

I. AUTOUR DES PROJECTEURS

I.A.

Q1. Il est évident que $F \subset \mathbf{R}^3$ et on a $0_{\mathbf{R}^3} = (0, 0, 0) \in F$ car $2 \cdot 0 + 0 - 0 = 0$. Enfin, soient (x, y, z) et (x', y', z') deux vecteurs de F et $\lambda \in \mathbf{R}$. Alors

$$2(x + \lambda x') + (y + \lambda y') - (z + \lambda z') = \underbrace{(2x + y - z)}_{=0 \text{ car } (x,y,z) \in F} + \lambda \underbrace{(2x' + y' - z')}_{=0 \text{ car } (x',y',z') \in F} = 0,$$

donc $(x, y, z) + \lambda(x', y', z') \in F$.

Ainsi, F est un sous-espace vectoriel de $\mathbf{R}^3$.

Q2. Soit $X = (x, y, z) \in \mathbf{R}^3$. Alors

$$X \in F \iff 2x + y - z = 0 \iff z = 2x + y \iff X = (x, y, 2x + y) \iff X = x \cdot \underbrace{(1, 0, 2)}_{e_1} + y \cdot \underbrace{(0, 1, 1)}_{e_2}$$

Ainsi, $F = \text{Vect}(e_1, e_2)$. La famille (e_1, e_2) étant visiblement libre, c'est donc une base de F .

De même, $G = \text{Vect}(e_3)$ avec $e_3 = (1, 1, 0) \neq 0$ donc (e_3) est une base de G .

Q3. On remarque d'ores et déjà, d'après la question précédente, que $\dim(F) + \dim(G) = 2 + 1 = 3 = \dim(\mathbf{R}^3)$.

Reste à montrer que F et G sont en somme directe. Soit $X = (x, y, z) \in F \cap G$. $X \in G = \text{Vect}(e_3)$ donc, d'après la question précédente, $x = y$ et $z = 0$. Par ailleurs, $X \in F$ donc $2x + y - z = 0$. En combinant ces trois équations, on trouve $x = 0$, puis $y = z = 0$ donc $X = 0$. On a ainsi montré que $F \cap G = \{0\}$, on en déduit que F et G sont en somme directe, donc que $F \oplus G = \mathbf{R}^3$.

Q4. On va calculer l'image par p des trois vecteurs de la base canonique de $\mathbf{R}^3$ en cherchant leur décomposition sur $F \oplus G$, c'est-à-dire en les écrivant comme somme d'un vecteur vérifiant $z = 2x + y$ et d'un vecteur de la forme $(x, x, 0)$. On obtient :

$$\begin{aligned} (1, 0, 0) &= \underbrace{\left(\frac{1}{3}, -\frac{2}{3}, 0\right)}_{\in F} + \underbrace{\left(\frac{2}{3}, \frac{2}{3}, 0\right)}_{\in G} \\ (0, 1, 0) &= \underbrace{\left(-\frac{1}{3}, \frac{2}{3}, 0\right)}_{\in F} + \underbrace{\left(\frac{1}{3}, \frac{1}{3}, 0\right)}_{\in G} \\ (0, 0, 1) &= \underbrace{\left(\frac{1}{3}, \frac{1}{3}, 1\right)}_{\in F} + \underbrace{\left(-\frac{1}{3}, -\frac{1}{3}, 0\right)}_{\in G} \end{aligned}$$

Pour obtenir le projeté sur F , on ne conserve que la composante suivant F dans cette décomposition. On en déduit la matrice :

$$\text{Mat}_{\mathcal{B}\mathcal{C}}(p) = \begin{pmatrix} 1/3 & -1/3 & 1/3 \\ -2/3 & 2/3 & 1/3 \\ 0 & 0 & 1 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 1 & -1 & 1 \\ -2 & 2 & 1 \\ 0 & 0 & 3 \end{pmatrix}.$$

Q5. Les familles (e_1, e_2) et (e_3) sont des familles libres respectives de F et G , qui sont en somme directe; on en déduit que $\mathcal{B} = (e_1, e_2, e_3)$ est une famille libre de $\mathbf{R}^3$. Or elle est composée de 3 vecteurs, ce qui est égal à la dimension de $\mathbf{R}^3$. C'est donc une base.

