DS4 (version B)

Exercice 1 (EML 2014)

On considère l'espace $\mathscr{M}_2(\mathbb{R})$ des matrices d'ordre 2 à coefficients réels. On définit :

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad C = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

$$\mathcal{E} = \{ \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \mid (a,b,c) \in \mathbb{R}^3 \}$$

1. Montrer que $\mathcal{E}$ est un espace vectoriel et que (A, B, C) est une base de $\mathcal{E}$.

Démonstration.

- Par définition de $\mathcal{E}$:

$$\begin{aligned} \mathcal{E} &= \left\{ \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \mid (a,b,c) \in \mathbb{R}^3 \right\} \\ &= \left\{ a \cdot \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + b \cdot \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + c \cdot \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \mid (a,b,c) \in \mathbb{R}^3 \right\} \\ &= \left\{ a \cdot A + b \cdot B + c \cdot C \mid (a,b,c) \in \mathbb{R}^3 \right\} \\ &= \operatorname{Vect}(A, B, C) \end{aligned}$$

$$\boxed{\text{Donc } \mathcal{E} \text{ est un espace vectoriel et la famille } (A,B,C) \text{ engendre } \mathcal{E}.}$$

- Montrons que la famille (A, B, C) est libre.
  Soit $(\lambda_1, \lambda_2, \lambda_3) \in \mathbb{R}^3$. Supposons : $\lambda_1 \cdot A + \lambda_2 \cdot B + \lambda_3 \cdot C = 0_{\mathscr{M}_2(\mathbb{R})}$ $(*)$.

$$\text{Or :} \quad (*) \Leftrightarrow \lambda_1 \cdot \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \lambda_2 \cdot \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + \lambda_3 \cdot \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

$$\Leftrightarrow \begin{pmatrix} \lambda_1 & \lambda_2 \\ 0 & \lambda_3 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

$$\Leftrightarrow \left\{ \lambda_1 = \lambda_2 = \lambda_3 = 0 \right.$$

$$\boxed{\text{La famille } (A,B,C) \text{ est donc libre.}}$$

- La famille (A, B, C) est :
  × libre,
  × génératrice de $\mathcal{E}$.

$$\boxed{\text{On en déduit que la famille } (A,B,C) \text{ est une base de } \mathcal{E}.}$$

Commentaire

- Il est relativement fréquent de trouver dans les sujets de concours des ensembles de matrices écrites à l'aide de paramètres. Lorsque c'est le cas, on trouve généralement une ou plusieurs questions consistant à démontrer la stabilité de ces ensembles (comme c'est le cas ici en question ***2.*** et ***3.***). Il faut donc être à l'aise sur la compréhension et la manipulation de tels ensembles. Ici, l'ensemble $\mathcal{E}$ n'est autre que l'ensemble des matrices carrées d'ordre 2 triangulaires supérieures.
- Dans l'énoncé, on demande de démontrer que $\mathcal{E}$ est un espace vectoriel engendrée par la famille (A, B, C). Cette famille étant fournie, l'écriture $\mathcal{E} = \text{Vect}(A, B, C)$ (c'est le caractère générateur de la famille) permet de démontrer que $\mathcal{E}$ est un espace vectoriel. Il faut privilégier cette démonstration à celle qui consiste à vérifier les propriétés axiomatiques de la notion d'espace vectoriel. Cependant cette manière de procéder doit aussi être connue car l'ensemble étudié ne se décrit pas toujours naturellement comme espace vectoriel engendré par une partie.
- Rappelons ci-dessous la rédaction.

(i) $\mathcal{E} \subset \mathscr{M}_2(\mathbb{R})$

(ii) $\mathcal{E} \neq \varnothing$ car $0_{\mathscr{M}_2(\mathbb{R})} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \in \mathcal{E}$.

(iii) Démontrons que $\mathcal{E}$ est stable par combinaisons linéaires.

Soit $(\lambda, \mu) \in \mathbb{R}^2$ et soit $(M, N) \in \mathcal{E}^2$.

× Comme $M \in \mathcal{E}$, il existe $(a_1, b_1, c_1) \in \mathbb{R}^3$ tel que $M = \begin{pmatrix} a_1 & b_1 \\ 0 & c_1 \end{pmatrix}$.

× Comme $N \in \mathcal{E}$, il existe $(a_2, b_2) \in \mathbb{R}^2$ tel que $N = \begin{pmatrix} a_2 & b_2 \\ 0 & c_2 \end{pmatrix}$.

Démontrons que $\lambda \cdot M + \mu \cdot N \in \mathcal{E}$. On a :

$$\lambda \cdot M + \mu \cdot N = \lambda \cdot \begin{pmatrix} a_1 & b_1 \\ 0 & c_1 \end{pmatrix} + \mu \cdot \begin{pmatrix} a_2 & b_2 \\ 0 & c_2 \end{pmatrix} = \begin{pmatrix} \lambda\, a_1 + \mu\, a_2 & \lambda\, b_1 + \mu\, b_2 \\ 0 & \lambda\, c_1 + \mu\, c_2 \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ 0 & \gamma \end{pmatrix} \in \mathcal{E}$$

avec : $(\alpha, \beta, \gamma) = (\lambda\, a_1 + \mu\, a_2, \lambda\, b_1 + \mu\, b_2, \lambda\, c_1 + \mu\, c_2) \in \mathbb{R}^3$.

$\boxed{\text{L'ensemble } \mathcal{E} \text{ est un sous-espace vectoriel de } \mathscr{M}_2(\mathbb{R}).}$

□

2. Établir que $\mathcal{E}$ est stable par multiplication, c'est à dire :

$$\forall (M, N) \in \mathcal{E}^2, MN \in \mathcal{E}$$

Démonstration.
Soit $(M, N) \in \mathcal{E}^2$. Il existe donc $(a_1, b_1, c_1) \in \mathbb{R}^3$ et $(a_2, b_2, c_2) \in \mathbb{R}^3$ tels que :

$$M = \begin{pmatrix} a_1 & b_1 \\ 0 & c_1 \end{pmatrix} \quad \text{et} \quad N = \begin{pmatrix} a_2 & b_2 \\ 0 & c_2 \end{pmatrix}$$

On obtient alors :

$$MN = \begin{pmatrix} a_1 & b_1 \\ 0 & c_1 \end{pmatrix} \begin{pmatrix} a_2 & b_2 \\ 0 & c_2 \end{pmatrix} = \begin{pmatrix} a_1\, a_2 & a_1\, b_2 + b_1\, c_2 \\ 0 & c_1\, c_2 \end{pmatrix} = a_1\, a_2 \cdot A + (a_1\, b_2 + b_1\, c_2) \cdot B + c_1\, c_2 \cdot C$$

Donc $MN \in \text{Vect}(A, B, C)$. D'où $MN \in \mathcal{E}$.

$\boxed{\text{L'ensemble } \mathcal{E} \text{ est bien stable par multiplication.}}$

Commentaire

On pouvait aussi rédiger autrement :

$$MN = \begin{pmatrix} a_1 & b_1 \\ 0 & c_1 \end{pmatrix} \begin{pmatrix} a_2 & b_2 \\ 0 & c_2 \end{pmatrix} = \begin{pmatrix} a_1\,a_2 & a_1\,b_2 + b_1\,c_2 \\ 0 & c_1\,c_2 \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ 0 & \gamma \end{pmatrix} \in \mathcal{E}$$

avec $(\alpha, \beta, \gamma) = (a_1\,a_2, a_1\,b_2 + b_1\,c_2, c_1\,c_2) \in \mathbb{R}^3$.

□

3. Montrer que, pour toute matrice M de $\mathcal{E}$, si M est inversible alors $M^{-1} \in \mathcal{E}$.

Démonstration.
Soit $M \in \mathcal{E}$. Il existe donc $(a, b, c) \in \mathbb{R}^3$ tel que $M = \begin{pmatrix} a & b \\ 0 & c \end{pmatrix}$.

- Rappelons tout d'abord :

$$M \text{ est inversible } \Leftrightarrow \det(M) \neq 0 \Leftrightarrow a \times c - 0 \times b \neq 0 \Leftrightarrow ac \neq 0$$

Commentaire

On se sert ici de la caractérisation de l'inversibilité des matrices carrées d'ordre 2. On pouvait aussi tout simplement remarquer que la matrice M est triangulaire (supérieure). Ainsi, elle est inversible si et seulement si ses coefficients diagonaux sont tous non nuls.

- Supposons M inversible, c'est-à-dire $ac \neq 0$. Dans ce cas :

$$M^{-1} = \frac{1}{ac} \begin{pmatrix} c & -b \\ 0 & a \end{pmatrix} = \frac{c}{ac} \cdot A - \frac{b}{ac} \cdot B + \frac{a}{ac} \cdot C \in \mathcal{E}$$

$$\boxed{\text{Ainsi, si } M \in \mathcal{E} \text{ est inversible, } M^{-1} \in \mathcal{E}.}$$

Commentaire

Attention, on ne montre pas dans cette question que toutes les matrices de $\mathcal{E}$ (c'est-à-dire les matrices carrées d'ordre 2 triangulaires supérieures) sont inversibles !
On démontre simplement que, si une matrice triangulaire supérieure est inversible, alors son inverse est aussi triangulaire supérieure. C'est à nouveau une propriété de stabilité : $\mathcal{E}$ est stable par passage à l'inverse.

□

Pour toute matrice de $\mathcal{E}$, on note $f(M) = TMT$.

4. Montrer que f est un endomorphisme de $\mathcal{E}$.

Démonstration.

- Démontrons que f est linéaire.
Soit $(\lambda_1, \lambda_2) \in \mathbb{R}^2$ et soit $(M_1, M_2) \in \mathcal{E}^2$.

$$\begin{array}{rcll} f(\lambda_1 \cdot M_1 + \lambda_2 \cdot M_2) & = & T\ (\lambda_1 \cdot M_1 + \lambda_2 \cdot M_2)\ T & \\ & = & (\lambda_1 \cdot TM_1 + \lambda_2 \cdot TM_2)\ T & \textit{(par distributivité à gauche de} \times \textit{sur} +) \\ & = & \lambda_1 \cdot TM_1T + \lambda_2 \cdot TM_2T & \textit{(par distributivité à droite de} \times \textit{sur} +) \\ & = & \lambda_1 \cdot f(M_1) + \lambda_2 \cdot f(M_2) & \end{array}$$

- Démontrons que f est à valeurs dans $\mathcal{E}$.
  Soit $M \in \mathcal{E}$.
  Remarquons tout d'abord : $T = A + B + C \in \text{Vect}(A, B, C)$. Ainsi $T \in \mathcal{E}$.
  On en déduit, par stabilité de $\mathcal{E}$ par multiplication (question ***3.***) : $TM \in \mathcal{E}$.
  Enfin : $f(M) = TMT = (TM)\,T \in \mathcal{E}$ en utilisant une nouvelle fois la question ***3.***

$$\boxed{\text{L'application } f \text{ est donc un endomorphisme de } \mathcal{E}.}$$

Commentaire

La rédaction choisie ici pour le deuxième point démontre une prise de recul par rapport aux questions précédentes. Cependant, il est aussi possible de résoudre cette question en effectuant un calcul. Plus précisément, si $M = \begin{pmatrix} a & b \\ 0 & c \end{pmatrix}$, alors :

$$f(M) \;=\; TMT \;=\; \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} a & b \\ 0 & c \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \;=\; \begin{pmatrix} a & b+c \\ 0 & c \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \;=\; \begin{pmatrix} a & a+b+c \\ 0 & c \end{pmatrix} \in \mathcal{E}$$

□

5. Vérifier que T est inversible et démontrer que f est un automorphisme de $\mathcal{E}$.

Démonstration.

- Tout d'abord : $\det(T) = 1 \times 1 - 0 \times 1 = 1 \neq 0$.

$$\boxed{\text{Ainsi, la matrice } T \text{ est inversible.}}$$

(l'inverse de T est : $T^{-1} = \dfrac{1}{1}\begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$)

- Déterminons $\text{Ker}(f)$.
  Soit $M \in \mathcal{E}$.

$$\begin{array}{rcl} M \in \text{Ker}(f) & \Leftrightarrow & f(M) \;=\; 0_{\mathscr{M}_2(\mathbb{R})} \\[1ex] & \Leftrightarrow & TMT \;=\; 0_{\mathscr{M}_2(\mathbb{R})} \\[1ex] & \Leftrightarrow & T^{-1}TMT \;=\; T^{-1} \times 0_{\mathscr{M}_2(\mathbb{R})} \quad \big(= 0_{\mathscr{M}_2(\mathbb{R})}\big) \\[1ex] & \Leftrightarrow & MT\,T^{-1} \;=\; 0_{\mathscr{M}_2(\mathbb{R})} \times T^{-1} \quad \big(= 0_{\mathscr{M}_2(\mathbb{R})}\big) \\[1ex] & \Leftrightarrow & M \;=\; 0_{\mathscr{M}_2(\mathbb{R})} \end{array}$$

  Donc $\text{Ker}(f) = \{0_{\mathscr{M}_2(\mathbb{R})}\}$ et l'endomorphisme f est injective.
- Enfin, comme f est un endomorphisme de $\mathcal{E}$ de dimension **finie** :

$$f \text{ est injective } \Leftrightarrow f \text{ est bijective}$$

$$\boxed{\text{L'application } f \text{ est donc un automorphisme de } \mathcal{E}.}$$

Commentaire

- On utilise ici un cas particulier de la proposition suivante :
  Soient E et F des espaces vectoriels de dimensions **finies** tels que $\dim(E) = \dim(F)$.
  Soit $f \in \mathscr{L}(E, F)$. Alors :
  $$f \text{ bijective } \Leftrightarrow f \text{ injective } \Leftrightarrow f \text{ surjective}$$
- Attention à ne pas confondre :
  × $T \in \mathscr{M}_2(\mathbb{R})$, matrice qui intervient dans la définition de f.
  × $F \in \mathscr{M}_3(\mathbb{R})$, matrice représentative de f dans la base (A, B, C) (*cf* question **7.**).

□

6. On note F la matrice de f dans la base (A, B, C) de $\mathcal{E}$.
Calculer $f(A), f(B), f(C)$ en fonction de (A, B, C) et en déduire F.

Démonstration.

- $f(A) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} = 1 \cdot A + 1 \cdot B + 0 \cdot C.$

  Ainsi : $\mathrm{Mat}_{(A,B,C)}\big(f(A)\big) = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$.

- $f(B) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = 0 \cdot A + 1 \cdot B + 0 \cdot C.$

  Ainsi : $\mathrm{Mat}_{(A,B,C)}\big(f(B)\big) = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$.

- $f(C) = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix} = 0 \cdot A + 1 \cdot B + 1 \cdot C.$

  Ainsi : $\mathrm{Mat}_{(A,B,C)}\big(f(C)\big) = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$.

$$\boxed{\text{Donc } F = \mathrm{Mat}_{(A,B,C)}(f) = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.}$$

□

7. Déterminer une base et la dimension de $\mathrm{Ker}(f - \mathrm{id})$.

Démonstration.

- Soit $M \in \mathcal{E}$. Alors il existe $(a, b, c) \in \mathbb{R}^3$ tel que $M = \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} = a \cdot A + b \cdot B + c \cdot C$.

