

Correction du devoir d'analyse et probabilités

Problème 1 :

Partie A : On appelle G_k l'évènement : « le lion a mangé une gazelle au k -ième repas observé » et Z_k l'évènement : « il a mangé un zèbre au k -ième repas observé ».

1. Probabilité pour que, lors des deux premiers repas observés, le lion ait mangé deux gazelles (et donc pas de zèbre) ?

On a $\mathbb{P}(G_1 \cap G_2) = \mathbb{P}(G_1) \cdot \mathbb{P}(G_2) = \frac{2}{3} \cdot \frac{2}{3}$ car les évènements sont indépendants.

Conclusion : $\mathbb{P}(G_1 \cap G_2) = \frac{4}{9}$.

2. Quelle est la probabilité pour que, lors des trois premiers repas, il ait mangé dans cet ordre, un zèbre puis deux gazelles ?

On a $\mathbb{P}(Z_1 \cap G_2 \cap G_3) = \mathbb{P}(Z_1) \cdot \mathbb{P}(G_2) \cdot \mathbb{P}(G_3) = \frac{1}{3} \cdot \frac{2}{3} \cdot \frac{2}{3}$ car les évènements sont indépendants.

Conclusion : $\mathbb{P}(G_1 \cap G_2) = \frac{4}{27}$.

3. Sur les quatre premiers repas observés, on considère l'évènement E : « il a mangé une gazelle deux fois de suite, pour la première fois au troisième et quatrième repas ».

$\mathbb{P}_{G_1}(E) = \frac{\mathbb{P}(E \cap G_1)}{\mathbb{P}(G_1)} = \frac{\mathbb{P}(G_1 \cap Z_2 \cap G_3 \cap G_4)}{\mathbb{P}(G_1)} = \mathbb{P}(Z_2) \cdot \mathbb{P}(G_3) \cdot \mathbb{P}(G_4) = \frac{1}{3} \left(\frac{2}{3}\right)^2$ car les évènements sont indépendants.

De même : $\mathbb{P}_{Z_1}(E) = \frac{\mathbb{P}(E \cap Z_1)}{\mathbb{P}(Z_1)} = \frac{\mathbb{P}(Z_1 \cap Z_2 \cap G_3 \cap G_4)}{\mathbb{P}(Z_1)} = \mathbb{P}(Z_2) \cdot \mathbb{P}(G_3) \cdot \mathbb{P}(G_4) = \frac{1}{3} \left(\frac{2}{3}\right)^2$.

Or $\{G_1, Z_1\}$ forme un système complet d'évènements, donc d'après la formule des probabilités totales :

$$\mathbb{P}(E) = \mathbb{P}_{G_1}(E) \cdot \mathbb{P}(G_1) + \mathbb{P}_{Z_1}(E) \cdot \mathbb{P}(Z_1) = \frac{4}{27} \cdot \frac{2}{3} + \frac{4}{27} \cdot \frac{1}{3} = \frac{4}{27}$$

Partie B : On observe le lion sur une assez longue période. On désigne par Y la variable aléatoire prenant pour valeur le nombre de repas nécessaires pour que, pour la première fois, le lion ait mangé deux gazelles à deux repas consécutifs.

On note $u_n = \mathbb{P}(Y = n)$ pour tout entier n , $n \geq 2$ et on pose $u_1 = 0$.

1. $u_2 = \mathbb{P}(Y = 2) = \mathbb{P}(G_1 \cap G_2) = \frac{4}{9}$.

$$u_3 = \mathbb{P}(Y = 3) = \mathbb{P}(Z_1 \cap G_2 \cap G_3) = \frac{4}{27}$$

$$u_4 = \mathbb{P}(Y = 4) = \mathbb{P}(E) = \frac{4}{27}.$$

2. Exprimons $\mathbb{P}_{G_1}(Y = n + 2)$ et $\mathbb{P}_{Z_1}(Y = n + 2)$ à l'aide des termes de la suite (u_n) .

Sachant que le lion a mangé une gazelle lors du premier repas observé, l'évènement $(Y = n + 2)$ est réalisé si et seulement si un zèbre a été mangé au 2ième repas et l'évènement $(Y = n)$ a été réalisé, à savoir lors des n repas qui ont suivi, les deux premières gazelles consécutives mangées l'ont été pour

la première fois lors de la n^{ime} observation.
Soit, les évènements étant indépendants :

$$P_{G_1}(Y = n + 2) = \mathbb{P}(Z_2 \cap (Y = n)) = \mathbb{P}(Z_2) \cdot \mathbb{P}(Y = n) = \frac{1}{3}u_n, n \geq 2$$

Sachant que le lion a mangé un zèbre lors du premier repas observé, l'évènement $(Y = n + 2)$ est réalisé si et seulement si l'évènement $(Y = n + 1)$ a été réalisé, à savoir si lors des $n + 1$ repas qui ont suivi le trucidage du zèbre, les deux premières gazelles consécutives mangées l'ont été pour la première fois lors de la $(n + 1)^{ime}$ observation.

Soit, les évènements étant indépendants :

$$P_{Z_1}(Y = n + 2) = \mathbb{P}(Y = n + 1) = u_{n+1}, n \geq 2$$

3. On applique la formule des probabilités totales en utilisant le système complet d'évènement $\{G_1, Z_1\}$.
Ce qui donne :

$$\begin{aligned} u_{n+2} &= \mathbb{P}(Y = n + 2) = \mathbb{P}_{G_1}(Y = n + 2)\mathbb{P}(G_1) + \mathbb{P}_{Z_1}(Y = n + 2)\mathbb{P}(Z_1) \\ &= \frac{1}{3}u_n \frac{2}{3} + u_{n+1} \frac{1}{3} \\ &= \frac{1}{3}u_{n+1} + \frac{2}{9}u_n, \forall n \geq 2 \end{aligned}$$

Vérifions que cette relation est encore vraie pour $n = 1$: Par hypothèse $u_1 = 0$.

