

Plaquelette 3 - Ne pas inverser le conditionnement : tests et faux positifs

Probabilités conditionnelles — Seconde

Corrections détaillées

Boîte à outils

Tests de dépistage. M : « la personne est malade » ; T : « le test est positif ».

- **Prévalence** : $P(M)$, proportion de malades dans la population testée.
- **Sensibilité** : $P_M(T)$; **spécificité** : $P_{\bar{M}}(\bar{T})$.
- **Faux positif** : sain et testé positif ; **faux négatif** : malade et testé négatif.
- Probabilité d'être malade quand le test est positif (valeur prédictive positive) :

$$P_T(M) = \frac{P(M \cap T)}{P(T)} = \frac{\text{vrais positifs}}{\text{vrais positifs} + \text{faux positifs}}$$

- Méthode : 10 000 personnes → malades et sains → positifs dans chaque groupe → tableau → $P_T(M)$.
- **Inversion du conditionnement** : $P_M(T)$ et $P_T(M)$ peuvent être très différentes, surtout quand la maladie est rare.

Correction de l'exercice 1 - Un test qui semble fiable

a) 100 malades, dont $0,9 \times 100 = 90$ positifs ; 9 900 personnes saines, dont 9 % positives à tort (spécificité 91 %) : $0,09 \times 9\,900 = 891$ faux positifs.

	Malades	Sains	Total
Test positif	90	891	981
Test négatif	10	9 009	9 019
Total	100	9 900	10 000

b) Référence : les 981 personnes positives.

$$P_T(M) = \frac{90}{981} \approx 0,092$$

Moins d'une personne positive sur dix est malade, alors que le test détecte 90 % des malades : les faux positifs, issus des très nombreuses personnes saines, sont dix fois plus nombreux que les vrais positifs.

Réponse. b) environ 0,092.

Correction de l'exercice 2 - Le vocabulaire des tests

a) « Détecte 98 % des personnes infectées » : parmi les infectés, 98 % ont un test positif.

$$P_I(T) = 0,98 \quad (\text{sensibilité}) \quad P_{\bar{I}}(\bar{T}) = 0,99 \quad (\text{spécificité})$$

b) Faux négatif chez une personne infectée : $P_I(\bar{T}) = 1 - 0,98 = 0,02$. Faux positif chez une personne non infectée : $P_{\bar{I}}(T) = 1 - 0,99 = 0,01$.

Attention. Aucune de ces probabilités ne dit ce que veut savoir une personne positive, $P_T(I)$: il faut aussi connaître la proportion d'infectés dans la population.

Réponse. a) $P_I(T) = 0,98$ (sensibilité), $P_{\bar{I}}(\bar{T}) = 0,99$ (spécificité) · b) 0,02 et 0,01.

Correction de l'exercice 3 - L'effet de la prévalence

Dans chaque cas, on calcule les vrais positifs ($0,9 \times \text{malades}$) et les faux positifs ($0,05 \times \text{sains}$), puis

$$P_T(M) = \frac{\text{vrais positifs}}{\text{vrais positifs} + \text{faux positifs}}$$

- a) 100 malades $\rightarrow$ 90 ; 9 900 sains $\rightarrow$ 495 : $P_T(M) = \frac{90}{585} \approx 0,154$.
 b) 1 000 malades $\rightarrow$ 900 ; 9 000 sains $\rightarrow$ 450 : $P_T(M) = \frac{900}{1350} \approx 0,667$.
 c) 3 000 malades $\rightarrow$ 2 700 ; 7 000 sains $\rightarrow$ 350 : $P_T(M) = \frac{2700}{3050} \approx 0,885$.

Le **même** test donne une valeur très différente selon la population : un résultat positif est bien plus parlant chez un patient qui a des symptômes que lors d'un dépistage de masse.

Réponse. a) $\approx 0,154$ · b) $\approx 0,667$ · c) $\approx 0,885$.

Correction de l'exercice 4 - Un test négatif rassure-t-il ?

Référence : les 9 019 personnes testées négatives, dont 9 009 saines et 10 malades (faux négatifs).

$$P_{\bar{T}}(\bar{M}) = \frac{9\,009}{9\,019} \approx 0,9989$$

et

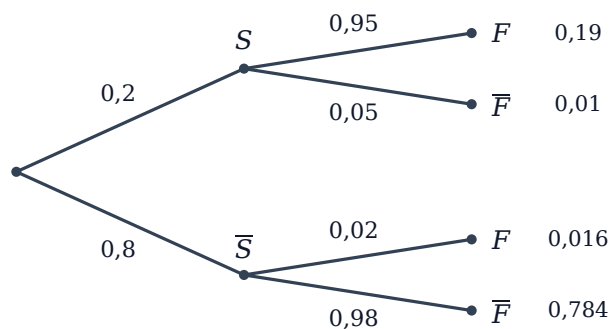
$$P_{\bar{T}}(M) = 1 - P_{\bar{T}}(\bar{M}) = \frac{10}{9\,019} \approx 0,0011$$

Un test négatif est très rassurant : environ une personne négative sur 900 est malade. Pour une maladie rare, ce test sert surtout à **écarter** la maladie ; un résultat positif doit être confirmé.