  Autrement dit : $U = \mathrm{Mat}_{(A,B,C)}(M) = \begin{pmatrix} a \\ b \\ c \end{pmatrix} \in \mathscr{M}_{3,1}(\mathbb{R})$.

$$\begin{array}{rcl} M \in \mathrm{Ker}(f - \mathrm{id}) & \Leftrightarrow & (f - \mathrm{id}_{\mathcal{E}})(M) = 0_{\mathcal{E}} \\ & \Leftrightarrow & (F - I_3)(U) = 0_{\mathscr{M}_3(\mathbb{R})} \\ & \Leftrightarrow & \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\ & \Leftrightarrow & \left\{ \begin{array}{rcrcl} & & 0 & = & 0 \\ a & + & c & = & 0 \\ & & 0 & = & 0 \end{array} \right. \\ & \Leftrightarrow & \left\{ \begin{array}{rcl} a & = & -c \end{array} \right. \end{array}$$

  On obtient alors :

$$\begin{array}{rcl} \mathrm{Ker}(f - \mathrm{id}) & = & \{M \in \mathcal{E} \mid (f - \mathrm{id}_{\mathcal{E}})(M) = 0_{\mathcal{E}}\} \\ & = & \{a \cdot A + b \cdot B + c \cdot C \mid a = -c\} \\ & = & \{-c \cdot A + b \cdot B + c \cdot C \mid (b, c) \in \mathbb{R}^2\} \\ & = & \{c \cdot (-A + C) + b \cdot B \mid (b, c) \in \mathbb{R}^2\} \\ & = & \mathrm{Vect}\,(-A + C, B) \end{array}$$

$$\boxed{\text{On en conclut : } \mathrm{Ker}(f - \mathrm{id}) = \mathrm{Vect}\,(-A + C, B).}$$

- La famille $(-A+C, B)$ est :
  - $\times$ génératrice de $\text{Ker}(f-\text{id})$,
  - $\times$ libre car constituée de deux vecteurs non colinéaires.

$$\boxed{\text{On en conclut que la famille } (-A+C,B) \text{ est une base de } \text{Ker}(f-\text{id}) \text{ et ainsi : } \dim\big(\text{Ker}(f-\text{id})\big) = \text{Card}(-A+C,B) = 2.}$$

Commentaire

- Il faut faire attention aux objets manipulés. On doit déterminer $E_1(f) = \text{Ker}(f - \text{id}_{\mathcal{E}})$, noyau d'un endomorphisme de $\mathcal{E}$. On doit donc obtenir un sous-espace vectoriel de $\mathcal{E}$. Si $M \in \mathscr{M}_2(\mathbb{R})$ et $U = \text{Mat}_{(A,B,C)}(M) \in \mathscr{M}_{3,1}(\mathbb{R})$ sont bien deux représentations différentes de la même matrice M, cela n'autorise pas pour autant à écrire l'égalité entre ces deux éléments :

$$\underbrace{a\cdot A + b\cdot B + c\cdot C}_{\in \mathscr{M}_2(\mathbb{R})} \ \cancel{=}\ \underbrace{\begin{pmatrix} a \\ b \\ c \end{pmatrix}}_{\in \mathscr{M}_{3,1}(\mathbb{R})} \quad \text{et} \quad \underbrace{\text{Vect}\,(-A+C,B)}_{E_1(f)} \ \cancel{=}\ \underbrace{\text{Vect}\left(\begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\right)}_{E_1(F)}$$

- Il faut s'habituer à déterminer les ensembles $E_\lambda(F)$ par lecture de la matrice $F - \lambda I$.

  Illustrons la méthode avec la matrice de l'exercice et $\lambda = 1$.

  On cherche les vecteurs $X = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$ de $E_1(F)$ c'est-à-dire les vecteurs tels que : $(F - I_3)\,X = 0_{\mathscr{M}_{3,1}(\mathbb{R})}$.

  Or :

$$\begin{aligned} \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} &= x\cdot C_1 + y\cdot C_2 + z\cdot C_3 \\ &= x\cdot\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + y\cdot\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} + z\cdot\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \end{aligned}$$

  Pour obtenir le vecteur $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ à l'aide de cette combinaison linéaire, il y a deux possibilités :

  - $\times$ si $x = 0$ alors forcément $z = 0$ car sinon on crée un coefficient non nul en $2^{\text{ème}}$ position de la combinaison linéaire. Enfin, si $x = z = 0$, alors toute valeur de y convient pour créer une combinaison linéaire nulle. On peut prendre par exemple $y = 1$. Ainsi :

$$E_1(F) \ \supset\ \text{Vect}\left(\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\right)$$

  - $\times$ si $x \neq 0$ alors forcément $z = -x$. En prenant par exemple $x = 1$ on obtient :

$$E_1(F) \ \supset\ \text{Vect}\left(\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}\right)$$

  Finalement : $E_1(F) \ \supset\ \text{Vect}\left(\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}\right)$.

  Et l'égalité est vérifiée pour des raisons de dimension (il est simple de démontrer : $\text{rg}(F - I_3) = 1$ ce qui permet de conclure, par théorème du rang : $\dim\big(E_1(F)\big) = 2$.

8. Soit λ un réel différent de 1. Résoudre l'équation $f(M) = \lambda\, M$, d'inconnue $M \in \mathcal{E}$.

Démonstration.
Soit $M \in \mathcal{E}$. Alors il existe $(a, b, c) \in \mathbb{R}^3$ tel que $M = \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} = a \cdot A + b \cdot B + c \cdot C$.

Autrement dit : $U = \mathrm{Mat}_{(A,B,C)}(M) = \begin{pmatrix} a \\ b \\ c \end{pmatrix} \in \mathscr{M}_{3,1}(\mathbb{R})$.

$$\begin{array}{rcl}
f(M) = \lambda \cdot M & \iff & F\,U = \lambda \cdot U \\
& \iff & (F - \lambda\, I_3)\, U = 0_{\mathscr{M}_{3,1}(\mathbb{R})} \\
& \iff & \begin{pmatrix} 1-\lambda & 0 & 0 \\ 1 & 1-\lambda & 1 \\ 0 & 0 & 1-\lambda \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \\
& \iff & \left\{ \begin{array}{rcrcrcl} (1-\lambda)\, a & & & & & = & 0 \\ a & + & (1-\lambda)\, b & + & c & = & 0 \\ & & & & (1-\lambda)\, c & = & 0 \end{array} \right. \\
& \underset{L_3 \leftarrow \frac{1}{1-\lambda} L_3}{\overset{L_1 \leftarrow \frac{1}{1-\lambda} L_1}{\iff}} & \left\{ \begin{array}{rcrcrcl} a & & & & & = & 0 \\ a & + & (1-\lambda)\, b & + & c & = & 0 \\ & & & & c & = & 0 \end{array} \right. \qquad \textit{(car } \lambda \neq 1\textit{)} \\
& \overset{L_2 \leftarrow L_2 - L_1}{\iff} & \left\{ \begin{array}{rcrcl} a & & & = & 0 \\ (1-\lambda)\, b & + & c & = & 0 \\ & & c & = & 0 \end{array} \right. \\
& \iff & \{ a = b = c = 0 \\
& & \textit{(par remontées successives)}
\end{array}$$

$$\boxed{\text{On en déduit que l'unique solution de l'équation } f(M) = \lambda \cdot M \text{ dans } \mathcal{E} \text{ est la matrice nulle } 0_{\mathscr{M}_2(\mathbb{R})}.}$$

□

On note $I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ et $H = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$.

9. Calculer H^2, puis pour tout a de $\mathbb{R}$ et tout n de $\mathbb{N}$, $(I + aH)^n$.

Démonstration.

- Tout d'abord :

$$H^2 = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

  On en déduit, par une récurrence immédiate, que pour tout $k \geqslant 2$, $H^k = 0_{\mathscr{M}_2(\mathbb{R})}$.
  (ou alors on remarque : $\forall k \geqslant 2,\ H^k = H^2\ H^{k-2} = 0_{\mathscr{M}_2(\mathbb{R})}\ H^{k-2} = 0_{\mathscr{M}_2(\mathbb{R})}$*)*

- Soit $a \in \mathbb{R}$.
  Les matrices I et $a\,H$ commutent car I commute avec toutes les matrices carrée du même ordre.
- Soit $n \geqslant 1$. D'après la formule du binôme de Newton :

$$\begin{array}{rcll}
(I+aH)^n & = & \sum\limits_{k=0}^{n} \binom{n}{k} I^{n-k}\,(aH)^k \;=\; \sum\limits_{k=0}^{n} \binom{n}{k} a^k\,H^k & \textit{(car : } \forall k \in [\![0,n]\!],\ I^{n-k} = I\textit{)} \\
& = & \sum\limits_{k=0}^{1} \binom{n}{k} a^k\,H^k \;+\; \sum\limits_{k=2}^{n} \binom{n}{k} a^k\,H^k & \textit{(cette décomposition est valide car } n \geqslant 1\textit{)} \\
& = & \sum\limits_{k=0}^{1} \binom{n}{k} a^k\,H^k & \textit{(car : } \forall k \geqslant 2,\ H^k = 0_{\mathscr{M}_2(\mathbb{R})}\textit{)} \\
& = & \binom{n}{0} a^0\,H^0 + \binom{n}{1} a^1\,H^1 & \\
& = & I + a\,n\,H \;=\; \begin{pmatrix} 1 & 0 & 0 \\ na & 1 & na \\ 0 & 0 & 1 \end{pmatrix} &
\end{array}$$

- De plus : $(I+aH)^0 = I$ et $I + a \times 0 \cdot H = I$.
  La formule précédente reste valable pour $n = 0$.

$$\boxed{\text{On en déduit que : } \forall n \in \mathbb{N},\ (I + a\,H)^n = I + a\,n\,H.}$$

Commentaire

- La relation de Chasles stipule que pour tout $(m,p,n) \in \mathbb{N}^3$ tel que $m \leqslant p \leqslant n$:

$$\sum_{k=m}^{n} u_k = \sum_{k=m}^{p} u_k + \sum_{k=p+1}^{n} u_k$$

  (la somme la plus à droite est nulle si $p = n$)
  où (u_n) est une suite quelconque de réels ou de matrices.
- Dans cette question, on est dans le cas où $m = 0$ et $p = 1$.
  L'argument $n \geqslant 1$ est donc essentiel pour découper la somme.
  Le cas $n = 0$ doit donc être traité à part.
- Ici, la matrice H vérifie : $\forall k \geqslant 2,\ H^k = 0_{\mathscr{M}_3(\mathbb{R})}$. Elle est dite nilpotente d'indice 2 (ce terme n'est pas au programme et il est préférable de ne pas l'utiliser dans une copie). Si elle avait été nilpotente d'ordre 3, il aurait fallu traiter à part les cas $n = 0$ et $n = 1$ (le découpage de la somme est alors valable pour $n \geqslant 2$). □

10. Calculer, pour tout n de $\mathbb{N}$, F^n.

Démonstration.
Soit $n \in \mathbb{N}$.
On remarque que : $F = I + H = I + 1 \cdot H$.
D'après la question ***9.*** appliquée à $a = 1$, on obtient : $F^n = I + 1 \times n \cdot H$.

$$\boxed{\text{Pour tout } n \in \mathbb{N},\ F^n = I + n\,H.}$$

□

11. Trouver une matrice G de $\mathscr{M}_3(\mathbb{R})$ telle que $G^3 = F$.
Existe-t-il un endomorphisme g de $\mathcal{E}$ tel que $g \circ g \circ g = f$?

Démonstration.

- D'après la question ***9.***, pour tout $a \in \mathbb{R}$: $(I + a\,H)^3 = I + 3\,a\,H$.
  Or $F = I + 1 \cdot H$. Donc, en choisissant $a = \dfrac{1}{3}$, on obtient :

$$\left(I + \frac{1}{3}\,H\right)^3 = I + 3 \times \frac{1}{3} \cdot H = I + H = F$$

$$\boxed{\text{Donc en posant } G = I + \frac{1}{3}\,H\text{, on obtient : } G^3 = F.}$$

- On rappelle que $F = \mathrm{Mat}_{(A,B,C)}(f)$.
  On considère alors l'endomorphisme $g \in \mathscr{L}(\mathcal{E})$ tel que $\mathrm{Mat}_{(A,B,C)}(g) = G$.
  D'après la relation du point précédent, on obtient :

$$\mathrm{Mat}_{(A,B,C)}(g^3) = \big(\mathrm{Mat}_{(A,B,C)}(g)\big)^3 = G^3 = F = \mathrm{Mat}_{(A,B,C)}(f)$$

  L'application $\mathrm{Mat}_{(A,B,C)}(.)$ étant un isomorphisme, on obtient, par injectivité : $g^3 = f$.

$$\boxed{\text{On a donc bien exhibé un endomorphisme } g \text{ de } \mathcal{E} \text{ tel que } g \circ g \circ g = f.}$$

Commentaire

Il faut retenir le schéma classique développé dans cette question :
(i) on démontre une propriété sous forme matricielle,
(ii) on en déduit une propriété sur les endomorphismes par la passerelle matrice / endomorphisme.

□

Exercice 2 (EML 2017)

On considère la fonction $f :]0,+\infty[\to \mathbb{R}$ définie par :

$$f : x \mapsto \mathrm{e}^x - \mathrm{e}\ln(x).$$

On admet les encadrements numériques suivants :

$$2,7 < \mathrm{e} < 2,8 \qquad 7,3 < \mathrm{e}^2 < 7,4 \qquad 0,6 < \ln(2) < 0,7.$$

Partie I : Étude de la fonction f

1. a) Montrer que f est deux fois dérivable sur $]0,+\infty[$ et calculer, pour tout x de $]0,+\infty[$, $f'(x)$ et $f''(x)$.

Démonstration.
Les fonctions $x \mapsto \mathrm{e}^x$ et $x \mapsto \ln(x)$ sont deux fois dérivables sur l'intervalle $]0,+\infty[$ en tant que fonctions usuelles.

La fonction f est deux fois dérivable sur $]0,+\infty[$ comme somme de fonctions deux fois dérivables sur $]0,+\infty[$.

$$\boxed{\forall x \in]0,+\infty[,\ f'(x) = \mathrm{e}^x - \frac{\mathrm{e}}{x} \quad \text{et} \quad f''(x) = \mathrm{e}^x + \frac{\mathrm{e}}{x^2}.}$$

□

b) Dresser le tableau de variations de f' avec la limite de f' en 0 et la limite de f' en $+\infty$ et préciser $f'(1)$.

Démonstration.

- Pour tout $x \in]0,+\infty[$, $f''(x) = \mathrm{e}^x + \dfrac{\mathrm{e}}{x^2} > 0$.
  La fonction f' est donc strictement croissante sur $]0,+\infty[$.
- Par ailleurs : $\lim\limits_{x\to 0} \mathrm{e}^x = 1$ et $\lim\limits_{x\to 0} \dfrac{\mathrm{e}}{x} = +\infty$.
  Donc $\lim\limits_{x\to 0} f'(x) = -\infty$.
- De plus : $\lim\limits_{x\to +\infty} \mathrm{e}^x = +\infty$ et $\lim\limits_{x\to +\infty} \dfrac{\mathrm{e}}{x} = 0$.
  Donc $\lim\limits_{x\to +\infty} f'(x) = +\infty$.
- $f'(1) = \mathrm{e}^1 - \dfrac{\mathrm{e}}{1} = \cancel{\mathrm{e}} - \cancel{\mathrm{e}} = 0$.

