HEC MI 2012

J.F. COSSUTTA jean-francois.cossutta@wanadoo.fr

Partie I. Quelques inégalités de concavité

► Dans toute cette première partie ψ est l'application de $]0,+\infty[$ dans $\mathbb{R}$ définie par $\forall x \in]0,+\infty[,\ \psi(x) = -x\ \ln x$.

1. a) $x \to -x$ et ln sont de classe $\mathcal{C}^2$ sur $]0,+\infty[$ donc, par produit, ψ est de classe $\mathcal{C}^2$ sur $]0,+\infty[$.

De plus $\forall x \in]0,+\infty[,\ \psi'(x) = -\ln x - x\,\dfrac{1}{x} = -\ln x - 1$ et $\psi''(x) = -\dfrac{1}{x}\cdot$

$x \to 1-x$ est strictement positive et de classe $\mathcal{C}^2$ sur $]-\infty,1[$. Par composition $\eta : x \to \psi(1-x)$ est de classe $\mathcal{C}^2$ sur $]-\infty,1[$.

De plus $\forall x \in]-\infty,1[,\ \eta'(x) = -\psi'(1-x) = \ln(1-x)+1$ et $\eta''(x) = \psi''(1-x) = -\dfrac{1}{1-x}\cdot$

Notons que $\forall x \in]0,1[,\ h(x) = \psi(x)+\eta(x)$. Par somme, h est de classe $\mathcal{C}^2$ sur $]0,1[$. De plus :

$\forall x \in]0,1[,\ h'(x) = \psi'(x)+\eta'(x) = -\ln x - 1 + \ln(1-x)+1 = -\ln x + \ln(1-x)$ et $h''(x) = \psi''(x)+\eta''(x) = -\dfrac{1}{x} - \dfrac{1}{1-x}\cdot$

Plus de doute h'' est négative sur $]0,1[$ donc h est concave sur $]0,1[$.

Soit x un élément de $]0,1[$. Alors $1-x$ est également un élément de $]0,1[$.

Ainsi $-x<0$, $\ln x < 0$, $-(1-x)<0$, $\ln(1-x)<0$.

Par produit $-x\ \ln x > 0$ et $-(1-x)\ \ln(1-x) > 0$. Par somme $h(x)>0$, et ceci pour tout élément x de $]0,1[$.

$$\boxed{h \text{ est (strictement) positive et concave sur }]0,1[.}$$

b) $\lim\limits_{x\to 0} \psi(x) = \lim\limits_{x\to 0} \big(-x\ \ln x\big) = 0$ (par croissance comparée) et $\lim\limits_{x\to 1} \psi(x) = \lim\limits_{x\to 1} \big(-x\ \ln x\big) = 0$.

Par composition $\lim\limits_{x\to 0} \eta(x) = \lim\limits_{x\to 0} \psi(1-x) = 0$ et $\lim\limits_{x\to 1} \eta(x) = \lim\limits_{x\to 1} \psi(1-x) = 0$.

Alors $\lim\limits_{x\to 0} h(x) = \lim\limits_{x\to 0} \big(\psi(x)+\eta(x)\big) = 0+0 = 0$ et $\lim\limits_{x\to 1} h(x) = \lim\limits_{x\to 1} \big(\psi(x)+\eta(x)\big) = 0+0=0$.

Ainsi h est continue sur $]0,1[$ et admet pour limite 0 en 0 et en 1. Alors :

$$\boxed{h \text{ est prolongeable en une fonction } \widehat{h} \text{ continue sur } [0,1].}$$

$$\boxed{\forall x \in [0,1],\ \widehat{h}(x) = \begin{cases} -x\ \ln(x) - (1-x)\ \ln(1-x) & \text{si } x \in]0,1[\\ 0 & \text{si } x=0 \text{ ou si } x=1 \end{cases}.}$$

$\forall x \in]0,1[,\ \dfrac{\widehat{h}(x)-\widehat{h}(0)}{x-0} = \dfrac{-x\ \ln(x) - (1-x)\ \ln(1-x)}{x} = -\ln x - \dfrac{(1-x)\ \ln(1-x)}{x}\cdot$

$-(1-x)\ \ln(1-x) \underset{x\to 0}{\sim} -1\times(-x) = x$ donc $-\dfrac{(1-x)\ \ln(1-x)}{x} \underset{x\to 0}{\sim} 1$. Alors $\lim\limits_{x\to 0}\left(-\dfrac{(1-x)\ \ln(1-x)}{x}\right) = 1$.

Comme $\lim\limits_{x\to 0}(-\ln x) = +\infty$: $\lim\limits_{x\to 0} \dfrac{\widehat{h}(x)-\widehat{h}(0)}{x-0} = \lim\limits_{x\to 0}\left(-\ln x - \dfrac{(1-x)\ \ln(1-x)}{x}\right) = +\infty$.

$\widehat{h}$ n'est pas dérivable en 0. Notons néanmoins que sa courbe représentative admet au point d'abscisse 0 une demi-tangente parallèle à l'axe $(y'y)$.

$$\boxed{\text{Le prolongement de } h \text{ continue sur } [0,1] \text{ n'est pas de classe } \mathcal{C}^1 \text{ sur } [0,1].}$$

Remarque *On montre de même que* $\displaystyle\lim_{x\to 1} \frac{\widehat{h}(x)-\widehat{h}(1)}{x-1} = -\infty.$

Ainsi $\widehat{h}$ *n'est pas dérivable en* 1 *et sa courbe représentative admet au point d'abscisse* 1 *une demi-tangente parallèle à l'axe* $(y'y)$.

c) Nous avons vu que $\forall x \in]0,1[,\ h'(x) = -\ln x + \ln(1-x)$.

$\forall x \in]0,1[,\ h'(x) \leqslant 0 \iff \ln(1-x) \leqslant \ln x \iff 1-x \leqslant x \iff x \geqslant \dfrac{1}{2}\cdot$

Ce suffit pour dire que h est croissante sur $\left]0, \dfrac{1}{2}\right]$ et décroisante sur $\left[\dfrac{1}{2}, 1\right[$.

Notons que $h\left(\dfrac{1}{2}\right) = -\dfrac{1}{2}\ln\dfrac{1}{2} - \dfrac{1}{2}\ln\dfrac{1}{2} = \ln 2.$

Rappelons que $\displaystyle\lim_{x\to 0} h(x) = 0$ et $\displaystyle\lim_{x\to 1} h(x) = 0$.