Par ailleurs, e_1 et e_2 appartiennent à F donc leur image par une projection sur F est $p(e_1) = e_1$ et $p(e_2) = e_2$. D'autre part, e_3 appartient à G donc son image par une projection parallèlement à G est $p(e_3) = 0$. On en déduit la matrice

$$\text{Mat}_{\mathcal{B}}(p) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

I.B.

Q6. Soit $x \in E$. On écrit $x = (x - p(x)) + p(x)$. Il est clair que $p(x) \in \text{Im}(p)$. Par ailleurs, $p(x - p(x)) = p(x) - (p \circ p)(x) = p(x) - p(x) = 0$, autrement dit $x - p(x) \in \text{Ker}(p)$. On a donc écrit x comme somme d'un élément de $\text{Ker}(p)$ et un élément de $\text{Im}(p)$; ceci étant vrai pour tout $x \in E$, on a donc $E = \text{Ker}(p) + \text{Im}(p)$ (on a en fait montré uniquement l'inclusion $\subset$, mais l'inclusion réciproque est triviale).

Reste à montrer que $\text{Ker}(p)$ et $\text{Im}(p)$ sont en somme directe. Soit $x \in \text{Ker}(p) \cap \text{Im}(p)$. Comme $x \in \text{Im}(p)$, il existe $y \in E$ tel qu'on puisse écrire $x = p(y)$. Par ailleurs, $x \in \text{Ker}(p)$ donc $p(x) = 0$. En combinant les deux écritures, on trouve $p(p(y)) = 0$, soit $p(y) = 0$ car $p \circ p = p$. Autrement dit $x = 0$.

Q7. On procède par double inclusion. Soit $x \in \text{Im}(p)$. Il existe donc y tel que $x = p(y)$. Mais alors $x - p(x) = p(y) - p(p(y)) = p(y) - p(y) = 0$. Ainsi $x \in \text{Ker}(\text{id}_E - p)$

Réciproquement, soit $x \in \text{Ker}(\text{id}_E - p)$, autrement dit $x - p(x) = 0$, autrement dit $x = p(x)$ et il appartient donc à $\text{Im}(p)$ (puisqu'il peut s'écrire comme image de lui-même). D'où l'égalité.

Q8. On s'inspire de la question **Q5**. Soient $\mathcal{B}_1$ et $\mathcal{B}_2$, respectivement, des bases de $\text{Im}(p)$ et $\text{Ker}(p)$. Comme $E = \text{Ker}(p) \oplus \text{Im}(p)$, on en déduit que la famille $\mathcal{B}$ obtenue en concaténant $\mathcal{B}_1$ et $\mathcal{B}_2$ est une base de E .

J'affirme que la matrice de p dans cette base est de la forme demandée. En effet, les colonnes de gauche correspondent aux vecteurs de la base $\mathcal{B}_1$; ce sont des vecteurs de $\text{Im}(p) = \text{Ker}(\text{id}_E - p)$, autrement dit ils vérifient $p(x) = x$. On observe donc des premières colonnes identiques à celles de la matrice identité. Au contraire, les colonnes de droite correspondent aux vecteurs de la base $\mathcal{B}_2$; ce sont des vecteurs de $\text{Ker}(p)$, ils vérifient donc $p(x) = 0$. Les colonnes correspondantes sont donc intégralement nulles.

