

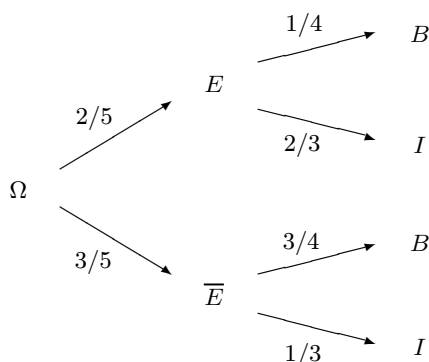
# PROBABILITÉS FINIES

## CORRECTION

### Exercice 1

Considérons les événements :

- $E$  : « il y a une embuscade »
- $B$  : « le blanc annonce une embuscade »
- $I$  : « l'indien annonce une embuscade »



Alors

$$\begin{aligned}
 P(E|\bar{B} \cap I) &= \frac{P(E \cap \bar{B} \cap I)}{P(\bar{B} \cap I)} && \text{par définition} \\
 &= \frac{P(E)P(\bar{B} \cap I|E)}{P(E)P(\bar{B} \cap I|E) + P(\bar{E})P(\bar{B} \cap I|\bar{E})} && \text{d'après la formule des probas totales à travers le s.c.e. } (E, \bar{E}) \\
 &= \frac{P(E)P(\bar{B}|E)P(I|E)}{P(E)P(\bar{B}|E)P(I|E) + P(\bar{E})P(\bar{B}|\bar{E})P(I|\bar{E})} && \text{car } \bar{B} \text{ et } I \text{ sont indépendants sachant } E \text{ et sachant } \bar{E} \\
 &= \frac{\frac{2}{5} \times \frac{3}{4} \times \frac{2}{3}}{\frac{2}{5} \times \frac{3}{4} \times \frac{2}{3} + \frac{3}{5} \times \frac{1}{4} \times \frac{1}{3}} \\
 &= \frac{4}{5},
 \end{aligned}$$

donc

la probabilité qu'il y ait une embuscade est égale à  $\frac{4}{5}$ .

### Exercice 2

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 qu'il est porté par un chromosome non sexuel. Chaque individu peut donc être 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 fait l'hypothèse que la population est idéale, autrement dit que : la population est de grande taille (ce qui permet de négliger les phénomènes de dérive génétique) ; il n'y a pas de sélection naturelle : les deux

allèles  $A$  et  $a$  ne confèrent ni avantage ni désavantage sélectif ; il n'y a pas de sélection comportementale : la population est panmictique, c'est-à-dire que les couples se forment indépendamment de leur génotype ; il n'y a pas de migration : aucune copie allélique n'est apportée de l'extérieur ; il n'y a pas de croisement entre des générations différentes ; on néglige les mutations.

On peut alors énoncer 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. L'objectif de cet exercice est démontrer cette loi.

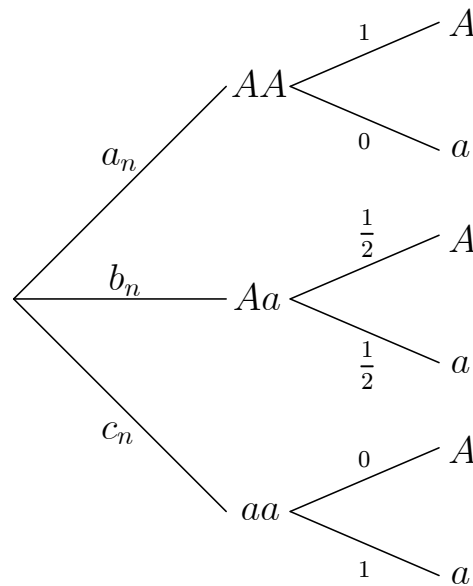
Dans la  $n$ -ème génération, les génotypes  $AA$ ,  $Aa$  et  $aa$  ont pour fréquences respectives  $a_n$ ,  $b_n$  et  $c_n$ . En outre, on note  $p_n$  et  $q_n$  les fréquences des allèles  $A$  et  $a$  dans cette  $n$ -ème génération.

