

# PROBABILITÉS FINIES

Traiter l'exercice 1 puis, au choix, l'exercice 2 ou l'exercice 3

## Exercice 1.

Custer a conduit son armée à l'entrée d'un canyon. Un doute lui vient à l'esprit. Il estime qu'il a 40 % de chances pour que Sitting Bull lui tende une embuscade dans ce défilé. Il consulte alors ses deux éclaireurs. Le premier, un visage pâle qui se trompe trois fois sur quatre, lui déclare qu'il n'y a pas de danger. Le deuxième, un indien qui communique une information fautive une fois sur trois, lui annonce un guet-apens. On suppose que les avis des éclaireurs sont indépendants sachant qu'il y a un guet-apens ou non.



Calculer, avec les éléments dont dispose Custer, la probabilité pour qu'il y ait une embuscade.

## Exercice 2. Loi de Hardy-Weinberg

Dans une population, on s'intéresse à la transmission d'un gène présent sous la forme de deux allèles  $A$  et  $a$ . On suppose que le gène est autosomique (c'est-à-dire porté par un chromosome non sexuel). Chaque individu est donc ou bien homozygote :  $AA$  ou  $aa$ , ou bien hétérozygote :  $Aa$ .

Au cours de la gamétogénèse, les cellules mâles ou femelles passent de  $2m$  chromosomes (cellule diploïde) à  $m$  chromosomes (cellule haploïde). Chez un hétérozygote, chacun des deux allèles est donc présent dans la moitié des gamètes. Pour cette raison, on considère qu'un parent hétérozygote transmet l'un ou l'autre allèle de façon équiprobable.

On suppose également que l'allèle transmis par un parent est indépendant de l'allèle transmis par l'autre parent.

On suppose la population idéale, c'est-à-dire de grande taille (on néglige les phénomènes de dérive génétique); sans sélection naturelle (les deux allèles  $A$  et  $a$  ne confèrent ni avantage ni désavantage sélectif); sans sélection comportementale (les couples se forment indépendamment de leur génotype); sans migration (aucune copie allélique n'est apportée de l'extérieur); sans croisement entre des générations différentes. Enfin, on néglige les mutations.

L'objectif de cet exercice est démontrer la loi de Hardy-Weinberg : « Dans une population diploïde idéale, les fréquences alléliques d'un gène sont constantes au fil des générations ». On en déduit, en particulier, que les relations de dominance entre allèles n'ont aucun effet sur l'évolution des fréquences alléliques.

Dans la  $n$ -ème génération, les génotypes  $AA$ ,  $Aa$  et  $aa$  ont pour fréquences respectives  $a_n$ ,  $b_n$  et  $c_n$ . En outre, on note  $p_n$  et  $q_n$  les fréquences des allèles  $A$  et  $a$  dans cette  $n$ -ème génération.

1. Pour tout  $n \in \mathbb{N}^*$ , exprimer  $p_n$  et  $q_n$  en fonction de  $a_n$ ,  $b_n$  et  $c_n$ .
2. Démontrer que, pour tout  $n \in \mathbb{N}^*$ , on a  $a_{n+1} = p_n^2$ ,  $b_{n+1} = 2p_nq_n$  et  $c_{n+1} = q_n^2$ .
3. En déduire, pour tout  $n \in \mathbb{N}^*$ , les expressions de  $p_{n+1}$  et  $q_{n+1}$  en fonction de  $p_n$  et  $q_n$ . Conclure.

### Exercice 3. Maladie autosomique récessive létale

Une maladie génétique est dite autosomique récessive lorsque, d'une part, le gène affecté est porté par un chromosome non sexuel (la maladie touche autant les hommes que les femmes) et, d'autre part, les mutations de ce gène sont récessives. C'est le cas, par exemple, de la drépanocytose et de la mucoviscidose.

Pour modéliser une telle maladie, on considère que, chez un individu, chacun des deux gènes en cause porte ou bien l'allèle normal  $A$  ou bien l'allèle muté  $a$ . En outre, on suppose que deux allèles mutés sont nécessaires pour que la maladie se manifeste. Chaque individu peut donc être ou bien  $AA$  (homozygote sain), ou bien  $Aa$  (hétérozygote sain) ou encore  $aa$  (homozygote malade).

On considère de plus que les homozygotes malades n'atteignent pas l'âge adulte ou qu'ils n'ont pas de descendance (les cas assez rares où cela se produit sont négligés<sup>(†)</sup>). Pour cette raison, on considèrera que la proportion de malades parmi les adultes est nulle.

Au cours de la gamétogénèse, les cellules mâles ou femelles passent de  $2m$  chromosomes (cellule diploïde) à  $m$  chromosomes (cellule haploïde). Chez un hétérozygote, chacun des deux allèles est donc présent dans la moitié des gamètes. Pour cette raison, on considère qu'un parent hétérozygote transmet l'un ou l'autre allèle de façon équiprobable.