On a donc bien $\frac{1}{3}u_2 + \frac{2}{9}u_1 = \frac{1}{3} \cdot \frac{4}{9} = \frac{4}{27} = u_3$

Conclusion :
$$u_{n+2} = \frac{1}{3}u_{n+1} + \frac{2}{9}u_n, \forall n \geq 1$$

4. En déduire l'expression de u_n en fonction de n pour tout $n \geq 1$:

On passe par l'équation caractéristique : (E) $r^2 - \frac{1}{3}r - \frac{2}{9} = 0$

$$\Delta = \frac{1}{9} + \frac{8}{9} = 1.$$

Donc (E) admet deux racines réelles distinctes : $r_1 = -\frac{1}{3}$ et $r_2 = \frac{2}{3}$

On peut donc assurer que : $\exists(\lambda, \mu) \in \mathbb{R}^2 / u_n = \lambda(-\frac{1}{3})^n + \mu(\frac{2}{3})^n, \forall n \in \mathbb{N}^*$

Déterminons λ et μ grâce aux valeurs de u_1 et de u_2 :

$$\begin{cases} u_1 = -\frac{\lambda}{3} + \frac{2\mu}{3} \\ u_2 = \frac{\lambda}{9} + \frac{4}{9}\mu = \frac{4}{9} \end{cases} \Leftrightarrow \begin{cases} \lambda = 2\mu \\ \frac{6\mu}{9} = \frac{4}{9} \end{cases} \begin{pmatrix} L_1 \\ L_2 + \frac{L_1}{3} \end{pmatrix} \Leftrightarrow \begin{cases} \lambda = \frac{4}{3} \\ \mu = \frac{2}{3} \end{cases}$$

Conclusion :
$$\forall n \geq 1, u_n = \frac{4}{3}(-\frac{1}{3})^n + \frac{2}{3}(\frac{2}{3})^n$$

5. Vérifions que la suite (u_n) définit bien une loi de probabilité :

a) Commençons par montrer que $u_n \geq 0$ pour tout n entier naturel strictement positif.

Etudions pour cela le signe de u_n (noté $sgn(u_n)$) :

$$sgn(u_n) = sgn\left(\frac{2}{3}(\frac{2}{3})^n[1 + 2(-\frac{1}{2})^n]\right) = sgn\left(1 + 2(-\frac{1}{2})^n\right)$$

— Si n est pair, il est clair que cette expression est positive.

— Si n est impair, alors

$$\operatorname{sgn}(u_n) = \operatorname{sgn}\left(1 - \frac{2}{2^n}\right) = \operatorname{sgn}\left(1 - \frac{1}{2^{n-1}}\right) \geq 0 \text{ car } \left(\frac{1}{2}\right)^{n-1} \leq 1, \forall n \geq 1$$

Conclusion : $u_n \geq 0, \forall n \geq 1$

b) Montrons que $\sum u_n$ converge :

$$\sum_{n \geq 1} \left(\frac{2}{3}\right)^{n-1} = \sum_{n \geq 0} \left(\frac{2}{3}\right)^n \text{ converge car série géométrique de raison } q_1 = \frac{2}{3} \in]-1, 1[.$$

$$\sum_{n \geq 1} \left(-\frac{1}{3}\right)^{n-1} = \sum_{n \geq 0} \left(-\frac{1}{3}\right)^n \text{ converge car série géométrique de raison } q_1 = -\frac{1}{3} \in]-1, 1[.$$

Conclusion : $\sum u_n$ converge comme combinaison linéaire de deux séries convergentes

c) Montrons que $\sum_{n=1}^{+\infty} u_n = 1$:

$$\begin{aligned} \sum_{n=1}^{+\infty} u_n &= \frac{4}{3} \cdot \frac{-1}{3} \sum_{n=1}^{\infty} \left(-\frac{1}{3}\right)^{n-1} + \frac{2}{3} \cdot \frac{2}{3} \sum_{n=1}^{\infty} \left(\frac{2}{3}\right)^{n-1} \\ &= \frac{-4}{9} \cdot \frac{1}{1 + \frac{1}{3}} + \frac{4}{9} \cdot \frac{1}{1 - \frac{2}{3}} = \frac{-4}{9} \cdot \frac{3}{4} + \frac{4}{9} \cdot 3 = \frac{-1}{3} + \frac{4}{3} = 1 \end{aligned}$$

Conclusion : La suite (u_n) définit bien une loi de probabilité.

6. Calculons, si elles existent, l'espérance et la variance de Y : $\sum_{n \geq 1} n \left(\frac{2}{3}\right)^{n-1}$ et $\sum_{n \geq 1} n \left(-\frac{1}{3}\right)^{n-1}$ convergent (séries géométriques dérivées de raisons comprises entre -1 et 1 . Donc $\mathbb{E}(Y)$ existe et vaut $\mathbb{E}(Y) = \left(\frac{2}{3}\right)^2 \sum_{n=1}^{\infty} n \left(\frac{2}{3}\right)^{n-1} + \frac{4}{3} \cdot \frac{-1}{3} \sum_{n=1}^{\infty} n \left(-\frac{1}{3}\right)^{n-1} = \frac{4}{9} \frac{1}{(1 - \frac{2}{3})^2} + \frac{-4}{9} \frac{1}{(1 + \frac{1}{3})^2} = \frac{15}{4}$

De même, la convergence de $\sum_{n \geq 1} n^2 \left(\frac{2}{3}\right)^{n-1}$ et $\sum_{n \geq 1} n^2 \left(-\frac{1}{3}\right)^{n-1}$ assurent l'existence de $\mathbb{E}(Y^2)$.

On obtient après calculs $\mathbb{E}(Y^2) = \frac{159}{8}$

Conclusion : $\mathbb{V}(Y)$ existe et vaut, d'après Koenig-Huygens : $\mathbb{V}(Y) = \mathbb{E}(Y^2) - \mathbb{E}^2(Y) = \frac{93}{16}$

Problème 2 :**Partie I :**

Pour tout p réel dans l'intervalle $]0, 1[$, on considère la suite récurrente $(v_n)_{n \in \mathbb{N}}$ définie par :

$$v_{n+1} = 1 - p + pv_n^2, \forall n \in \mathbb{N} \text{ et } v_0 \in]0, 1[$$

- ① Écrivons une fonction Python permettant de déterminer v_n en fonction de p et de n :

Analyse : Nous choisissons de passer les paramètres p et n en variable d'entrée et demandons à l'utilisateur de choisir le premier terme v_0 dans l'intervalle $]0, 1[$.