On essaiera de donner des noms les plus intuitifs possibles aux événements utilisés et on ne se privera pas de faire des arbres de probabilités, si besoin.

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

Soit  $n \in \mathbb{N}$ . Dans la  $n$ -ème génération, on choisit au hasard un individu puis l'un de ses allèles. On note  $A$  et  $a$  les événements « l'allèle choisi est  $A$  » et « l'allèle choisi est  $a$  ». Par ailleurs, on note  $AA$ ,  $Aa$  et  $aa$  les événements « le génotype de l'individu est  $AA$  », « le génotype de l'individu est  $Aa$  » et « le génotype de l'individu est  $aa$  ».

Les hypothèses de l'énoncé se traduisent alors par l'arbre suivant :



D'après la formule des probabilités totales appliquée à travers le système complet d'événements  $(AA, Aa, aa)$ , on a

$$P(A) = P(AA)P(A | AA) + P(Aa)P(A | Aa) + P(aa)P(A | aa).$$

Il est clair que

$$P(A | AA) = 1 \quad \text{et} \quad P(A | aa) = 0.$$

Par ailleurs, le choix de l'un ou l'autre allèle étant équiprobable, on a

$$P(A | Aa) = \frac{1}{2}.$$

Donc

$$P(A) = P(AA) + \frac{1}{2}P(Aa).$$

Or  $P(A) = p_n$ ,  $P(AA) = a_n$  et  $P(Aa) = b_n$ , donc

$$\boxed{\forall n \in \mathbb{N}^*, \quad p_n = a_n + \frac{1}{2}b_n.}$$

De même, on démontre que

$$\boxed{\forall n \in \mathbb{N}^*, \quad q_n = \frac{1}{2}b_n + c_n.}$$

2. Démontrer que, pour tout  $n \in \mathbb{N}^*$ ,  $a_{n+1} = p_n^2$ ,  $b_{n+1} = 2p_nq_n$  et  $c_{n+1} = q_n^2$ .

Pour tout  $n \in \mathbb{N}^*$ , dans la  $n$ -ème génération, on considère un couple et l'un de ses enfants (qui appartient donc à la  $n+1$ -ème génération).

Pour tout  $n \in \mathbb{N}^*$ , on a

$$\begin{aligned} a_{n+1} &= P(\text{« l'enfant est } AA \text{ »}) \\ &= P(\text{« le père transmet } A \text{ »} \cap \text{« la mère transmet } A \text{ »}) \\ &= P(\text{« le père transmet } A \text{ »})P(\text{« la mère transmet } A \text{ »}) \quad \text{par } \perp \\ &= p_n \times p_n, \end{aligned}$$

c'est-à-dire

$$\boxed{\forall n \in \mathbb{N}^*, \quad a_{n+1} = p_n^2.}$$

De même, on démontre que, pour tout  $n \in \mathbb{N}^*$ ,

$$\boxed{\forall n \in \mathbb{N}^*, \quad c_{n+1} = q_n^2.}$$

Enfin, on a

$$\begin{aligned} b_{n+1} &= P(\text{« l'enfant est } Aa \text{ »}) \\ &= P(\text{« le père transmet } A \text{ »} \cap \text{« la mère transmet } a \text{ »}) \\ &\quad + P(\text{« le père transmet } a \text{ »} \cap \text{« la mère transmet } A \text{ »}) \\ &= P(\text{« le père transmet } A \text{ »})P(\text{« la mère transmet } a \text{ »}) \\ &\quad + P(\text{« le père transmet } a \text{ »})P(\text{« la mère transmet } A \text{ »}) \quad \text{par } \perp \\ &= p_nq_n + q_np_n, \end{aligned}$$

ce qui donne

$$\boxed{\forall n \in \mathbb{N}^*, \quad b_{n+1} = 2p_nq_n.}$$

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.