On obtient alors le tableau de variations suivant :

x	0		1		$+\infty$
Signe de $f''(x)$	‖	$+$		$+$	
Variations de f'	‖ $-\infty$	↗	0	↗	$+\infty$

□

2. Dresser le tableau de variations de f avec la limite de f en 0 et la limite de f en $+\infty$ et préciser $f(1)$.

Démonstration.

- La fonction f' est strictement croissante sur $]0,+\infty[$. Ainsi :
  - $\times$ $\forall x \in]0,1[$, $f'(x) < f'(1) = 0$.
  - $\times$ $\forall x \in]1,+\infty[$, $f'(x) > f'(1) = 0$.

  Donc f est strictement décroissante sur $]0,1]$ et strictement croissante sur $[1,+\infty[$.
- Déterminons la limite de f en 0. Comme : $\lim\limits_{x\to 0} \mathrm{e}^x = 1$ et $\lim\limits_{x\to 0} \ln(x) = -\infty$, alors $\lim\limits_{x\to 0} f'(x) = +\infty$.
- Déterminons alors la limite de f en $+\infty$. On écrit :

$$f(x) = \mathrm{e}^x - \mathrm{e}\ln(x) = \mathrm{e}^x\left(1 - \mathrm{e}\,\frac{\ln(x)}{\mathrm{e}^x}\right)$$

  Par croissances comparées : $\lim\limits_{x\to+\infty} \dfrac{\ln(x)}{\mathrm{e}^x} = 0$.
  Donc $\lim\limits_{x\to+\infty} f(x) = +\infty$.
- $f(1) = \mathrm{e}^1 - \mathrm{e}\ln(1) = \mathrm{e}$.

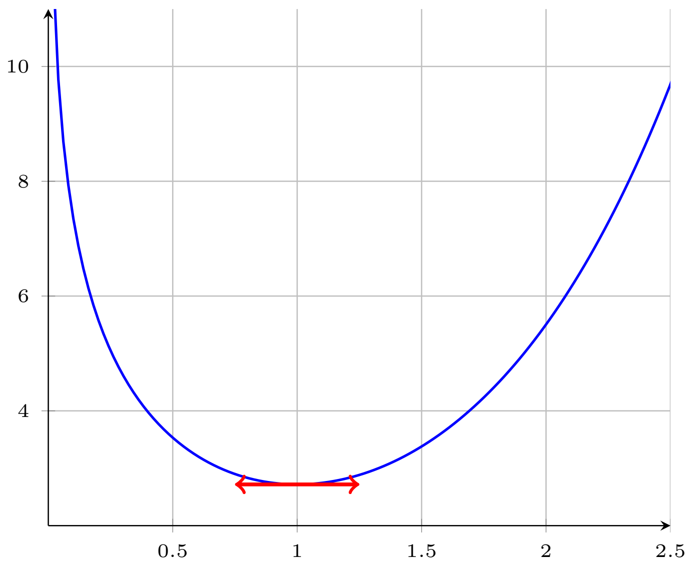
On obtient alors le tableau de variations suivant :

x	0		1		$+\infty$
Signe de $f'(x)$	‖	$-$	0	$+$	
Variations de f	‖ $+\infty$	$\searrow$	e	$\nearrow$	$+\infty$

□

3. Tracer l'allure de la courbe représentative de f.

Démonstration.



Commentaire

On s'efforcera de faire les questions de tracé de courbe.
En effet, ce type de question consiste uniquement à faire apparaître sur un graphique toutes les informations qu'on a recueillies dans les questions précédentes. D'ailleurs, il ne faut oublier de faire apparaître les tangentes horizontales et les points d'inflexion si on les a déterminés auparavant.

□

4. a) Étudier les variations de la fonction $u : \begin{array}{ccc}]0,+\infty[& \to & \mathbb{R} \\ x & \mapsto & f'(x) - x \end{array}$

Démonstration.

- La fonction u est dérivable sur $]0,+\infty[$ car elle est la somme $f_1 + f_2$ de :
  - $\times$ $f_1 = f'$ dérivable sur $]0,+\infty[$ car f est deux fois dérivable sur $]0,+\infty[$,
  - $\times$ $f_2 : x \mapsto x$ dérivable sur $]0,+\infty[$.

$$\boxed{\text{La fonction } u \text{ est dérivable sur }]0,+\infty[.}$$

- Soit $x \in]0,+\infty[$:

$$u'(x) \;=\; f''(x) - 1 \;=\; \mathrm{e}^x + \frac{\mathrm{e}}{x^2} - 1 \;=\; \frac{x^2\ \mathrm{e}^x + \mathrm{e} - x^2}{x^2} \;=\; \frac{x^2\ (\mathrm{e}^x - 1) + \mathrm{e}}{x^2}$$

  Comme $x^2 > 0$, on en déduit que le signe de $u'(x)$ est celui de $x^2\ (\mathrm{e}^x - 1) + \mathrm{e}$.
  Or :
  - $\times$ comme $x > 0$, alors : $\mathrm{e}^x > \mathrm{e}^0 = 1$. Ainsi : $\mathrm{e}^x - 1 > 0$.
  - $\times$ de plus : $x^2 > 0$ et $\mathrm{e} > 0$.

  On en déduit : $u'(x) > 0$.

$$\boxed{\text{La fonction } u \text{ est donc strictement croissante sur }]0,+\infty[.}$$

- D'après ***1.b)*** : $\lim\limits_{x \to 0} f'(x) = -\infty$. De plus : $\lim\limits_{x \to 0} x = 0$.
  D'où : $\lim\limits_{x \to 0} u(x) = -\infty$.
- Déterminons alors la limite de f en $+\infty$. Soit $x \in]0,+\infty[$:

$$u(x) \;=\; \mathrm{e}^x - \frac{\mathrm{e}}{x} - x \;=\; \mathrm{e}^x \left(1 - \frac{\mathrm{e}}{x\mathrm{e}^x} - \frac{x}{\mathrm{e}^x}\right)$$

  Par croissances comparées, $\lim\limits_{x \to +\infty} \dfrac{x}{\mathrm{e}^x} = 0$. De plus : $\lim\limits_{x \to +\infty} \dfrac{\mathrm{e}}{x\mathrm{e}^x} = 0$ et $\lim\limits_{x \to +\infty} \mathrm{e}^x = +\infty$.
  D'où : $\lim\limits_{x \to +\infty} u(x) = +\infty$.

On obtient le tableau de variations suivant :

x	0	$+\infty$
Signe de $u'(x)$	‖	$+$
Variations de u	‖	$-\infty$ ↗ $+\infty$

□

b) En déduire que l'équation $f'(x) = x$, d'inconnue $x \in \,]0, +\infty[$, admet une solution et une seule, notée α, et montrer : $1 < \alpha < 2$.

Démonstration.

- Tout d'abord, soit $x \in \,]0, +\infty[$:

$$f'(x) = x \;\Leftrightarrow\; f'(x) - x = 0 \;\Leftrightarrow\; u(x) = 0$$

On cherche donc à montrer que l'équation $u(x) = 0$ admet une unique solution sur $]0, +\infty[$.

- La fonction u est :
  - × continue sur $]0, +\infty[$ (car dérivable sur $]0, +\infty[$ d'après la question ***4.a)***),
  - × strictement croissante sur $]0, +\infty[$.

  Ainsi la fonction u réalise une bijection de $]0, +\infty[$ sur $u(]0, +\infty|)$.

$$u(]0, +\infty[) = \left] \lim_{x \to 0} u(x), \lim_{x \to +\infty} u(x) \right[=]-\infty, +\infty[$$

Or $0 \in \,]-\infty, +\infty[$, donc l'équation $u(x) = 0$ admet une unique solution α sur $]0, +\infty[$.

$\boxed{\text{On en déduit que l'équation } f'(x) = x \text{ admet une unique solution } \alpha \text{ sur }]0, +\infty[.}$

- Tout d'abord :
  - × $u(\alpha) = 0$
  - × $u(1) = f'(1) - 1 = 0 - 1 = -1 < 0$
  - × $u(2) = f'(2) - 2 = e^2 - \dfrac{e}{2} - 2 > 0$

    En effet, d'après les encadrements donnés par l'énoncé, on obtient :
    - d'une part : $7,3 < e^2 < 7,4$. Donc : $5,3 < e^2 - 2 < 5,4$
    - d'autre part : $2,7 < e < 2,8$. Donc : $1,35 < \dfrac{e}{2} < 1,4$. D'où : $-1,4 < -\dfrac{e}{2} < -1,35$

    Ainsi : $3,9 < e^2 - 2 - \dfrac{e}{2} < 4,05$. On en conclut : $u(2) > 3,9 > 0$.

  On en déduit :

$$u(1) \;<\; u(\alpha) \;<\; u(2)$$

Or, d'après le théorème de la bijection, $u^{-1} :]-\infty, +\infty[\to]0, +\infty[$ est strictement croissante sur $]-\infty, +\infty[$. En appliquant u^{-1}, on obtient alors :

$$\begin{array}{ccccc} u^{-1}\big(u(1)\big) & < & u^{-1}\big(u(\alpha)\big) & < & u^{-1}\big(u(e)\big) \\ \shortparallel & & \shortparallel & & \shortparallel \\ 1 & & \alpha & & e \end{array}$$

$$\boxed{1 < \alpha < 2}$$

□

Partie II : Étude d'une suite, étude d'une série

On considère la suite réelle $(u_n)_{n\in\mathbb{N}}$ définie par :

$$\begin{cases} u_0 = 2 \\ \forall n \in \mathbb{N},\ u_{n+1} = f(u_n) \end{cases}$$

5. Montrer que, pour tout n de $\mathbb{N}$, u_n existe et $u_n \geqslant 2$.

Démonstration.
Démontrons par récurrence que : $\forall n \in \mathbb{N}$, $\mathcal{P}(n)$, où $\mathcal{P}(n)$: u_n existe et $u_n \geqslant 2$.

- **Initialisation** :
  D'après l'énoncé : $u_0 = 2 \geqslant 2$.
  D'où $\mathcal{P}(0)$.
- **Hérédité** : soit $n \in \mathbb{N}$.
  Supposons $\mathcal{P}(n)$ et démontrons $\mathcal{P}(n+1)$ (c'est-à-dire u_{n+1} existe et $u_{n+1} \geqslant 2$).
  Par hypothèse de récurrence, u_n existe et $u_n \geqslant 2$.
  - La fonction f est définie sur $]0,+\infty[$.
    Or $u_n \geqslant 2$, donc $u_n \in \,]0,+\infty[$. Ainsi $u_{n+1} = f(u_n)$ existe.
  - D'après la question ***2.***, e est le minimum de f sur $]0,+\infty[$, c'est-à-dire :
    $$\forall x \in \,]0,+\infty[,\ f(x) \geqslant \mathrm{e}$$
    En appliquant cette inégalité à $x = u_n \in \,]0,+\infty[$, on obtient :
    $$u_{n+1} \geqslant \mathrm{e} > 2$$

  D'où $\mathcal{P}(n+1)$.

$$\boxed{\text{Par principe de récurrence, pour tout } n \in \mathbb{N},\ u_n \text{ existe et } u_n \geqslant 2.}$$

□

6. a) Étudier les variations, puis le signe, de la fonction $g : \begin{array}{rcl} [2,+\infty[& \to & \mathbb{R} \\ x & \mapsto & f(x) - x \end{array}$

Démonstration.

- La fonction g est dérivable sur $[2,+\infty[$ car elle est la somme $g_1 + g_2$ de :
  × $g_1 = f$ dérivable (car deux fois dérivable d'après ***1.a)***) sur $]0,+\infty[$,x donc sur $[2,+\infty[$,
  × $g_2 : x \mapsto x$ dérivable sur $[2,,+\infty[$.

$$\boxed{\text{La fonction } g \text{ est dérivable sur } [2,+\infty[.}$$

- Soit $x \in [2,+\infty[$, $g'(x) = f'(x) - 1$.
  D'après la question ***1.b)***, la fonction f' est croissante sur $[2,+\infty[\,\subset\,]0,+\infty[$. Ainsi :
  $$f'(x) \ \geqslant\ f'(2)$$
  Or, avec les encadrements donnés par l'énoncé : $f'(2) = \mathrm{e}^2 - \dfrac{\mathrm{e}}{2} \geqslant 7,3 - 1,4 = 5,9 > 1$.
  On en déduit : $f'(x) > 1$. D'où : $g'(x) > 0$.
  On obtient le tableau de variations suivant :

x	2	$+\infty$
Signe de $g'(x)$	$+$	
Variations de g	$g(2)$ $\nearrow$	$+\infty$

- Déterminons maintenant le signe de g. Comme g est croissante sur $[2, +\infty[$, on a :

$$\forall x \geqslant 2, \quad g(x) \geqslant g(2)$$

Il suffit alors de démontrer : $g(2) > 0$ pour pouvoir conclure. Calculons :

$$g(2) \;=\; f(2) - 2 \;=\; \mathrm{e}^2 - \mathrm{e}\ln(2) - 2$$

D'après les encadrements de l'énoncé :
× d'une part : $7,3 < \mathrm{e}^2 < 7,4$
× d'autre part : $2,7 < \mathrm{e} < 2,8$ et $0,6 < \ln(2) < 0,7$. D'où :

$$\underset{\substack{\shortparallel \\ 1,62}}{2,7 \times 0,6} \quad < \quad \mathrm{e}\ln(2) \quad < \quad \underset{\substack{\shortparallel \\ 1,96}}{2,8 \times 0,7}$$

Ainsi : $-1,96 < -\mathrm{e}\ln(2) < -1,62$.
On en déduit : $g(2) = \mathrm{e}^2 - \mathrm{e}\ln(2) - 2 \;>\; 7,3 - 1,96 - 2 = 3,34$.
On obtient alors : $g(2) > 0$.

$$\boxed{\forall x \in [2, +\infty[,\ g(x) > 0}$$

□

b) En déduire que la suite $(u_n)_{n\in\mathbb{N}}$ est croissante.

Démonstration.

- D'après la question ***6.a)*** : $\forall x \geqslant 2$, $g(x) > 0$. Autrement dit, pour tout $x \geqslant 2$:

$$f(x) > x$$

- Soit $n \in \mathbb{N}$. En appliquant cette inégalité à $x = u_n \geqslant 2$ (question ***5.***), on obtient :

$$u_{n+1} = f(u_n) > u_n$$

$$\boxed{\text{Ainsi, pour tout } n \in \mathbb{N},\ u_{n+1} > u_n. \text{ La suite } (u_n) \text{ est donc (strictement) croissante.}}$$

□

7. Démontrer que la suite $(u_n)_{n\in\mathbb{N}}$ admet $+\infty$ pour limite.

Démonstration.