La trace de cette matrice est donc égale au nombre de 1 présents sur la diagonale, c'est-à-dire au nombre de vecteurs de $\mathcal{B}_1$, autrement dit à la dimension de $\text{Im}(p)$, c'est-à-dire au rang de p . Comme la trace d'une application linéaire est égale à la trace de sa matrice dans n'importe quelle base, on a donc montré que $\text{Tr}(p) = \text{rg}(p)$.

Q9. On raisonne par récurrence sur $k \in \mathbb{N}^*$.

Si $k = 1$, il est trivialement vrai que $\dim(E_1) \leq \dim(E_1)$.

Soit maintenant $k \in \mathbb{N}^*$. On suppose que, si $E_1, \dots, E_k$ sont k sous-espaces de E , alors $\dim(E_1 + \dots + E_k) \leq \dim(E_1) + \dots + \dim(E_k)$. Soient maintenant $E_1, \dots, E_{k+1}$ $k+1$ sous-espaces de E . On a $E_1 + \dots + E_{k+1} = (E_1 + \dots + E_k) + E_{k+1}$ donc, d'après la formule de Grassmann,

$$\begin{aligned} \dim(E_1 + \dots + E_{k+1}) &= \dim(E_1 + \dots + E_k) + \dim(E_{k+1}) - \dim((E_1 + \dots + E_k) \cap E_{k+1}) \\ &\leq \dim(E_1 + \dots + E_k) + \dim(E_{k+1}) \\ &\leq \dim(E_1) + \dots + \dim(E_k) + \dim(E_{k+1}) \quad \text{par HR,} \end{aligned}$$

ce qui démontre bien le résultat souhaité.

I.C.

Q10. Il s'agit de voir si $q_k \circ q_k = q_k$. On calcule donc $q_k \circ q_k$ et on développe par linéarité :

$$q_k \circ q_k = (p_1 + \dots + p_k) \circ (p_1 + \dots + p_k) = \sum_{i=1}^n \sum_{j=1}^n p_i \circ p_j$$

Or $p_i \circ p_j = 0$ si $i \neq j$, par hypothèse, tandis que $p_i \circ p_i = p_i$ car p_i est un projecteur. Après simplification, il ne reste donc que $q_k \circ q_k = \sum_{i=1}^n p_i$, autrement dit $q_k \circ q_k = q_k$.

Q11. (a) Soit $x \in \text{Im}(q_k)$. Par définition, il existe $y \in E$ tel que $x = q_k(y)$, autrement dit $x = p_1(y) + \dots + p_k(y)$. Mais $p_1(y) \in \text{Im}(p_1), \dots, p_k(y) \in \text{Im}(p_k)$, on a donc bien $x \in \text{Im}(p_1) + \dots + \text{Im}(p_k)$.

(b) D'après la partie précédente, q_k , et $p_1, \dots, p_k$ étant des projecteurs, leur trace est égale à leur rang. On a donc, par linéarité de la trace,

$$\begin{aligned} \text{rg}(q_k) &= \text{Tr}(q_k) = \text{Tr}(p_1 + \dots + p_k) = \text{Tr}(p_1) + \dots + \text{Tr}(p_k) = \text{rg}(p_1) + \dots + \text{rg}(p_k) \\ &= \dim(\text{Im}(p_1)) + \dots + \dim(\text{Im}(p_k)) \geq \dim(\text{Im}(p_1) + \dots + \text{Im}(p_k)) \quad \text{d'après Q9} \end{aligned}$$

Par ailleurs, d'après la question précédente, $\text{Im}(q_k) \subset \text{Im}(p_1) + \dots + \text{Im}(p_k)$ donc

$$\text{rg}(q_k) = \dim(\text{Im}(q_k)) \leq \dim(\text{Im}(p_1) + \dots + \text{Im}(p_k)).$$

On a donc l'égalité des dimensions par double inégalité. Les sous-espaces $\text{Im}(q_k)$ et $\text{Im}(p_1) + \dots + \text{Im}(p_k)$ sont ainsi inclus l'un dans l'autre et de même dimension, ils sont donc égaux.