Pour tout  $n \in \mathbb{N}^*$ , on a

$$p_{n+1} = a_{n+1} + \frac{1}{2}b_{n+1} = p_n^2 + \frac{1}{2}2p_nq_n = p_n(p_n + q_n) = p_n$$

car  $p_n + q_n = 1$ . De même, pour tout  $n \in \mathbb{N}^*$ , on a

$$q_{n+1} = q_n.$$

Les suites  $(p_n)$  et  $(q_n)$  sont donc constantes, ce qui démontre que

dans une population idéale, les fréquences alléliques d'un gène sont constantes au fil des générations.

Note d'un biologiste :

Assez rarement réalisée dans la nature, la loi de Hardy-Weinberg constitue l'une des bases de la biologie évolutive moderne.

Elle énonce donc que sous certaines conditions (les conditions de Hardy-Weinberg : population infinie, panmixie, etc. . . ), une population n'évolue pas, ses fréquences alléliques restant constantes à travers les générations successives. Il faut bien comprendre que cette stabilité est très forte : la ségrégation aléatoire (lois de Mendel) des allèles au cours de la reproduction, les « brassages génétiques » n'y font rien : les fréquences ne changent pas si la population est « idéale ».

C'est ainsi lorsque la loi de Hardy-Weinberg n'est pas respectée que le biologiste est interpellé. Seule une (ou des) violation(s) des conditions de Hardy-Weinberg permet(tent) de s'écarter de la stabilité.

La marche à suivre est alors simple : On relève les fréquences des génotypes ( $AA$ ,  $aa$ ,  $Aa$ , etc. . . ). On en déduit les fréquences d'allèles. On calcule les nouvelles fréquences génotypiques selon la loi de Hardy-Weinberg et on compare avec les fréquences relevées. Si il y a égalité, c'est qu'on a affaire à un système stable et à une population que l'on peut considérer comme idéale. Si il n'y a pas égalité, nous avons une population qui fluctue, et il faut commencer à chercher les raisons de cette fluctuation. Ce pourra être la mise en évidence d'une consanguinité, d'une sélection (naturelle ou artificielle) de certains génotypes, de mutations, d'un phénomène de dérive génétique . . .

### Exercice 3

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 la présence de deux allèles mutés est nécessaire 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. Pour cette raison, on considérera 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  $a_n$ ,  $b_n$ .

Soit  $n \in \mathbb{N}$ . Dans la  $n$ -ème génération, on choisit au hasard un individu puis l'un de ses allèles. On note  $A$  et  $a$  les événements « l'allèle choisi est  $A$  » et « l'allèle choisi est  $a$  ». Par ailleurs, on note  $Hom$  et  $Het$  les événements « l'individu est homozygote  $AA$  » et « l'individu est hétérozygote  $Aa$  ».

D'après la formule des probabilités totales appliquée à l'événement  $A$  à travers le système complet d'événements  $(Hom, Het)$ , on a

$$P(A) = P(Hom)P(A | Hom) + P(Het)P(A | Het).$$

Il est clair que

$$P(A | Hom) = 1.$$

Par ailleurs, le choix de l'un ou l'autre allèle étant équiprobable, on a

$$P(A | Het) = \frac{1}{2}.$$

Donc

$$P(A) = P(Hom) + \frac{1}{2}P(Het).$$

Or  $P(A) = p_n$ ,  $P(Hom) = a_n$  et  $P(Het) = b_n$ , donc

$$\boxed{\forall n \in \mathbb{N}, \quad p_n = a_n + \frac{1}{2}b_n.}$$

Pour déterminer  $q_n$ , on peut adapter le raisonnement ci-dessus ou alors on peut utiliser le fait que  $\forall n \in \mathbb{N}$ ,  $p_n + q_n = 1$ . Cela donne, pour tout  $n \in \mathbb{N}$ ,

$$q_n = 1 - p_n = 1 - a_n - \frac{1}{2}b_n = b_n - \frac{1}{2}b_n,$$

c'est-à-dire

$$\boxed{\forall n \in \mathbb{N}, \quad q_n = \frac{1}{2}b_n.}$$

B. Transmission sans mutation

1. On fixe  $n \in \mathbb{N}$ . Dans la génération  $n$ , un couple 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 : Dans le cas de la mucoviscidose, on sait que 4% des adultes occidentaux sont porteurs sains. Calculer la probabilité d'avoir un enfant malade pour un couple d'occidentaux.