Soit x un élément de $]0,1[$. Notons que $1-x \in]0,1[$.

De plus $h(1-x) = \psi(1-x) + \eta(1-x) = \eta(x) + \psi\big(1-(1-x)\big) = \eta(x) + \psi(x) = h(x)$.

Alors la courbe représentative de h, dans un repère orthogonal, admet la droite d'équation $x = \dfrac{1}{2}$ comme axe de symétrie.

Rappelons pour finir que la courbe représentative de $\widehat{h}$ admet aux points d'abscisses 0 et 1 une demi-tangente parallèle à l'axe $(y'y)$.

Toutes ces informations permettent de donner fière allure à la courbe représentaive de h. Je vous laisse le tracé....

2. ► $\boxed{\text{Dans toute la suite } L \text{ est l'application de }]0,+\infty[\text{ dans } \mathbb{R} \text{ définie par : } \forall u \in]0,+\infty[,\ L(u) = u - 1 - \ln u.}$

L est dérivable sur $]0,+\infty[$ et $\forall u \in]0,+\infty[,\ L'(u) = 1 - \dfrac{1}{u} = \dfrac{u-1}{u}\cdot$

L est continue sur $]0,1]$ et $\forall u \in]0,1[,\ L'(u) < 0$. Ceci suffit pour dire que L est strictement décroissante sur $]0,1]$.

L est continue sur $[1,+\infty[$ et $\forall u \in]1,+\infty[,\ L'(u) > 0$. Ceci suffit pour dire que L est strictement croissante sur $[1,+\infty[$.

Notons que $L(1) = 0$. Alors $\forall u \in]0,1[,\ L(u) > L(1) = 0$ et $\forall u \in]1,+\infty[,\ L(u) > L(1) = 0$. Donc :

$$\boxed{\forall u \in]0,+\infty[,\ L(u) = u-1-\ln u \geqslant 0 \text{ et } \forall u \in]0,+\infty[,\ L(u) = u-1-\ln u = 0 \iff u = 1.}$$

Ce qui donne encore :

$$\boxed{\forall u \in]0,+\infty[,\ \ln u \leqslant u-1 \text{ et } \forall u \in]0,+\infty[,\ \ln u = u-1 \iff u = 1.}$$

► *Remarque* *Le concepteur aurait été bien inspiré de nous faire déduire de cela que si* x *et* y *sont deux réels strictement positifs :* $x \ln\dfrac{y}{x} \leqslant y - x$ *avec égalité si et seulement si* $x = y$.

3. Soient x et y deux éléments de $]0,1[$. Notons que $\dfrac{y}{x}$ et $\dfrac{1-y}{1-x}$ sont deux réels strictement positifs.

Alors $d(x,y)=x\,\ln\left(\dfrac{y}{x}\right)+(1-x)\,\ln\left(\dfrac{1-y}{1-x}\right)$ est parfaitement défini.

$d(x,y)=x\,\left[\ln\left(\dfrac{y}{x}\right)-\dfrac{y}{x}+1\right]+(1-x)\,\left[\ln\left(\dfrac{1-y}{1-x}\right)-\dfrac{1-y}{1-x}+1\right]-x\,\left[-\dfrac{y}{x}+1\right]-(1-x)\,\left[-\dfrac{1-y}{1-x}+1\right].$

$d(x,y)=-x\,L\left(\dfrac{y}{x}\right)-(1-x)\,L\left(\dfrac{1-y}{1-x}\right)+y-x+(1-y)-(1-x)=-x\,L\left(\dfrac{y}{x}\right)-(1-x)\,L\left(\dfrac{1-y}{1-x}\right)+y-x+1-y-1+x.$

Ainsi $d(x,y)=-x\,L\left(\dfrac{y}{x}\right)-(1-x)\,L\left(\dfrac{1-y}{1-x}\right)=-\left[x\,L\left(\dfrac{y}{x}\right)+(1-x)\,L\left(\dfrac{1-y}{1-x}\right)\right].$

$x>0$, $y>0$, $L\left(\dfrac{y}{x}\right)\geqslant 0$ et $L\left(\dfrac{1-y}{1-x}\right)\geqslant 0$ (d'après **Q2**). Alors $x\,L\left(\dfrac{y}{x}\right)+(1-x)\,L\left(\dfrac{1-y}{1-x}\right)\geqslant 0$. Ainsi $d(x,y)\leqslant 0$.

- Supposons que $d(x,y)=0$. Alors $x\,L\left(\dfrac{y}{x}\right)+(1-x)\,L\left(\dfrac{1-y}{1-x}\right)=0$.

Ainsi $x\,L\left(\dfrac{y}{x}\right)$ et $(1-x)\,L\left(\dfrac{1-y}{1-x}\right)$ sont deux réels positifs ou nuls de somme nulle. Alors ces deux réels sont nuls.

Donc $x\,L\left(\dfrac{y}{x}\right)=(1-x)\,L\left(\dfrac{1-y}{1-x}\right)=0$. Comme x et $1-x$ ne sont pas nuls : $L\left(\dfrac{y}{x}\right)=L\left(\dfrac{1-y}{1-x}\right)=0$.

L'un des résultat de la question 2 donne alors : $\dfrac{y}{x}=\dfrac{1-y}{1-x}=1$. Plus de doute : $x=y$.

- Réciproquement supposons que $x=y$.

$d(x,y)=-x\,L\left(\dfrac{y}{x}\right)-(1-x)\,L\left(\dfrac{1-y}{1-x}\right)=-x\,L(1)-(1-x)\,L(1)=0$ car $L(1)=0$.

Finalement $d(x,y)$ est nul si et seulement si x et y sont égaux.

$$\boxed{\text{Si } x \text{ et } y \text{ sont deux éléments de }]0,1[,\ d(x,y)\leqslant 0, \text{ et } d(x,y)=0 \text{ si et seulement si } x=y.}$$

4. a) Soit b un réel quelconque (b pour éviter un télescopage de notations...).

ℓ est concave sur $\mathbb{R}$ donc sa courbe représentative est en dessous de toutes ses tangentes en particulier de celle au point d'abscisse b dont une des équations est $y=\ell'(b)\,(x-b)+\ell(b)$.

Alors $\forall x\in\mathbb{R},\ \ell(x)\leqslant \ell'(b)\,(x-b)+\ell(b)$ donc $\forall x\in\mathbb{R},\ \ell(x)-\ell(b)\leqslant \ell'(b)\,(x-b)$ et ceci pour tout réel b.