Puisqu'on nous demande de calculer v_n connaissant v_0 , il est nécessaire de faire n appels successifs à la relation de récurrence $v_{n+1} = 1 - p + pv_n^2$.

Nous utilisons donc une structure répétitive « Pour ».

```
def calculV(p,n):
    v0 = float(input('Donner le premier terme (0<v0<1) :'))
    for k in range(n):
        v = 1-p+p*v0**2
        v0 = v
    return v
```

- ② Soit $f : x \mapsto 1 - p + px^2$ telle que $v_{n+1} = f(v_n)$.

- a) Montrons que les solutions de l'équation $f(x) = x$ sont 1 et $\frac{1-p}{p}$:

$$f(x) = x \Leftrightarrow 1 - p + px^2 = x \Leftrightarrow px^2 - x + 1 - p = 0 \quad (*)$$

Rédaction 1 : Il s'agit d'un trinôme du second degré qui admet au plus deux racines réelles. On commence par noter que 1 est racine évidente. D'où :

$$f(x) = x \Leftrightarrow p \left(x^2 - \frac{1}{p}x + \frac{1-p}{p} \right) = p(x-1)(x-a) = 0$$

avec, par identification,

$$1 - p = pa \Leftrightarrow a = \frac{1-p}{p}$$

Soit $f(x) = x \Leftrightarrow p(x-1) \left(x - \frac{1-p}{p} \right) = 0$

En notant que $\frac{1-p}{p} = 1 \Leftrightarrow p = \frac{1}{2}$, on conclut que $f(x) = x$ admet deux racines réelles distinctes qui sont 1 et $\frac{1-p}{p}$ si $p \neq 1/2$ et une racine double égale à 1 si $p = 1/2$.

Rédaction 2 :

Ce trinôme du second degré a pour discriminant $\Delta = 1 - 4p(1-p) = 1 - 4p + 4p^2 = (2p-1)^2 \geq 0$.

— *Premier cas :* Si $p = 1/2$, $\Delta = 0$ et $f(x) = x$ admet $x_0 = 1 = \frac{1-p}{p}$ pour racine double.

(on peut vérifier que sous cette condition : $f(x) = x \Leftrightarrow \frac{1}{2}(x^2 - 2x + 1) = \frac{1}{2}(x-1)^2$)

— *Deuxième cas :* Si $p \in]0, 1[\setminus \{1/2\}$ alors $\Delta > 0$ et $f(x) = x$ admet deux racines réelles :

$$x_1 = \frac{1 - |2p - 1|}{2p} \text{ et } x_2 = \frac{1 + |2p - 1|}{2p}$$

avec $|2p - 1| = \begin{cases} 2p - 1 & \text{si } p > 1/2 \\ 1 - 2p & \text{si } p < 1/2 \end{cases}$.

Pour toute valeur de p dans $]0, 1[\setminus \{1/2\}$ les deux racines sont donc 1 et $\frac{1-p}{p}$.

Conclusion : les solutions de $f(x) = x$ sont 1 et $\frac{1-p}{p}$

b) Montrons que la restriction de f à l'intervalle $[0, 1]$ prend toutes ses valeurs dans l'intervalle $[0, 1]$: f est continue et dérivable sur $[0, 1]$ en tant que fonction polynôme.

$\forall x \in [0, 1], f'(x) = 2px > 0$ donc f est strictement croissante sur $[0, 1]$.

Donc $0 \leq x \leq 1 \Rightarrow f(0) = 1 - p \leq f(x) \leq f(1) = 1$ avec $1 - p > 0$ car $p \in]0, 1[$.

f étant continue sur $[0, 1]$, l'image d'un segment est un segment et on a : $f([0, 1]) = [1 - p, 1] \subset [0, 1]$.

Conclusion : $\forall x \in [0, 1], f(x) \in [0, 1]$

On en déduit que $v_n \in [0, 1], \forall n \in \mathbb{N}$ par récurrence. En effet :

➤ $v_0 \in [0, 1]$

➤ supposons que $v_n \in [0, 1]$ pour n fixé ($n \geq 0$).

➤ Alors $v_{n+1} = f(v_n) \in [0, 1]$ puisque $f([0, 1]) \subset [0, 1]$

➤ **Conclusion :** $\forall n \in \mathbb{N}, v_n \in [0, 1]$

③ Montrons que si $p \leq 1/2$ alors la suite (v_n) est croissante et convergente vers une limite qu'on déterminera :

— On commence par noter, comme demandé dans l'énoncé, que :

$$\frac{1-p}{p} \geq 1 \Leftrightarrow 1-p \geq p \Leftrightarrow 2p \leq 1 \Leftrightarrow p \leq \frac{1}{2}$$

— Nous venons de voir que $v_n \in I = [0, 1]$ pour tout entier naturel n .

Étudions le signe de $f(x) - x$ sur I :

D'après la question 2.a) nous savons que $f(x) - x = 0$ admet deux racines réelles 1 et $\frac{1-p}{p}$.

Dès lors, puisque $f(x) - x$ est du signe de p à l'extérieur de ses racines, il découle que :

$$f(x) - x \geq 0, \forall x \in [0, 1]$$

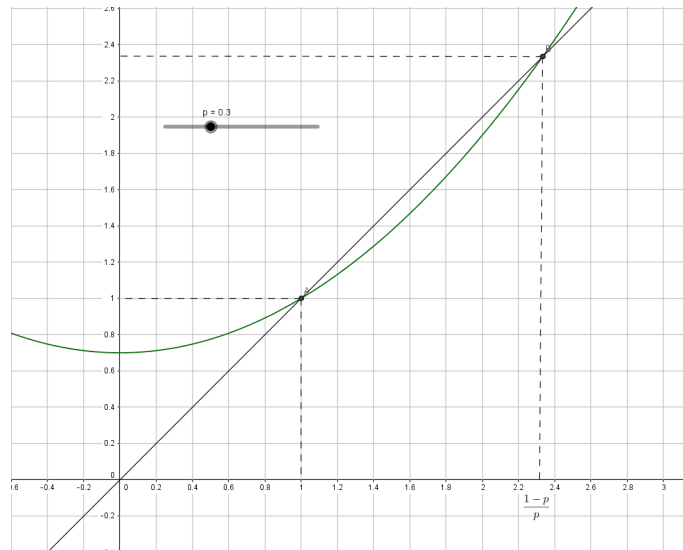
— Il suffit alors de prendre $x = u_n$ pour montrer que : $\forall n \in \mathbb{N}, v_{n+1} = f(v_n) \geq v_n$.

