

Test génétique et probabilités

Principe d'un test génétique

Test
génétique et
probabilités

Chaque être vivant est caractérisé par l'ADN contenu dans ces cellules.

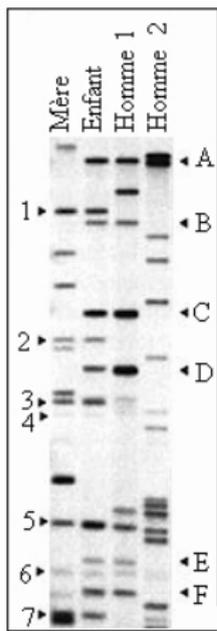
L'empreinte génétique est une technique pour identifier des individus en utilisant un petit échantillon de leur ADN. A certains emplacements de l'ADN (loci), on a remarqué des répétitions de motifs. La longueur de la répétition définit un allèle pour ce loci.

Des loci ont été choisis pour faire les tests ADN (16 en France). Pour chaque individu, on recherche les deux allèles présentes sur chaque loci.

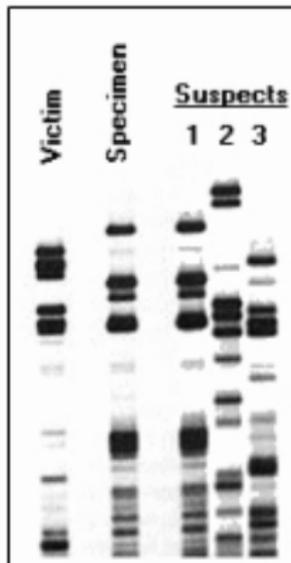
Utilisation des tests génétiques

Test
génétique et
probabilités

Test de filiation



Test judiciaire



Avec le test Quad STR analyse quatre marqueurs génétiques.
La fréquence moyenne des profils génétiques possibles dans ce système est de 0.01%. Cette fréquence paraît faible.

Utilisation erronée

Test
génétique et
probabilités

Un expert indique au tribunal :

Il n'y a pas de différences entre les profils génétiques du suspect et de la trace de sang retrouvée sur les lieux du crime.

L'accusation argumente :

- *La fréquence d'apparition du profil est de 0,01%.*
- *La probabilité que quelqu'un d'autre ait laissé cette trace est de 0,01%.*
- *On est sûr à 99,99% que le suspect a laissé cette trace.*

Explication

Test
génétique et
probabilités

Le raisonnement précédent est faux.

L'expert a estimé la probabilité que les deux empreintes génétiques concordent sachant que le suspect n'est pas l'auteur de la trace. Alors que l'accusation la considère égale à la probabilité que le suspect n'est pas l'auteur de la trace sachant que les empreintes coïncident.

Probabilité d'une autre correspondance

Test
génétique et
probabilités

On va calculer la probabilité p que quelqu'un d'autre ait laissé cette trace. Supposons que n personnes aient pu accéder à la scène du crime (on considère par ex. les habitants du quartier), alors

$$p = 1 - \left(1 - \frac{1}{10000}\right)^n.$$

Si on considère que 1000 personnes ont pu se trouver sur le lieu du crime, il y a 10% de chance pour que la trace ait pu être laissée par quelqu'un d'autre.

Paradoxe des anniversaires

On va calculer la probabilité p_k pour que dans un groupe de k personnes toutes aient une date d'anniversaire différente.

Pour un groupe de 2 personnes : La première peut avoir son anniversaire n'importe quel jour. La seconde a le choix entre toutes les dates d'anniversaire sauf une.

$$p_2 = 1 - \frac{1}{365}$$

Pour un groupe de 3 personnes : la troisième a le choix entre toutes les dates sauf les deux dates d'anniversaire des deux premières personnes.

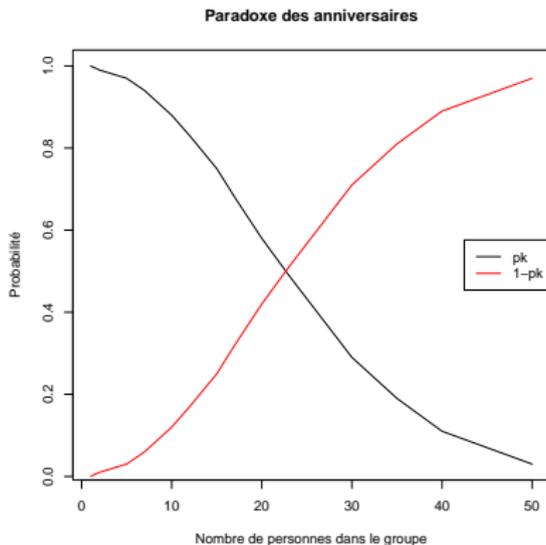
$$p_3 = \left(1 - \frac{1}{365}\right) \left(1 - \frac{2}{365}\right)$$

Pour un groupe de k personnes :

$$p_k = \left(1 - \frac{1}{365}\right) \left(1 - \frac{2}{365}\right) \dots \left(1 - \frac{k-1}{365}\right)$$

Suite du paradoxe

Test
génétique et
probabilités



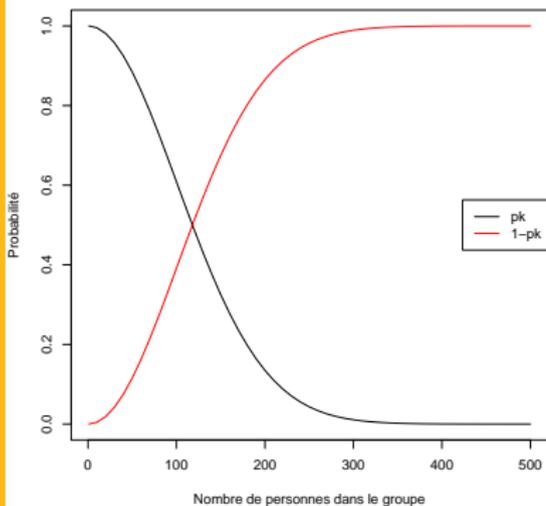
p_k est la probabilité que dans un groupe de k personnes, toutes aient une date d'anniversaire différente.

$1 - p_k$ est la probabilité que dans un groupe de k personnes, il y en ait au moins deux qui aient la même date d'anniversaire.

A partir de 23 personnes, on a une chance sur deux de trouver deux personnes ayant la même date d'anniversaire.

Retour sur les tests ADN

Test
génétique et
probabilités



p_k est la probabilité que dans un groupe de k personnes, toutes aient une empreinte génétique différente
 $1 - p_k$ est la probabilité que dans un groupe de k personnes, il y en ait au moins 2 avec la même empreinte.

Sur ce test la probabilité d'avoir deux personnes ayant le même profil génétique est de 50% dès qu'on atteint un groupe de 120 personnes. Ce test ne peut donc pas servir à faire un fichier d'empreinte génétique.