider ou à rejeter une hypothèse. On s'intéresse alors à différents indicateurs de ce test :

- La **sensibilité** (ou sélectivité) : c'est la probabilité que le test donne un résultat positif lorsque l'hypothèse est vérifiée.
- La **spécificité** : c'est la probabilité que le test donne un résultat négatif lorsque l'hypothèse n'est pas vérifiée.
- La **valeur prédictive positive** (VPP) : C'est la probabilité que l'hypothèse soit vérifiée lorsque le test est positif.
- La **valeur prédictive négative** (VPN) : C'est la probabilité que l'hypothèse ne soit pas vérifiée lorsque le test est négatif.

En épidémiologie, on appelle **prévalence** d'une maladie la proportion de personnes atteintes par cette maladie dans une population donnée. Dans cet article, **la prévalence désignera la proportion de personnes infectées par le SARS-CoV-2 dans une population.**

3/ Hypothèses :

Concernant les tests sérologiques de dépistages d'une infection par le virus SARS-CoV-2, les valeurs seuils minimales de sensibilité et de spécificité des tests sérologiques ,évoqués dans le deuxième paragraphe, ont été fixées par la Haute Autorité de Santé à **90 % pour la sensibilité et 98 % pour la spécificité.** (Source : *Rapport de la HAS du 02/05/20*)

Ce document ainsi que sa référence bibliographique sont téléchargeables sur www.has-sante.fr
Haute Autorité de santé – Service communication information
5 avenue du Stade de France – 93218 Saint-Denis la Plaine Cedex. Tél. : +33 (0)1 55 93 70 00
© Haute Autorité de santé – mai 2020 – ISBN :

Caractéristiques d'un test : sensibilité, spécificité, valeurs prédictives, faux positifs...

Valeur prédictive positive (VPP) : Le test est positif et la personne est grippée. Le test fonctionne. La VPP est la proportion de personnes grippées parmi celles qui ont un test positif.

Valeur prédictive négative (VPN) : Le test est négatif et la personne n'est pas grippée. Le test fonctionne. La VPN est la proportion de personnes non grippées parmi celles qui ont un test négatif.

Faux test positif : Le résultat du test est positif. Pourtant la personne n'est pas grippée ! Le test échoue.

Faux test négatif : Le résultat du test est négatif. Pourtant la personne est grippée ! Le test échoue.

Note : cet exercice peut être donné sans préambule. L'énoncé donne toutes les informations utiles à son traitement (Prérequis : calcul littéral, inéquation du premier degré, notion de fonctions, probabilité d'évènement contraires)

« Taux de Faux Positifs » d'un test sérologique en fonction de la prévalence

Dans cet exercice, la « **prévalence** » désigne la proportion de personnes infectées par un virus dans une population.

On note x la prévalence (par exemple $x = 0,03$ lorsque 3 % de la population ont été infectés par le virus) La valeur de x est encadrée par 0 et 1. On a $x \in [0;1]$

On choisit une personne au hasard dans une population. Cette personne passe un test sérologique permettant de déterminer si elle a produit des anticorps en réponse immunitaire après avoir été infectée par le virus.

La valeur prédictive positive du test (VPP) est la probabilité conditionnelle que la personne ait produit des anticorps sachant que le résultat du test est positif.

Une entreprise de l'industrie pharmaceutique produit et commercialise des tests sérologiques. Son laboratoire a rédigé une notice jointe à tout contrat de vente. Cette notice indique que la VPP en fonction de la prévalence x est donnée par la fonction f définie pour $x \in [0;1]$ par : $f(x) = \frac{45x}{44x+1}$

1/ Dans cette question, on suppose que la prévalence x est égale à 0,05. Cela signifie que 5 % de la population ont été infectés par le virus.

Calculer $f(0,05)$. Puis interpréter le résultat dans le contexte de l'exercice.

2/ Compléter le tableau de valeurs des VPP ci-dessous : (On arrondira les valeurs au 1/1000)

x	0,02	0,04	0,06	0,08	0,10	0,12	0,14	0,16	0,18	0,20
$f(x)$										

3/ On appelle « **Taux de faux positifs** » la proportion de personnes testées positives mais **n'ayant PAS produit d'anticorps** en réponse immunitaire après avoir été infectée par le virus.

On note g la fonction définie pour $x \in [0;1]$ qui donne le « **Taux de faux positifs** » en fonction de la prévalence x .

$$\text{On a : } g(x) = 1 - f(x) = 1 - \frac{45x}{44x+1}$$

Car si E est un évènement et $p(E)$ la probabilité que l'évènement se réalise, alors $p(\bar{E}) = 1 - p(E)$

a. Montrer qu'on peut aussi exprimer $g(x)$ sous la forme : $g(x) = \frac{1-x}{44x+1}$

b. Calculer $g(0,05)$ puis interpréter le résultat dans le contexte de l'exercice.