Le tout pouvant encore s'écrire de la manière suivante :

$$\boxed{\text{pour tout couple } (x,y) \text{ de } \mathbb{R}^2 \text{ on a l'inégalité } \ell(x)-\ell(y)\leqslant \ell'(y)(x-y).}$$

Ou encore :

$$\boxed{\text{pour tout couple } (x,y) \text{ de } \mathbb{R}^2 \text{ on a l'inégalité } \ell(x)\leqslant \ell(y)+\ell'(y)(x-y).}$$

Remarque *Si l'on ne veut pas utiliser la position de la courbe représentative de ℓ par rapport à ses tangentes on peut facilement obtenir ce résultat, par exemple en utilisant le théorème des accroissements finis ou une inégalité des accroissements finis ou encore la définition de la concavité.*

b) Soit y un réel quelconque.

D'après ce qui précède : $\forall x\in\mathbb{R},\ \ell\big(r(x)\big)\leqslant \ell(y)+\ell'(y)\,(r(x)-y)=\ell(y)+\ell'(y)\,r(x)-\ell'(y)\,y$ et $f(x)\geqslant 0$.

Donc $\forall x\in\mathbb{R},\ \ell\big(r(x)\big)\,f(x)\leqslant \ell(y)\,f(x)+\ell'(y)\,r(x)\,f(x)-\ell'(y)\,y\,f(x)$.

Par hypothèses : $\displaystyle\int_{-\infty}^{+\infty}\ell\big(r(x)\big)\,f(x)\,\mathrm{d}x$, $\displaystyle\int_{-\infty}^{+\infty}r(x)\,f(x)\,\mathrm{d}x$ et $\displaystyle\int_{-\infty}^{+\infty}f(x)\,\mathrm{d}x$ convergent. De plus $-\infty\leqslant+\infty$ (!).

Donc en intégrant l'inégalité précédente on obtient :

$$\int_{-\infty}^{+\infty} \ell\big(r(x)\big)\, f(x)\,\mathrm{d}x \leqslant \ell(y) \int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x + \ell'(y) \int_{-\infty}^{+\infty} r(x)\, f(x)\,\mathrm{d}x - \ell'(y)\, y \int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x.$$

En remarquant que $\displaystyle\int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x = 1$ il vient : $\displaystyle\int_{-\infty}^{+\infty} \ell\big(r(x)\big)\, f(x)\,\mathrm{d}x \leqslant \ell(y) + \ell'(y) \int_{-\infty}^{+\infty} r(x)\, f(x)\,\mathrm{d}x - \ell'(y)\, y.$

Soit encore : $\displaystyle\int_{-\infty}^{+\infty} \ell\big(r(x)\big)\, f(x)\,\mathrm{d}x \leqslant \ell(y) + \ell'(y) \left(\int_{-\infty}^{+\infty} r(x)\, f(x)\,\mathrm{d}x - y\right).$

$$\boxed{\text{Pour tout réel } y : \int_{-\infty}^{+\infty} \ell\big(r(x)\, f(x)\big)\,\mathrm{d}x \leqslant \ell(y) + \ell'(y) \left(\int_{-\infty}^{+\infty} r(x)\, f(x)\,\mathrm{d}x - y\right).}$$

c) En appliquant le résultat précédent pour $y = \displaystyle\int_{-\infty}^{+\infty} r(x)\, f(x)\,\mathrm{d}x$ on obtient :

$$\int_{-\infty}^{+\infty} \ell\big(r(x)\, f(x)\big)\,\mathrm{d}x \leqslant \ell\left(\int_{-\infty}^{+\infty} r(x)\, f(x)\,\mathrm{d}x\right) + \ell'\left(\int_{-\infty}^{+\infty} r(x)\, f(x)\,\mathrm{d}x\right) \left(\int_{-\infty}^{+\infty} r(x)\, f(x)\,\mathrm{d}x - \int_{-\infty}^{+\infty} r(x)\, f(x)\,\mathrm{d}x\right).$$

Le dernier terme étant nul, on obtient :

$$\boxed{\int_{-\infty}^{+\infty} \ell\big(r(x)\, f(x)\big)\,\mathrm{d}x \leqslant \ell\left(\int_{-\infty}^{+\infty} r(x)\, f(x)\,\mathrm{d}x\right).}$$

5. Pour obtenir le résultat, nous allons reprendre la démarche de Q4. en remplaçant $\displaystyle\int$ par $\sum$...

Soit y un élément de $\mathbb{R}$.

$\forall n \in \mathbb{N}^*,\ \ell\big(r(x_n)\big) \leqslant \ell(y) + \ell'(y)\,(r(x_n) - y) = \ell(y) + \ell'(y)\, r(x_n) - \ell'(y)\, y$ et $p_n \geqslant 0$.

Donc $\forall n \in \mathbb{N}^*,\ \ell\big(r(x_n)\big)\, p_n \leqslant \ell(y)\, p_n + \ell'(y)\, r(x_n)\, p_n - \ell'(y)\, y\, p_n$.

Rappelons que les séries de termes généraux $\ell\big(r(x_n)\big)\, p_n$, $r(x_n)\, p_n$ et p_n convergent.

En sommant l'inégalité précédente il vient : $\displaystyle\sum_{n=1}^{+\infty} \Big(\ell\big(r(x_n)\big)\, p_n\Big) \leqslant \ell(y) \sum_{n=1}^{+\infty} p_n + \ell'(y) \sum_{n=1}^{+\infty} \Big(r(x_n)\, p_n\Big) - \ell'(y)\, y \sum_{n=1}^{+\infty} p_n.$

Comme $\displaystyle\sum_{n=1}^{+\infty} p_n = 1 : \sum_{n=1}^{+\infty} \Big(\ell\big(r(x_n)\big)\, p_n\Big) \leqslant \ell(y) + \ell'(y) \sum_{n=1}^{+\infty} \Big(r(x_n)\, p_n\Big) - \ell'(y)\, y.$

Cette inégalité est vraie pour tout réel y donc en particulier pour $y = \displaystyle\sum_{n=1}^{+\infty} \Big(r(x_n)\, p_n\Big).$

Dans ce cas : $\displaystyle\sum_{n=1}^{+\infty} \Big(\ell\big(r(x_n)\big)\, p_n\Big) \leqslant \ell\left(\sum_{n=1}^{+\infty} \Big(r(x_n)\, p_n\Big)\right).$

$$\boxed{\sum_{n=1}^{+\infty} \Big(\ell\big(r(x_n)\big)\, p_n\Big) \leqslant \ell\left(\sum_{n=1}^{+\infty} \Big(r(x_n)\, p_n\Big)\right).}$$

Remarque 1 Ceci vaut encore si l'on suppose que ℓ concave et de classe $\mathcal{C}^1$ sur un intervalle I de $\mathbb{R}$ si l'on dit que $(x_n)_{n\in\mathbb{N}^*}$ est une suite de réels telle que $\forall n \in \mathbb{N}^*, r(x_n) \in I$.

