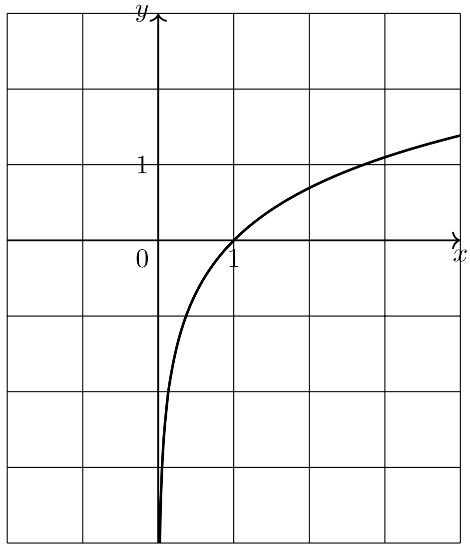
CCS 2017 TSI MATHS 1

Correction

I Questions préliminaires

A -



B - On considère la fonction $k : x \mapsto x - 1 - \ln(x)$ définie sur $\mathbb{R}^{+*}$.

La fonction k est dérivable sur $\mathbb{R}^{+*}$ et $k'(x) = 1 - \dfrac{1}{x}$.

On a donc le tableau de variation suivant :

x	0		1		$+\infty$
$k(x)$	|| $+\infty$	$\searrow$	0	$\nearrow$	$+\infty$

Ainsi, pour tout $x > 0$, $k(x) \geqslant 0$ avec égalité si, et seulement si, $x = 1$.

On a donc montré que $\ln(x) \leqslant x - 1$ avec égalité si, et seulement si, $x = 1$.

C - Graphiquement, la question précédente signifie que la courbe représentative de la fonction ln est en dessous de la droite d'équation $y = x - 1$ et que ces deux courbes n'ont qu'un seul point commun, le point d'abscisse $x = 1$.

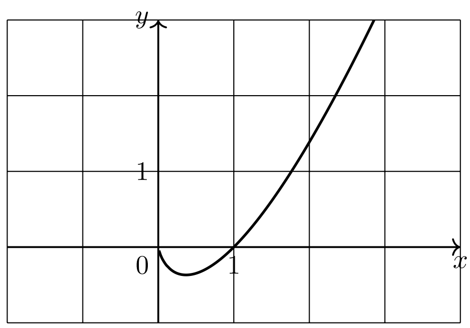
D - Par produit de fonctions usuelles continues et dérivables, la fonction g est continue et dérivable sur $]0; 1]$.

De plus on sait que $\lim\limits_{x \to 0^+} x\ln(x) = 0$ donc $\lim\limits_{x \to 0^+} g(x) = g(0)$ ce qui signifie que g est continue en 0.

Pour tout $x > 0$, $g'(x) = \ln(x) + 1$ donc on a le tableau de variation suivant :

x	0		$1/\mathrm{e}$		$+\infty$
$g(x)$	0	$\searrow$	$-1/\mathrm{e}$	$\nearrow$	$+\infty$

Et la courbe suivante :



II Mathématisation de l'effet de surprise

A - L'événement certain est l'ensemble Ω. Il vérifie $P(\Omega) = 1$.
On a donc $S(\Omega) = f(P(\Omega)) = f(1) = 0$.
L'événement certain a donc une quantité d'information nulle, ce qui est cohérent car l'événement certain ne provoque pas d'effet de surprise.

B - Lorsque A et B sont indépendants, $P(A \cap B) = P(A) \times P(B)$. D'après les propriétés de la fonction f on a donc :

$$S(A \cap B) = f(P(A \cap B)) = f(P(A)P(B)) = f(P(A)) + f(P(B)) = S(A) + S(B).$$

Les effets de surprise de deux événements indépendants s'ajoutent lorsqu'on prend en compte les deux événements en même temps.

C - La fonction $f : x \mapsto -\ln(x)$ vérifie bien toutes les conditions demandées.