On en déduit que la suite (v_n) est croissante et majorée par 1. Elle converge par application du théorème de la limite monotone.

— Soit L sa limite. f étant continue sur $[0, 1]$, on a : $f(L) = L$ et donc $L = 1$ ou $L = \frac{1-p}{p}$.

Or on vient de voir que $\frac{1-p}{p} \geq 1$, donc la seule limite possible est $L = 1$.

Conclusion : (u_n) est croissante et converge vers $L = 1$

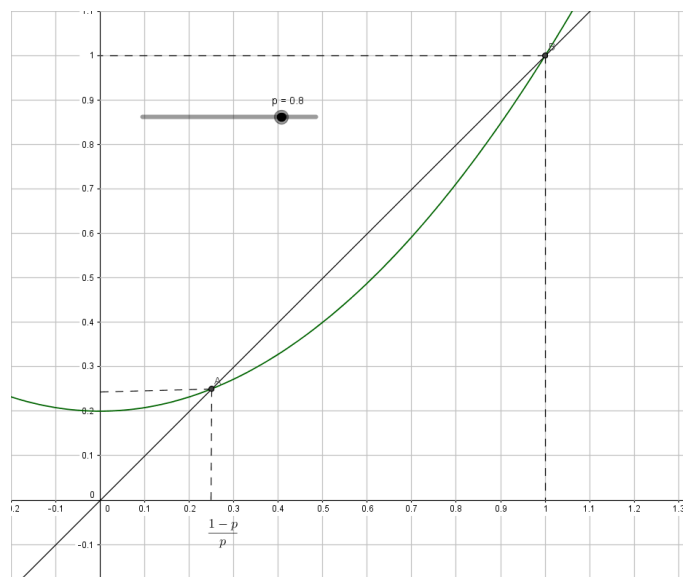


④ Que pouvons-nous dire dans le cas $p > 1/2$?

Si $p > 1/2$ alors $0 < \frac{1-p}{p} < 1$.

E comme $f(x) - x$, trinôme du second degré, est positif à l'extérieur de ses racines, on a cette fois :

$$f(x) - x \geq 0 \text{ si } x \leq \frac{1-p}{p} \text{ et } f(x) - x \leq 0 \text{ si } \frac{1-p}{p} < x < 1$$



— *Premier cas* : Supposons que $0 < v_0 < \frac{1-p}{p}$ et posons $J_1 =]0, \alpha[$ où $\alpha = \frac{1-p}{p}$.

f est croissante sur $I = [0, 1]$ et donc sur J_1 qui est un intervalle stable par f puisque :

$$x \in J_1 \Leftrightarrow 0 < x < \alpha \Rightarrow f(0) = 1-p \leq f(x) \leq f(\alpha) = \alpha$$

Dès lors, on montre par récurrence que $\forall n \in \mathbb{N}$, $v_n \in J_1$ et comme $f(x) \geq x$ sur J_1 , on a :

$$f(v_n) = v_{n+1} \geq v_n \quad \forall n \in \mathbb{N}.$$

On en déduit que (v_n) est une suite croissante et majorée par $\alpha = \frac{1-p}{p}$. Elle converge par application du théorème de la limite monotone vers $L \in J_1$ tel que $f(L) = L$.

Conclusion : $(v_n)_{n \geq 0}$ est croissante et converge vers $\alpha = \frac{1-p}{p}$

— *Deuxième cas :* Supposons que $\frac{1-p}{p} < v_0 < 1$ et posons $J_2 =]\alpha, 1[$ où $\alpha = \frac{1-p}{p}$.

f est croissante sur $I = [0, 1]$ et donc sur J_2 qui est aussi un intervalle stable par f puisque :

$$x \in J_2 \Leftrightarrow \alpha < x < 1 \Rightarrow f(\alpha) = \alpha \leq f(x) \leq f(1) = 1$$

Dès lors, $\forall n \in \mathbb{N}$, $v_n \in J_2$ avec $f(v_n) = v_{n+1} \leq v_n$.

On en déduit que (v_n) est une suite décroissante et minorée par $\alpha = \frac{1-p}{p}$. Elle converge par application du théorème de la limite monotone vers $L \in J_2$ tel que $f(L) = L$.

Conclusion : $(v_n)_{n \geq 0}$ est décroissante et converge vers $\alpha = \frac{1-p}{p}$

Partie II : Application

On considère une population de bactéries pour laquelle une bactérie a une probabilité p de donner 2 cellules filles avant de mourir, et une probabilité $q = 1 - p$ de mourir sans se reproduire.

Toutes les bactéries suivent la même loi et leurs reproductions sont considérées comme indépendantes les unes des autres.

Soit X_n la taille de la population à la n -ième génération. On suppose qu'il n'y a qu'une bactérie au début de l'expérience et donc que X_0 vaut 1 de façon certaine.

① Première génération :

a) Déterminons $X_1(\Omega)$ et $\mathbb{P}(X_1 = k)$ pour tout $k \in X_1(\Omega)$:

X_1 ne peut prendre que deux valeurs : 0 si la cellule meurt sans se reproduire et 2 si elle se divise en deux cellules filles. Dès lors, $X_1(\Omega) = \{0, 2\}$ et, d'après l'énoncé

$$\mathbb{P}(X_1 = 0) = 1 - p, \mathbb{P}(X_1 = 2) = p$$

b) Calculons l'espérance et la variance de X_1 . On utilise les formule qui nous sont rappelées :

$$\mathbb{E}(X_1) = \sum_{k \in X_1(\Omega)} k \mathbb{P}(X_1 = k) = 0 \mathbb{P}(X_1 = 0) + 2 \mathbb{P}(X_1 = 2) = 2p;$$

$$\mathbb{E}(X_1^2) = \sum_{k \in X_1(\Omega)} k^2 \mathbb{P}(X_1 = k) = 0^2 \mathbb{P}(X_1 = 0) + 2^2 \mathbb{P}(X_1 = 2) = 4p$$

$$\text{D'où } \mathbb{V}(X_1) = \mathbb{E}(X_1^2) - \mathbb{E}^2(X_1) = 4p - 4p^2 = 4p(1 - p).$$

Conclusion : $\mathbb{E}(X_1) = 2p$ et $\mathbb{V}(X_1) = 4p(1 - p)$

② On pose $u_n = \mathbb{P}(X_n = 0)$. Interprétons u_n et explicitons u_0 et u_1 :

$(x_n = 0)$ est l'événement : « Il n'y a plus de bactéries à la génération n ». Donc u_n est la probabilité d'extinction de la population de bactéries à partir de la n -ième génération.