On a

$$\begin{aligned} P(AA) &= P(\text{« le père transmet } A \text{ »} \cap \text{« la mère transmet } A \text{ »}) \\ &= P(\text{« le père transmet } A \text{ »})P(\text{« la mère transmet } A \text{ »}) \quad \text{par } \perp\!\!\!\perp \\ &= p_n \times p_n, \end{aligned}$$

c'est-à-dire

$$\boxed{P(AA) = p_n^2.}$$

De même, on démontre que

$$P(aa) = q_n^2.$$

Enfin, on a

$$\begin{aligned} P(Aa) &= P(\text{« le père transmet } A \text{ »} \cap \text{« la mère transmet } a \text{ »}) \\ &\quad + P(\text{« le père transmet } a \text{ »} \cap \text{« la mère transmet } A \text{ »}) \\ &= P(\text{« le père transmet } A \text{ »})P(\text{« la mère transmet } a \text{ »}) \\ &\quad + P(\text{« le père transmet } a \text{ »})P(\text{« la mère transmet } A \text{ »}) \quad \text{par } \perp\!\!\!\perp \\ &= p_n q_n + q_n p_n, \end{aligned}$$

ce qui donne

$$P(Aa) = 2p_n q_n.$$

Application : La donnée de l'énoncé, appliquée pour la  $n$ -ème génération, nous dit que  $b_n = 0,04$  c'est-à-dire  $q_n = 0,02$ . Dans ce cas, on a donc  $P(aa) = q_n^2 = (0,02)^2 = 0,0004$  ce qui signifie que

l'incidence de la mucoviscidose chez les occidentaux est d'une naissance sur 2500.

- b) À l'aide de probas conditionnelles, établir que  $a_{n+1} = p_n^2 / (1 - q_n^2)$  et  $b_{n+1} = 2p_n q_n / (1 - q_n^2)$  puis simplifier ces relations en tenant compte du fait que  $p_n + q_n = 1$ .

Lorsqu'on s'intéresse à un individu adulte de la  $(n+1)$ -ème génération, on sait que celui-ci ne porte pas le génotype létal  $aa$ . La probabilité qu'il soit de génotype  $AA$  est donc la probabilité qu'il ait été un enfant de génotype  $AA$  sachant que  $aa$  ne s'est pas produit. Autrement dit, on a

$$a_{n+1} = P(AA | \overline{aa}).$$

Le calcul de cette probabilité conditionnelle nous donne alors

$$\begin{aligned} a_{n+1} &= \frac{P(AA \cap \overline{aa})}{P(\overline{aa})} \quad \text{par déf.} \\ &= \frac{P(AA)}{1 - P(aa)} \quad \text{car } AA \text{ implique } \overline{aa} \\ &= \frac{p_n^2}{1 - q_n^2} \quad \text{d'après a).} \end{aligned}$$

De la même façon, on a

$$b_{n+1} = P(Aa | \overline{aa}),$$

d'où

$$\begin{aligned} b_{n+1} &= \frac{P(Aa \cap \overline{aa})}{P(\overline{aa})} \quad \text{par déf.} \\ &= \frac{P(Aa)}{1 - P(aa)} \quad \text{car } Aa \text{ implique } \overline{aa} \\ &= \frac{2p_n q_n}{1 - q_n^2} \quad \text{d'après a).} \end{aligned}$$

On a donc bien

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

Comme  $p_n + q_n = 1$ , on constate alors que

$$a_{n+1} = \frac{p_n^2}{1 - q_n^2} = \frac{p_n^2}{(1 - q_n)(1 + q_n)} = \frac{p_n^2}{p_n(1 + q_n)} = \frac{p_n}{1 + q_n}$$

et

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

En conclusion

$$a_{n+1} = \frac{p_n}{1 + q_n} \quad \text{et} \quad b_{n+1} = \frac{2q_n}{1 + q_n}.$$

- c) En déduire que  $p_{n+1} = 1/(2 - p_n)$  et  $q_{n+1} = q_n/(1 + q_n)$ .