D - 1) On considère le changement de variable $t = pu$. On a donc $\mathrm{d}t = p\,\mathrm{d}u$ et u varie entre $\frac{1}{2}$ et 1. Ainsi :

$$\frac{1}{p}\int_{p/2}^{p} f(t)\,\mathrm{d}t = \frac{1}{p}\int_{1/2}^{1} f(pu)\;p\,\mathrm{d}u = \int_{1/2}^{1} (f(p) + f(u))\,\mathrm{d}u = \frac{1}{2}f(p) + \int_{1/2}^{1} f(u)\,\mathrm{d}u.$$

2) D'après la question précédente, pour tout $p \in]0;1]$:

$$f(p) = \frac{2}{p}\int_{p/2}^{p} f(t)\,\mathrm{d}t - 2\int_{1/2}^{1} f(u)\,\mathrm{d}u.$$

La fonction $p \mapsto -2\int_{1/2}^{1} f(u)\,\mathrm{d}u$ est constante donc dérivable sur $]0;1]$.

La fonction $p \mapsto \frac{2}{p}$ est dérivable sur $]0;1]$ (fonction usuelle).

Il nous reste à montrer que la fonction $p \mapsto \int_{p/2}^{p} f(t)\,\mathrm{d}t$ est dérivable sur $]0;1]$.

Notons, pour $x \in]0;1]$, $F(x) = \int_{1}^{x} f(t)\,\mathrm{d}t$. Comme la fonction f est continue sur $]0;1]$, on sait que la fonction F est de classe $\mathscr{C}^1$ sur $]0;1]$.

En remarquant que $\int_{p/2}^{p} f(t)\,\mathrm{d}t = F(p) - F(p/2)$, on en déduit que la fonction $p \mapsto \int_{p/2}^{p} f(t)\,\mathrm{d}t$ est de classe $\mathscr{C}^1$ donc dérivable sur $]0;1]$.

En conclusion, la fonction f est dérivable sur $]0;1]$.

3) En dérivant par rapport à q l'égalité iii, on obtient $pf'(pq) = f'(q)$, pour tous p et q appartenant à $]0;1]$.
Ainsi, pour $q = 1$, on obtient, pour tout $p \in]0;1]$, $f'(p) = \frac{f'(1)}{p}$.
On a bien $f'(p) = \frac{a}{p}$ avec $a = f'(1)$.

4) On suppose maintenant que $f'(p) = \frac{a}{p}$. On a donc $f(p) = a\ln(p) + b$ avec b une constante réelle.
Pour que la fonction f vérifie le point i il faut : $f(1) = 0 \Leftrightarrow b = 0$.
Pour que la fonction f vérifie le point ii il faut : $f'(p) \leqslant 0$, donc $a \leqslant 0$.
Pour que la fonction f vérifie le point iii il faut : $f(pq) = f(p) + f(q) \Leftrightarrow a\ln(pq) + b = a\ln(p) + b + a\ln(q) + b \Leftrightarrow b = 0$.
Quelles que soient les valeurs de a et b la fonction f vérifie le point iv.
En conclusion, les fonctions f qui vérifient $f'(p) = \frac{a}{p}$ et les quatre points de l'énoncé sont les fonctions f de la forme $f : p \mapsto a\ln(p)$ avec $a \leqslant 0$.

5) Parmi les fonctions de la question précédente, $f(1/\mathrm{e}) = 1$ équivaut à $-a = 1$ et donc $a = -1$.
La seule fonction vérifiant les conditions de la question suivante et la contrainte $f(1/\mathrm{e}) = 1$ est la fonction $h : p \mapsto -\ln(p)$.
On a alors $\lim\limits_{p \to 0^+} h(p) = +\infty$. On peut interpréter ce résultat en disant que lorsque la probabilité d'un événement devient très petite, l'effet de surprise devient grand.

6) Les deux lancers sont indépendants.

On a $P(E) = \dfrac{3}{6} = \dfrac{1}{2}$, donc $S(E) = h(1/2) = -\ln(1/2) = \ln(2)$.

L'événement M signifie que l'un des deux numéros obtenu est le chiffre 4 et l'autre numéro est inférieur à 4. On peut donc avoir $\{1;4\}$, ou $\{2;4\}$, ou $\{3;4\}$, ou $\{4;4\}$, ou $\{4;3\}$, ou $\{4;2\}$ ou encore $\{4;1\}$. Les résultats étant équiprobables on a $P(M) = \dfrac{7}{36}$.