- On sait que la suite (u_n) est croissante (d'après la question ***6.b)***). Deux cas se présentent alors :
  × soit (u_n) est de plus majorée, et alors elle converge.
  × soit (u_n) n'est pas majorée, et alors elle diverge vers $+\infty$.
- Démontrons par l'absurde que (u_n) n'est pas majorée.
  Supposons que la suite (u_n) est majorée.
  × Dans ce cas, la suite (u_n) est croissante et majorée donc elle converge vers un réel ℓ.
  D'après la question ***5.*** : $\forall n \in \mathbb{N}$, $u_n \geqslant 2$.
  Par passage à la limite dans cette inégalité, on obtient : $\ell \geqslant 2$.
  × D'autre part, par définition : $\forall n \in \mathbb{N}$, $u_{n+1} = f(u_n)$.
  La fonction f est continue sur $[2, +\infty[$, elle est donc en particulier continue en ℓ. Par passage à la limite dans l'égalité précédente, on obtient : $\ell = f(\ell)$, *i.e.* $g(\ell) = 0$.

  Absurde ! En effet, d'après la question ***6.a)*** : $\forall x \geqslant 2$, $g(x) > 0$. En particulier : $\forall x \geqslant 2$, $g(x) \neq 0$.

$$\boxed{\text{Ainsi, la suite } (u_n) \text{ n'est pas majorée et diverge donc vers } +\infty}$$

Commentaire

Il faut prendre le réflexe de penser à un raisonnement par l'absurde lorsque le résultat à démontrer est formulé sous forme de négation (et pas d'affirmation comme c'est le cas en général). À titre d'illustration, il faut penser à ce type de raisonnement pour :

× montrer qu'une suite `N`'est `PAS` majorée,

× montrer qu'une matrice admettant une seule valeur propre `N`'est `PAS` diagonalisable.

□

8. Écrire un programme en **Scilab** qui, étant donné un réel A, renvoie un entier naturel N tel que $u_N \geqslant A$.

Démonstration.
On propose le programme suivant :

```
1  A = input('Entrez un réel A : ')
2  N = 0
3  u = 2
4  while u < A
5      N = N + 1
6      u = exp(u) - %e * log(u)
7  end
8  disp(N)
```

Détaillons les éléments de ce script.

- **Début du programme**
La valeur de `A` est choisie par l'utilisateur à l'aide de la fonction `input`.

```
1  A = input('Entrez un réel A : ')
```

On initialise ensuite la variable `N` à 0.

```
2  N = 0
```

La variable `u` qui contiendra les valeurs successives de la suite (u_n) est initialisée à 2 : la valeur de u_0.

```
3  u = 2
```

- **Structure itérative**
Les lignes 4 à 7 consistent à :

1) déterminer un entier n tel que : $u_n \geqslant A$,

2) calculer les valeurs succesives de u_n.

On doit donc :

1) incrémenter la variable `N` de 1 jusqu'à ce que : $u_n \geqslant A$. Autrement dit, on doit incrémenter la variable `N` de 1 tant que : $u_n < A$. Pour cela on met en place une structure while :

```
4  while u < A
```

Puis on met à jour la variable `N`.

```
5      N = N + 1
```

2) calculer les valeurs successives de u_n :

```
6     u = exp(u) - %e * log(u)
```

- **Fin du programme**
À l'issue de cette boucle, la variable `N` contient le premier entier n tel que : $u_n \geqslant A$. On finit donc ce programme en affichant la valeur de cette variable.

```
8  disp(N)
```

Commentaire

Afin de permettre une bonne compréhension des mécanismes en jeu, on a détaillé la réponse à cette question. Cependant, proposer un programme **Scilab** correct démontre la bonne compréhension de ces mécanismes et permet certainement d'obtenir la majorité des points alloués à cette question. □

9. a) Démontrer : $\forall x \in [2, +\infty[,\ 2\ln(x) \leqslant x \leqslant \dfrac{e^x}{3}$.

Démonstration.

- Démontrons tout d'abord : $\forall x \in [2, +\infty[,\ 2\ln(x) \leqslant x$.
La fonction $h : x \mapsto \ln(x)$ est concave sur $]0, +\infty[$.
Sa courbe représentative $\mathcal{C}_h$ est donc située en dessous de ses tangentes, notamment celle au point d'abscisse 2. Or cette tangente est la droite d'équation :

$$\begin{aligned} y &= h(2) + h'(2)\,(x-2) \\ &= \ln(2) + \frac{1}{2}\,(x-2) \end{aligned}$$

On en déduit :

$$\forall x \in [2, +\infty[, \quad \ln(x) \ \leqslant\ \ln(2) + \frac{1}{2}\,(x-2)$$

Ainsi, pour tout $x \in [2, +\infty[$:

$$2\ln(x) \ \leqslant\ 2\ln(2) + (x-2) = x + 2(\ln(2) - 1)$$

Or, comme : $0,6 < \ln(2) < 0,7$, alors : $2(\ln(2) - 1) < 0$. D'où : $x + 2(\ln(2) - 1) < x$.

$$\boxed{\text{Finalement, par transitivité : } \forall x \in [2, +\infty[,\ 2\ln(x) \leqslant x}$$

- Démontrons maintenant : $\forall x \in [2, +\infty[,\ x \leqslant \dfrac{e^x}{3}$.

Considérons la fonction $h : x \mapsto \dfrac{e^x}{3} - x$.
  - × La fonction h est dérivable sur $[2, +\infty[$ comme somme de fonctions dérivables sur $[2, +\infty[$.
  - × Soit $x \in [2, +\infty[$.

$$h'(x) \ =\ \frac{e^x}{3} - 1 \ =\ \frac{e^x - 3}{3}$$

Comme $3 > 0$, la quantité $h'(x)$ est du signe de $e^x - 3$.
Or, comme $x \geqslant 2$, par croissance de la fonction exp sur $\mathbb{R}$:

$$e^x \ \geqslant\ e^2 > 7,3$$

Et ainsi : $e^x - 3 > 4,3 > 0$. D'où : $h'(x) > 0$.

× On obtient le tableau de variations suivant :

x	2 $\qquad\qquad$ $+\infty$
Signe de $h'(x)$	$+$
Variations de h	$h(2) \nearrow +\infty$

× De plus : $h(2) = \dfrac{e^2}{3} - 2 = \dfrac{e^2 - 6}{3} > 0$ car $e^2 > 7,3$.
Or h est strictement croissante sur $[2, +\infty[$. Ainsi, pour tout $x \in [2, +\infty[$:

$$h(x) \ \geqslant \ h(2) \ \geqslant \ 0$$

$$\boxed{\text{On en conclut : } \forall x \in [2, +\infty[,\ x \leqslant \dfrac{e^x}{3}.}$$

Commentaire

Pour démontrer la première inégalité, il est bien évidemment possible de passer par l'étude de la fonction $\varphi : x \mapsto x - 2\ln(x)$.

- La fonction φ est dérivable sur $[2, +\infty[$ en tant que somme de fonctions dérivables sur $[2, +\infty[$.
- Soit $x \in [2, +\infty[$.
$$\varphi'(x) \ = \ 1 - \frac{2}{x} \ = \ \frac{x-2}{x}$$
Comme $x \geqslant 2$, on en déduit : $\varphi'(x) \geqslant 0$.
- On obtient le tableau de variations suivant :

x	2 $\qquad\qquad$ $+\infty$
Signe de $\varphi'(x)$	$+$
Variations de φ	$\varphi(2) \nearrow +\infty$

- Comme la fonction φ est croissante sur $[2, +\infty[$:
$$\forall x \in [2, +\infty, \quad \varphi(x) \geqslant \varphi(2) = 2 - 2\ \ln(2)$$
Or : $2 - 2\ \ln(2) > 0$. Ainsi : $\forall x \in [2, +\infty[$, $\varphi(x) \geqslant 0$. D'où :
$$\forall x \in [2, +\infty[, \quad x - 2\ \ln(x) \ \geqslant \ 0$$

□

b) En déduire : $\forall n \in \mathbb{N},\ u_{n+1} \geqslant \dfrac{6-\mathrm{e}}{2} u_n$.

Démonstration.
Soit $n \in \mathbb{N}$.

- Tout d'abord :
$$u_{n+1} \ = \ f(u_n) \ = \ \mathrm{e}^{u_n} - \mathrm{e}\ \ln(u_n)$$
- De plus, d'après la question précédente :
$$\forall x \in [2, +\infty[,\ 2\ln(x) \leqslant x \leqslant \frac{\mathrm{e}^x}{3}$$
En appliquant cette double inégalité à $x = u_n \in [2, +\infty[$, on obtient :
$$2\ln(u_n) \ \leqslant \ u_n \ \leqslant \ \frac{\mathrm{e}^{u_n}}{3}$$
On en déduit :
  × d'une part : $2\ln(u_n) \leqslant u_n$. Ainsi : $-\mathrm{e}\ln(u_n) \geqslant -\dfrac{\mathrm{e}}{2} u_n$.
  × d'autre part : $u_n \leqslant \dfrac{\mathrm{e}^{u_n}}{3}$. Ainsi : $\mathrm{e}^{u_n} \geqslant 3u_n$.

On obtient alors :
$$\underbrace{\mathrm{e}^{u_n} - \mathrm{e}\ln(u_n)}_{= \, u_{n+1}} \ \geqslant \ \underbrace{3u_n - \frac{\mathrm{e}}{2}u_n}_{= \, \frac{6-\mathrm{e}}{2}\, u_n}$$
$$\boxed{\forall n \in \mathbb{N},\ u_{n+1} \geqslant \frac{6-\mathrm{e}}{2} u_n}$$

□

c) Déterminer la nature de la série de terme général $\dfrac{1}{u_n}$.

Démonstration.
Soit $n \in \mathbb{N}$.

- D'après la question ***9.b)*** :
$$\begin{array}{rclcl}
u_{n+1} & \geqslant & \frac{6-\mathrm{e}}{2} \ u_n & & \\
& \geqslant & \frac{6-\mathrm{e}}{2} \ \frac{6-\mathrm{e}}{2} \ u_{n-1} & = & \left(\frac{6-\mathrm{e}}{2}\right)^2 u_{n-1} \\
& \geqslant & \frac{6-\mathrm{e}}{2} \ \frac{6-\mathrm{e}}{2} \ \frac{6-\mathrm{e}}{2} \ u_{n-2} & = & \left(\frac{6-\mathrm{e}}{2}\right)^3 u_{n-2} \\
& \dots & & & \\
& \geqslant & \frac{6-\mathrm{e}}{2} \ \frac{6-\mathrm{e}}{2} \ \frac{6-\mathrm{e}}{2} \ \frac{6-\mathrm{e}}{2} \ \dots \ \frac{6-\mathrm{e}}{2} \ u_0 & = & \left(\frac{6-\mathrm{e}}{2}\right)^{n+1} u_0
\end{array}$$
On peut démontrer rigoureusement ce résultat par récurrence (voir remarque en page suivante).
$$\boxed{\text{Comme } u_0 = 2, \text{ on obtient : } \forall n \in \mathbb{N},\ u_n \geqslant 2\left(\frac{6-\mathrm{e}}{2}\right)^n.}$$

- Soit $n \in \mathbb{N}$.

$$\begin{array}{rcll} u_n & \geqslant & 2\left(\dfrac{6-\mathrm{e}}{2}\right)^n & \\ \text{donc} \quad \dfrac{1}{u_n} & \leqslant & \dfrac{1}{2}\left(\dfrac{2}{6-\mathrm{e}}\right)^n & \begin{array}{l}\textit{(par stricte décroissance de la}\\ \textit{fonction inverse sur }]0,+\infty[,\\ \textit{car } u_n \geqslant 2 > 0 \textit{ d'après } \textbf{\textit{5.}}\textit{)}\end{array} \end{array}$$

- On obtient alors :
  × $\forall n \in \mathbb{N},\ 0 \leqslant \dfrac{1}{u_n} \leqslant \left(\dfrac{2}{6-\mathrm{e}}\right)^n$
  × la série $\sum \left(\dfrac{2}{6-\mathrm{e}}\right)^n$ est une série géométrique de raison $\dfrac{2}{6-\mathrm{e}} \in]-1,1[$.
  (en effet : $3 < 3,2 < 6-\mathrm{e} < 3,3 < 4$. D'où : $-1 < \dfrac{2}{4} < \dfrac{2}{6-\mathrm{e}} < \dfrac{2}{3} < 1$)
  Elle est donc convergente.

$\boxed{\text{Par critère de comparaison de séries à termes positifs, la série } \sum \dfrac{1}{u_n} \text{ est convergente.}}$

Commentaire

- On pouvait aussi résoudre cette question de la façon suivante.
  On sait que les termes de la suite (u_n) sont non nuls, donc, d'après la question ***9.b)*** :

$$\forall k \in \mathbb{N},\ \frac{2}{6-\mathrm{e}} \geqslant \frac{u_k}{u_{k+1}}$$

  En effectuant le produit de ces inégalités pour k variant de 0 à $n-1$, on obtient (tous les termes considérés sont positifs) :

$$\prod_{k=0}^{n-1} \frac{2}{6-\mathrm{e}} \geqslant \prod_{k=1}^{n-1} \frac{u_k}{u_{k+1}} \quad \text{donc} \quad \left(\frac{2}{6-\mathrm{e}}\right)^n \geqslant \frac{u_0}{\cancel{u_1}}\,\frac{\cancel{u_1}}{\cancel{u_2}} \cdots \frac{\cancel{u_{n-1}}}{u_n} = \frac{u_0}{u_n} = \frac{2}{u_n}$$

  On conclut en utilisant le critère de comparaison de séries à termes positifs.
- La question portait ici sur l'application du critère de comparaison des séries à termes positifs. On pouvait se permettre de simplement citer le principe de récurrence. Détaillons quand même cette rédaction.

Commentaire

- Démontrons par récurrence : $\forall n \in \mathbb{N}$, $\mathcal{P}(n)$, où $\mathcal{P}(n)$: $u_n \geqslant 2\left(\dfrac{6-\mathrm{e}}{2}\right)^n$

  - **Initialisation** :
    D'une part, $u_0 = 2$.
    D'autre part, $2\left(\dfrac{6-\mathrm{e}}{2}\right)^0 = 2$.
    D'où $\mathcal{P}(0)$.
  - **Hérédité** : soit $n \in \mathbb{N}$.
    Supposons $\mathcal{P}(n)$ et démontrons $\mathcal{P}(n+1)$ (c'est-à-dire $u_{n+1} \geqslant 2\left(\dfrac{6-\mathrm{e}}{2}\right)^{n+1}$)
    D'après la question ***9.b)***, $u_{n+1} \geqslant \dfrac{6-\mathrm{e}}{2} u_n$.
    Par hypothèse de récurrence, on a : $u_n \geqslant 2\left(\dfrac{6-\mathrm{e}}{2}\right)^n$. Ainsi, par transitivité :

$$u_{n+1} \geqslant \frac{6-\mathrm{e}}{2}\, u_n \geqslant \frac{6-\mathrm{e}}{2}\, 2\left(\frac{6-\mathrm{e}}{2}\right)^n = 2\left(\frac{6-\mathrm{e}}{2}\right)^{n+1}$$

    D'où $\mathcal{P}(n+1)$.

  Par principe de récurrence, on obtient : $\forall n \in \mathbb{N}$, $u_n \geqslant 2\left(\dfrac{6-\mathrm{e}}{2}\right)^n$. □

Problème (ESSEC II 2003)

L'objectif du problème est d'étudier les rudiments de la théorie de la communication *- ou théorie de l'information - introduite en 1948 par Claude Shannon.*

Définitions et notations

$(\Omega, \mathcal{A}, \mathbb{P})$ désigne un espace probabilisé.