En particulier :

$u_0 = \mathbb{P}(X_0 = 0) = 0$ car $(X_0 = 0)$ est un événement impossible.

$u_1 = \mathbb{P}(X_1 = 0) = 1 - p$ d'après les hypothèses.

③ a) Commentons les lignes 2 à 4 :

On crée une liste P formée de $n + 1$ zéros qui doit contenir, à terme, le nombre d'individus jusqu'à la génération n .

$P[0]$ doit contenir le nombre nB de bactérie à la génération 0, à savoir $nB=1$.

b) *Justifions la structure conditionnelle en ligne 6. Pourquoi l'absence d'un `else` ?* Si $nB > 0$ détermine s'il reste des bactéries à la k -ième génération. Si c'est le cas, chacune d'entre elle se dédouble. Sinon, ça veut dire qu'il n'en reste plus. La population est éteinte et $P[k]=0$. Mais comme la liste P a été initialisée avec des valeurs nulles, il n'y a dans ce cas rien à faire.

c) *Complétons les lignes 5, 8, 9 et 11 en le justifiant :*

- Ligne 5 : Les bactéries vont se reproduire n fois. Le nombre de répétitions étant connu, on opte pour une structure répétitive « Pour », avec k allant de 1 à n ou k allant de 0 à $n - 1$.
- Lignes 8, 9 et 11 : Pour chacune des nB bactéries, on les fait se dédoubler avec une probabilité p et mourir sans reproduction avec une probabilité $1 - p$. On utilise pour ça la fonction `random()` de la bibliothèque `random` qui retourne un réel aléatoire entre 0 et 1.
Dans le cas où elle se reproduit, le nombre nB de bactéries augmente de 1 (elle donne naissance à deux filles et meurt) et, si elle meurt avant de se reproduire, le nombre nB de bactéries diminue de 1.

En complétant la fonction proposée, on obtient :

```
def simulPopulation(p,n):
    nB = 1
    P = [0]*(n+1)
    P[0] = nB
    for k in range(1,n+1):
        if nB > 0: # il y a des bactéries
            for i in range(nB):
                if rdm.random()<p:
                    nB += 1
                else:
                    nB -= 1
            P[k]=nB
    return P
```

④ *Justifions que que : $\mathbb{P}_{(X_1=2)}(X_{n+1} = 0) = \mathbb{P}(X_n = 0)^2$:* Si il y a deux bactéries à la première génération, alors la probabilité de ne plus avoir de bactéries à la $(n + 1)$ -ième génération est la probabilité que chacune de ces deux bactéries n'ait plus de descendance à leur n -ième génération.
Autrement dit :

$$\mathbb{P}_{(X_1=2)}(X_{n+1} = 0) = \mathbb{P}((X_n = 0) \cap (X_n = 0)) = \mathbb{P}(X_n = 0) \cdot \mathbb{P}(X_n = 0)$$

car on nous a dit que les reproductions étaient considérées indépendantes.

⑤ *A l'aide du système complet d'événements $\{(X_1 = 0), (X_1 = 2)\}$ établissons une relation entre u_{n+1} et u_n :* Commençons par noter que $\{(X_1 = 0), (X_1 = 2)\}$ est bien un système complet d'événements car : $(X_1 = 0) \cup (X_1 = 2) = \Omega$ et $(X_1 = 0) \cap (X_1 = 2) = \emptyset$

Note : Il faut penser que cette partie est une application de la Partie I du sujet. Autrement dit, le but du jeu est de montrer que : $u_{n+1} = 1 - p + pu_n^2$... voyons comment :

D'après la formule des probabilités totales :

$$\begin{aligned} u_{n+1} &= \mathbb{P}(X_{n+1} = 0) = \mathbb{P}_{(X_1=0)}(X_{n+1} = 0)\mathbb{P}(X_1 = 0) + \mathbb{P}_{(X_1=2)}(X_{n+1} = 0)\mathbb{P}(X_1 = 2) \\ &= 1 \cdot (1 - p) + p \cdot \mathbb{P}_{(X_1=2)}(X_{n+1} = 0)\mathbb{P}(X_1 = 2) \\ &= 1 - p + p \cdot [\mathbb{P}(X_n = 0)]^2 \\ &= f(u_n) \end{aligned}$$

En effet, si il n'y a plus de cellules à la première génération, alors il n'y en aura pas à la génération $n + 1$ et donc $\mathbb{P}_{(X_1=0)}(X_{n+1} = 0) = 1$.

⑥ interpréter pour $p \leq 1/2$:

a) Puisque $u_{n+1} = f(u_n)$ il est possible d'utiliser les résultats de la première partie.

Si $p \leq 1/2$ alors $\lim_{n \rightarrow \infty} u_n = \lim_{n \rightarrow \infty} \mathbb{P}(X_n = 0) = 1$.

Si $p > 1/2$, alors $\lim_{n \rightarrow \infty} u_n = \lim_{n \rightarrow \infty} \mathbb{P}(X_n = 0) = \frac{1-p}{p} \in]0, 1[$.