4/ a. Sachant que $44x + 1$ est toujours positif pour $x \in [0;1]$, résoudre l'inéquation $\frac{1-x}{44x+1} \leq 0,1$.

b. En déduire un encadrement de la prévalence x dans lequel **le taux de faux positif est inférieur à 10 %**.

5/ Pour quelles prévalences x le taux de faux positif est inférieur à 1 % ?

(Toute trace de recherche sera prise en compte)

Probabilités : Fiabilité de tests sérologiques en fonction de la prévalence

La prévalence désigne la proportion de personnes infectées par un virus dans une population.

On note x la prévalence (par exemple $x = 0,03$ lorsque 3 % de la population ont été infectés par le virus)

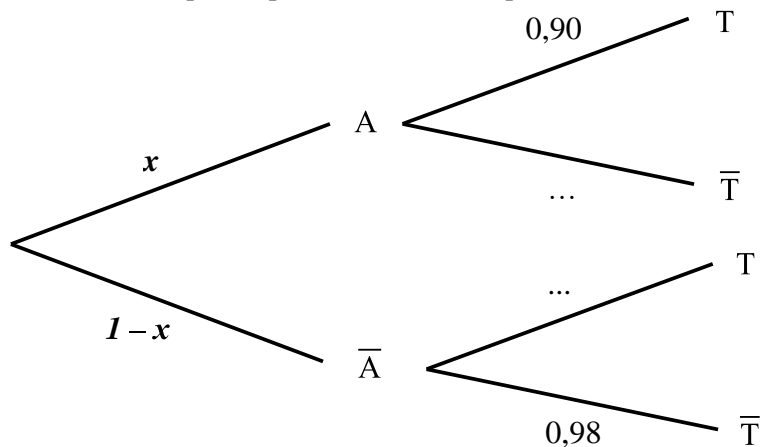
La valeur de x est encadrée par 0 et 1. On a $x \in [0;1]$

On choisit une personne au hasard dans une population. Cette personne passe un test sérologique permettant de déterminer si elle a produit des anticorps en réponse immunitaire après avoir été infectée par le virus.

On note T le caractère : « le Test sérologique de dépistage d'anticorps est positif » et \bar{T} le caractère contraire.

On note A le caractère : « la personne testée est porteuse d'Anticorps » et \bar{A} le caractère contraire.

L'arbre ci-dessous indique les probabilités sur chaque branche :



La probabilité de \bar{T} sachant \bar{A} vaut 0,90 car la sensibilité du test est fixée à 90 %.

La probabilité de T sachant A vaut 0,98 car la spécificité du test est fixée à 98 %.

1/ a. On note $P(T \cap A)$ la probabilité que $T \cap A$ se réalise. Exprimer $p(T \cap A)$ en fonction de x

b. On note $p(T \cap \bar{A})$ la probabilité que $T \cap \bar{A}$ se réalise. Exprimer $p(T \cap \bar{A})$ en fonction de x

c. On note $p(T)$ la probabilité que T se réalise. Montrer que $p(T) = 0,88x + 0,02$.

Une rédaction exhaustive est attendue.

d. En déduire que $p(T) = 0,98 - 0,88x$

2/ On rappelle que la valeur prédictive positive du test (VPP) est la probabilité conditionnelle de A sachant T.

Montrer que l'expression de la VPP en fonction de x peut s'écrire sous la forme : $VPP = \frac{45x}{44x+1}$

3/ On rappelle que la valeur prédictive négative du test (VPN) est la probabilité conditionnelle de \bar{A} sachant \bar{T}

[Tapez ici]

Montrer que l'expression de la VPN en fonction de x peut s'écrire sous la forme : $VPN = \frac{49(1-x)}{49-44x}$

4/ Compléter le tableau de valeurs des VPP et VPN ci-dessous : (On arrondira les résultats au 1/1000)

x	0,02	0,04	0,06	0,08	0,10	0,12	0,14	0,16	0,18	0,20
VPP										
VPN										

5/ Dresser le tableau de variation de la VPP en fonction de x pour $x \in [0;1]$

6/ Dresser le tableau de variation de la VPN en fonction de x pour $x \in [0;1]$

7/ Que peut-on dire de la VPP lorsque la prévalence est faible ? Interpréter dans le contexte.

*Note : cet exercice peut être donné sans préambule. L'énoncé donne toutes les informations utiles à son traitement
(Prérequis : Probabilité conditionnelle, notions de fonction)*

Fiabilité d'un test sérologique performant avec une prévalence très faible

Le 3 mai 2020, le groupe pharmaceutique suisse Roche affirme dans un communiqué de presse avoir développé le test sérologique Elecsys® Anti-SARS-CoV-2 pour détecter les anticorps SARS-CoV-2 avec une sensibilité de 100% et une spécificité de 99.8%. Le test a été approuvé pour être commercialisé aux USA et dans l'Union Européenne (Label CE).