Réponse. $P_{\bar{T}}(\bar{M}) \approx 0,9989$; $P_{\bar{T}}(M) \approx 0,0011$.

Correction de l'exercice 5 - Filtre anti-spam

On construit l'arbre : $P(S) = 0,2$, $P_S(F) = 0,95$, $P_{\bar{S}}(F) = 0,02$.



S : courriel indésirable ; F : filtré

a) Deux chemins mènent à F :

$$P(F) = 0,2 \times 0,95 + 0,8 \times 0,02 = 0,19 + 0,016 = 0,206$$

b) $P_F(S) = \frac{0,19}{0,206} \approx 0,922$.

Ici les spams sont fréquents (20 %) : la grande majorité des courriels bloqués sont bien

des spams. Environ 8 % des courriels bloqués sont tout de même légitimes : il faut vérifier le dossier des indésirables.

Réponse. a) $0,206$ · b) environ $0,922$.

Correction de l'exercice 6 - Contrôle antidopage

On note D : « dopé » et T : « contrôle positif ». Sur 10 000 coureurs : 500 dopés, dont $0,8 \times 500 = 400$ positifs ; 9 500 non dopés, dont $0,01 \times 9 500 = 95$ positifs.

a) Référence : les 495 positifs.

$$P_T(D) = \frac{400}{495} \approx 0,808$$

Environ 81 % des positifs sont dopés : un contrôle positif seul ne suffit pas à condamner, d'où les contre-expertises.

b) $P_D(\bar{T}) = 1 - 0,8 = 0,2$: un coureur dopé sur cinq n'est pas détecté.

Réponse. a) environ $0,808$ · b) $0,2$.

Correction de l'exercice 7 - Confirmer par un second test

a) 2 000 malades, dont $0,95 \times 2 000 = 1 900$ positifs ; 98 000 sains, dont $0,1 \times 98 000 = 9 800$ positifs. Positifs : 11 700.

$$P_T(M) = \frac{1 900}{11 700} \approx 0,162$$

b) Parmi les 1 900 malades positifs, $0,95 \times 1 900 = 1 805$ le sont à nouveau ; parmi les 9 800 sains positifs, $0,1 \times 9 800 = 980$. Après deux tests positifs :

$$\frac{1 805}{1 805 + 980} = \frac{1 805}{2 785} \approx 0,648$$

Le second test fait passer la probabilité d'être malade d'environ 16 % à environ 65 % : il est appliqué à une population où la maladie est beaucoup plus fréquente.

Réponse. a) 11 700 ; $\approx 0,162$ · b) 2 785 ; $\approx 0,648$.

Correction de l'exercice 8 - Alarme incendie

On note I : « incendie » et A : « l'alarme sonne ». Les chemins qui mènent à A :

$$P(I \cap A) = 0,001 \times 0,98 = 0,000 98 \quad P(\bar{I} \cap A) = 0,999 \times 0,03 = 0,029 97$$

d'où $P(A) = 0,030 95$ et

$$P_A(I) = \frac{0,000 98}{0,030 95} \approx 0,032$$

Seule une alarme sur trente environ correspond à un vrai incendie. Il ne faut pas l'ignorer pour autant : l'erreur « ne pas évacuer pendant un vrai incendie » est bien plus grave que celle « évacuer pour rien ». La probabilité ne dit pas tout, il faut aussi peser les conséquences.

Réponse. Environ $0,032$; non, l'enjeu justifie d'évacuer.

Correction de l'exercice 9 - Une trace ADN

a) Parmi environ 5 000 000 d'innocents, une proportion $\frac{1}{1 000 000}$ correspond par hasard :

$$5 000 000 \times \frac{1}{1 000 000} = 5$$

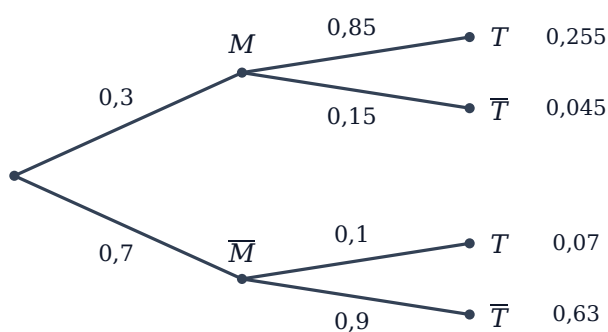
b) Les personnes dont le profil correspond sont donc environ 6 : le coupable et 5 innocents. Sans autre élément, la probabilité que le suspect soit coupable vaut environ $\frac{1}{6}$.