Conclusion : Si $p \leq 1/2$, la population de bactérie va s'éteindre de façon certaine.

b) En nous inspirant de la fonction `simulPopulation()`, écrivons une fonction `tempsJusquAExtinction(p)` qui retourne le nombre de générations nécessaire pour que la population s'éteigne : Il suffit pour ça de faire se reproduire les bactéries tant que la population en contient au moins une car dès qu'elle est éteinte, il n'y en aura jamais plus d'autre par absence de génération spontanée...

```
def tempsJusquAExtinction(p):
    nB = 1
    k = 0
    while nB > 0:
        for i in range(nB):
            if rdm.random() < p:
                nB += 1
            else:
                nB -= 1
        k += 1
    return k
```

c) Écrivons une fonction `ListeTpsExtinction(p,m)` qui répète un nombre m de fois (supposé grand) la fonction précédente et retourne une liste L formée des temps nécessaires à l'extinction de la population pour chacune de m répétition :

```
def TempsExtinction(p,m):
    L = []
    for k in range(m):
        L.append(tempsJusquAExtinction(p))
    return L
```

d) En exécutant `L.count(1)/m` on obtient successivement pour $p = 0.4$ et $m = 1000$ les valeurs 0.592 et 0.604. Cela semble cohérent puisqu'on rappelle que `L.count(1)/m` calcul la fréquence des 1 dans la liste L , à savoir estime la probabilité qu'il faille 1 génération pour que la population s'éteigne, ou encore $u_1 = \mathbb{P}(X_1 = 0) = 1 - p = 0.6$ d'après II.2

- e) Comment ferions-nous, connaissant L , pour estimer la valeur de u_2 ?

u_2 désigne la probabilité que la population soit éteinte à la deuxième génération. Il suffit donc, en appelant la fonction `TempsExtinction(p,m)`, de dénombrer au sein de la liste L , les 1 et les 2 puisque, si la population est éteinte dès la première génération, alors elle sera éteinte à la seconde... Soit :

Conclusion : $u_2 \approx (L.\text{count}(1) + L.\text{count}(2)) / m$

- f) Écrire, sans recours à la bibliothèque `numpy` une fonction `moyenne(L)` et `ecartType(L)` retournant respectivement la moyenne et l'écart-type d'une liste L .

A écrire (fait plusieurs fois) et à vérifier en comparant avec les résultats de `np.mean()` et `np.std()`... Lequel de ces deux résultats vous semble le plus probable lorsque $p = 0.4$? $m1 = 2.35$ et $s1 = 2.67$ ou $m2 = 14.25$ et $s2 = 3.24$? On a montré précédemment que $u_1 = \mathbb{P}(X_1 = 0) = 0.6$ et on peut calculer que $u_2 = f(u_1) = 0.744$.

Il y a donc 60% de chances que la population soit éteinte au bout d'une génération et 74% de chances qu'elle soit éteinte au bout de 2 génération. On peut donc juger plus probable que la population s'éteigne en moyenne en un peu plus de 2 génération ($m1$), plutôt qu'en un peu plus de 14 générations ($m2$).

Partie III : Le cas particulier $p > 1/2$

Nous nous plaçons désormais dans le cas $p > 1/2$. Sous cette condition, la probabilité que la population s'éteigne n'est plus nulle et on peut imaginer qu'elle peut converger vers un état stable ou bien tendre vers l'infini.

On donne la fonction génératrice g_X de la variable aléatoire X définie par :

$$\forall t \in \mathbb{R}, g_X(t) = \sum_{k \in X(\Omega)} \mathbb{P}(X = k) t^k$$

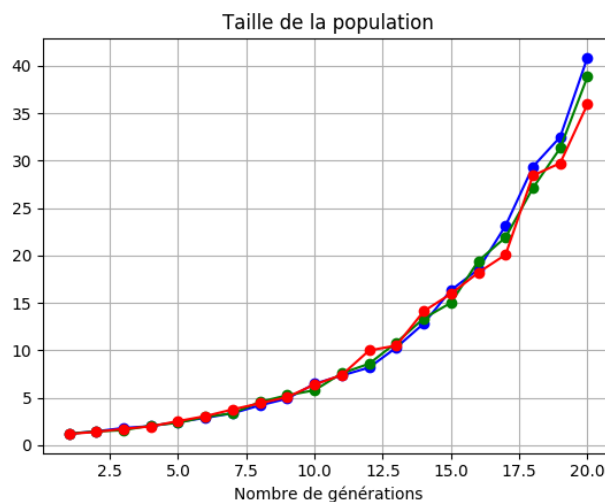
- ① Écrivons une fonction python `tailleMoyenne(n,p,m)` permettant de calculer en fonction de p le nombre moyen d'individu à la génération n lors de la répétition un nombre m de fois (supposé grand) du processus de reproduction à partir d'une bactérie :

```
def tailleMoyenne(p,n,m):
    # retourne la taille moyenne à la generation n
    # dans le cas où p > 1/2 (on n'est pas sûr que la population s'éteigne).
    S = 0
    for k in range(m):
        P = simulPopulation(p,n)
        S += P[-1]
    return S/m
```

En traçant les valeurs obtenues pour n allant de 1 à 20 on obtient les tailles moyennes ci-dessous. *Quelle conjecture pouvons-nous faire ?* Que la population, en moyenne tend vers l'infini de façon polynomiale...

Pour le tracé, on pourra faire :

```
def graphe(p,m):
    generations = np.arange(1,21)
    LM = [tailleMoyenne(p, n, m) for n in generations]
    plt.plot(generation, LM, 'o-', label='p = '+str(p))
    plt.grid()
    plt.show()
```



② a) Montrons que $g_{X_1} = f$:

Puisque $X_1(\Omega) = \{0, 2\}$, on a par définition,

$$\forall t \in \mathbb{R}, g_{X_1}(t) = \mathbb{P}(X_1 = 0)t^0 + \mathbb{P}(X_1 = 2)t^2 = 1 - p + pt^2 = f(t)$$

Conclusion : $\boxed{g_{X_1} = f}$

On en déduit que $\boxed{g'_{X_1}(1) = \mathbb{E}(X_1) = f'(1) = 2p}$.

b) Pour tout $n \in \mathbb{N}^*$, déterminons $X_n(\Omega)$:

➤ A la génération 1 nous venons de voir que $X_1(\Omega) = \{0, 2\}$.