Remarque 2 *Supposons ici ℓ est concave et $\mathcal{C}^1$ sur un intervalle I de $\mathbb{R}$. En prenant pour fonction r la fonction constante égale à 1, on peut dire que si $(x_n)_{n\in\mathbb{N}^*}$ est une suite d'éléments de I et si $(p_n)_{n\in\mathbb{N}^*}$ est une suite de réels positifs telle que $\sum\limits_{n=1}^{+\infty} p_n = 1$ on a :*

$$\sum_{n=1}^{+\infty} l(x_n)\, p_n \leqslant \ell\left(\sum_{n=1}^{+\infty} x_n\, p_n\right).$$

Remarque 3 *Supposons ici encore que ℓ est concave et $\mathcal{C}^1$ sur un intervalle I de $\mathbb{R}$.*

Soient N un élément de $\mathbb{N}^*$, $(x_1, x_2, \ldots, x_N)$ un N-uplet d'éléments de I et $(p_1, p_2, \ldots, p_N)$ un N-uplet de réels positifs tel que $\sum\limits_{n=1}^{N} p_n = 1$. n posant $\forall n \in [\![N+1, +\infty[\![,\ x_k = x_1$ et $p_n = 0$ et en appliquant ce qui précède on obtient :

$$\sum_{n=1}^{N} l(x_n)\, p_n \leqslant \ell\left(\sum_{n=1}^{N} x_n\, p_n\right).$$

On prenant $p_1 = p_2 = \ldots = p_N = \dfrac{1}{N}$ il vient : $\dfrac{1}{N}\sum\limits_{n=1}^{N} l(x_n) \leqslant \ell\left(\sum\limits_{n=1}^{N} \dfrac{1}{N}\, x_n\right)$ ou :

$$\frac{\ell(x_1) + \ell(x_2) \cdots + \ell(x_N)}{N} \leqslant \ell\left(\frac{x_1 + x_2 + \cdots + x_N}{N}\right).$$

Air très connu qui aurait pu servir pour traiter **III Q8. c)** *(on en reparlera) et qui vaut encore en supposant simplement ℓ concave sur I...*

Partie II. Entropie dans le cas continu

6. a) LA densité continue de Z est la fonction φ définie par : $\forall x \in \mathbb{R},\ \varphi(x) = \dfrac{1}{\sqrt{2\pi}}\, e^{-\frac{x^2}{2}}$.

Notons que φ est continue sur $\mathbb{R}$, à valeurs strictement positives et vérifie $\displaystyle\int_{-\infty}^{+\infty} \varphi(x)\, \mathrm{d}x = 1$.

$\forall x \in \mathbb{R},\ \varphi(x)\, \ln\big(\varphi(x)\big) = \varphi(x)\, \ln\left(\dfrac{1}{\sqrt{2\pi}}\, e^{-\frac{x^2}{2}}\right) = \varphi(x)\left[\ln\dfrac{1}{\sqrt{2\pi}} + \ln e^{-\frac{x^2}{2}}\right] = -\ln\left(\sqrt{2\pi}\right)\, \varphi(x) - \dfrac{1}{2}\, x^2\, \varphi(x)$.

Or $\displaystyle\int_{-\infty}^{+\infty} \varphi(x)\, \mathrm{d}t$ converge et vaut 1. De plus Z possède un moment d'ordre 2 donc $\displaystyle\int_{-\infty}^{+\infty} x^2\, \varphi(x)\, \mathrm{d}x$ converge.

Alors $\displaystyle\int_{-\infty}^{+\infty} \varphi(x)\, \ln\big(\varphi(x)\big)\, \mathrm{d}x$ converge comme combinaison linéaire de deux intégrales convergentes.

Ceci achève de montrer que φ appartient à $\mathcal{F}$ et permet de définir l'entropie de X.

De plus $\displaystyle\int_{-\infty}^{+\infty} \varphi(x)\, \ln\big(\varphi(x)\big)\, \mathrm{d}x = -\ln\sqrt{2\pi}\int_{-\infty}^{+\infty} \varphi(x)\, \mathrm{d}x - \frac{1}{2}\int_{-\infty}^{+\infty} x^2\, \varphi(x)\, \mathrm{d}x = -\ln\sqrt{2\pi} - \frac{1}{2}\, E(Z^2)$.

Or $E(Z^2) = V(Z) + \big(E(Z)\big)^2 = 1 + 0^2 = 1$. Ainsi $\displaystyle\int_{-\infty}^{+\infty} \varphi(x)\, \ln\big(\varphi(x)\big)\, \mathrm{d}x = -\ln\sqrt{2\pi} - \frac{1}{2}\cdot$

Donc $H(Z) = -\displaystyle\int_{-\infty}^{+\infty} \varphi(x)\, \ln\big(\varphi(x)\big)\, \mathrm{d}x = \ln\sqrt{2\pi} + \frac{1}{2} = \frac{\ln(2\pi) + 1}{2}\cdot$

On a encore $H(Z) = \ln\sqrt{2\pi} + \dfrac{1}{2} = \ln\sqrt{2\pi} + \ln\sqrt{e} = \ln\sqrt{2\pi\, e}$.

Une variable aléatoire Z qui suit la loi normale centrée réduite possède une entropie qui vaut $\dfrac{\ln(2\pi) + 1}{2}$ ou $\ln\sqrt{2\pi\, e}$.

Remarque φ *est l'unique densité de* Z *qui appartient à* $\mathcal{F}$.

b) Notons F_X (resp. F_Y) la fonction de répartition de X (resp. Y).

$\forall x \in \mathbb{R},\ F_Y(x) = P(Y \leqslant x) = P\,(a\,X + b \leqslant x) = P\left(X \leqslant \dfrac{x-b}{a}\right) = F_X\left(\dfrac{x-b}{a}\right)$ car a est strictement positif.