Donc $S(M) = h(7/36) = \ln(36/7) = \ln(5 + \dfrac{1}{7})$.

L'événement N signifie que l'on a obtenu : $\{1;6\}$, ou $\{2;5\}$, ou $\{3;4\}$, ou $\{4;3\}$, ou $\{5;2\}$ ou encore $\{6;1\}$. Les résultats étant équiprobables on a $P(N) = \dfrac{6}{36} = \dfrac{1}{6}$.

Donc $S(N) = h(1/6) = \ln(6)$.

Ainsi $S(E) < S(M) < S(N)$.

L'événement N est donc plus surprenant que l'événement M, lui-même plus surprenant que l'événement E.

III Entropie d'une variable aléatoire

A - 1) Par définition, $H(X) = \displaystyle\sum_{k=0}^{n} -P(X=k)\ln(P(X=k)) = \sum_{k=0}^{n} S([X=k])P(X=k)$.

On peut donc voir $H(X)$ comme l'espérance d'une variable aléatoire Z qui prendrait ses valeurs dans $\{-\ln(p_k)\}$ et telle que $P(Z = -\ln(p_k)) = P(X=k)$. Cela correspondrait à faire la moyenne pondérée des quantités d'information contenues dans les événements $[X=k]$.

2) Pour tout $k \in [\![0;n]\!]$, $p_k \in [0;1]$ donc soit $p_k \ln(p_k) = 0$, soit $\ln(p_k) \leqslant 0$.

Dans tous les cas, $-p_k \ln(p_k) \geqslant 0$, donc $H(X) \geqslant 0$.

$H(X)$ étant une somme de nombres positifs, $H(X) = 0$ si, et seulement si, tous les termes de la somme sont nuls.

Ainsi $H(X) = 0 \Leftrightarrow \forall k \in [\![0;n]\!], p_k\ln(p_k) = 0 \Leftrightarrow \forall k \in [\![0;n]\!], p_k = 0$ ou $p_k = 1$.

Comme $\displaystyle\sum_{k=0}^{n} p_k = 1$, on a donc nécessairement un seul des p_k qui vaut 1 et tous les autres sont nuls.

Ainsi $H(X) = 0$ si, et seulement si, X est une variable aléatoire certaine.

3) a) On a donc $P(X_0 = k) = \dfrac{1}{n+1}$, pour tout $k \in [\![0;n]\!]$.

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

b) En appliquant l'inégalité de la question I.B pour $x = \dfrac{1}{(n+1)p_k}$, on obtient :

$$\ln\left(\frac{1}{(n+1)p_k}\right) \leqslant \frac{1}{(n+1)p_k} - 1 \Leftrightarrow p_k\left(\ln\left(\frac{1}{n+1}\right) - \ln(p_k)\right) \leqslant \frac{1}{n+1} - p_k$$
$$\Leftrightarrow -p_k\ln(p_k) + p_k\ln\left(\frac{1}{n+1}\right) \leqslant \frac{1}{n+1} - p_k.$$

c) On somme l'inégalité précédente pour k allant de 0 à n. On obtient :

$$H(X) - \ln(n+1)\sum_{k=0}^{n} p_k \leqslant 1 - \sum_{k=0}^{n} p_k \Leftrightarrow H(X) - H(X_0) \leqslant 0 \Leftrightarrow H(X) \leqslant H(X_0).$$

D'après la question I.B, il y aura égalité dans notre inégalité lorsque, pour tout $k \in [\![0;n]\!]$, $\dfrac{1}{(n+1)p_k} = 1$, c'est-à-dire lorsque $p_k = \dfrac{1}{n+1}$.

Ainsi le cas d'égalité est possible si, et seulement si, X suit la même loi que X_0.

On peut interpréter ce résultat en disant que pour toute variable aléatoire finie, la quantité d'information moyenne des événements $[X=k]$ est majorée par $\ln(n+1)$, avec égalité uniquement pour la loi uniforme sur $[\![0;n]\!]$.

B - 1) On sait que, pour tout $k \in \mathbb{N}^*$, $P(X_1 = k) = (1-p)^{k-1} \times p$. De plus $E(X) = \dfrac{1}{p}$.