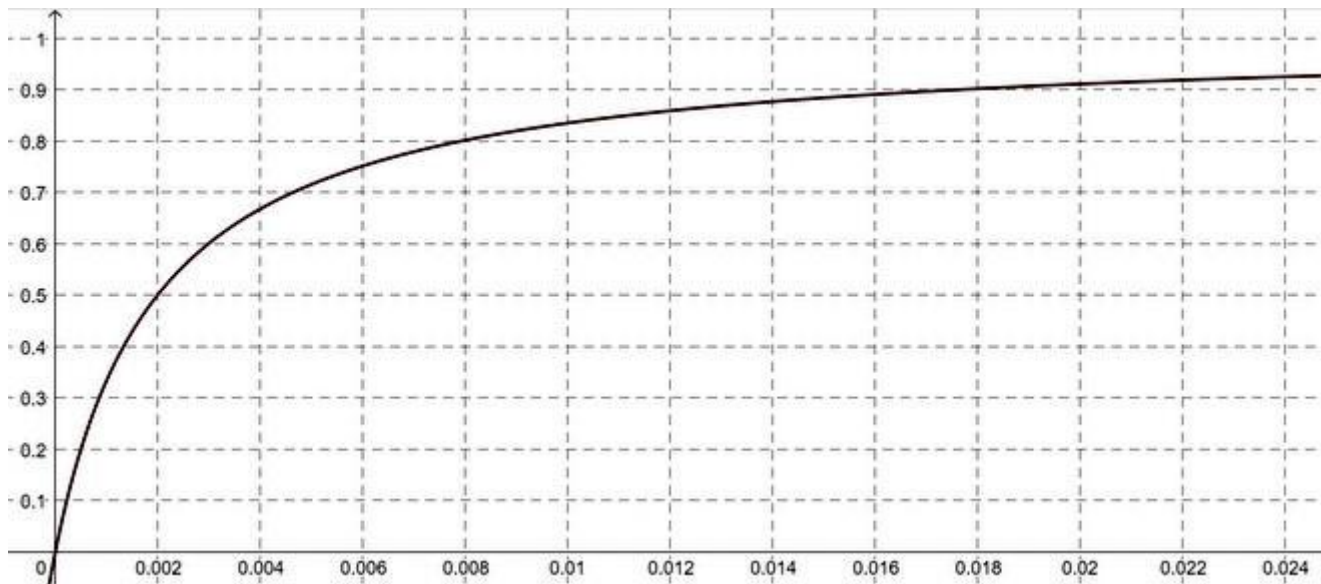
On choisit une personne au hasard. Cette personne passe le test sérologique du groupe Roche afin de déterminer si elle a produit des anticorps en réponse immunitaire après avoir été infectée par le virus SARS-CoV-2.

La valeur **prédictive positive du test** (VPP) est la probabilité conditionnelle que la personne ait développé des anticorps sachant que le résultat du test est positif.

La **prévalence** désigne ici la proportion de personnes infectées par le SARS-CoV-2 dans une population.

Si on note x la prévalence et $f(x)$ la VPP du test sérologique, alors pour $x \in [0;1]$, on a $f(x) = \frac{500x}{499x+1}$

Le graphique ci-dessous représente les variations de $f(x)$ en fonction de x pour de petites valeurs de la prévalence :



1/ a. Calculer $f(0,022)$ puis vérifier le résultat avec le graphique ci-dessus.

b. Interpréter le résultat dans le contexte. Formuler une interprétation explicite et précise.

2/ a. Calculer $f(0,002)$ puis vérifier le résultat avec le graphique ci-dessus.

b. Interpréter le résultat dans le contexte. Formuler une interprétation explicite et précise.

3/ A L'île de La Réunion, en mai 2020, on estime que la proportion de personnes infectées par le SARS-CoV-2 dans la population est inférieure à 0,002 (soit moins de 0,2 % de la population infectée par le virus).

Le test sérologique du groupe pharmaceutique Roche peut-il être jugé fiable à La Réunion en mai 2020 ?

Justifier votre réponse.