φ est la fonction définie sur $]0,1]$ par $x \mapsto \varphi(x) = -\dfrac{\ln(x)}{\ln(2)}$.

Pour un événement A de probabilité non nulle, on pose $i(A) = \varphi(\mathbb{P}(A))$. h est la fonction définie sur $[0,1]$ par :

$$h(0) = 0 \qquad \text{et} \qquad \forall x \in \,]0,1], \quad h(x) = -x\,\frac{\ln(x)}{\ln(2)}$$

Pour une variable aléatoire X discrète définie sur $(\Omega, \mathscr{A}, \mathbb{P})$ à valeurs réelles, on pose sous réserve d'existence :

$$H(X) \ = \sum_{x \in X(\Omega)} h(\mathbb{P}([X = x]))$$

Si X est à valeurs dans un ensemble fini $\{x_1, x_2, \dots, x_n\}$, alors $H(X)$ existe et, en notant $p_k = \mathbb{P}([X = x_k])$, on a :

$$H(X) \ = \ \sum_{k=1}^{n} h(\mathbb{P}([X = x_k])) \ = \ \sum_{k=1}^{n} h(p_k)$$

Remarque : En théorie de l'information, $i(A)$ est appelé incertitude de l'événement *A et $H(X)$ est l'*incertitude moyenne *- ou* entropie *- de X.*

Partie I : Incertitude des événements

1. On choisit une carte au hasard dans un jeu de 32 cartes.
Soit A l'événement « la carte tirée est la dame de cœur ».
Que valent $\mathbb{P}(A)$ et $i(A)$?

Démonstration.

- Déterminons tout d'abord $\mathbb{P}(A)$.
  - × L'univers Ω est la liste des 32 cartes pouvant être choisies. En particulier, $\text{Card}(\Omega) = 32$. *(comme Ω est un ensemble fini, on choisit de le munir de la tribu $\mathscr{A} = \mathcal{P}(\Omega)$)*
  - × On munit enfin $(\Omega, \mathscr{A})$ de la probabilité uniforme notée $\mathbb{P}$.
  - × Comme il n'y a qu'une seule dame de coeur dans le jeu, on en déduit : $\text{Card}(A) = 1$.
    On obtient alors :
$$\mathbb{P}(A) \ = \ \frac{\text{Card}(A)}{\text{Card}(\Omega)} \ = \ \frac{1}{32}$$
$$\boxed{\mathbb{P}(A) = \frac{1}{32}}$$
- Par définition de l'incertitude, on obtient :

$$i(A) \ = \ \varphi(\mathbb{P}(A)) \ = \ -\frac{\ln\left(\frac{1}{32}\right)}{\ln(2)} \ = \ -\frac{-\ln(32)}{\ln(2)} \ = \ \frac{\ln(32)}{\ln(2)} \ = \ \frac{\ln(2^5)}{\ln(2)} \ = \ \frac{5\,\cancel{\ln(2)}}{\cancel{\ln(2)}} \ = \ 5$$

$$\boxed{i(A) = 5}$$

□

2. Soit $n \in \mathbb{N}^*$. On lance n fois une pièce équilibrée.
A est l'événement « obtenir n fois Pile ».
Préciser $i(A)$.

Démonstration.

- Déterminons tout d'abord $\mathbb{P}(A)$.
  - × L'univers Ω est l'ensemble des n-uplets d'éléments de l'ensemble $\{\text{Pile}, \text{Face}\}$. En particulier : $\text{Card}(\Omega) = 2^n$.
    (comme Ω est un ensemble fini, on choisit de le munir de la tribu $\mathscr{A} = \mathcal{P}(\Omega)$)
  - × On munit enfin $(\Omega, \mathscr{A})$ de la probabilité uniforme notée $\mathbb{P}$.
  - × Seul le n-lancer $(\text{Pile}, \text{Pile}, \dots, \text{Pile})$ réalise l'événement A. On en déduit : $\text{Card}(A) = 1$. On obtient alors :

$$\mathbb{P}(A) = \frac{\text{Card}(A)}{\text{Card}(\Omega)} = \frac{1}{2^n}$$

$$\boxed{\mathbb{P}(A) = \frac{1}{2^n}}$$

- Par définition de l'incertitude, on obtient :

$$i(A) = -\frac{\ln\left(\frac{1}{2^n}\right)}{\ln(2)} = \frac{\ln(2^n)}{\ln(2)} = \frac{n\ \cancel{\ln(2)}}{\cancel{\ln(2)}} = n$$

$$\boxed{i(A) = n}$$

Commentaire

On aurait pu traîter cette question en décomposant l'événement A à l'aide événements élémentaires (détails ci-dessous). Ce n'était cependant pas ce qui était attendu à cette question : c'est l'objet de la question *4.* Détaillons cette décomposition.
Pour tout $k \in [\![1, n]\!]$, on introduit les événements suivants :

$$P_k = \text{« obtenir Pile au } k^{\text{ème}} \text{ lancer »}$$

Par définition de l'événement A :

$$A = P_1 \cap P_2 \cap \dots \cap P_n = \bigcap_{k=1}^{n} P_k$$

On obtient :

$$\begin{aligned} \mathbb{P}(A) &= \mathbb{P}\left(\bigcap_{k=1}^{n} P_k\right) \\ &= \prod_{k=1}^{n} \mathbb{P}(P_k) && \textit{(par indépendance des lancers)} \\ &= \prod_{k=1}^{n} \frac{1}{2} && \textit{(car la pièce est équilibrée)} \\ &= \left(\frac{1}{2}\right)^n = \frac{1}{2^n} \end{aligned}$$

□

3. Vérifier les points suivants :

(i) Pour un événement Ω' quasi-certain : $i(\Omega') = 0$.

Démonstration.
Soit Ω' un événement quasi-certain. Alors : $\mathbb{P}(\Omega') = 1$. On obtient donc :

$$i(\Omega') = -\frac{\ln(\mathbb{P}(\Omega'))}{\ln(2)} = -\frac{\ln(1)}{\ln(2)} = -\frac{0}{\ln(2)} = 0$$

$$\boxed{\text{Pour un événement } \Omega' \text{ quasi-certain : } i(\Omega') = 0.}$$

□

(ii) Si A et l'événement contraire $\overline{A}$ sont équiprobables, alors $i(A) = 1$.

Démonstration.
Soit A un événement. Supposons que A et $\overline{A}$ sont équiprobables.
Alors : $\mathbb{P}(A) = \mathbb{P}(\overline{A})$.

- Comme $\mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A)$, on obtient :

$$\begin{array}{lrcl} & \mathbb{P}(A) & = & 1 - \mathbb{P}(A) \\ \text{donc} & 2\,\mathbb{P}(A) = 1 \\ \text{d'où} & \mathbb{P}(A) & = & \dfrac{1}{2} \end{array}$$

$$\boxed{\text{On en déduit : } \mathbb{P}(A) = \frac{1}{2} \text{ (et } \mathbb{P}(\overline{A}) = \frac{1}{2}\text{).}}$$

- Enfin, par définition de l'incertitude :

$$i(A) = -\frac{\ln(\mathbb{P}(A))}{\ln(2)} = -\frac{\ln\left(\frac{1}{2}\right)}{\ln(2)} = -\frac{-\cancel{\ln(2)}}{\cancel{\ln(2)}} = 1$$

$$\boxed{\text{Si } A \text{ et } \overline{A} \text{ sont équiprobables, alors } i(A) = 1.}$$

□

(iii) Si A et B sont indépendants pour la probabilité $\mathbb{P}$ et si $\mathbb{P}(A \cap B) \neq 0$, alors $i(A \cap B) = i(A) + i(B)$.

Démonstration.
Soient A et B deux événements. Supposons que A et B sont indépendants et : $\mathbb{P}(A \cap B) \neq 0$.
Alors : $\mathbb{P}(A \cap B) > 0$. Ainsi $i(A \cap B)$ est bien définie. De plus :

$$\begin{array}{rcll} i(A \cap B) & = & -\dfrac{\ln(\mathbb{P}(A \cap B))}{\ln(2)} & \\ & = & -\dfrac{\ln(\mathbb{P}(A)\ \mathbb{P}(B))}{\ln(2)} & \textit{(car } A \textit{ et } B \textit{ sont indépendants)} \\ & = & -\dfrac{\ln(\mathbb{P}(A)) + \ln(\mathbb{P}(B))}{\ln(2)} & \\ & = & -\dfrac{\ln(\mathbb{P}(A))}{\ln(2)} - \dfrac{\ln(\mathbb{P}(B))}{\ln(2)} & \\ & = & i(A) + i(B) & \end{array}$$

$$\boxed{\text{Si } A \text{ et } B \text{ sont indépendants et } \mathbb{P}(A \cap B) \neq 0, \text{ alors } i(A \cap B) = i(A) + i(B).}$$

□

4. Préciser $i(A_1 \cap A_2 \cap \ldots \cap A_n)$ quand les événements $A_1, A_2, \ldots, A_n$ sont mutuellement indépendants et $\mathbb{P}(A_1 \cap A_2 \cap \ldots \cap A_n) \neq 0$.
En déduire une nouvelle démonstration de *2*.

Démonstration.

- Soient A_1, A_2, ..., A_n des événements. Supposons qu'ils sont mutuellement indépendants et : $\mathbb{P}(A_1 \cap A_2 \cap \cdots \cap A_n) \neq 0$.

  $\boxed{\text{Par récurrence : } i(A_1 \cap A_2 \cap \cdots \cap A_n) = i(A_1) + i(A_2) + \cdots + i(A_n).}$

Commentaire

L'énoncé demandait ici simplement de **préciser** un résultat. Il n'était donc pas nécessaire de faire une démonstration complète. Démontrons maintenant ce point proprement.
Démontrons par récurrence : $\forall n \in \mathbb{N}^*$, $\mathcal{P}(n)$,
où $\mathcal{P}(n)$: pour tous événements A_1, A_2, ..., A_n, si A_1, ..., A_n sont mutuellement indépendants et $\mathbb{P}(A_1 \cap A_2 \cap \cdots \cap A_n) \neq 0$, alors : $i(A_1 \cap A_2 \cap \cdots \cap A_n) = i(A_1) + i(A_2) + \cdots + i(A_n)$.

▸ **Initialisation** :
Soit A_1 un événement. Supposons : $\mathbb{P}(A_1) \neq 0$. Alors : $i(A_1) = i(A_1)$.
D'où $\mathcal{P}(1)$.

▸ **Hérédité** : soit $n \in \mathbb{N}^*$.
Supposons $\mathcal{P}(n)$ et démontrons $\mathcal{P}(n+1)$ (*i.e.* pour tous événements A_1, A_2, ..., A_n, A_{n+1}, si A_1, ..., A_{n+1} sont mutuellement indépendants et $\mathbb{P}(A_1 \cap \cdots \cap A_n \cap A_{n+1}) \neq 0$, alors : $i(A_1 \cap \cdots \cap A_n \cap A_{n+1}) = i(A_1) + \cdots + i(A_n) + i(A_{n+1})$).
Soient A_1, ..., A_n, A_{n+1} des événements. Supposons qu'ils sont mutuellement indépendants et : $\mathbb{P}(A_1 \cap \cdots \cap A_n \cap A_{n+1}) \neq 0$.

× Alors tout d'abord : $\mathbb{P}(A_1 \cap \cdots \cap A_n \cap A_{n+1}) > 0$. Ainsi $i(A_1 \cap A_2 \cap \cdots \cap A_n)$ est bien définie.
De plus, soit $k \in [\![1, n+1]\!]$: $A_1 \cap \cdots \cap A_{n+1} \subset A_k$. On en déduit :

$$\mathbb{P}(A_1 \cap \cdots \cap A_{n+1}) \leqslant \mathbb{P}(A_k)$$

D'où, par transitivité : $\mathbb{P}(A_k) > 0$. Ainsi, pour tout $k \in [\![1, n+1]\!]$, $i(A_k)$ est bien définie.

× On remarque ensuite :

$$\begin{aligned} \mathbb{P}(A_1 \cap \cdots \cap A_n \cap A_{n+1}) &= \mathbb{P}(A_1) \cdots \mathbb{P}(A_n)\, \mathbb{P}(A_{n+1}) && \textit{(par indépendance mutuelle de } A_1, \cdots, A_n, A_{n+1}\textit{)} \\ &= \mathbb{P}(A_1 \cap \cdots \cap A_n)\, \mathbb{P}(A_{n+1}) && \textit{(par indépendance mutuelle de } A_1, \cdots, A_n\textit{)} \end{aligned}$$

On en déduit que les événements $B = A_1 \cap \cdots \cap A_n$ et A_{n+1} sont indépendants.

× Ains les événements B et A_{n+1} sont indépendants et :

$$\mathbb{P}(B \cap A_{n+1}) = \mathbb{P}\big((A_1 \cap \cdots \cap A_n) \cap A_{n+1}\big) \neq 0$$

On en déduit :

$$\begin{aligned} i(A_1 \cap \cdots \cap A_n \cap A_{n+1}) &= i(B \cap A_{n+1}) \\ &= i(B) + i(A_{n+1}) && \textit{(d'après } \textbf{\textit{3.(iii)}}\textit{)} \\ &= i(A_1 \cap \cdots \cap A_n) + i(A_{n+1}) \\ &= i(A_1) + \cdots + i(A_n) + i(A_{n+1}) && \textit{(par hypothèse de récurrence)} \end{aligned}$$

D'où $\mathcal{P}(n+1)$.
Par principe de récurrence, on a bien résolu la question.

- Démontrons de nouveau ***2.***
  - $\times$ Pour tout $k \in [\![1, n]\!]$, on introduit les événements suivants :

$$P_k \; = \; \text{« obtenir Pile au } k^{\text{ème}} \text{ lancer »}$$

Par définition de l'événement A :

$$A \; = \; P_1 \cap P_2 \cap \cdots \cap P_n$$

  - $\times$ Les lancers sont indépendants, donc P_1, P_2, ..., P_n sont mutuellement indépendants. De plus :

$$\begin{array}{rcll} \mathbb{P}(P_1 \cap \cdots \cap P_n) & = & \mathbb{P}(P_1) \times \cdots \times \mathbb{P}(P_n) & \textit{(par indépendance mutuelle de } P_1, \ldots, P_n\textit{)} \\ & = & \left(\dfrac{1}{2}\right)^n & \textit{(car la pièce est équilibrée)} \end{array}$$

On en déduit : $\mathbb{P}(P_1 \cap \cdots \cap P_n) \neq 0$.
On peut donc appliquer le point précédent. On obtient :

$$i(A) \; = \; i(P_1 \cap P_2 \cap \cdots \cap P_n) \; = \; i(P_1) + i(P_2) + \cdots + i(P_n)$$

  - $\times$ Soit $k \in [\![1, n]\!]$. Comme $\mathbb{P}(P_k) = \dfrac{1}{2}$, les événements P_k et $\overline{P_k}$ sont équiprobables. Ainsi, d'après ***3.(ii)*** : $i(P_1) = 1$.

$$\boxed{\text{On en déduit : } i(A) = 1 + \cdots + 1 = n}$$

□

5. Soit A et B deux événements tels que $A \subset B$ et $P(A) \neq 0$. Comparer $i(A)$ et $i(B)$.

Démonstration.