On suppose que les porteurs sains ignorent qu'ils ont l'allèle muté et donc que les couples se forment indépendamment du fait que l'un des deux soit hétérozygote ou non. On suppose également que l'allèle transmis par un parent est indépendant de l'allèle transmis par l'autre parent.

Chez les adultes de la  $n$ -ème génération, les génotypes  $AA$  et  $Aa$  ont pour fréquences respectives  $a_n$  et  $b_n$ . Par ailleurs, on note  $p_n$  et  $q_n$  les fréquences des allèles  $A$  et  $a$  dans cette  $n$ -ème génération.

A. Pour tout  $n \in \mathbb{N}^*$ , exprimer  $p_n$  et  $q_n$  en fonction de  $a_n$  et  $b_n$ .

#### B. Transmission sans mutation

1. Un couple de la  $n$ -ème génération ( $n \in \mathbb{N}$ ) décide d'avoir un enfant. On note  $AA$  l'événement « l'enfant est  $AA$  »,  $Aa$  l'événement « l'enfant est  $Aa$  » et  $aa$  l'événement « l'enfant est  $aa$  ».

a) Démontrer que  $P(AA) = p_n^2$ ,  $P(Aa) = 2p_nq_n$  et  $P(aa) = q_n^2$ .

Application : Pour la mucoviscidose, 4% des adultes occidentaux sont porteurs sains. Calculer la probabilité d'avoir un enfant malade pour un couple d'occidentaux.

b) Établir que

$$a_{n+1} = \frac{p_n^2}{1 - q_n^2} \quad \text{et} \quad b_{n+1} = \frac{2p_nq_n}{1 - q_n^2},$$

puis simplifier ces relations en tenant compte du fait que  $p_n + q_n = 1$ .

c) En déduire que

$$p_{n+1} = \frac{1}{2 - p_n} \quad \text{et} \quad q_{n+1} = \frac{q_n}{1 + q_n}.$$

2. Déterminer  $q_n$  en fonction de  $n$  et  $q_0$ . Quel est le sens de variation et la limite de  $(q_n)_{n \geq 0}$  ?

Avec ce modèle, que suffirait-il de faire pour voir disparaître une maladie autosomique récessive ? Cela vous paraît-il crédible ?

#### C. Transmission avec mutation

On suppose désormais que des mutations entraînent l'altération d'une proportion  $k$  d'allèles  $A$  en allèles  $a$ . Ces mutations (dites germinales) ont lieu à la formation des gamètes.

1. On fixe  $n \in \mathbb{N}$ . On conserve les notations de la question B. 1.

a) On note  $p'_n$  et  $q'_n$  les fréquences des allèles  $A$  et  $a$  dans les gamètes des individus de la génération  $n$ . Justifier que

$$p'_n = (1 - k)p_n \quad \text{et} \quad q'_n = q_n + kp_n$$

et en déduire les expressions de  $P(Aa)$  et  $P(aa)$  en fonction de  $k$ ,  $p_n$  et  $q_n$ .

---

(†) En fait, pour de nombreuses maladies génétiques, les progrès de la médecine permettent aujourd'hui aux malades d'atteindre l'âge adulte et d'avoir des enfants. Pour la mucoviscidose par exemple, l'espérance de vie est passée de 5 ans dans les années 50 à plus de 47 ans aujourd'hui.

b) Établir que

$$b_{n+1} = \frac{2(1-k)p_n(q_n + kp_n)}{1 - (q_n + kp_n)^2},$$

puis démontrer que l'on peut simplifier cette fraction par  $(1-k)p_n$ .

c) En déduire que

$$q_{n+1} = \frac{k + (1-k)q_n}{1 + k + (1-k)q_n}.$$

2. On pose

$$f(x) = \frac{k + (1-k)x}{1 + k + (1-k)x}.$$

a) Soient  $x, y \in \mathbb{R}$ . Démontrer que

$$\frac{f(x) - f(y)}{x - y} = \frac{1 - k}{(1 + k + (1-k)x)(1 + k + (1-k)y)},$$

lorsque les termes de cette relation sont définis.

b) Démontrer que les points fixes de la fonction  $f$  sont les nombres

$$\alpha = \frac{\sqrt{k}}{1 + \sqrt{k}} \quad \text{et} \quad \beta = -\frac{\sqrt{k}}{1 - \sqrt{k}}.$$

c) À l'aide de la suite  $(u_n)_{n \geq 0}$  de terme général

$$u_n = \frac{q_n - \alpha}{q_n - \beta}.$$

démontrer que la suite  $(q_n)_{n \geq 0}$  tend vers  $\alpha$ .

Application : La mucoviscidose est une maladie suffisamment ancienne pour que l'on puisse considérer que les 4% de porteurs sains correspondent à la limite de la suite  $(b_n)_{n \geq 0}$ . En déduire la valeur de  $k$  pour cette maladie.