En combinant les résultats de la question A et de la question précédente, on obtient

$$p_{n+1} = a_{n+1} + \frac{1}{2}b_{n+1} = \frac{p_n}{1 + q_n} + \frac{q_n}{1 + q_n} = \frac{p_n + q_n}{1 + q_n} = \frac{1}{1 + q_n} = \frac{1}{2 - p_n}$$

et

$$q_{n+1} = \frac{1}{2}b_{n+1} = \frac{q_n}{1 + q_n}.$$

Donc

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

Remarque : On peut aussi déterminer  $q_{n+1}$  et en déduire  $p_{n+1}$  car  $p_{n+1} = 1 - q_{n+1}$ .

2. À l'aide de la suite  $(1/q_n)_{n \geq 0}$ , déterminer l'expression de  $q_n$  en fonction de  $n$  et  $q_0$ . Quel est le sens de variation et la limite de la suite  $(q_n)_{n \geq 0}$  ? Avec ce modèle, que suffirait-il de faire pour voir disparaître une maladie autosomique récessive ? Cela vous paraît-il crédible ?

La relation de récurrence démontrée à la question précédente dit que

$$\forall n \in \mathbb{N}, \quad \frac{1}{q_{n+1}} = \frac{1 + q_n}{q_n} = \frac{1}{q_n} + 1,$$

ce qui signifie que la suite  $(1/q_n)_{n \geq 0}$  est arithmétique de raison 1. Il s'ensuit que, pour tout  $n \in \mathbb{N}$ , on a

$$\forall n \in \mathbb{N}, \quad \frac{1}{q_n} = \frac{1}{q_0} + n,$$

c'est-à-dire

$$\boxed{\forall n \in \mathbb{N}, \quad q_n = \frac{q_0}{1 + nq_0}.}$$

On constate alors que

$$\boxed{\text{la suite } (q_n)_{n \geq 0} \text{ tend vers 0 en décroissant.}}$$

On en déduit qu'

$$\boxed{\text{avec ce modèle, il suffit d'attendre pour qu'une maladie autosomique récessive disparaisse.}}$$

Ce modèle n'est pas réaliste car si la proportion d'allèle muté  $a$  ne peut que décroître, il est impossible que cette mutation ait pu apparaître et perdurer. Par conséquent,

$$\boxed{\text{le modèle développé dans cette partie n'est pas crédible.}}$$

### 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 au moment de 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  $p'_n = (1 - k)p_n$  et  $q'_n = q_n + kp_n$  et en déduire les expressions de  $P(Aa)$  et  $P(aa)$  en fonction de  $k$ ,  $p_n$  et  $q_n$ .

La mutation d'une proportion  $k$  d'allèles  $A$  en allèles  $a$  diminue de  $kp_n$  la fréquence de l'allèle  $A$  et augmente celle de l'allèle  $a$  de la même quantité, donc

$$\boxed{p'_n = p_n - kp_n \quad \text{et} \quad q'_n = q_n + kp_n.}$$

On peut alors reprendre le raisonnement de la question B. 1. a), ce qui donne

$$P(Aa) = 2p'_nq'_n \quad \text{et} \quad P(aa) = q_n'^2$$

ou encore

$$\boxed{P(Aa) = 2(1 - k)p_n(q_n + kp_n) \quad \text{et} \quad P(aa) = (q_n + kp_n)^2.}$$

- b) Établir que  $b_{n+1} = (2(1-k)p_n(q_n + kp_n))/(1 - (q_n + kp_n)^2)$ , puis démontrer que l'on peut simplifier cette fraction par  $(1-k)p_n$ .