Donc $\forall x \in \mathbb{R},\ F_Y(x) = F_X\left(\dfrac{x-b}{a}\right)$.

f est une densité de X continue sur $\mathbb{R}$ donc F_X est de classe $\mathcal{C}^1$ sur $\mathbb{R}$ et $\forall x \in \mathbb{R},\ F_X'(x) = f(x)$.

Comme $x \to \dfrac{x-b}{a}$ est de classe $\mathcal{C}^1$ sur $\mathbb{R}$, par composition F_Y est de classe $\mathcal{C}^1$ sur $\mathbb{R}$.

Ceci suffit largement pour dire que Y est une variable aléatoire à densité et que F_Y' en est une densité définie sur $\mathbb{R}$.

Posons $f_Y = F_Y'$. $\forall x \in \mathbb{R},\ f_Y(x) = F_Y'(x) = \dfrac{1}{a}\,F_X'\left(\dfrac{x-b}{a}\right) = \dfrac{1}{a}\,f\left(\dfrac{x-b}{a}\right)$.

Notons déjà que :

1) f_Y est continue sur $\mathbb{R}$ car F_Y est de classe $\mathcal{C}^1$ sur $\mathbb{R}$;

2) f_Y est strictement positive sur $\mathbb{R}$ car $a > 0$ et $f \in \mathcal{F}$;

3) $\displaystyle\int_{-\infty}^{+\infty} f_Y(x)\,\mathrm{d}x$ converge et vaut 1 car f_Y est une densité de probabilité.

Pour montrer que f_Y appartient à $\mathcal{F}$ il ne reste plus qu'à montrer la convergence de $\displaystyle\int_{-\infty}^{+\infty} f_Y(x)\,\ln(f_Y(x))\,\mathrm{d}x$ c'est à dire la convergence de $\displaystyle\int_{-\infty}^{+\infty} \frac{1}{a}\,f\left(\frac{x-b}{a}\right)\,\ln\left(\frac{1}{a}\,f\left(\frac{x-b}{a}\right)\right)\mathrm{d}x$.

Notons que $x \to \dfrac{x-b}{a}$ définit une bijection strictement croissante ($a > 0$) de $\mathbb{R}$ sur $\mathbb{R}$, de classe $\mathcal{C}^1$ sur $\mathbb{R}$.
Alors le théorème de changement de variable sur les intégrales impropres permet de dire que
$\displaystyle\int_{-\infty}^{+\infty} \frac{1}{a}\,f\left(\frac{x-b}{a}\right)\,\ln\left(\frac{1}{a}\,f\left(\frac{x-b}{a}\right)\right)\mathrm{d}x$ est de même nature que $\displaystyle\int_{-\infty}^{+\infty} \frac{1}{a}\,f(u)\,\ln\left(\frac{1}{a}\,f(u)\right)a\,\mathrm{d}u$ et qu'en cas de convergence ces deux intégrales sont égales.

$\forall u \in \mathbb{R},\ \dfrac{1}{a}\,f(u)\,\ln\left(\dfrac{1}{a}\,f(u)\right)a = f(u)\,\ln\left(\dfrac{1}{a}\,f(u)\right) = \left(\ln\dfrac{1}{a}\right)f(u) + f(u)\,\ln\big(f(u)\big) = -(\ln a)\,f(u) + f(u)\,\ln\big(f(u)\big)$.

$\displaystyle\int_{-\infty}^{+\infty} f(u)\,\mathrm{d}u$ converge et vaut 1 et $\displaystyle\int_{-\infty}^{+\infty} f(u)\,\ln\big(f(u)\big)\,\mathrm{d}u$ converge et vaut $-H(X)$ ($f \in \mathcal{F}$ et X possède une entropie).

Alors $\displaystyle\int_{-\infty}^{+\infty} \frac{1}{a}\,f(u)\,\ln\left(\frac{1}{a}\,f(u)\right)a\,\mathrm{d}u$ converge comme combinaison linéaire de deux intégrales convergentes.

De plus $\displaystyle\int_{-\infty}^{+\infty} \frac{1}{a}\,f(u)\,\ln\left(\frac{1}{a}\,f(u)\right)a\,\mathrm{d}u = -\ln a\int_{-\infty}^{+\infty} f(u)\,\mathrm{d}u + \int_{-\infty}^{+\infty} f(u)\,\ln\big(f(u)\big)\,\mathrm{d}u = -\ln a - H(X)$. Ainsi

1) $\displaystyle\int_{-\infty}^{+\infty} f_Y(x)\,\ln\big(f_Y(x)\big)\,\mathrm{d}x$ converge ;

2) $\displaystyle\int_{-\infty}^{+\infty} f_Y(x)\,\ln\big(f_Y(x)\big)\,\mathrm{d}x = \int_{-\infty}^{+\infty} \frac{1}{a}\,f(u)\,\ln\left(\frac{1}{a}\,f(u)\right)a\,\mathrm{d}u = -\ln a - H(X)$.

Ce qui achève de montrer que f_Y appartient à $\mathcal{F}$. Mais cela permet aussi de dire que :

1) Y possède une entropie ;

2) $\displaystyle H(Y) = -\int_{-\infty}^{+\infty} f_Y(x)\,\ln\big(f_Y(x)\big)\,\mathrm{d}x = \ln a + H(X) = H(X) + \ln a$.

$\boxed{\text{La variable aléatoire } Y = a\,X + b \text{ possède une densité qui appartient à } \mathcal{F} \text{ et une entropie qui vaut } H(X) + \ln a.}$

Remarque $x \to \dfrac{1}{a}\, f\left(\dfrac{x-b}{a}\right)$ *est l'unique densité de Y qui appartient à $\mathcal{F}$.*

c) Soit T une variable aléatoire qui suit la loi normale d'espérance μ et d'écart-type $\sigma > 0$.

Posons $T^* = \dfrac{T-\mu}{\sigma} = \dfrac{T - E(T)}{\sqrt{V(T)}}$. T^* suit la loi normale centrée réduite.

Donc, d'après **Q6. a)**, T^* possède une densité qui appartient à $\mathcal{F}$ et une entropie qui vaut $\dfrac{\ln(2\pi)+1}{2}$.

Comme $T = \sigma\, T^* + \mu$ et que σ est strictement positif, **Q6. b)** montre que T possède une densité qui appartient à $\mathcal{F}$ et une entropie qui vaut $H(T^*) + \ln \sigma$.