- (c) Supposons qu'il existe $x_1 \in \text{Im}(p_1), \dots, x_k \in \text{Im}(p_k)$ tels que $x_1 + \dots + x_k = 0$. Raisonnons par l'absurde et supposons que l'un des x_i n'est pas nul, disons x_1 pour faciliter les notations. On peut alors écrire $x_1 = -x_2 - \dots - x_k$, autrement dit $x_1 \in \text{Im}(p_2) + \dots + \text{Im}(p_k)$. Mais simultanément $x_1 \in \text{Im}(p_1)$ par définition. On en déduit que $\text{Im}(p_1) \cap (\text{Im}(p_2) + \dots + \text{Im}(p_k)) \neq \{0\}$; en particulier, cette intersection est de dimension strictement positive.

Mais alors, d'après la formule de Grassmann,

$$\begin{aligned} \dim(\text{Im}(p_1) + \dots + \text{Im}(p_k)) &= \dim(\text{Im}(p_1)) + \dim(\text{Im}(p_2) + \dots + \text{Im}(p_k)) \\ &\quad - \dim(\text{Im}(p_1) \cap (\text{Im}(p_2) + \dots + \text{Im}(p_k))) \\ &< \dim(\text{Im}(p_1)) + \dim(\text{Im}(p_2) + \dots + \text{Im}(p_k)) \\ &< \dim(\text{Im}(p_1)) + \dim(\text{Im}(p_2)) + \dots + \dim(\text{Im}(p_k)) \text{ d'après Q9.} \end{aligned}$$

Ceci contredit la question précédente, d'après laquelle ces deux nombres devraient être égaux. On en déduit que tous les x_i sont nuls et que les $\text{Im}(p_i)$ sont en somme directe.

- Q12.** (a) Soit $j \in \llbracket 1; k \rrbracket$. Soit $x \in E$. Alors $p_j(x) \in \text{Im}(p_j) \subset \text{Im}(p_1) + \dots + \text{Im}(p_k) = \text{Im}(q_k) = \text{Ker}(\text{id}_E - q_k)$ d'après les questions précédentes. Autrement dit, $q_k(p_j(x)) = p_j(x)$.
 (b) Soient $j \in \llbracket 1 : k \rrbracket$ et $x \in E$. D'après la question précédente, $q_k(p_j(x)) = p_j(x)$, autrement dit

$$\sum_{i=1}^n p_i(p_j(x)) = p_j(x).$$

On sépare le terme $i = j$ de la somme et on le passe du côté droit :

$$\sum_{1 \leq i \leq n, i \neq j} p_i(p_j(x)) = p_j(x) - p_j(p_j(x)) = 0 \text{ car } p_j \text{ est un projecteur.}$$

- (c) Soit $x \in E$. Dans l'égalité ci-dessus, chacun des termes $p_i(p_j(x))$ est un élément de $\text{Im}(p_i)$. L'écriture $\sum_{1 \leq i \leq n, i \neq j} p_i(p_j(x)) = 0$ est donc une décomposition du vecteur nul sur $\text{Im}(p_1) \oplus \dots \oplus \text{Im}(p_k)$. Par unicité de cette décomposition (la somme étant directe), on en déduit directement que $p_i(p_j(x)) = 0$ pour tout $i \neq j$.

- Q13.** D'après les questions **Q10** (sens direct) et **Q11** et **Q12** (sens réciproque), on en déduit que q_k est un projecteur si et seulement si $p_i \circ p_j = 0$ pour tous $i, j \in \llbracket 1; n \rrbracket$ distincts.

II. ENTROPIE ET « EFFET DE SURPRISE » EN PROBABILITÉS

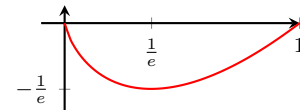
II.A.