On reprend cette fois le raisonnement de la question B.1. b), ce qui donne

$$\begin{aligned} b_{n+1} &= \frac{P(Aa \cap \overline{aa})}{P(\overline{aa})} && \text{par déf.} \\ &= \frac{P(Aa)}{1 - P(aa)} && \text{car } Aa \text{ implique } \overline{aa}, \end{aligned}$$

d'où, d'après a),

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

On constate que

$$\begin{aligned} b_{n+1} &= \frac{2(1-k)p_n(q_n + kp_n)}{(1 - q_n - kp_n)(1 + q_n + kp_n)} \\ &= \frac{2(1-k)p_n(q_n + kp_n)}{(p_n - kp_n)(1 + q_n + kp_n)} \\ &= \frac{2(1-k)p_n(q_n + kp_n)}{(1-k)p_n(1 + q_n + kp_n)}, \end{aligned}$$

ce qui donne, après simplification par  $(1-k)p_n$ ,

$$b_{n+1} = \frac{2(q_n + kp_n)}{1 + q_n + kp_n}.$$

- c) En déduire que  $q_{n+1} = (k + (1-k)q_n)/(1 + k + (1-k)q_n)$ .

Comme  $q_{n+1} = b_{n+1}/2$  d'après A, on a

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

c'est-à-dire

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

2. On pose  $f(x) = (k + (1-k)x)/(1 + k + (1-k)x)$ .

- a) Soient  $x, y \in \mathbb{R}$ . Calculer, en fonction de  $k$ , la quantité  $(f(x) - f(y))/(x - y)$  lorsque celle-ci est définie.

On a

$$\begin{aligned} &\frac{f(x) - f(y)}{x - y} \\ &= \frac{1}{x - y} \left( \frac{k + (1-k)x}{1 + k + (1-k)x} - \frac{k + (1-k)y}{1 + k + (1-k)y} \right) \\ &= \frac{(k + (1-k)x)(1 + k + (1-k)y) - (k + (1-k)y)(1 + k + (1-k)x)}{(x - y)(1 + k + (1-k)x)(1 + k + (1-k)y)} \\ &= \frac{k(1+k) + k(1-k)y + (1-k^2)x + (1-k)^2xy - k(1+k) - k(1-k)x - (1-k^2)y - (1-k)^2xy}{(x - y)(1 + k + (1-k)x)(1 + k + (1-k)y)} \\ &= \frac{(1-k^2)(x - y) - k(1-k)(x - y)}{(x - y)(1 + k + (1-k)x)(1 + k + (1-k)y)} \\ &= \frac{1 - k^2 - k(1-k)}{(1 + k + (1-k)x)(1 + k + (1-k)y)}, \end{aligned}$$

donc

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

- b) Démontrer que les points fixes de  $f$  sont les nombres  $\alpha = \sqrt{k}/(1 + \sqrt{k})$  et  $\beta = -\sqrt{k}/(1 - \sqrt{k})$ . Parmi ces points fixes, lequel est une limite possible de la suite  $(q_n)_{n \geq 0}$  ?

Les points fixes de la fonction  $f$  sont les nombres réels  $\ell$  qui sont solutions de l'équation

$$\begin{aligned}\ell &= f(\ell) \\ \iff \ell &= \frac{k + (1 - k)\ell}{1 + k + (1 - k)\ell} \\ \iff (1 + k)\ell + (1 - k)\ell^2 &= k + (1 - k)\ell \\ \iff (1 - k)\ell^2 + 2k\ell - k &= 0.\end{aligned}$$

Le discriminant de cette équation du second degré vaut  $\Delta = 4k^2 + 4(1 - k)k = 4k > 0$ , donc ses solutions sont les deux nombres réels

$$\ell_1 = \frac{-2k + 2\sqrt{k}}{2(1 - k)} = \frac{2\sqrt{k}(1 - \sqrt{k})}{2(1 - k)} = \frac{\sqrt{k}}{1 + \sqrt{k}}$$

et

$$\ell_2 = \frac{-2k - 2\sqrt{k}}{2(1 - k)} = -\frac{2\sqrt{k}(1 + \sqrt{k})}{2(1 - k)} = -\frac{\sqrt{k}}{1 - \sqrt{k}}.$$

Donc

les points fixes de la fonction  $f$  sont les nombres  $\alpha = \frac{\sqrt{k}}{1 + \sqrt{k}}$  et  $\beta = -\frac{\sqrt{k}}{1 - \sqrt{k}}$ .