- Tout d'abord, comme $\mathbb{P}(A) \neq 0$, alors : $\mathbb{P}(A) > 0$. Ainsi $i(A)$ est bien définie.
  De plus, comme $A \subset B$, alors :

$$\mathbb{P}(A) \; \leqslant \; \mathbb{P}(B)$$

  D'où, par transitivité : $\mathbb{P}(B) > 0$. Ainsi $i(B)$ est bien définie.
- De plus :

$$\begin{array}{llcll} & \mathbb{P}(A) \leqslant \mathbb{P}(B) & & & \\ \text{donc} & \ln(\mathbb{P}(A)) & \leqslant & \ln(\mathbb{P}(B)) & \textit{(par croissance de } \ln \textit{ sur }]0, +\infty[\textit{)} \\ \text{d'où} & -\ln(\mathbb{P}(A)) & \geqslant & -\ln(\mathbb{P}(B)) & \\ \text{ainsi} & -\dfrac{\ln(\mathbb{P}(A))}{\ln(2)} & \geqslant & -\dfrac{\ln(\mathbb{P}(B))}{\ln(2)} & \textit{(car } \ln(2) > 0\textit{)} \\ \text{ainsi} & i(A) \geqslant i(B) & & & \end{array}$$

$$\boxed{i(A) \geqslant i(B)}$$

□

6. Que vaut $\lim\limits_{x \to 0^+} \varphi(x)$ et quelle interprétation peut-on donner de ce résultat ?

Démonstration.

- On sait : $\lim\limits_{x \to 0^+} \ln(x) = -\infty$.

$$\boxed{\text{Ainsi : } \lim_{x \to 0^+} \varphi(x) = +\infty.}$$

- Soit A un événement. Par définition de l'incertitude : $i(A) = \varphi(\mathbb{P}(A))$.
Ainsi, d'après le point précédent, si $\mathbb{P}(A)$ tend vers 0, alors, alors $i(A)$ tend vers $+\infty$.
Autrement dit, si l'événements A a peu de chances d'être réalisé, alors son incertitude est grande.

□

Partie II : Incertitude d'une variable aléatoire discrète

7. Soit $n \in \mathbb{N}^*$. Si U_n suit la loi uniforme sur $\{1, 2, \ldots, n\}$, que vaut $H(U_n)$?

Démonstration.

- Tout d'abord, la v.a.r. U_n admet une entropie car c'est une v.a.r. finie.
- De plus :

$$\begin{array}{rcll} H(U_n) & = & \sum\limits_{k=1}^{n} h\big(\mathbb{P}([U_n = k])\big) & \textit{(car : } U_n(\Omega) = [\![1, n]\!]\textit{)} \\ & = & \sum\limits_{k=1}^{n} h\left(\dfrac{1}{n}\right) & \textit{(car } U_n \hookrightarrow \mathcal{U}([\![1, n]\!])\textit{)} \\ & = & n\ h\left(\dfrac{1}{n}\right) & \\ & = & \cancel{n}\left(-\dfrac{1}{\cancel{n}}\dfrac{\ln\left(\frac{1}{n}\right)}{\ln(2)}\right) & \\ & = & -\dfrac{-\ln(n)}{\ln(2)} & \end{array}$$

$$\boxed{\text{Finalement : } H(U_n) = \frac{\ln(n)}{\ln(2)}}$$

□

8. Si on suppose $\mathbb{P}([Z = 1]) = \dfrac{1}{4}$, $\mathbb{P}([Z = 2]) = \dfrac{1}{4}$ et $\mathbb{P}([Z = 3]) = \dfrac{1}{2}$, que vaut $H(Z)$?
Comparer $H(Z)$ et $H(U_3)$.

Démonstration.

- Tout d'abord, la v.a.r. Z admet une entropie car c'est une v.a.r. finie.
- De plus :

$$\begin{array}{rcl} H(Z) & = & h\big(\mathbb{P}([Z = 1])\big) + h\big(\mathbb{P}([Z = 2])\big) + h\big(\mathbb{P}([Z = 3])\big) \\ & = & h\left(\dfrac{1}{4}\right) + h\left(\dfrac{1}{2}\right) + h\left(\dfrac{1}{4}\right) \\ & = & 2h\left(\dfrac{1}{4}\right) + h\left(\dfrac{1}{2}\right) \end{array}$$

On en déduit, par définition de h :

$$\begin{aligned} H(Z) &= 2\left(-\frac{1}{4}\frac{\ln\left(\frac{1}{4}\right)}{\ln(2)}\right) - \frac{1}{2}\frac{\ln\left(\frac{1}{2}\right)}{\ln(2)} \\ &= \frac{1}{2}\,\frac{\ln(4)}{\ln(2)} + \frac{1}{2}\,\frac{\cancel{\ln(2)}}{\cancel{\ln(2)}} \\ &= \frac{1}{2}\,\frac{\ln(2^2)}{\ln(2)} + \frac{1}{2} \\ &= \frac{\cancel{2\ln(2)}}{\cancel{2\ln(2)}} + \frac{1}{2} \\ &= 1 + \frac{1}{2} \end{aligned}$$

$$\boxed{\text{On en déduit : } H(Z) = \frac{3}{2}}$$

- D'après la question ***7.*** : $H(U_3) = \dfrac{\ln(3)}{\ln(2)}$.

  Pour comparer $H(Z)$ et $H(U_3)$, on étudie le signe de leur différence.

$$H(Z) - H(U_3) = \frac{3}{2} - \frac{\ln(3)}{\ln(2)} = \frac{3\ln(2) - 2\ln(3)}{\ln(2)} = \frac{\ln(2^3) - \ln(3^2)}{\ln(2)} = \frac{\ln(8) - \ln(9)}{\ln(2)}$$

  Or :
  - × comme $8 \leqslant 9$, par croissance de ln sur $]0, +\infty[$: $\ln(8) \leqslant \ln(9)$. D'où : $\ln(8) - \ln(9) \leqslant 0$.
  - × de plus : $\ln(2) > 0$

$$\boxed{\text{On en déduit : } H(Z) - H(U_3) \leqslant 0. \text{ Ainsi : } H(Z) \leqslant H(U_3)}$$

□

9. On se propose de simuler informatiquement une variable aléatoire.
On supposera que `grand(1,1,'uin',1,3)` fournit au hasard un nombre élément de $\{1, 2, 3\}$ et que `grand(1,1,'uin',1,2)` fournit au hasard un élément de $\{1, 2\}$.

```
y = 0
ini = grand(1,1,'uin',1,3)
if ini == 3 then
    y = grand(1,1,'uin',1,2)
else
    y = 3
end
```

On appelle Y le contenu de `y` après exécution de ce programme.
Donner la loi de Y, calculer son espérance $\mathbb{E}(Y)$ et son incertitude $H(Y)$.

Démonstration.

- **Début du programme**
  On commence par initialiser la variable `y` à 0.
  (on rappelle qu'en fin de programme la variable `y` *contient une simulation d'une v.a.r. Y dont on doit déterminer la loi)*

```
1  y = 0
```

  On stocke ensuite dans la variable `ini` un nombre choisi au hasard parmi 1, 2 et 3. Autrement dit, la variable `ini` contient une simulation d'une v.a.r. U de loi $\mathcal{U}(⟦1, 3⟧)$.

```
2  ini = grand(1, 1, 'uin', 1, 3)
```

- **Structure conditionnelle**
  Les lignes 3 à 7 consistent à mettre à jour la variable `y` selon la valeur de la variable `ini`.
  × si la variable `ini` contient la valeur 3, alors on stocke dans la variable `y` le résultat de l'appel `grand(1, 1, 'uin', 1, 2)`. Autrement dit, la variable `y` contient une simulation d'une v.a.r. V de loi $\mathcal{U}(⟦1, 2⟧)$.

```
3  if ini == 3 then
4      y = grand(1,1,'uin',1,2)
```

  Avec les v.a.r. introduites précédemment, cela signifie que la loi de la v.a.r. Y conditionnellement à l'événement $[U = 3]$ est la loi $\mathcal{U}(⟦1, 2⟧)$. Ainsi :

$$\mathbb{P}_{[U=3]}([Y = 1]) = \frac{1}{2} \quad \text{et} \quad \mathbb{P}_{[U=3]}([Y = 2]) = \frac{1}{2}$$

  Notons que les probabilités conditionnelles ci-dessus sont bien définies car : $\mathbb{P}([U = 3]) = \dfrac{1}{3} \neq 0$

  × sinon, lorsque la variable `ini` contient une valeur différente de 3, alors on stocke dans la variable `y` la valeur 3.

```
5  else
6      y = 3
```

  Avec les v.a.r. introduites précédemment, cela signifie que la loi de la v.a.r. Y conditionnellement à l'événement $[U \neq 3]$ est la loi certaine égale à 3. Ainsi :

$$\mathbb{P}_{[U\neq 3]}([Y = 3]) \ = \ 1$$

  Notons que cette probabilité conditionnelle est bien définie car :

$$\mathbb{P}([U \neq 3]) \ = \ 1 - \mathbb{P}([U = 3]) \ = \ 1 - \frac{1}{3} \ = \ \frac{2}{3} \ \neq \ 0$$

  × **Loi de Y**
  D'après les explications précédentes :
  × on considère : $Y(\Omega) = \{1, 2, 3\}$.
  × déterminons $\mathbb{P}([Y = 1])$.
  La famille $([U = 3], [U \neq 3])$ forme un système complet d'événements.
  Ainsi, par formule des probabilité totales :

$$\begin{aligned}
\mathbb{P}([Y = 1]) &= \mathbb{P}([U = 3] \cap [Y = 1]) + \mathbb{P}([U \neq 3] \cap [Y = 1]) \\
&= \mathbb{P}([U = 3])\ \mathbb{P}_{[U=3]}([Y = 1]) + \mathbb{P}([U \neq 3])\ \mathbb{P}_{[U\neq 3]}([Y = 1]) \\
&= \frac{1}{3} \times \frac{1}{2} + \frac{2}{3} \times 0 \ = \ \frac{1}{6}
\end{aligned}$$

× Déterminons $\mathbb{P}([Y=2])$.
La famille $([U=3],[U\neq 3])$ forme un système complet d'événements.
Ainsi, par formule des probabilité totales :

$$\begin{array}{rcl}\mathbb{P}([Y=2]) & = & \mathbb{P}([U=3]\cap[Y=2])+\mathbb{P}([U\neq 3]\cap[Y=2]) \\ & = & \mathbb{P}([U=3])\ \mathbb{P}_{[U=3]}([Y=2])+\mathbb{P}([U\neq 3])\ \mathbb{P}_{[U\neq 3]}([Y=2]) \\ & = & \dfrac{1}{3}\times\dfrac{1}{2}+\dfrac{2}{3}\times 0 \ = \ \dfrac{1}{6}\end{array}$$

× Comme $Y(\Omega)=\{1,2,3\}$, la famille $([Y=1],[Y=2],[Y=3])$ forme un système complet d'événements. Ainsi :

$$\mathbb{P}([Y=3]) \ = \ 1-\mathbb{P}([Y=1])-\mathbb{P}([Y=2] \ = \ 1-\frac{1}{6}-\frac{1}{6} \ = \ 1-\frac{1}{3} \ = \ \frac{2}{3}$$

$$\boxed{\begin{array}{c}Y(\Omega)=\{1,2,3\} \\ \mathbb{P}([Y=1])=\dfrac{1}{6},\quad \mathbb{P}([Y=2])=\dfrac{1}{6},\quad \mathbb{P}([Y=3])=\dfrac{2}{3}\end{array}}$$

Commentaire

On pouvait bien sûr déterminer $\mathbb{P}([Y=3])$ de nouveau avec la formule des probabilités totales appliquée avec le système complet d'événements $([U=3],[U\neq 3])$:

$$\begin{array}{rcl}\mathbb{P}([Y=3]) & = & \mathbb{P}([U=3]\cap[Y=3])+\mathbb{P}([U\neq 3]\cap[Y=3]) \\ & = & \mathbb{P}([U=3])\ \mathbb{P}_{[U=3]}([Y=3])+\mathbb{P}([U\neq 3])\ \mathbb{P}_{[U\neq 3]}([Y=3]) \\ & = & \dfrac{1}{3}\times 0+\dfrac{2}{3}\times 1 \ = \ \dfrac{2}{3}\end{array}$$

- **Espérance de Y**

× La v.a.r. Y admet une espérance, car c'est une v.a.r. finie.

× Par définition de l'espérance :

$$\begin{array}{rcl}\mathbb{E}(Y) & = & 1\times\mathbb{P}([Y=1])+2\times\mathbb{P}([Y=2])+3\times\mathbb{P}([Y=3]) \\ & = & \dfrac{1}{6}+2\times\dfrac{1}{6}+3\times\dfrac{2}{3} \\ & = & \dfrac{15}{6}\end{array}$$

$$\boxed{\text{D'où : } \mathbb{E}(Y)=\frac{5}{2}.}$$

- **Entropie de Y**
  - × La v.a.r. Y admet une entropie car c'est une v.a.r. finie.
  - × Par définition de l'entropie :

$$\begin{array}{rcl} H(Y) & = & h\big(\mathbb{P}([Y=1])\big) + h\big(\mathbb{P}([Y=2])\big) + h\big(\mathbb{P}([Y=3])\big) \\ & = & h\left(\frac{1}{6}\right) + h\left(\frac{1}{6}\right) + h\left(\frac{2}{3}\right) \\ & = & 2h\left(\frac{1}{6}\right) + h\left(\frac{2}{3}\right) \\ & = & 2\left(-\frac{1}{6}\frac{\ln\left(\frac{1}{6}\right)}{\ln(2)}\right) - \frac{2}{3}\,\frac{\ln\left(\frac{2}{3}\right)}{\ln(2)} \\ & = & \frac{1}{3}\,\frac{\ln(6)}{\ln(2)} - \frac{2}{3}\,\frac{\ln\left(\frac{2}{3}\right)}{\ln(2)} \\ & = & \frac{\ln(6) - 2(\ln(2)-\ln(3))}{3\ln(2)} \\ & = & \frac{3\ln(3)-\ln(2)}{\ln(2)} \\ & = & \frac{\ln(3)}{\ln(2)} - \frac{1}{3} \end{array}$$

$$\boxed{H(Y) = \frac{\ln(3)}{\ln(2)} - \frac{1}{3}}$$

□

10. Vérifier que h est continue et positive sur $[0,1]$.
Est-elle dérivable en 0 ? Étudier h et dessiner sa courbe représentative.

Démonstration.