A la génération 2, il y a quatre possibilités :

- les deux cellules filles meurent avant de se reproduire et alors $(X_2 = 0)$ est réalisé.
 - L'une seulement des deux cellules se reproduit avant de mourir et $(X_2 = 2)$ est réalisé.
 - Les deux cellules filles se reproduisent et $(X_2 = 4)$ est réalisé.
- donc $X_2(\Omega) = \{0, 2, 4\}$.

A la génération 3, on trouve par le même raisonnement que $X_3(\Omega) = \{0, 2, 4, 6, 8\}$

➤ Supposons que $X_n(\Omega) = \{0, 2, \dots, 2^n\} = \{2j, 0 \leq j \leq 2^{n-1}\}$ pour n fixé, $n \geq 1$.

➤ alors, à la génération $n + 1$, il y aura aucune bactéries s'il n'y en avait déjà plus à la génération précédente et sinon, chacune des $2j$ bactéries de la génération n peut se reproduire et donner naissance à 2 cellules filles, soit mourir sans se reproduire.

Il y aura donc au maximum $2 * (2^n) = 2^{n+1}$ bactéries à la génération $n + 1$ et chaque valeur paire comprise entre 0 et 2^{n+1} peut être atteinte.

➤ **Conclusion** : $\boxed{\forall n \in \mathbb{N}^*, X_n(\Omega) = \{2j, 0 \leq j \leq 2^{n-1}\}}$

c) Donnons la valeur de $g_{X_n}(1)$: Il suffit d'écrire que, par définition,

$$g_{X_n}(1) = \sum_{k \in X(\Omega)} \mathbb{P}(X_n = k). \text{ Donc } \boxed{g_{X_n}(1) = 1}$$

par définition d'une loi de probabilité.

Justifions la dérivabilité de g_{X_n} et déterminons $g'_{X_n}(1)$?

Par définition : $g_{X_n}(t) = \sum_{k \in X(\Omega)} \mathbb{P}(X_n = k)t^k$ donc, d'après ce qui précède :

$$g_{X_n}(t) = \sum_{j=0}^{2^{n-1}} \mathbb{P}(X_n = 2j)t^{2j}$$

g_{X_n} est donc un polynôme de degré 2^n à coefficients réels.

Ce qui justifie sa dérivabilité sur $\mathbb{R}$.

Par ailleurs :

$$g'_{X_n}(t) = \sum_{k \in X(\Omega)} k \cdot \mathbb{P}(X_n = k)t^{k-1}$$

D'où

$$\boxed{g'_{X_n}(1) = \sum_{k \in X(\Omega)} k \cdot \mathbb{P}(X_n = k)1^{k-1} = \sum_{k \in X(\Omega)} k \cdot \mathbb{P}(X_n = k) = \mathbb{E}(X_n)}$$

Remarque : Ce résultat est d'ailleurs validé par la réponse à la question 7.a) puisque $\mathbb{E}(X_1) = 2p$.

d) Montrons que $\mathbb{P}_{(X_n=2j)}(X_{n+1}=k) = 0$ si j impair ou k impair et que, sinon :

$$\mathbb{P}_{(X_n=2j)}(X_{n+1}=2k) = \binom{2j}{k} p^k q^{2j-k} \text{ pour tout } 0 \leq k \leq 2j$$

Puisque pour tout $n \in \mathbb{N}^*$, X_n ne prend que des valeurs paires, il découle immédiatement que $\mathbb{P}_{(X_n=2j)}(X_{n+1}=k) = 0$ si j impair ou k impair.

Sinon, sachant qu'à la génération n il y a $2j$ bactéries, le nombre de bactéries à la génération $n+1$ découle du nombre de bactéries qui se sont dédoublées à la génération précédente. Il s'agit de $2j$ épreuves de Bernoulli indépendantes (puisque les reproductions sont indépendantes) de même probabilité de succès « se reproduire » égale à p .

Or, $(X_{n+1}=2k)$ est réalisé signifie que k bactéries de la génération précédente se sont reproduites tandis que les autres sont mortes avant de pouvoir le faire. Donc, sachant que $(X_n=2j)$ est réalisé, $(X_{n+1}=2k)$ dénombre k succès au cours de $2j$ épreuves de Bernoulli indépendantes.

Il s'agit du modèle d'une loi binomiale.

Conclusion : $\mathbb{P}_{(X_n=2j)}(X_{n+1}=2k) = \binom{2j}{k} p^k q^{2j-k}, \forall 0 \leq k \leq 2j$

Remarque : Il est par ailleurs clair que $(X_{n+1}=2k) = \emptyset$ si $k > 2j$.

e) Dédisons-en que $g_{X_{n+1}}(t) = g_{X_n}(g_{X_1}(t))$ et montrons que $\mathbb{E}(X_{n+1}) = \mathbb{E}(X_n)\mathbb{E}(X_1)$:

On a par définition : $g_{X_{n+1}}(t) = \sum_{i \in X_{n+1}(\Omega)} \mathbb{P}(X_{n+1}=i) t^i = \sum_{k=0}^{2^n} \mathbb{P}(X_{n+1}=2k) t^{2k}$.

Or, d'après la formule des probabilités totales, sachant que $\{(X_n=2j), 0 \leq j \leq 2^{n-1}\}$ est un système complet d'événements :

$$\mathbb{P}(X_{n+1}=2k) = \sum_{j=0}^{2^{n-1}} \mathbb{P}_{(X_n=2j)}(X_{n+1}=2k) \mathbb{P}(X_n=2j)$$

d'où, en utilisant la question précédente :

$$\begin{aligned} g_{X_{n+1}}(t) &= \sum_{k=0}^{2^n} \sum_{j=0}^{2^{n-1}} \mathbb{P}_{(X_n=2j)}(X_{n+1}=2k) \mathbb{P}(X_n=2j) t^{2k} \\ &= \sum_{j=0}^{2^{n-1}} \sum_{k=0}^{2^n} \mathbb{P}_{(X_n=2j)}(X_{n+1}=2k) \mathbb{P}(X_n=2j) t^{2k} \\ &= \sum_{j=0}^{2^{n-1}} \mathbb{P}(X_n=2j) \sum_{k=0}^{2^n} \mathbb{P}_{(X_n=2j)}(X_{n+1}=2k) t^{2k} \\ &\quad \text{avec } \mathbb{P}_{(X_n=2j)}(X_{n+1}=2k) = 0 \text{ si } 2k > 2 \times 2j \text{ ou encore } k > 2j \\ &= \sum_{j=0}^{2^{n-1}} \mathbb{P}(X_n=2j) \sum_{k=0}^{2j} \binom{2j}{k} p^k q^{2j-k} t^{2k} = \sum_{j=0}^{2^{n-1}} \mathbb{P}(X_n=2j) \sum_{k=0}^{2j} \binom{2j}{k} (pt^2)^k q^{2j-k} \\ &= \sum_{j=0}^{2^{n-1}} \mathbb{P}(X_n=2j) (pt^2 + q)^{2j} \\ &= g_{X_n}(pt^2 + q) \end{aligned}$$

Or, d'après la question 7.a) : $g_{X_1}(t) = pt^2 + q$.