«Les choses ne sont pas difficiles parce que l'on n'ose pas les faire mais c'est parce que l'on n'ose pas les faire qu'elles sont difficiles» Sénèque

« Un bonheur que rien n'a entamé succombe à la moindre atteinte ; mais quand on doit se battre

contre les difficultés incessantes ,on s'aguerrit dans l'épreuve ,on résiste à n'importe quels maux,

et même si l'on trébuche , on lutte encore à genoux» Sénèque

OMJS

[Tapez ici]

Note : cet exercice peut être donné sans préambule. L'énoncé donne toutes les informations utiles à son traitement

Fiabilité d'un test sérologique performant avec une prévalence très faible

Le 3 mai 2020, le groupe pharmaceutique suisse Roche affirme dans un communiqué de presse avoir développé le test sérologique Elecsys® Anti-SARS-CoV-2 pour détecter les anticorps SARS-CoV-2.

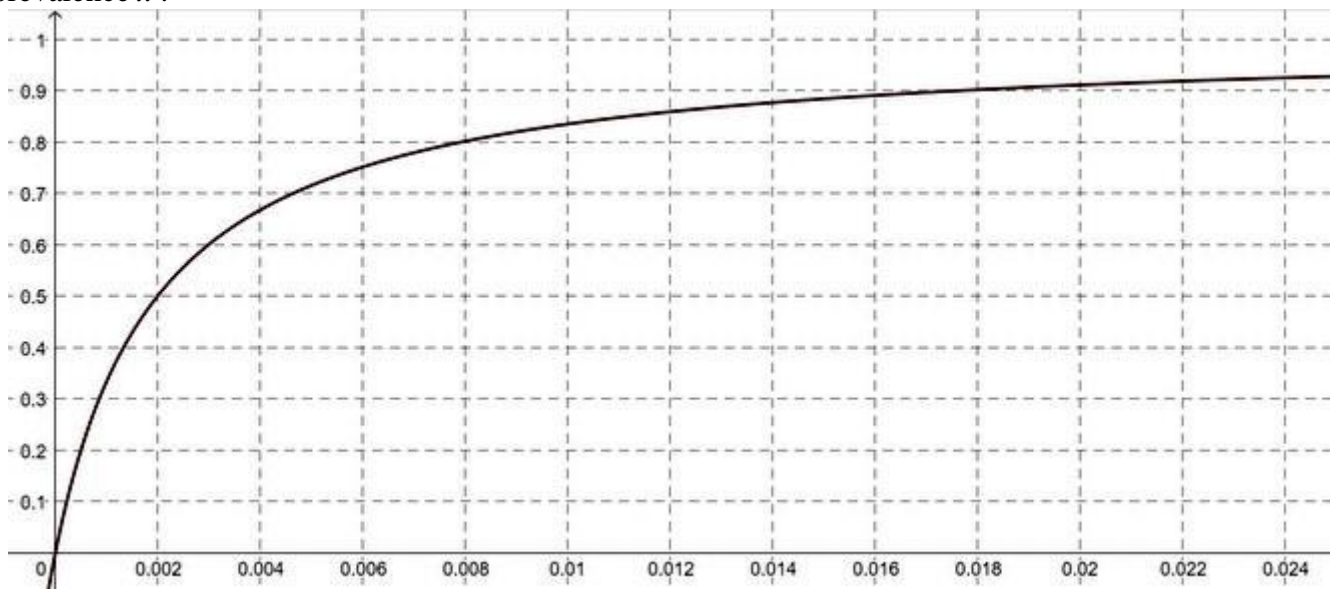
On choisit une personne au hasard dans une population. Cette personne passe le test sérologique du groupe Roche afin de déterminer si elle a produit des anticorps en réponse immunitaire après avoir été infectée par le virus SARS-CoV-2.

La valeur **prédictive positive du test** (VPP) est la probabilité conditionnelle que la personne ait développé des anticorps sachant que le résultat du test est positif.

La **prévalence** désigne la proportion de personnes infectées par le SARS-CoV-2 dans une population. Une prévalence de 0,02 signifie que 2 % de la population ont été infectés par le virus.

Si on note x la prévalence et $f(x)$ la VPP du test sérologique, alors pour $x \in [0; 1]$, on a $f(x) = \frac{500x}{499x+1}$

Le graphique ci-dessous représente les variations de $f(x)$ en fonction de x pour de petites valeurs de la prévalence x :



- 1/ Quelle est approximativement la valeur prédictive positive du test pour une prévalence de 0,018 ?
- 2/ Quelle est approximativement la prévalence lorsque la valeur prédictive positive de ce test vaut 50 % ?
- 3/ On remarque que $499x+1$ est toujours strictement positif lorsque $x \in [0; 1]$.
Résoudre l'inéquation du premier degré $\frac{500x}{499x+1} \geq 0,95$

[Tapez ici]

Note : cet exercice peut être donné sans préambule. L'énoncé donne toutes les informations utiles à son traitement

4/ Interpréter le résultat de la question 3/ dans le contexte de l'exercice. Formuler une interprétation explicite en faisant référence à la prévalence (en%) et à la valeur prédictive positive du test.