- c) À l'aide de la suite  $(u_n)_{n \geq 0}$  de terme général  $u_n = (q_n - \alpha)/(q_n - \beta)$ , 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. Les mutations permettent-elles d'expliquer la persistance de cette maladie ?

Pour tout  $n \in \mathbb{N}$ , on a

$$\begin{aligned}u_{n+1} &= \frac{q_{n+1} - \alpha}{q_{n+1} - \beta} \\ &= \frac{f(q_n) - f(\alpha)}{f(q_n) - f(\beta)} \quad \begin{array}{l} \text{car } q_{n+1} = f(q_n) \\ \text{et } \alpha \text{ et } \beta \text{ sont fixes} \end{array} \\ &= \frac{f(q_n) - f(\alpha)}{f(q_n) - f(\beta)} \times \frac{q_n - \alpha}{q_n - \beta} \\ &= \frac{1 - k}{\frac{(1 + k + (1 - k)q_n)(1 + k + (1 - k)\alpha)}{1 - k}} \times u_n \\ &= \frac{1 + k + (1 - k)\beta}{1 + k + (1 - k)\alpha} \times u_n.\end{aligned}$$

Or

$$r = \frac{1 + k + (1 - k)\beta}{1 + k + (1 - k)\alpha} = \frac{1 + k - (1 - k)\frac{\sqrt{k}}{1 - \sqrt{k}}}{1 + k + (1 - k)\frac{\sqrt{k}}{1 + \sqrt{k}}} = \frac{1 + k - (1 + \sqrt{k})\sqrt{k}}{1 + k + (1 - \sqrt{k})\sqrt{k}} = \frac{1 - \sqrt{k}}{1 + \sqrt{k}},$$

donc

la suite  $(u_n)_{n \geq 0}$  est géométrique de raison  $\frac{1 - \sqrt{k}}{1 + \sqrt{k}}$ .

Comme

$$0 < \frac{1 - \sqrt{k}}{1 + \sqrt{k}} < 1,$$

on en déduit que

la suite  $(u_n)_{n \geq 0}$  tend vers 0.

Pour tout  $n \in \mathbb{N}$ , on a

$$q_n - \alpha = u_n(q_n - \beta).$$

Or  $(q_n)_{n \geq 0}$  reste bornée entre 0 et 1 (puisque les nombres  $q_n$  sont des probabilités) donc la suite  $(q_n - \beta)_{n \geq 0}$  est également bornée. Comme le produit d'une suite bornée par une suite qui tend vers 0 tend vers 0, on en déduit que la suite  $(q_n - \alpha)_{n \geq 0}$  tend vers 0, ce qui signifie que

$$\boxed{\text{la suite } (q_n)_{n \geq 0} \text{ tend vers } \alpha = \frac{\sqrt{k}}{1 + \sqrt{k}}.}$$

Application : L'énoncé nous dit que  $(b_n)_{n \geq 0}$  tend vers 0,04. Comme  $\forall n \in \mathbb{N}$ ,  $q_n = b_n/2$  d'après A, on en déduit que  $(q_n)_{n \geq 0}$  tend vers 0,02. On a donc

$$\begin{aligned} \alpha &= 0,02 \\ \iff \frac{\sqrt{k}}{1 + \sqrt{k}} &= 0,02 \\ \iff \sqrt{k} &= \frac{0,02}{1 - 0,02} \\ \iff \sqrt{k} &= \frac{1}{49} \\ \iff k &= \frac{1}{49^2}, \end{aligned}$$

ce qui donne

$$\boxed{k \approx 4,2 \times 10^{-4}.}$$