- La fonction h est continue :
  - × sur $]0,1]$ car elle est le produit $h = h_1 \times h_2$ de :
    - $h_1 : x \mapsto -x$ continue sur $]0,1]$ en tant que fonction polynomiale,
    - $h_2 : x \mapsto \dfrac{\ln(x)}{\ln(2)}$ continue sur $]0,1]$.
  - × en 0. En effet, par croissances comparées : $\lim\limits_{x\to 0^+} x\ln(x) = 0$. Ainsi :

$$\lim_{x\to 0^+} h(x) \;=\; 0 \;=\; h(0)$$

$$\boxed{\text{La fonction } h \text{ est continue sur } [0,1].}$$

- Soit $x \in \mathbb{R}$. Deux cas se présentent :
  - × si $x = 0$, alors : $h(0) = 0 \geqslant 0$.
  - × si $x \in]0,1]$, alors :
    - d'une part : $-x < 0$,
    - d'autre part : $\ln(x) < 0$

    Ainsi, comme $\ln(2) > 0$: $h(x) = -x\dfrac{\ln(x)}{\ln(2)} > 0$.

$$\boxed{\forall x \in [0,1],\ h(x) \geqslant 0}$$

- Soit $x \in]0,1]$.

$$\tau_0(h)(x) \;=\; \frac{h(x)-h(0)}{x-0} \;=\; \frac{-x\frac{\ln(x)}{\ln(2)}-0}{x-0} \;=\; -\frac{\ln(x)}{\ln(2)}$$

Or : $\lim\limits_{x\to 0^+} \ln(x) = -\infty$. D'où : $\lim\limits_{x\to 0^+} \tau_0(h)(x) = +\infty$.

$$\boxed{\text{La fonction } h \text{ n'est donc pas dérivable en } 0.}$$

- La fonction h est dérivable sur $]0,1]$ en tant que produit de fonctions dérivables sur $]0,1]$.
  - × Soit $x \in]0,1]$.

$$h'(x) \;=\; -\frac{\ln(x)}{\ln(2)} - \frac{\cancel{x}}{\ln(2)} \times \frac{1}{\cancel{x}} \;=\; -\frac{\ln(x)+1}{\ln(2)}$$

  - × On en déduit :

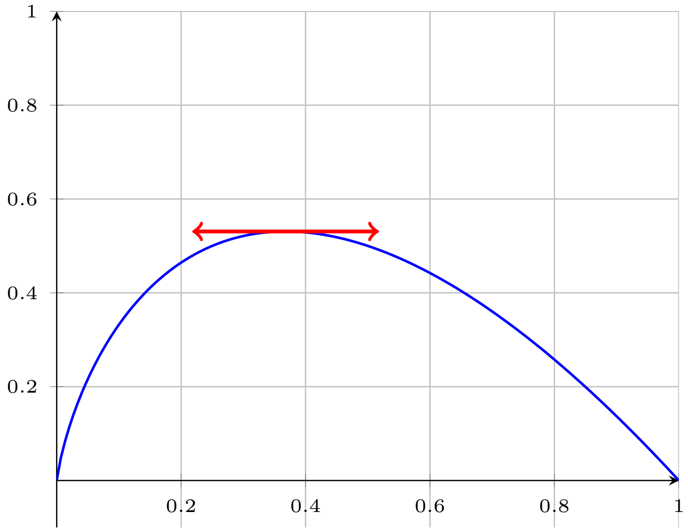
$$\begin{aligned} h'(x) \geqslant 0 \;&\Leftrightarrow\; \ln(x)+1 \geqslant 0 \\ &\Leftrightarrow\; \ln(x) \geqslant -1 \\ &\Leftrightarrow\; x \geqslant \mathrm{e}^{-1} \end{aligned}$$

(par stricte croissance de la fonction exp *sur* $\mathbb{R}$*)*

On obtient le tableau de variations suivant :

x	0		e^{-1}		1
Signe de $h'(x)$		$+$	0	$-$	
Variations de h	0	↗	$\frac{1}{\mathrm{e}\ln(2)}$	↘	0

  - × Détaillons l'obtention des valeurs $h(\mathrm{e}^{-1})$ et $h(1)$.
    - ▸ $h(\mathrm{e}^{-1}) \;=\; -\mathrm{e}^{-1}\,\dfrac{\ln(\mathrm{e}^{-1})}{\ln(2)} \;=\; \dfrac{1}{\mathrm{e}}\,\dfrac{\ln(\mathrm{e})}{\ln(2)} \;=\; \dfrac{1}{\mathrm{e}\ln(2)}$
    - ▸ $h(1) \;=\; -1\,\dfrac{\ln(1)}{\ln(2)} \;=\; 0$



$\square$

11. Soit X une variable aléatoire à valeurs dans un ensemble fini.
Montrer que $H(X) \geqslant 0$ avec égalité si, et seulement si, X est quasi-certaine.

Démonstration.

- La v.a.r. X admet une entropie car c'est une v.a.r. finie.
  Notons : $X(\Omega) = \{x_1, \dots, x_n\}$.
- Par définition de l'entropie :
  $$H(X) \ = \ \sum_{k=1}^{n} h\big(\mathbb{P}([X = x_k])\big)$$
  Or d'après la question ***10.*** : $\forall x \in [0,1]$, $h(x) \geqslant 0$.
  Donc : $\forall k \in [\![1,n]\!]$, $h\big(\mathbb{P}([X = x_k]\,)\big) \geqslant 0$.
  Ainsi $H(X)$ est une somme de termes positifs.
  $$\boxed{\text{On en déduit : } H(X) \geqslant 0}$$
- Démontrons que $H(X) = 0$ si et seulement si X est quasi-certaine.
  On procède par double implication.
  ($\Rightarrow$) Supposons : $H(X) = 0$.
  - × Alors :
    $$\sum_{k=1}^{n} h\big(\mathbb{P}([X = x_k])\big) \ = \ 0$$
    Or on a déjà vu dans le point précédent :
    $$\forall k \in [\![1,n]\!], \quad h\big(\mathbb{P}([X = x_k])\big) \ \geqslant \ 0$$
    Comme la somme précédente est nulle et que tous les termes de cette somme sont positifs, on en déduit qu'ils sont tous nuls :
    $$\forall k \in [\![1,n]\!], \quad h\big(\mathbb{P}([X = x_k])\big) \ = \ 0$$
  - × Démontrons maintenant : $\forall k \in [\![1,n]\!], \mathbb{P}([X = x_k]) \in \{0,1\}$.
    D'après la question précédente :
    - La fonction h est strictement croissante sur $[0, \mathrm{e}^{-1}]$. Ainsi :
      $$\forall x \in \,]0, \mathrm{e}^{-1}], \ h(x) > h(0) = 0$$
    - La fonction h est strictement décroissante sur $[\mathrm{e}^{-1}, 1]$. Ainsi :
      $$\forall x \in \ [\mathrm{e}^{-1}, 1[, \ h(x) > h(1) = 0$$
    - De plus : $h(0) = 0$ et $h(1) = 0$.

    On en déduit :
    $$h(x) = 0 \quad \Leftrightarrow \quad x = 0 \ \texttt{OU} \ x = 1$$
    Soit $k \in [\![1,n]\!]$. D'après le point précédent : $h\big(\mathbb{P}([X = x_k])\big) = 0$.
    On en déduit : $\mathbb{P}([X = x_k]) \in \{0,1\}$.
  - × De plus, la famille $([X = x_k])_{k \in [\![1,n]\!]}$ forme un système complet d'événements. Ainsi :
    $$\sum_{k=1}^{n} \mathbb{P}([X = x_k]) \ = \ 1$$
    Avec le point précédent, on en déduit qu'une seule de ces probabilités vaut 1 et que les autres sont nulles. Autrement dit, il existe $k_0 \in [\![1,n]\!]$ tel que :
    $$\left\{ \begin{array}{l} \mathbb{P}([X = x_{k_0}]) = 1 \\ \forall k \neq k_0, \ \mathbb{P}([X = x_k]) = 0 \end{array} \right.$$

  On en déduit que X est une v.a.r. quasi-certaine (égale à x_{k_0}).

($\Leftarrow$) Supposons que la v.a.r. X est quasi-certaine. Alors il existe $m \in \mathbb{R}$ tel que :

× on peut considérer : $X(\Omega) = \{m\}$,

× $\mathbb{P}([X = m]) = 1$.

On en déduit :

$$H(X) \ = \ h\big(\mathbb{P}([X = m])\big) \ = \ h(1) \ = \ 0$$

$$\boxed{\text{Finalement } H(X) = 0 \text{ si et seulement si } X \text{ est quasi-certaine.}}$$

□

Partie III : Maximalité de l'entropie

12. *Étude pour $n = 2$.*
Pour $x \in [0, 1]$, on pose $h_2(x) = h(x) + h(1 - x)$.

a) Pour $x \in [0, 1]$, on a clairement $h_2(x) = h_2(1 - x)$.
Que signifie ce résultat quant à la courbe de h_2 dans un repère orthonormé ?

Démonstration.

- D'après l'énoncé : $\forall y \in [0, 1]$, $h_2(x) = h_2(1 - x)$.
- Soit $x \in [0, \frac{1}{2}]$.
En appliquant cette égalité à $y = \dfrac{1}{2} + x$, on obtient :

$$h_2\left(\frac{1}{2} + x\right) \ = \ h_2\left(1 - \left(\frac{1}{2} + x\right)\right) \ = \ h_2\left(\frac{1}{2} - x\right)$$

On en déduit que la courbe représentative de h_2 est symétrique par rapport à la droite d'équation $x = \dfrac{1}{2}$.

Commentaire

De manière plus générale, la courbe représentative d'une fonction f admet la droite $x = a$ comme axe de symétrie si :

$$\forall x \in \mathcal{D}_f, \ f(a + x) = f(a - x)$$

□

b) Étudier h_2 et donner son graphe.

Démonstration.
On étudie la fonction h_2 uniquement sur $[0, \frac{1}{2}]$. On déduira l'étude de h_2 sur $[\frac{1}{2}, 1]$ par symétrie.

- La fonction h_2 est continue sur $[0, \frac{1}{2}]$ car elle est la somme $h_2 = h + g_2$ de :
  × h qui est continue sur $[0, \frac{1}{2}]$ d'après la question ***10***,
  × $g_2 : x \mapsto h(1 - x)$ continue sur $[0, \frac{1}{2}]$ car elle est la composée $g_2 = h \circ f_1$ de :
    - $f_1 : x \mapsto 1 - x$ qui est :
      ▸ continue sur $[0, \frac{1}{2}]$,
      ▸ telle que : $f_1([0, \frac{1}{2}]) \subset [0, 1]$.
    - h qui est continue sur $[0, 1]$.

- La fonction h_2 est dérivable sur $]0, \frac{1}{2}]$ par des arguments similaires à ceux de la continuité sur cet intervalle.
  - $\times$ Soit $x \in]0, \frac{1}{2}[$.

$$\begin{array}{rcl} h_2'(x) & = & h'(x) - h'(1-x) \\ & = & -\dfrac{\ln(x) + \cancel{1}}{\ln(2)} + \dfrac{\ln(1-x) + \cancel{1}}{\ln(2)} \quad \textit{(d'après \textbf{10.})} \\ & = & \dfrac{\ln(1-x) - \ln(x)}{\ln(2)} \end{array}$$

Comme $\ln(2) > 0$, on en déduit :

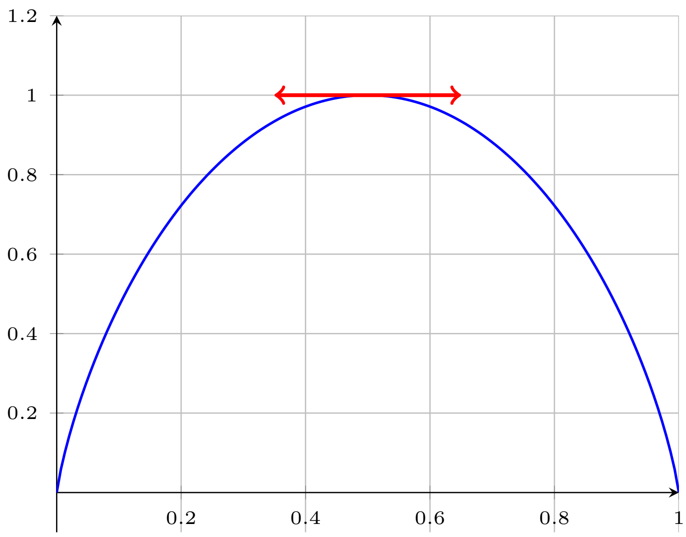
$$\begin{array}{rcl} h_2'(x) \geqslant 0 & \Leftrightarrow & \ln(1-x) - \ln(x) \geqslant 0 \\ & \Leftrightarrow & \ln(1-x) \geqslant \ln(x) \\ & \Leftrightarrow & 1 - x \geqslant x \quad \textit{(par stricte croissance de la fonction } \exp \textit{ sur } \mathbb{R}\textit{)} \\ & \Leftrightarrow & 1 \geqslant 2x \\ & \Leftrightarrow & \dfrac{1}{2} \geqslant x \end{array}$$

  - $\times$ On obtient le tableau de variations suivant, en complétant sur $[\frac{1}{2}, 1]$ par symétrie par rapport à l'axe $x = \frac{1}{2}$:

x	0		$\frac{1}{2}$		1
Signe de $h_2'(x)$		$+$	0	$-$	
Variations de h_2	0	$\nearrow$	1	$\searrow$	0

- Détaillons l'obtention de $h_2(\frac{1}{2})$.

$$h_2\left(\frac{1}{2}\right) = h\left(\frac{1}{2}\right) + h\left(1 - \frac{1}{2}\right) = 2\,h\left(\frac{1}{2}\right) = \cancel{2}\left(-\frac{1}{\cancel{2}}\frac{\ln\left(\frac{1}{2}\right)}{\ln(2)}\right) = -\frac{-\cancel{\ln(2)}}{\cancel{\ln(2)}} = 1$$



$\square$

c) Soit X une variable aléatoire suivant une loi de Bernoulli de paramètre $p \in \,]0,1[$.
Montrer que $H(X) \leqslant 1$ avec égalité si, et seulement si, $p = \dfrac{1}{2}$.

Démonstration.

- La v.a.r. X admet une entropie car c'est une v.a.r. finie.
- Démontrons : $H(X) \leqslant 1$.
  - × Par définition de l'entropie :

$$H(X) \;=\; h\big(\mathbb{P}([X=0])\big) + h\big(\mathbb{P}([X=1])\big) \;=\; h(1-p) + h(p) \;=\; h_2(p)$$

  - × Or, d'après la question précédente, la fonction h_2 admet un maximum égal à 1. D'où :

$$\forall x \in [0,1], \quad h_2(x) \;\leqslant\; 1$$

$$\boxed{\text{On en déduit : } H(X) \leqslant 1.}$$

- Démontrons : $H(X) = 1 \Leftrightarrow p = \dfrac{1}{2}$.
  D'après la question précédente :
  - × la fonction h_2 est strictement croissante sur $[0, \frac{1}{2}]$. Ainsi :

$$\forall x \in [0, \frac{1}{2}[, \quad h_2(x) \;<\; h_2(\frac{1}{2}) = 1$$

  - × la fonction h_2 est strictement décroissante sur $[\frac{1}{2}, 0]$. Ainsi :

$$\forall x \in]\frac{1}{2}, 1], \quad h_2(x) \;<\; h_2(\frac{1}{2}) = 1$$

  - × $h(\frac{1}{2}) = 1$.