On a donc ici $p_k\ln(p_k) = (1-p)^{k-1} \times p\ln((1-p)^{k-1} \times p) = p\ln(1-p)(k-1)(1-p)^{k-1} + p\ln(p)(1-p)^{k-1}$.

La série $\sum x^k$ est absolument convergente pour $x \in]-1;1[$, et par théorème de dérivation termes à termes la série $\sum kx^{k-1}$ est absolument convergente pour $x \in]-1;1[$.
Avec $x = 1-p$, on obtient donc que la série $\sum p_k \ln(p_k)$ est absolument convergente, donc X_1 est d'entropie finie.
Et sachant que $\displaystyle\sum_{k=0}^{+\infty} x^k = \frac{1}{1-x}$ et $\displaystyle\sum_{k=1}^{+\infty} kx^{k-1} = \frac{1}{(1-x)^2}$ on a :

$$\begin{aligned}
H(X_1) &= -p\ln(1-p)\left(\sum_{k=1}^{n} k(1-p)^{k-1} - \sum_{k=1}^{+\infty}(1-p)^{k-1}\right) - p\ln(p)\sum_{k=1}^{+\infty}(1-p)^{k-1}\\
&= -p\ln(1-p)\left(\frac{1}{p^2} - \frac{1}{p}\right) - p\ln(p)\frac{1}{p}\\
&= -\frac{1-p}{p}\ln(1-p) - \ln(p).
\end{aligned}$$

2) a) Comme on sait que la série $\sum p_k$ doit être convergente et valoir 1, on a $\displaystyle\lim_{k\to+\infty} p_k = 0$.

b) On sait que, au voisinage de 0, $\ln(x) = o(x^\alpha)$ quel que soit $\alpha < 0$. Donc, comme $\displaystyle\lim_{k\to+\infty} p_k = 0$, on a bien $\displaystyle\lim_{k\to+\infty} \sqrt{p_k}\ln(p_k) = 0$.
Par définition d'une limite, il existe $k_0 \in \mathbb{N}^*$, tel que pour tout $k \geqslant k_0$, $|\sqrt{p_k}\ln(p_k) - 0| \leqslant 1$.
Or $|\sqrt{p_k}\ln(p_k)| = -\sqrt{p_k}\ln(p_k)$ donc on obtient bien $0 \leqslant -\sqrt{p_k}\ln(p_k) \leqslant 1$, pour $k \geqslant k_0$.

c) D'après la question précédente $0 \leqslant -p_k\ln(p_k) \leqslant \sqrt{p_k}$.
Donc, si $p_k \leqslant \dfrac{1}{k^3}$, on a bien $0 \leqslant -p_k\ln(p_k) \leqslant \dfrac{1}{k^{3/2}}$.
Si $p_k \geqslant \dfrac{1}{k^3}$ alors $\ln(p_k) \geqslant -3\ln(k)$ et donc, en multipliant par $-p_k$ qui est négatif, $0 \leqslant -p_k\ln(p_k) \leqslant 3p_k\ln(k)$.

d) D'après la question I.B, on a $\ln(k) \leqslant k-1 \leqslant k$.
La série $\sum \dfrac{1}{k^{3/2}}$ est une série de Riemann convergente car $\dfrac{3}{2} > 1$.
De plus, $0 \leqslant 3p_k\ln(k) \leqslant 3kp_k$. Or, on sait que la série $\sum kp_k$ est convergente (car X admet un espérance), donc d'après les critères de comparaison sur les séries à termes positifs la série $\sum 3p_k\ln(k)$ est convergente.
Ainsi, la série $\displaystyle\sum_{k\geqslant 1}\left(\frac{1}{k^{3/2}} + 3p_k\ln(k)\right)$ est convergente.

e) D'après la question c), quelle que soit la valeur de k, on a donc $0 \leqslant -p_k\ln(p_k) \leqslant \dfrac{1}{k^{3/2}} + 3p_k\ln(k)$.
D'après les critères de comparaison sur les séries à termes positifs et d'après la question précédente, la série $\displaystyle\sum_{k\geqslant 1} p_k\ln(p_k)$ est donc absolument convergente.
Ainsi X est d'entropie finie.