- Q1.** Pour tout $x > 0$, on pose $\varphi(x) = x - 1 - \ln(x)$. C'est une fonction dérivable par composition et, pour tout $x > 0$, $\varphi'(x) = 1 - \frac{1}{x} = \frac{x-1}{x}$. On en déduit que la fonction φ est strictement décroissante sur $]0; 1]$ et strictement croissante sur $[1; +\infty[$; elle atteint donc un unique minimum en $x = 1$, de valeur $\varphi(1) = 0$. On a donc $\varphi(x) \geq 0$ pour tout $x > 0$, avec égalité ssi $x = 1$, ce qui est exactement la question posée.

Graphiquement, cela signifie que la courbe représentative de la fonction logarithme est toujours en-dessous de sa tangente au point d'abscisse 1.

- Q2.** La fonction g est continue et dérivable sur $]0; 1]$ par composition. De plus, $x \ln(x)$ tend vers 0 lorsque $x \rightarrow 0$ (croissance comparée), on en déduit que le prolongement $g(0) = 0$ définit bien une fonction continue sur $[0; 1]$.

Pour représenter graphiquement g , on peut constater que la fonction s'annule en 0 et en 1; sa dérivée vaut $\ln(x) + 1$, g est donc décroissante puis croissante avec un minimum en $\frac{1}{e} \simeq 0,37$.



II.B.

- Q3.** L'événement certain vérifie $\mathbf{P}(\Omega) = 1$. On a donc $S(\Omega) = 0$, d'après la propriété (i). En termes d'effet de surprise, un événement qui se produit de manière certaine ne crée aucune surprise.
Q4. Si A et B sont indépendants, alors $\mathbf{P}(A \cap B) = \mathbf{P}(A)\mathbf{P}(B)$ donc $S(A \cap B) = S(A) + S(B)$. La surprise de voir simultanément les événements A et B est l'addition des surprises des événements séparés.

Q5. La fonction $f = -\ln$ convient.

Q6. (a) Soit $p \in]0; 1]$. En faisant le changement de variables $t = pu$ ($dt = p du$), on a donc

$$\int_{p/2}^p f(t) dt = \int_{1/2}^1 f(pu) p du = p \int_{1/2}^1 (f(p) + f(u)) du = \frac{p}{2} f(p) + p \int_{1/2}^1 f(u) du.$$

(b) Notons F une primitive de f sur $]0; 1]$. D'après la question précédente, pour tout $p \in]0; 1]$,

$$f(p) = \frac{2}{p} (F(p) - F(p/2)) - 2 \int_{1/2}^1 f(u) du.$$

La fonction F est dérivable (par définition), ainsi que la fonction constante égale à $\int_{1/2}^1 f(u) du$. On en déduit que f est dérivable par théorème de composition.

(c) Soit $p \in]0; 1]$. Dérivons (iii) par rapport à q (on peut le faire car f est dérivable, d'après la question précédente) : pour tout $q \in]0; 1]$, $pf'(pq) = f'(q)$. En particulier, pour $q = 1$: $f'(p) = \frac{f'(1)}{p}$, ce qui répond à la question avec $a = f'(1)$.

(d) On déduit de l'égalité $f'(p) = \frac{a}{p}$ que f est une primitive de la fonction $p \mapsto \frac{a}{p}$ sur $]0; 1]$. Autrement dit, il existe $C \in \mathbf{R}$ tel que, pour tout $p \in]0; 1]$, $f(p) = a \ln(p) + C$. D'après la condition (i), on a nécessairement $C = 0$. Enfin, la propriété de décroissance de f impose d'avoir $a \leq 0$. Réciproquement, on constate que toutes les fonctions $p \mapsto a \ln(p)$ avec $a \leq 0$ vérifient les quatre conditions. L'ensemble des solutions est donc $\mathcal{S} = \{p \mapsto a \ln(p) \mid a \leq 0\}$.