  On en déduit :

$$h_2(x) \;=\; 1 \quad \Leftrightarrow \quad x = \frac{1}{2}$$

$$\boxed{\text{Or } H(X) = h_2(p). \text{ On en déduit : } H(X) = 1 \Leftrightarrow p = \frac{1}{2}.}$$

□

13. Soit $p \in \,]0,1[$ et G une variable aléatoire suivant une loi géométrique de paramètre p.
On pose $m = \mathbb{E}(G)$ et pour $k \in \mathbb{N}^*$, $p_k = \mathbb{P}([G=k])$.

a) Rappeler la valeur de m, montrer que $H(G)$ existe et la calculer.

Démonstration.

$$\boxed{\text{Comme } G \hookrightarrow \mathcal{G}(p), \text{ alors : } m = \mathbb{E}(G) = \frac{1}{p}.}$$

- La v.a.r. G admet une entropie si et seulement si la série $\sum\limits_{k \geqslant 1} h\big(\mathbb{P}([G=k])\big)$ est convergente.

- Tout d'abord, soit $k \in \mathbb{N}^*$.

$$\begin{array}{rcll} h\big(\mathbb{P}([G=k])\big) & = & h(p_k) \ = \ -p_k \dfrac{\ln(p_k)}{\ln(2)} & \\ & = & -p_k \ \dfrac{\ln\big((1-p)^{k-1}p\big)}{\ln(2)} & \textit{(car } G \hookrightarrow \mathcal{G}(p)\textit{)} \\ & = & -\dfrac{p_k}{\ln(2)} \ \big((k-1)\ln(1-p) + \ln(p)\big) & \\ & = & -\dfrac{p_k}{\ln(2)} \ \big(k\ln(1-p) - \ln(1-p) + \ln(p)\big) & \\ & = & -\dfrac{p_k}{\ln(2)} \left(k\ln(1-p) + \ln\left(\dfrac{p}{1-p}\right)\right) & \end{array}$$

- Soit $n \in \mathbb{N}^*$. On obtient alors :

$$\begin{array}{cl} & \sum\limits_{k=1}^{n} h\big(\mathbb{P}([G=k])\big) \\ = & \sum\limits_{k=1}^{n} \left(-\dfrac{p_k}{\ln(2)} \left(k\ln(1-p) + \ln\left(\dfrac{p}{1-p}\right)\right)\right) \\ = & -\dfrac{1}{\ln(2)} \sum\limits_{k=1}^{n} \left(k\, p_k \ \ln(1-p) + p_k \ \ln\left(\dfrac{p}{1-p}\right)\right) \\ = & -\dfrac{1}{\ln(2)} \left(\ln(1-p) \sum\limits_{k=1}^{n} k \ p_k + \ln\left(\dfrac{p}{1-p}\right) \sum\limits_{k=1}^{n} p_k\right) \\ = & -\dfrac{1}{\ln(2)} \left(\ln(1-p) \sum\limits_{k=1}^{n} k \ \mathbb{P}([G=k]) + \ln\left(\dfrac{p}{1-p}\right) \sum\limits_{k=1}^{n} \mathbb{P}([G=k])\right) \end{array}$$

Or :

× la famille $(\,[G=k]\,)_{k\in\mathbb{N}^*}$ forme un système complet d'événements. Ainsi la série $\sum\limits_{k\geqslant 1} \mathbb{P}([G=k])$ est convergente et :

$$\sum_{k=1}^{+\infty} \mathbb{P}([G=k]) \ = \ 1$$

× la v.a.r. G admet une espérance. Ainsi la série $\sum\limits_{k\geqslant 1} k\,\mathbb{P}([G=k])$ est convergente (elle est même absolument convergente) et :

$$\sum_{k=1}^{+\infty} k\,\mathbb{P}([G=k]) \ = \ \mathbb{E}(G) \ = \ \frac{1}{p}$$

On en déduit que la série $\sum\limits_{k\geqslant 1} h\big(\mathbb{P}([G=k])\big)$ est convergente.

$$\boxed{\text{On en conclut que } G \text{ admet une entropie.}}$$

- De plus :

$$H(G) \ = \ \sum_{k=1}^{+\infty} h\big(\mathbb{P}([G=k])\big) \ = \ -\frac{1}{\ln(2)} \left(\ln(1-p) \times \frac{1}{p} + \ln\left(\frac{p}{1-p}\right) \times 1\right)$$

$$\boxed{H(G) = -\frac{1}{\ln(2)} \left(\frac{\ln(p)}{p} + \ln\left(\frac{p}{1-p}\right)\right)}$$

□

b) Justifier par un argument de convexité :

$$\forall u > 0, \quad \ln(u) \leqslant u - 1$$

▶ *Dans la suite, on pourra utiliser sans démonstration que* $\ln(u) = u - 1$ *si, et seulement si,* $u = 1$.

Démonstration.
La fonction $f : u \mapsto \ln(u)$ est concave sur $]0, +\infty[$.
Sa courbe représentative $\mathcal{C}_f$ est donc située en dessous de ses tangentes, notamment celle au point d'abscisse 1. Or cette tangente est la droite d'équation :

$$\begin{aligned} y &= f(1) + f'(1)\,(x-1) \\ &= \ln(1) + \frac{1}{1}\,(x-1) \;=\; x - 1 \end{aligned}$$

$$\boxed{\text{On en déduit : } \forall u \in\,]0, +\infty[,\ \ln(u) \leqslant u - 1}$$ □

c) Soit X une variable aléatoire telle que $X(\Omega) = \mathbb{N}^*$, $\mathbb{E}(X) = m$ et $H(X)$ existe.
Pour $k \in \mathbb{N}^*$, on pose $q_k = \mathbb{P}([X = k])$ et on supposera : $q_k > 0$.
En utilisant ***13.b)***, vérifier que pour tout $k \in \mathbb{N}^*$, on a :

$$q_k \ln(p) + (k-1)q_k \ln(1-p) - q_k \ln(q_k) \leqslant p_k - q_k$$

et établir : $H(X) \leqslant H(G)$ avec égalité si, et seulement si, X suit la même loi que G.

Démonstration.

- Soit $k \in \mathbb{N}^*$.

$$\begin{aligned} & q_k \ln(p) + (k-1)q_k \ln(1-p) - q_k \ln(q_k) \\ =\ & q_k \ln(p) + q_k \ln\left((1-p)^{k-1}\right) - q_k \ln(q_k) \\ =\ & q_k \left(\ln(p) + \ln\left((1-p)^{k-1}\right) - \ln(q_k)\right) \\ =\ & q_k \left(\ln\left(p(1-p)^{k-1}\right) - \ln(q_k)\right) \\ =\ & q_k \left(\ln(p_k) - \ln(q_k)\right) && \textit{(avec les notations de } \textbf{\textit{13.a)}} \textit{ car } G \hookrightarrow \mathcal{G}(p)\textit{)} \\ =\ & q_k \ \ln\left(\frac{p_k}{q_k}\right) \end{aligned}$$

On applique alors la question ***13.b)*** à $u = \dfrac{p_k}{q_k} \in\,]0, +\infty[$. On obtient :

$$\ln\left(\frac{p_k}{q_k}\right) \leqslant \frac{p_k}{q_k} - 1 \qquad (\star)$$

Comme $q_k \geqslant 0$, on en déduit :

$$q_k \ \ln\left(\frac{p_k}{q_k}\right) \leqslant \underbrace{q_k \left(\frac{p_k}{q_k} - 1\right)}_{=\ p_k - q_k}$$

$$\boxed{\text{Ainsi : } \forall k \in \mathbb{N}^*,\ q_k \ln(p) + (k-1)q_k \ln(1-p) - q_k \ln(q_k) \leqslant p_k - q_k}$$

- Démontrons : $H(X) \leqslant H(G)$.
Pour comparer $H(G)$ et $H(X)$, on étudie le signe de leur différence.
  × Tout d'abord :

$$\begin{array}{rcl} H(G) - H(X) & = & \sum\limits_{k=1}^{+\infty} h(p_k) - \sum\limits_{k=1}^{+\infty} h(q_k) \\ & = & \sum\limits_{k=1}^{+\infty} \big(h(p_k) - h(q_k)\big) \\ & = & \sum\limits_{k=1}^{+\infty} \left(-p_k \,\dfrac{\ln(p_k)}{\ln(2)} + q_k \,\dfrac{\ln(q_k)}{\ln(2)}\right) \\ & = & -\dfrac{1}{\ln(2)} \sum\limits_{k=1}^{+\infty} \big(p_k \ln(p_k) - q_k \ln(q_k)\big) \end{array}$$

  × Soit $k \in \mathbb{N}^*$. D'après le point précédent :

$$\begin{array}{lrcl} & q_k \ \ln(p) + (k-1)\, q_k \ \ln(1-p) - q_k \ \ln(q_k) & \leqslant & p_k - q_k \\ \text{donc} & -q_k \ \ln(q_k) & \leqslant & p_k - q_k - q_k \ \ln(p) - (k-1)\, q_k \ \ln(1-p) \\ \text{d'où} & p_k \ln(p_k) - q_k \ln(q_k) & \leqslant & p_k \ln(p_k) + p_k - q_k - q_k \ln(p) - (k-1) q_k \ln(1-p) \end{array}$$

  Simplifions le terme de droite de cette inégalité.

$$\begin{array}{rl} & p_k \ln(p_k) + p_k - q_k - q_k \ln(p) - (k-1) q_k \ln(1-p) \\ = & p_k \ \ln\big((1-p)^{k-1} p\big) + p_k - q_k\big(1 + \ln(p) + (k-1)\ \ln(1-p)\big) \qquad \textit{(car } G \hookrightarrow \mathcal{G}\,(p)\textit{)} \\ = & p_k\big((k-1)\ \ln(1-p) + \ln(p)\big) + p_k - q_k\big(1 + \ln(p) + (k-1)\ \ln(1-p)\big) \\ = & p_k \ \big(1 + \ln(p) + (k-1)\ \ln(1-p)\big) - q_k\big(1 + \ln(p) + (k-1)\ \ln(1-p)\big) \\ = & (p_k - q_k)\ \big(1 + \ln(p) + (k-1)\ \ln(1-p)\big) \\ = & (p_k - q_k)\ \big(1 + \ln(p) - \ln(1-p) + k\ \ln(1-p)\big) \\ = & (p_k - q_k) \left(1 + \ln\left(\dfrac{p}{1-p}\right) + k\ \ln(1-p)\right) \\ = & \left(1 + \ln\left(\dfrac{p}{1-p}\right)\right) (p_k - q_k) + \ln(1-p)\, k\ p_k - \ln(1-p)\, k\ q_k \end{array}$$

On en déduit :

$$\boxed{p_k \ \ln(p_k) - q_k \ \ln(q_k) \ \leqslant \ \left(1 + \ln\left(\dfrac{p}{1-p}\right)\right) (p_k - q_k) + \ln(1-p)\, k\ p_k - \ln(1-p)\, k\ q_k.}$$

  × Or :
    - les famille $(p_k)_{k\in\mathbb{N}^*}$ et $(q_k)_{k\in\mathbb{N}^*}$ forment des systèmes complets d'événements (par définition de (p_k) et (q_k)). Ainsi les séries $\sum\limits_{k\geqslant 1} p_k$ et $\sum\limits_{k\geqslant 1} q_k$ sont convergentes et :

$$\sum_{k=1}^{+\infty} p_k \ = \ 1 \quad \text{et} \quad \sum_{k=1}^{+\infty} q_k \ = \ 1$$

- les v.a.r. X et G admettent une espérance. Ainsi les séries $\sum\limits_{k\geqslant 1} k\, p_k$ et $\sum\limits_{k\geqslant 1} k\, q_k$ sont convergentes. De plus, d'après l'énoncé : $\mathbb{E}(X) = m = \mathbb{E}(G) = \dfrac{1}{p}$. Ainsi :

$$\sum_{k=1}^{+\infty} k\ p_k \ = \ \mathbb{E}(G) \ = \ \frac{1}{p} \quad \text{et} \quad \sum_{k=1}^{+\infty} k\ q_k \ = \ \mathbb{E}(X) \ = \ \frac{1}{p}$$

$\times$ On en déduit que la série $\sum\limits_{k\geqslant 1} \left(\left(1+\ln\left(\dfrac{p}{1-p}\right)\right)(p_k - q_k) + \ln(1-p)\, k\ p_k - \ln(1-p)\, k\ q_k\right)$ est convergente en tant que combinaison linéaire de séries convergentes. De plus :

$$\begin{aligned}
& \sum_{k=1}^{+\infty} \left(\left(1+\ln\left(\frac{p}{1-p}\right)\right)(p_k - q_k) + \ln(1-p)\, k\ p_k - \ln(1-p)\, k\ q_k\right) \\
= \ & \left(1+\ln\left(\frac{p}{1-p}\right)\right)\left(\sum_{k=1}^{+\infty} p_k - \sum_{k=1}^{+\infty} q_k\right) + \ln(1-p)\sum_{k=1}^{+\infty} k\ p_k - \ln(1-p)\sum_{k=1}^{+\infty} k\ q_k \\
= \ & \left(1+\ln\left(\frac{p}{1-p}\right)\right)(\cancel{1} - \cancel{1}) + \cancel{\ln(1-p)\frac{1}{p}} - \cancel{\ln(1-p)\frac{1}{p}} \\
= \ & 0
\end{aligned}$$

$\times$ On en conclut :

$$\sum_{k=1}^{+\infty} \big(p_k \ln(p_k) - q_k \ln(q_k)\big) \ \leqslant \ 0$$

Comme $-\dfrac{1}{\ln(2)} < 0$, on obtient :

$$\underbrace{-\frac{1}{\ln(2)} \sum_{k=1}^{+\infty} \big(p_k \ln(p_k) - q_k \ln(q_k)\big)}_{=\, H(G) - H(X)} \ \geqslant \ 0$$

$$\boxed{\text{Finalement : } H(G) \geqslant H(X)}$$

- Démontrons que $H(G) = H(X)$ si et seulement si X suit la même loi que G.

$\times$ Dans toute la démonstration qui précède. les seules majorations sont les inégalités $(\star)$:

$$\forall k \in \mathbb{N}^*, \quad \ln\left(\frac{p_k}{q_k}\right) \ \leqslant \ \frac{p_k}{q_k} - 1$$

Ainsi :

$$H(G) = H(X) \quad \Leftrightarrow \quad \forall k \in \mathbb{N}^*, \quad \ln\left(\frac{p_k}{q_k}\right) \ = \ \frac{p_k}{q_k} - 1$$

Or, d'après l'indication de l'énoncé en question précédente, pour tout $k \in \mathbb{N}^*$:

$$\ln\left(\frac{p_k}{q_k}\right) \ = \ \frac{p_k}{q_k} - 1 \quad \Leftrightarrow \quad \frac{p_k}{q_k} \ = \ 1 \quad \Leftrightarrow \quad p_k \ = \ q_k \quad \Leftrightarrow \quad \mathbb{P}([G=k]) \ = \ \mathbb{P}([X=k])$$

On en déduit : $H(G) = H(X) \quad \Leftrightarrow \quad \forall k \in \mathbb{N}^*, \mathbb{P}([G=k]) \ = \ \mathbb{P}([X=k])$.

$$\boxed{\begin{array}{c}\text{Or } X(\Omega) = \mathbb{N}^* = G(\Omega). \text{ On en conclut } H(G) = H(X) \\ \text{si et seulement si } X \text{ suit la même loi que } G.\end{array}}$$

$\square$