L'avocat général confond $P_{\text{innocent}}(\text{correspondance}) = \frac{1}{1\,000\,000}$ et $P_{\text{correspondance}}(\text{innocent})$, qui vaut ici environ $\frac{5}{6}$. C'est l'erreur d'inversion du conditionnement, appelée « sophisme du procureur ».

Réponse. a) environ 5 · b) environ $\frac{1}{6}$; il inverse le conditionnement.

Correction de l'exercice 10 - Patients avec symptômes

a) $P(M) = 0,3$, $P_M(T) = 0,85$, $P_{\bar{M}}(T) = 0,1$.



M : malade ; T : test positif

Deux chemins mènent à T :

$$P(T) = 0,3 \times 0,85 + 0,7 \times 0,1 = 0,255 + 0,07 = 0,325$$

b) $P_T(M) = \frac{0,255}{0,325} \approx 0,785$.

Dans cette population où la maladie est fréquente, un test positif est assez fiable (environ 78 %) : le médecin peut prescrire un antibiotique.

Réponse. a) 0,325 · b) environ 0,785.

Correction de l'exercice 11 - Les faux négatifs

a) Chemin « malade et test négatif » : $P(M \cap \bar{T}) = 0,3 \times 0,15 = 0,045$. Deux chemins mènent à $\bar{T}$:

$$P(\bar{T}) = 0,045 + 0,7 \times 0,9 = 0,045 + 0,63 = 0,675$$

b) $P_{\bar{T}}(M) = \frac{0,045}{0,675} \approx 0,067$.

Environ un patient négatif sur quinze est malade. Le test négatif rend l'angine bactérienne peu probable, mais pas impossible : si les symptômes persistent, le médecin peut revoir le patient. Plus la sensibilité est faible, plus les faux négatifs sont fréquents.

Réponse. a) 0,045 ; 0,675 · b) environ 0,067.

Correction de l'exercice 12 - « Fiable à 99 % »

a) 100 malades, dont $0,99 \times 100 = 99$ positifs ; 9 900 sains, dont $0,01 \times 9\,900 = 99$ positifs à tort.

b) Référence : les $99 + 99 = 198$ positifs.

$$P_T(M) = \frac{99}{198} = 0,5$$

Non : une personne positive n'a qu'une chance sur deux d'être malade. Le slogan « fiable à 99 % » parle de $P_M(T)$ et $P_{\overline{M}}(\overline{T})$, pas de $P_T(M)$. Quand la maladie est rare, même un excellent test produit autant de faux positifs que de vrais positifs.

Réponse. a) 99 et 99 · b) Non : $P_T(M) = 0,5$.

Correction de l'exercice 13 - Vérifier une affirmation

Chaque phrase a sa population de référence.

— « 90 % des personnes positives ne sont pas malades » : référence, les 981 positifs.

$$P_T(\overline{M}) = \frac{891}{981} \approx 0,908$$

La phrase est juste (environ 91 %).

— « 90 % des malades sont positifs » : référence, les 100 malades. $P_M(T) = \frac{90}{100} = 0,9$: juste aussi, c'est la sensibilité.

Les deux phrases, presque identiques en apparence, décrivent deux probabilités conditionnelles différentes ; ici elles valent toutes deux environ 90 %, par coïncidence.

Réponse. Les deux sont justes : $P_T(\overline{M}) \approx 0,908$ et $P_M(T) = 0,9$.

Correction de l'exercice 14 - Lots contaminés

a) 500 lots contaminés, dont 490 positifs ; 9 500 lots sains, dont $0,04 \times 9\,500 = 380$ positifs.

	Contaminés	Sains	Total
Positif	490	380	870
Négatif	10	9 120	9 130
Total	500	9 500	10 000

b) Positifs :

$$P_T(C) = \frac{490}{870} \approx 0,563$$

Négatifs : $P_T(\overline{C}) = \frac{9\,120}{9\,130} \approx 0,9989$.

Les lots négatifs peuvent être livrés presque sans risque ; les lots positifs doivent être analysés au laboratoire, car près de la moitié sont sains.

Réponse. b) environ 0,563 ; environ 0,9989.

Correction de l'exercice 15 - Choisir entre deux tests

200 malades et 9 800 personnes saines.

— Test A : 198 vrais positifs, 2 faux négatifs, $0,1 \times 9\,800 = 980$ faux positifs.

$$P_T(M) = \frac{198}{1\,178} \approx 0,168$$

— Test B : 180 vrais positifs, 20 faux négatifs, $0,01 \times 9\,800 = 98$ faux positifs, donc $P_T(M) = \frac{180}{278} \approx 0,647$.

Pour un **dépistage**, on veut ne manquer presque aucun malade : le test A (seulement 2

faux négatifs). Pour **confirmer** un diagnostic, on veut qu'un positif soit fiable : le test B. En pratique, on enchaîne souvent les deux.

Réponse. A : 2 faux négatifs, $P_T(M) \approx 0,168$; B : 20 faux négatifs, $P_T(M) \approx 0,647$.