

Document du GEPS-Version provisoire.

Octobre 2001.

● Appariement aléatoire en génétique

Dans les cas simples, un gène peut prendre deux formes (ou allèles) A et a et un individu peut avoir l'un des trois génotypes suivants : AA, Aa, aa. Considérons une population (génération 0) dont les proportions respectives de ces génotypes sont p_0, q_0, r_0 . Un enfant lors de sa naissance hérite d'un gène de chaque parent, chacun d'eux étant choisi au hasard. On admet que les couples se forment au hasard quant aux génotypes considérés (appariement aléatoire). On veut voir comment évoluent les proportions de génotypes dans la population à chaque génération. Soient p_n, q_n, r_n les proportions des génotypes AA, Aa, aa à la n-ème génération.

On peut démontrer qu'il y a stabilité à partir de la première génération : ce résultat est appelé en biologie la loi de Hardy-Weinberg.

Parmi les méthodes possibles pour étudier cette question, on peut utiliser la suivante :

- Sachant qu'un parent est AA, on détermine la loi du génotype de l'enfant.

- Sachant qu'un parent est Aa, on détermine la loi du génotype de l'enfant.

- On détermine alors la probabilité p_1 qu'un enfant de la première génération ait le génotype AA. La probabilité r_1 qu'un enfant de la première génération ait le génotype aa s'en déduit en intervertissant p_0 et r_0 . On peut vérifier que $p_1 - r_1 = p_0 - r_0$. Soit α la valeur commune de ces deux différences. On peut alors exprimer p_1, q_1, r_1 en fonction de α , en déduire p_2, q_2, r_2 et conclure.

Le calcul apporte une information supplémentaire : l'état stable ne dépend que de $p_0 - r_0$.

● A propos de la problématique de tests de dépistage

Le fabricant d'un test spécifique d'une maladie fournit les caractéristiques suivantes :

La probabilité qu'un individu malade ait un test positif est 0,99.

La probabilité qu'un individu non malade ait un test négatif est 0,99.

Désignons par p le pourcentage de gens malades dans la population ciblée par le test et par $P(M/T)$ la valeur diagnostique du test dans cette population, où $P(M/T)$ est la probabilité qu'un individu de cette population dont le test est positif soit malade. De la formule $P(M/T) = 99p / (98p + 1)$, on déduit :

p	0,001	0,010	0,100	0,300	0,500	0,800
P(M/T)	0,090	0,500	0,912	0,977	0,990	0,997

On remarque que :

- la valeur diagnostique n'est pas une notion intrinsèque au test lui-même : elle varie fortement selon la population cible.
- dans les cas où p est faible, la valeur diagnostique du test l'est aussi ; cependant, par exemple pour $p = 0,01$, la connaissance de la positivité du test multiplie par 50 la probabilité d'être atteint de la maladie : ce résultat est un élément à prendre en compte dans le processus diagnostique.
- si la population cible est une population d'individus présentant des symptômes évocateurs de la maladie, ou une population dite à risque, p n'est pas faible : la positivité du test sera un élément important du diagnostic. Par contre, pour une maladie rare, un test de dépistage systématique de toute une population aura l'inconvénient majeur de fournir beaucoup de faux positifs (individus non malades dont le test est positif). Pour ces derniers, l'inquiétude liée à la découverte d'un test positif peut-être grande : c'est là un des problèmes éthiques liés à la mise en place des tests de dépistage systématique d'une maladie rare.
- Lorsque $p < 0,10$, des calculs analogues montrent que la probabilité qu'un individu dont le test est négatif ne soit pas atteint de la maladie étudiée est supérieure à 0,999.