3) a) La série $\displaystyle\sum_{k\geqslant 1} kp_k$ est convergente car l'espérance de X existe. La série $\displaystyle\sum_{k\geqslant 1} p_k$ est convergente par définition d'une variable aléatoire.
Donc la série $\displaystyle\sum_{k\geqslant 1}(k-1)p_k$ est convergente et :

$$\sum_{k=1}^{+\infty}(k-1)p_k = \sum_{k=1}^{+\infty} kp_k - \sum_{k=1}^{+\infty} p_k = E(X) - 1.$$

b) On sait que $q_k = p(1-p)^{k-1}$, donc $p_k\ln(q_k) = p_k\ln(p) + (k-1)p_k\ln(1-p)$.
Ainsi, d'après la question précédente et la convergence de la série $\displaystyle\sum_{k\geqslant 1} p_k$, la série $\displaystyle\sum_{k\geqslant 1} p_k\ln(q_k)$ est convergente et :

$$\sum_{k=1}^{+\infty} p_k\ln(q_k) = \sum_{k=1}^{+\infty}\left(p_k\ln(p) + (k-1)p_k\ln(1-p)\right) = \ln(p) + \ln(1-p)(E(X)-1).$$

c) On sait que $-H(X_1) = \dfrac{1-p}{p}\ln(1-p) + \ln(p)$. De plus, on a l'hypothèse $E(X) \leqslant \dfrac{1}{p}$ donc $E(X) - 1 \leqslant \dfrac{1-p}{p}$.
En multipliant par $\ln(1-p)$ qui est négatif, on obtient : $\ln(1-p)(E(X)-1) \geqslant \dfrac{1-p}{p}\ln(1-p)$.

Puis en ajoutant $\ln(p)$ de chaque côté, on a bien $\displaystyle\sum_{k=1}^{+\infty} p_k \ln(q_k) \geqslant -H(X_1)$.

d) On sait que $H(X) = \displaystyle\sum_{k=1}^{+\infty} - p_k \ln(p_k)$. On ajoute donc de chaque côté de l'inégalité précédente cette valeur et on obtient :

$$H(X) - H(X_1) \leqslant \sum_{k=1}^{+\infty} (-p_k \ln(p_k) + p_k \ln(q_k)) \Rightarrow H(X) - H(X_1) \leqslant \sum_{k=1}^{+\infty} p_k \ln\left(\frac{q_k}{p_k}\right).$$

D'après la question I.B, on a $\ln\left(\dfrac{q_k}{p_k}\right) \leqslant \dfrac{q_k}{p_k} - 1$.

Donc $H(X) - H(X_1) \leqslant \displaystyle\sum_{k=1}^{+\infty} (q_k - p_k) = 1 - 1 = 0$.

On a donc bien $H(X) \leqslant H(X_1)$.

Dans cette partie la quantité d'information moyenne des événements $[X = k]$, pour une variable alétoire discrète, est majorée par la quantité moyenne d'information d'une variable suivant une loi géométrique.

IV Quatrième partie

A - *Propriétés de l'information entre deux couples*

1) On sait que $\displaystyle\sum_{i=0}^{n}\sum_{j=0}^{m} \lambda_{ij} = \sum_{i=0}^{n}\sum_{j=0}^{m} \lambda'_{ij} = 1$.

On a donc :

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

2) D'après la question I.B appliquée pour $x = \dfrac{\lambda'_{ij}}{\lambda_{ij}}$, on a $\ln\left(\dfrac{\lambda'_{ij}}{\lambda_{ij}}\right) - \dfrac{\lambda'_{ij}}{\lambda_{ij}} + 1 \leqslant 0$.

De plus $\lambda_{ij} \geqslant 0$, donc $-\lambda_{ij}\left(\ln\left(\dfrac{\lambda'_{ij}}{\lambda_{ij}}\right) - \dfrac{\lambda'_{ij}}{\lambda_{ij}} + 1\right) \geqslant 0$ et ainsi $K(X, Y, X', Y') \geqslant 0$.