(e) D'après la question précédente, $f(1/e) = 1 \iff a = -1$. Il existe donc une unique fonction $h = -\ln$ répondant à cette condition supplémentaire.

(f) On a $h(p) \rightarrow +\infty$ lorsque p tend vers 0 ; un événement extrêmement improbable provoque une surprise extrêmement élevée (autrement dit, contient énormément d'information).

*

Q7. Pour cette expérience, l'univers est l'ensemble des couples $(i, j) \in \llbracket 1; 6 \rrbracket^2$, qui sont tous équiprobables. En énumérant tous les couples correspondant à chaque événement, on trouve $\mathbf{P}(E) = \frac{1}{2}$, $\mathbf{P}(M) = \frac{7}{36}$ et $\mathbf{P}(N) = \frac{1}{6}$. On a donc $\mathbf{P}(N) < \mathbf{P}(M) < \mathbf{P}(E)$, d'où (par la propriété (ii)) $S(E) < S(M) < S(N)$. L'événement le moins surprenant est E , l'événement le plus surprenant est N .

II.C.

Q8. Par définition, $H(X) = -\sum_{k=0}^n \mathbf{P}(X = k) \ln(\mathbf{P}(X = k)) = \sum_{k=0}^n \mathbf{P}(X = k) S(X = k)$. On peut considérer $H(X)$ comme l'espérance de la variable aléatoire dont les valeurs sont les $-\ln(p_k)$, prises chacune avec probabilité p_k . On peut aussi l'interpréter comme la moyenne pondérée de l'information contenue dans chacun des événements $(X = k)$.

Q9. Pour tout $k \in \llbracket 0; n \rrbracket$, $p_k \in [0; 1]$ donc $p_k \ln(p_k) \leq 0$. On en déduit que $H(X) \geq 0$. Pour avoir égalité dans cette inégalité, il faut avoir $p_k \ln(p_k) = 0$ pour tout $k \in \llbracket 0; n \rrbracket$, autrement dit tous les p_k doivent valoir 0 ou 1. Par définition d'une variable aléatoire, on doit évidemment avoir $p_0 + \dots + p_n = 1$. On en déduit que tous les p_k sont nuls, sauf un qui vaut 1 ; ce qui veut dire que X est une variable aléatoire certaine.

Q10. (a) Dans ce cas, on a $\mathbf{P}(X_0 = k) = \frac{1}{n+1}$ pour tout k . On a donc

$$H(X_0) = -\sum_{k=0}^n \frac{1}{n+1} \ln\left(\frac{1}{n+1}\right) = \ln(n+1).$$

(b) Soit $k \in \llbracket 0; n \rrbracket$. Soit $x = \frac{1}{p_k(n+1)} > 0$. D'après **Q1**,

$$-\ln(p_k) + \ln\left(\frac{1}{n+1}\right) \leq \frac{1}{p_k(n+1)} - 1,$$

et on obtient le résultat demandé en multipliant des deux côtés de l'inégalité par $p_k > 0$.

(c) On remarque que $p_0 + \dots + p_k = 1$. On peut alors écrire, d'après la question précédente :

$$\begin{aligned} H(X) - H(X_0) &= \sum_{k=0}^n -p_k \ln(p_k) - \ln(n+1) = \sum_{k=0}^n -p_k \ln(p_k) + \sum_{k=0}^n p_k \ln\left(\frac{1}{n+1}\right) \\ &\leq \sum_{k=0}^n \left(\frac{1}{n+1} - p_k\right) = \frac{n+1}{n+1} - 1 = 0. \end{aligned}$$

Le cas d'égalité correspond au cas d'égalité dans l'inégalité de la question **Q1** : autrement dit si $p_k = \frac{1}{n+1}$ pour tout k , c'est-à-dire si X suit une loi uniforme.

Ce résultat signifie qu'une variable aléatoire finie donnée contient moins d'information en moyenne qu'une variable aléatoire uniforme de même univers-image.