Notons que $H(T) = H(T^*) + \ln\sigma = \dfrac{\ln(2\pi)+1}{2} + \ln\sigma = \dfrac{\ln(2\pi\,\sigma^2)+1}{2} = \ln\left(\sqrt{2\pi\, e}\,\sigma\right)$.

Une variable aléatoire qui suit la loi normale d'espérance μ et d'écart-type $\sigma > 0$ possède une entropie qui vaut $\dfrac{\ln(2\pi\,\sigma^2)+1}{2}$ ou $\ln\left(\sqrt{2\pi\, e}\,\sigma\right)$.

Remarque $x \to \dfrac{1}{\sqrt{2\pi}\,\sigma}\, e^{-\frac{(x-\mu)^2}{2\,\sigma^2}}$ *est l'unique densité de T qui appartient à $\mathcal{F}$.*

7. a) Rappelons que $\forall u \in]0, +\infty[,\ L(u) = u - 1 - \ln u \geqslant 0$ avec égalité si et seulement si $u = 1$.

$\forall x \in \mathbb{R},\ f(x)\,\ln\left(\dfrac{g(x)}{f(x)}\right) = f(x)\,\left[\ln\left(\dfrac{g(x)}{f(x)}\right) - \dfrac{g(x)}{f(x)} + 1\right] - f(x)\,\left[-\dfrac{g(x)}{f(x)} + 1\right] = -f(x)\,L\left(\dfrac{g(x)}{f(x)}\right) + g(x) - f(x)$.

Donc $\forall x \in \mathbb{R},\ f(x)\,\ln\left(\dfrac{g(x)}{f(x)}\right) = -f(x)\,L\left(\dfrac{g(x)}{f(x)}\right) + g(x) - f(x)$ et $f(x)\,L\left(\dfrac{g(x)}{f(x)}\right) = -f(x)\,\ln\left(\dfrac{g(x)}{f(x)}\right) + g(x) - f(x)$.

Or $\displaystyle\int_{-\infty}^{+\infty} f(x)\,\ln\left(\frac{g(x)}{f(x)}\right)\mathrm{d}x$, $\displaystyle\int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x$ et $\displaystyle\int_{-\infty}^{+\infty} g(x)\,\mathrm{d}x$ convergent. Alors :

1) $\displaystyle\int_{-\infty}^{+\infty} f(x)\,L\left(\frac{g(x)}{f(x)}\right)\mathrm{d}x$ converge comme combinaison linéaire de trois intégrales convergentes ;

2) $\displaystyle\int_{-\infty}^{+\infty} f(x)\,\ln\left(\frac{g(x)}{f(x)}\right)\mathrm{d}x = -\int_{-\infty}^{+\infty} f(x)\,L\left(\frac{g(x)}{f(x)}\right)\mathrm{d}x + \int_{-\infty}^{+\infty} g(x)\,\mathrm{d}x - \int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x$.

Rappelons que $\displaystyle\int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x = \int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x = 1$. Alors $\displaystyle\int_{-\infty}^{+\infty} f(x)\,\ln\left(\frac{g(x)}{f(x)}\right)\mathrm{d}x = -\int_{-\infty}^{+\infty} f(x)\,L\left(\frac{g(x)}{f(x)}\right)\mathrm{d}x$.

Ainsi $D(f,g) = -\displaystyle\int_{-\infty}^{+\infty} f(x)\,\ln\left(\frac{g(x)}{f(x)}\right)\mathrm{d}x = \int_{-\infty}^{+\infty} f(x)\,L\left(\frac{g(x)}{f(x)}\right)\mathrm{d}x$.

$\forall x \in \mathbb{R},\ f(x) > 0$ et $L\left(\dfrac{g(x)}{f(x)}\right) \geqslant 0$. Donc $\forall x \in \mathbb{R},\ f(x)\,L\left(\dfrac{g(x)}{f(x)}\right) \geqslant 0$ (résultat à retenir...).

Alors $D(f,g) = \displaystyle\int_{-\infty}^{+\infty} f(x)\,L\left(\frac{g(x)}{f(x)}\right)\mathrm{d}x \geqslant 0$ car $-\infty \leqslant +\infty$!

$$\boxed{D(f,g) \geqslant 0.}$$

b) Supposons $D(f,g) = 0$. Alors $\displaystyle\int_{-\infty}^{+\infty} f(x)\,L\left(\frac{g(x)}{f(x)}\right)\mathrm{d}x = D(f,g) = 0$.

Remarque *Pour faire plaisir au concepteur notons que* $\displaystyle\int_{-\infty}^{+\infty} f(x)\, L\left(\frac{g(x)}{f(x)}\right) \mathrm{d}x = 0$ *donne :*

$\displaystyle\int_{-\infty}^{+\infty} f(x)\, \left(\frac{g(x)}{f(x)} - 1 - \ln\left(\frac{g(x)}{f(x)}\right)\right) \mathrm{d}x = 0.$ *En multipliant par* -1 *on obtient :*

$$\boxed{\int_{-\infty}^{+\infty} \left(\ln\left(\frac{g(x)}{f(x)}\right) + 1 - \left(\frac{g(x)}{f(x)}\right)\right) f(x)\, \mathrm{d}x = 0.}$$

f et g sont deux fonctions continues et strictement positives sur $\mathbb{R}$. Alors $\dfrac{g}{f}$ est continue et strictement positive sur $\mathbb{R}$. Comme L est continue sur $]0, +\infty[$, par composition $x \to L\left(\dfrac{g(x)}{f(x)}\right)$ est continue sur $\mathbb{R}$.

Par produit $x \to f(x)\, L\left(\dfrac{g(x)}{f(x)}\right)$ est continue sur $\mathbb{R}$.

Ajoutons comme nous l'avons vu plus haut que cette fonction est positive ou nulle sur $\mathbb{R}$ et résumons.

1) $x \to f(x)\, L\left(\dfrac{g(x)}{f(x)}\right)$ est continue sur $\mathbb{R}$.

2) $x \to f(x)\, L\left(\dfrac{g(x)}{f(x)}\right)$ est positive sur $\mathbb{R}$.

3) $\displaystyle\int_{-\infty}^{+\infty} f(x)\, L\left(\frac{g(x)}{f(x)}\right) \mathrm{d}x = D(f,g) = 0.$

4) $-\infty \neq +\infty$!