Conclusion : $g_{X_{n+1}}(t) = g_{X_n}(g_{X_1}(t))$

- f) A la question 7.b) nous avons obtenu que $\mathbb{E}(X_{n+1}) = g'_{X_{n+1}}(1)$.
Par dérivation d'une fonction composée, nous pouvons écrire :

$$g'_{X_{n+1}}(t) = g'_{X_1}(t) \cdot g'_{X_n}(g_{X_1}(t)) \Rightarrow g'_{X_{n+1}}(1) = g'_{X_1}(1) \cdot g'_{X_n}(g_{X_1}(1))$$

Conclusion : $\boxed{\mathbb{E}(X_{n+1}) = \mathbb{E}(X_n)\mathbb{E}(X_1)}$

- g) *Concluons sur l'expression de $\mathbb{E}(X_n)$ pour tout $n \in \mathbb{N}^*$:*
D'après ce qui précède, $\mathbb{E}(X_{n+1}) = (2p) \cdot \mathbb{E}(X_n)$, $\forall n \in \mathbb{N}^*$.
La suite $(\mathbb{E}(X_n))_{n \geq 0}$ est donc une suite géométrique de raison $q = 2p$ et de premier terme $\mathbb{E}(X_0) = 1$ (puisque $X = 0$ est la variable aléatoire certaine égale à 1...)

Conclusion : $\boxed{\forall n \in \mathbb{N}^*, \mathbb{E}(X_n) = (2p)^n}$

Nous avons désormais une bonne idée du comportement de la population bactérienne dans les cas où elle ne s'éteint pas. En moyenne, elle croît infiniment puisque $p > \frac{1}{2} \Rightarrow 2p > 1$.

- ③ On reprend les notations de la question II.2 en notant $u_n = \mathbb{P}(X_n = 0)$ et on rappelle que :

$$u_{n+1} = 1 - p + pu_n^2, \forall n \geq 1$$

- a) *Déterminons $\lim_{n \rightarrow \infty} u_n$ et interprétons ce résultat :*

D'après la partie I, question 4., dans le cas $p > 1/2$, la limite de la suite (u_n) vaut $\frac{1-p}{p} = \frac{q}{p}$.

Conséquence : Il existe une génération à partir de laquelle la probabilité que la population s'éteigne s'approche de $\frac{1-p}{p} = \frac{q}{p}$ aussi près qu'on le souhaite.

- b) On note pour tout n entier naturel, D_n l'événement : « la population disparaît exactement à l'issue de l'étape n ». *Montrons que :* $\forall n \in \mathbb{N}^*, \mathbb{P}(D_n) = u_n - u_{n-1}$:
 D_n signifie qu'il y a encore au moins une bactérie à l'issue de l'étape $n-1$ et qu'il n'y en a plus à l'issue de l'étape n .
Donc $D_n = (X_n = 0) \cap (X_{n-1} \neq 0)$.
 $\{(X_{n-1} \neq 0), (X_{n-1} = 0)\}$ est un système complet d'événements, donc par application de la formule des probabilités totales :

$$\mathbb{P}(X_n = 0) = \mathbb{P}((X_n = 0) \cap (X_{n-1} \neq 0)) + \mathbb{P}((X_n = 0) \cap (X_{n-1} = 0))$$

On a : $\mathbb{P}(X_n = 0) = u_n$ et $\mathbb{P}((X_n = 0) \cap (X_{n-1} \neq 0)) = \mathbb{P}(D_n)$.

Par ailleurs, $\mathbb{P}((X_n = 0) \cap (X_{n-1} = 0)) = \mathbb{P}(X_{n-1} = 0) = u_{n-1}$ car $(X_{n-1} = 0)$ implique $(X_n = 0)$.

Conclusion : $\boxed{\forall n \in \mathbb{N}^*, \mathbb{P}(D_n) = u_n - u_{n-1}}$

- c) Soit R l'événement : « la population de bactéries finit par s'éteindre ».

Justifions le fait que $R = \bigcup_{n=1}^{+\infty} D_n$: Il suffit de dire que la population s'éteint si, et seulement si, elle s'éteint au moins une fois exactement à l'issue de l'étape n où $n \in \mathbb{N}^*$.

Dès lors, la probabilité que la population de bactérie s'éteigne vaut : $\mathbb{P}(R) = \mathbb{P}(\bigcup_{n=1}^{+\infty} D_n)$.

Les événements D_n sont deux à deux incompatibles, donc par σ -additivité :

$$\sum \mathbb{P}(D_n) \text{ converge et } \sum_{n=1}^{\infty} \mathbb{P}(D_n) = \mathbb{P}(\bigcup_{n=1}^{+\infty} D_n) = \mathbb{P}(R)$$

Mais $\mathbb{P}(D_n) = u_n - u_{n-1}$. Il s'agit donc d'une série télescopique :

Pour tout $n \in \mathbb{N}^*$, on considère la somme partielle et on pose $S_n = \sum_{k=1}^n (u_k - u_{k-1}) = u_n - u_0$.

Or $u_0 = \mathbb{P}(X_0 = 0) = 0$, donc $S_n = u_n$.

Et comme $\lim_{n \rightarrow \infty} u_n = q/p$, on en déduit que $\lim_{n \rightarrow \infty} S_n = q/p$.

Conclusion : La probabilité que la population s'éteigne vaut q/p