

MCR 2021 - Équilibre de Hardy-Weinberg

Soient p_1, p_2, p_3 trois réels strictement positifs tels que $p_1 + p_2 + p_3 = 1$. On considère N variables aléatoires $T_1, \dots, T_N$ indépendantes de même loi donnée par :

$$T_1(\Omega) = \{1, 2, 3\}, \quad P(T_1 = 1) = p_1, \quad P(T_1 = 2) = p_2, \quad P(T_1 = 3) = p_3.$$

Cette suite permet de modéliser la répartition d'une population de N individus de type AA (type 1), aa (type 2) ou Aa (type 3) en fonction de leur génotype à un locus génétique donné. Pour tout $i \in \llbracket 1; 3 \rrbracket$, on note N_i la variable aléatoire égale au nombre d'individus de type i . On a donc $N_1 + N_2 + N_3 = N$.

1. Déterminer la loi de N_1 .
2. Donner, sans justification, l'espérance et la variance de N_1 .
3. Soit $(a, b) \in \mathbb{R}^2$ tel que $a \leq b$. Exprimer la quantité

$$\lim_{N \rightarrow +\infty} \mathbb{P} \left(a \leq \frac{N_1 - Np_1}{\sqrt{Np_1(1-p_1)}} \leq b \right)$$

sous forme d'une intégrale.

4. On considère la matrice $W = \begin{pmatrix} \mathbb{V}(N_1) & \text{Cov}(N_1, N_2) \\ \text{Cov}(N_2, N_1) & \mathbb{V}(N_2) \end{pmatrix}$.
 - a. Justifier que W est une matrice diagonalisable.
 - b. Soit $(a, b) \in \mathbb{R}^2$. Prouver que

$$\mathbb{V}(aN_1 + bN_2) = (a \quad b)W \begin{pmatrix} a \\ b \end{pmatrix},$$

où l'on a identifié $\mathcal{M}_1(\mathbb{R})$ et $\mathbb{R}$.

- c. En déduire que les valeurs propres de W sont positives.
- d. Prouver que les valeurs propres de W sont strictement positives.
- e. En déduire l'existence de deux matrices P et D appartenant à $\mathcal{M}_2(\mathbb{R})$, telles que :
 - D soit diagonale et inversible,
 - P soit inversible d'inverse P^T ,
 - $W = PD^2P^{-1}$.
5. On note $A = D^{-1}P^{-1}$ et on considère les variables aléatoires Y_1, Y_2 telles que :

$$\begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} = A \begin{pmatrix} N_1 - Np_1 \\ N_2 - Np_2 \end{pmatrix}.$$

On admet que l'on a :

$$\begin{pmatrix} \mathbb{V}(Y_1) & \text{Cov}(Y_1, Y_2) \\ \text{Cov}(Y_2, Y_1) & \mathbb{V}(Y_2) \end{pmatrix} = AW A^T.$$

En déduire $\mathbb{V}(Y_1)$, $\mathbb{V}(Y_2)$ et $\text{Cov}(Y_1, Y_2)$.

6. a. Déterminer la variance de $N_1 + N_2$.
- b. En déduire $\text{Cov}(N_1, N_2)$ la covariance de N_1 et N_2 .
- c. Calculer l'inverse de W .
- d. Déterminer $A^T A$ en fonction de W et en déduire que :

$$Y_1^2 + Y_2^2 = \frac{(N_1 - Np_1)^2}{Np_1} + \frac{(N_2 - Np_2)^2}{Np_2} + \frac{(N_3 - Np_3)^2}{Np_3}.$$

* *
*