Comme $K(X, Y, X', Y')$ est une somme de termes positifs, on a $K(X, Y, X', Y') = 0$ si, et seulement si, tous les termes de la somme sont nul ce qui signifie que $-\lambda_{ij}\left(\ln\left(\dfrac{\lambda'_{ij}}{\lambda_{ij}}\right) - \dfrac{\lambda'_{ij}}{\lambda_{ij}} + 1\right) = 0$ pour tous i et j.

Or on a supposé que $\lambda_{ij} \neq 0$, donc on se retrouve avec le cas d'égalité de la question I.B qui a déjà été résolu.

On a donc $K(X, Y, X', Y') = 0 \Leftrightarrow \dfrac{\lambda'_{ij}}{\lambda_{ij}} = 1 \Leftrightarrow \lambda'_{ij} = \lambda_{ij}$ pour tous i et j.

Pour conclure $K(X, Y, X', Y') \geqslant 0$, avec égalité si, et seulement si, les couples (X, Y) et (X', Y') on la même loi conjointe.

3) Comme X' et Y' sont indépendantes, on a $\lambda'_{ij} = P(X' = i) \times P(Y' = j) = p_i q_i$ car X' a la même loi que X et Y' la même loi que Y.

On a donc :

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

Or, d'après la formule des probabilités totales appliquée avec le système complet d'événement $([Y = j])_{j \in [\![0;m]\!]}$, on sait que

$$P(X = i) = \sum_{j=0}^{m} P([X = i] \cap [Y = j]) \Leftrightarrow p_i = \sum_{j=0}^{m} \lambda_{ij}.$$

De même $q_j = \sum_{i=0}^{n} \lambda_{ij}$.

On a donc

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

Ayant montré que $K(X,Y,X',Y') \geqslant 0$ on en déduit que $H(X) + H(Y) \geqslant H(X,Y)$.

De plus le cas d'égalité correspond à $K(X,Y,X',Y') = 0$ ce qui signifie que (X,Y) et (X',Y') ont la même loi conjointe.

B - ***Entropie conditionnelle***

1) D'après la question IV.A.3) $H(X,Y) - H(X) \leqslant H(Y)$. Donc on a bien $H_X(Y) \leqslant H(Y)$.
   Ainsi, la quantité d'information moyenne contenue dans les événements $[Y = k]$ est supérieur à celle mesurée lorsqu'on fixe la connaissance d'une autre variable aléatoire.

2) a) Les a_i étant tous positifs, on a $0 < a_j \leqslant a_0 + \ldots + a_m$ et par croissance de la fonction ln on obtient :

$$\ln(a_j) \leqslant \ln(a_0 + \ldots + a_m).$$

On multiplie l'inégalité ci-dessus par $a_j > 0$ et on somme ensuite pour j allant de 0 à m. On a donc :

$$\sum_{j=0}^{m} a_j \ln(a_j) \leqslant \sum_{j=0}^{m} a_j \ln(a_0 + \ldots + a_m) = \ln(a_0 + \ldots + a_m) \sum_{j=0}^{m} a_j$$
$$\Leftrightarrow \sum_{j=0}^{m} g(a_j) \leqslant g(a_0 + \ldots + a_m) = g\left(\sum_{j=0}^{m} a_j\right).$$

b) Étant donné que $g(0) = 0$, $\sum_{j=0}^{m} g(a_j) = \sum_{j=0/a_j \neq 0}^{m} g(a_j)$.

De plus $g\left(\sum_{j=0}^{m} a_j\right) = g\left(\sum_{j=0/a_j \neq 0}^{m} a_j\right)$.

L'égalité (IV.2) reste donc vraie pour $(a_0, \ldots, a_m) \in [0;1]^m$.

c) S'il existe au plus un indice j tel que $a_j \neq 0$, on voit facilement que l'inégalité (IV.2) devient une égalité.
Réciproquement, supposons que l'inégalité (IV.2) est une égalité.
S'il existe au moins deux indices tels que $a_j \neq 0$, quitte à réordonner, on peut supposer que $a_0 \neq 0$, ..., $a_r \neq 0$ et pour $j > r$, $a_j = 0$ avec $r \geqslant 1$.