II.D.

Q11. Les deux sommes sont égales à 1. On peut donc écrire :

$$\begin{aligned} K(X, Y, X', Y') &= - \sum_{i=0}^n \sum_{j=0}^m \lambda_{ij} \ln\left(\frac{\lambda'_{ij}}{\lambda_{ij}}\right) + 1 - 1 = - \sum_{i=0}^n \sum_{j=0}^m \lambda_{ij} \ln\left(\frac{\lambda'_{ij}}{\lambda_{ij}}\right) + \sum_{i=0}^n \sum_{j=0}^m \lambda'_{ij} - \sum_{i=0}^n \sum_{j=0}^m \lambda_{ij} \\ &= - \sum_{i=0}^n \sum_{j=0}^m \left(\lambda_{ij} \ln\left(\frac{\lambda'_{ij}}{\lambda_{ij}}\right) - \lambda'_{ij} + \lambda_{ij}\right) = - \sum_{i=0}^n \sum_{j=0}^m \lambda_{ij} \left(\ln\left(\frac{\lambda'_{ij}}{\lambda_{ij}}\right) - \frac{\lambda'_{ij}}{\lambda_{ij}} + 1\right). \end{aligned}$$

Q12. Pour tous $(i, j) \in \llbracket 0; n \rrbracket \times \llbracket 0; m \rrbracket$, on peut appliquer l'inégalité de la question **Q1** au nombre $x = \frac{\lambda'_{ij}}{\lambda_{ij}} > 0$; on en déduit que tous les termes de la somme ci-dessus sont négatifs, donc que $K(X, Y, X', Y') \geq 0$. Il y a égalité si et seulement si chaque terme de la somme est nul, c'est-à-dire (compte tenu du cas d'égalité dans **Q1**) dit si $\lambda_{ij} = \lambda'_{ij}$ pour tout couple (i, j) ; autrement dit si (X, Y) et (X', Y') ont la même loi conjointe.

Q13. Notons, pour tous $(i, j) \in \llbracket 0; n \rrbracket \times \llbracket 0; m \rrbracket$, $p_i = \mathbf{P}(X = i) = \mathbf{P}(X' = i)$ et $q_j = \mathbf{P}(Y = j) = \mathbf{P}(Y' = j)$. Les variables X' et Y' étant supposées indépendantes, on a donc $\lambda'_{ij} = p_i q_j$ pour tout couple (i, j) . De plus, $p_i = \sum_{j=0}^m \lambda_{ij}$ et $q_j = \sum_{i=0}^n \lambda_{ij}$. On a alors :

$$\begin{aligned} K(X, Y, X', Y') &= - \sum_{i=0}^n \sum_{j=0}^m \lambda_{ij} (\ln(p_i) + \ln(q_j) - \ln(\lambda_{ij})) \\ &= - \sum_{i=0}^n \sum_{j=0}^m \lambda_{ij} \ln(p_i) - \sum_{j=0}^m \sum_{i=0}^n \lambda_{ij} \ln(q_j) - H(X, Y) \\ &= - \sum_{i=0}^n p_i \ln(p_i) - \sum_{j=0}^m q_j \ln(q_j) - H(X, Y) = H(X) + H(Y) - H(X, Y). \end{aligned}$$

Chacun des nombres λ_{ij} étant dans $]0; 1]$, on a $\lambda_{ij} \ln(\lambda_{ij}) \leq 0$ donc $H(X, Y) \geq 0$; on en déduit que $K(X, Y, X', Y') \leq H(X) + H(Y)$. Il y a égalité si et seulement si tous les $\lambda_{ij} \ln(\lambda_{ij})$ sont nuls, autrement dit si tous les λ_{ij} valent 0 ou 1. De la même manière qu'à la **Q9**, il y a donc égalité si et seulement si le couple (X, Y) suit une loi certaine.