Alors $x \to f(x)\, L\left(\dfrac{g(x)}{f(x)}\right)$ est identiquement nulle sur $\mathbb{R}$.

Comme f ne s'annule pas sur $\mathbb{R}$: $\forall x \in \mathbb{R},\ L\left(\dfrac{g(x)}{f(x)}\right) = 1.$

Ce que nous avons vu dans **I Q2.** permet de dire que $\forall x \in \mathbb{R},\ \dfrac{g(x)}{f(x)} = 1.$

Donc $\forall x \in \mathbb{R},\ g(x) = f(x)$. Alors $f = g$.

$$\boxed{\text{Si } D(f,g) = 0 \text{ alors } f = g.}$$

Remarque *Supposons que* $f = g$. *Alors* $\forall x \in \mathbb{R},\ f(x)\, \ln\left(\dfrac{g(x)}{f(x)}\right) = f(x)\, \ln 1 = 0.$

Donc $\displaystyle\int_{-\infty}^{+\infty} f(x)\, \ln\left(\frac{g(x)}{f(x)}\right) \mathrm{d}x$ *converge et vaut* 0. *Alors* $D(f,g)$ *existe et vaut* 0.

Partie III. Entropie dans le cas discret

Dans toute la mesure du possible nous éviterons la convention $x \ln x = 0$ si $x = 0$.

► Pour cela posons : $\boxed{\boxed{\forall x \in [0, +\infty[,\ t(x) = \begin{cases} x \ln x & \text{si } x > 0 \\ 0 & \text{si } x = 0 \end{cases}}}$.

t est continue sur $]0, +\infty[$. De plus $\lim\limits_{x \to 0} x \ln x = 0 = t(0)$ (par croissance comparée) donc t est continue en 0.

Ainsi t est continue sur son domaine de définition.

Notons alors que dans la question 8, $H(X) = -\sum\limits_{k=1}^{N} t(p_k)$... et ceci sans convention ni concession.

Terminons en remarquant que t est de classe $\mathcal{C}^2$ sur $]0,+\infty[$ et que, $\forall x \in]0,+\infty[,\ t'(x) = \ln x + 1$ et $t''(x) = \dfrac{1}{x}\cdot$

8. Dans cette question je pose $\mathcal{O} = \big(]0,1[\big)^N$. Notons que $\mathcal{O}$ est un ouvert de $\mathbb{R}^N$ comme produit de N ouverts de $\mathbb{R}$.

Nous avons aussi : $\forall x = (x_1, x_2, \ldots, x_N) \in \mathcal{O},\ h_N(x) = -\displaystyle\sum_{k=1}^{N} t(x_k)$.

a) Posons $\forall k \in [\![1,N]\!],\ \forall x = (x_1, x_2, \ldots, x_N) \in \mathcal{O},\ C_k(x) = x_k$.

Alors $\forall x = (x_1, x_2, \ldots, x_N) \in \mathcal{O},\ h_N(x) = -\displaystyle\sum_{k=1}^{N} (t \circ C_k)(x)$ ou $h_N = -\displaystyle\sum_{k=1}^{N} t \circ C_k$.

Pour tout k dans $[\![1,\mathbb{N}]\!]$, C_k est de classe $\mathcal{C}^2$ sur $\mathcal{O}$ (comme fonction polynômiale) et prend ses valeurs dans $]0,+\infty[$.

Alors, par composition, pour tout k dans $[\![1,N]\!]$, $t \circ C_k$ est de classe $\mathcal{C}^2$ sur $\mathcal{O}$ car t est de classe $\mathcal{C}^2$ sur $]0,+\infty[$.

Par somme $\displaystyle\sum_{k=1}^{N} t \circ C_k$ est de classe $\mathcal{C}^2$ sur $\mathcal{O}$. Ainsi $h_N = -\displaystyle\sum_{k=1}^{N} t \circ C_k$ est de classe $\mathcal{C}^2$ sur $\mathcal{O}$.

$\forall i \in [\![1,N]\!],\ \dfrac{\partial h_N}{\partial x_i} = -\displaystyle\sum_{k=1}^{N} \dfrac{\partial (t \circ C_k)}{\partial x_i} = -\displaystyle\sum_{k=1}^{N} \left(\dfrac{\partial C_k}{\partial x_i}\, t' \circ C_k\right)$.

Or si i et k sont deux éléments de $[\![1,N]\!]$, $\forall x \in \mathcal{O},\ \dfrac{\partial C_k}{\partial x_i}(x) = \begin{cases} 1 & \text{si } k = i \\ 0 & \text{sinon} \end{cases}$.

Donc $\forall i \in [\![1,N]\!],\ \dfrac{\partial h_N}{\partial x_i} = -t' \circ C_i$. Ainsi $\forall i \in [\![1,N]\!],\ \forall x = (x_1, x_2, \ldots, x_N) \in \mathcal{O},\ \dfrac{\partial h_N}{\partial x_i}(x) = -\ln x_i - 1$.

$$\boxed{\forall x = (x_1, x_2, \ldots, x_N) \in \big(]0,1[\big)^N,\ \nabla H_N(x) = (-\ln x_1 - 1, -\ln x_2 - 1, \ldots, -\ln x_N - 1).}$$

$\forall i \in [\![1,N]\!],\ \forall j \in [\![1,N]\!],\ \dfrac{\partial^2 h_N}{\partial x_j\, \partial x_i} = \dfrac{\partial \frac{\partial h_N}{\partial x_i}}{\partial x_j} = \dfrac{\partial(-t' \circ C_i)}{\partial x_j} = -\dfrac{\partial C_i}{\partial x_j}\, t'' \circ C_i$.

Alors $\forall i \in [\![1,N]\!],\ \forall j \in [\![1,N]\!], \forall x \in \mathcal{O},\ \dfrac{\partial^2 H_N}{\partial x_j\, \partial x_i}(x) = -\dfrac{\partial C_i}{\partial x_j}(x)\,(t'' \circ C_i)(x) = \begin{cases} -(t'' \circ C_i)(x) & \text{si } i = j \\ 0 & \text{sinon} \end{cases}$.