On a alors $\sum_{j=0}^{m} g(a_j) = \sum_{j=0}^{r} g(a_j)$. De plus, pour $0 \leqslant j \leqslant r$, on a $\ln(a_j) < \ln(\sum_{j=0}^{r} a_i)$ (car il y a au moins deux termes dans la somme donc l'inégalité est stricte).

En reprenant le raisonnement précédent, on obtient alors $\sum_{j=0}^{r} g(a_j) < g(a_0 + \ldots + a_r)$.

Puis en rajoutant les termes nuls on voit alors que l'inégalité (IV.2) est alors stricte donc n'est pas une égalité.
En conclusion l'inégalité (IV.2) est une égalité si, et seulement s'il existe au plus un indice $j \in [\![0;m]\!]$ pour lequel $a_j \neq 0$.

3) En appliquant l'inégalité (IV.2) pour $a_j = \lambda_{ij}$ et en utilisant le fait que $p_i = \sum_{j=0}^{m} \lambda_{ij}$ on obtient que $\sum_{j=0}^{m} g(\lambda_{ij}) \leqslant g(p_i)$.

On somme alors pour i allant de 0 à n et on obtient :

$$\sum_{i=0}^{n} \sum_{j=0_{ij}}^{\lambda} \ln(\lambda_{ij}) \leqslant \sum_{i=0}^{n} p_i \ln(p_i)$$
$$\Leftrightarrow -H(X,Y) \leqslant -H(X)$$
$$\Leftrightarrow 0 \leqslant H(X,Y) - H(X) = H_X(Y).$$

On obtient bien $H_X(Y) \geqslant 0$.

V Une application

A - D'après la question III.A.3)a) on a $H(Y) = \ln(2016)$.

B - On utilise ici un principe de dichotomie et comme la valeur cherchée est un entier, on finira forcément par tomber sur la bonne valeur.

Il lui faudra 11 questions pour trouver la valeur de Y car $2^{11} = 2048$ donc en divisant à chaque fois l'intervalle en 2, 11 fois de suite on aura forcément la réponse.

C - On sait que les variables X_i prennent comme valeur soit 0 soit 1.

Donc X prends pour valeurs des entiers entre 0 et $\displaystyle\sum_{i=1}^{N} 2^{i-1} = \frac{1-2^N}{1-2} = 2^N - 1$.

X prend donc bien ses valeurs dans $[\![0; 2^N - 1]\!]$.

D - Soit X_0 la variable aléatoire suivant la loi uniforme sur $[\![0; 2^N - 1]\!]$.

D'après la question III.A.3) on sait que $H(X) \leqslant H(X_0) = \ln(2^N - 1)$. Or $\ln(2^N - 1) \leqslant \ln(2^N) = N\ln(2)$.

On a donc bien $H(X) \leqslant N\ln(2)$.

E - Lorsqu'on connait la valeur de Y, les variables X_i valent toutes 1 (car on connait les questions à poser !) et donc il n'y a pas d'incertitude sur la valeur de X. On peut donc dire que $H_Y(X) = 0$.

Comme $H_Y(X) = H(Y, X) - H(Y)$ on en déduit que $H(Y, X) = H(Y)$. Mais par définition de l'entropie du couple $H(X, Y) = H(Y, X)$.

Donc $H(X, Y) = H(Y)$.

F - On a $H_X(Y) = H(X, Y) - H(X) = H(Y) - H(X) = \ln(2016) - H(X)$. Or $H(X) \leqslant N\ln(2)$, donc

$$H_X(Y) \geqslant \ln(2016) - N\ln(2).$$

G - On cherche ici à savoir pour quelle valeur de N l'incertitude restant sur la valeur de Y peut devenir nulle.

On veut donc avoir N tel que $\ln(2016) - N\ln(2) \leqslant 0$ pour que $H_X(Y)$ puisse prendre la valeur 0.

Il faut donc avoir $N \geqslant \dfrac{\ln(2016)}{\ln(2)} \approx 10,97$. À partir de 11 questions B est certain de trouver la valeur de Y.

On retrouve le résultat de la question V.B qui est la méthode qui marche à tous les coups mais qui n'est pas forcément la plus astucieuse.