Alors $\forall i \in [\![1,N]\!],\ \forall j \in [\![1,N]\!], \forall x = (x_1, x_2, \ldots, x_N) \in \mathcal{O},\ \dfrac{\partial^2 H_N}{\partial x_j\, \partial x_i}(x) = \begin{cases} -\dfrac{1}{x_i} & \text{si } i = j \\ 0 & \text{sinon} \end{cases}$.

$$\boxed{\forall x = (x_1, x_2, \ldots, x_N) \in \big(]0,1[\big)^N,\ \nabla^2 h_N(x) = \mathrm{Diag}\left(-\frac{1}{x_1}, -\frac{1}{x_2}, \ldots, -\frac{1}{x_N}\right) = -\mathrm{Diag}\left(\frac{1}{x_1}, \frac{1}{x_2}, \ldots, \frac{1}{x_N}\right).}$$

b) Posons $\mathcal{C} = \{(x_1, x_2, \ldots, x_N) \in \mathcal{O} \mid \displaystyle\sum_{k=1}^{N} x_k = 1\}$ et $\forall x = (x_1, x_2, \ldots, x_N) \in \mathbb{R}^N, s(x) = \displaystyle\sum_{k=1}^{N} x_k$.

s est une forme linéaire sur $\mathbb{R}^N$ et $\forall z \in \mathbb{R}^n,\ \nabla s(z) = (1, 1, \ldots, 1)$.

Rappelons que $(\mathrm{Ker}\, s)^{\perp} = \mathrm{Vect}\big(\nabla(z)\big)$ pour tout z dans $\mathbb{R}^N$. Ainsi $(\mathrm{Ker}\, s)^{\perp} = \mathrm{Vect}\, 1, 1, \ldots, 1)$.

Soit $x = (x_1, x_2, \ldots, x_N)$ un élément de $\mathbb{R}^N$.

x est un point critique de h_N dans l'optimisation sous la contrainte $\mathcal{C}$ si et seulement si $x \in \mathcal{C}$ et $\nabla h_N(x) \in (\mathrm{Ker}\, s)^{\perp} = \mathrm{Vect}(1, 1, \ldots, 1)$.

Notons que $x \in \mathcal{C}$ et $\nabla h_N(x) \in \mathrm{Vect}(1, 1, \ldots, 1)$ si et seulement si :

$$\begin{cases} \forall k \in [\![1,N]\!],\ x_k \in]0,1[\\ \displaystyle\sum_{k=1}^{N} x_k = 1 \\ \exists \lambda \in \mathbb{R},\ (-\ln x_1 - 1, -\ln x_2 - 1, \dots, -\ln x_N - 1) = \lambda\,(1,1,\dots,1) \end{cases} \qquad (1).$$

$$(1) \Longleftrightarrow \begin{cases} \forall k \in [\![1,N]\!],\ x_k \in]0,1[\\ \displaystyle\sum_{k=1}^{N} x_k = 1 \\ -\ln x_1 - 1 = -\ln x_2 - 1 = \ldots = -\ln x_N - 1 \end{cases} \Longleftrightarrow \begin{cases} \forall k \in [\![1,N]\!],\ x_k \in]0,1[\\ \displaystyle\sum_{k=1}^{N} x_k = 1 \\ \ln x_1 = \ln x_2 = \ldots = \ln x_N \end{cases}.$$

$$(1) \Longleftrightarrow \begin{cases} \forall k \in [\![1,N]\!],\ x_k \in]0,1[\\ \displaystyle\sum_{k=1}^{N} x_k = 1 \\ x_1 = x_2 = \ldots = x_N \end{cases} \Longleftrightarrow \begin{cases} \forall k \in [\![1,N]\!],\ x_k \in]0,1[\\ N\,x_1 = 1 \\ x_1 = x_2 = \ldots = x_N \end{cases} \Longleftrightarrow \begin{cases} \forall k \in [\![1,N]\!],\ x_k \in]0,1[\\ x_1 = x_2 = \cdots = x_N = \dfrac{1}{N} \end{cases}.$$

$(1) \Longleftrightarrow x_1 = x_2 = \cdots = x_N = \dfrac{1}{N}$ puisque $\dfrac{1}{N} \in]0,1[$ $(N \geqslant 2)$.

Pour l'optimisation de h_N sous la contrainte $\displaystyle\sum_{k=1}^{N} x_k = 1$ il existe un unique point critique que nous notons x^*.

x^* est l'élément de $\big(]0,1[\big)^N$ égal à $\left(\dfrac{1}{N}, \dfrac{1}{N}, \dots, \dfrac{1}{N}\right)$.

c) Soit $x = (x_1, x_2, \dots, x_N)$ un élément de $\mathcal{C}$. Montrons que $h_N(x) \leqslant h_N(x^*)$.

Posons $\delta = x - x^*$. $x = x^* + \delta$.

$\forall \lambda \in [0,1], x^* + \lambda\,\delta = x^* + \lambda\,(x - x^*) = \lambda\,x + (1-\lambda)\,x^* = \left(\lambda\,x_1 + (1-\lambda)\,\dfrac{1}{N}, \lambda\,x_2 + (1-\lambda)\,\dfrac{1}{N}, \dots, \lambda\,x_N + (1-\lambda)\,\dfrac{1}{N}\right)$.

Or $]0,1[$ est convexe et, $\forall k \in]0,1[,\ x_k \in]0,1[$ et $\dfrac{1}{N} \in]0,1[$. Donc $\forall k \in]0,1[,\ \forall \lambda \in [0,1],\ \lambda\,x_k + (1-\lambda)\,\dfrac{1}{N} \in]0,1[$.

Par conséquent : $\forall \lambda \in [0,1],\ x^* + \lambda\,\delta = \lambda\,x + (1-\lambda)\,x^* \in \mathcal{O}$. Donc le segment $[x^*, x^* + \delta]$ est contenu dans $\mathcal{O}$.

h_N étant de classe $\mathcal{C}^2$ sur $\mathcal{O}$ nous pouvons lui appliquer la formule de Taylor-Lagrange à l'ordre 1.

Ainsi il existe un élément θ appartenant à $]0,1[$ tel que $h_N(x^* + \delta) = h_N(x^*) + <\nabla(x^*), \delta> + \dfrac{1}{2}\,q_{x^* + \theta\,\delta}(\delta)$, $q_{x^* + \theta\,\delta}$ étant la forme quadratique sur $\mathbb{R}^N$ associée à la Hessienne de h_N en $x^* + \theta\,\delta$.

x et x^* sont dans $\mathcal{C}$ donc $s(\delta) = s(x - x^*) = s(x) - s(x^*) = 1 - 1 = 0$. Alors δ appartient au noyau de s.

Comme x^* est un point critique de h_N dans l'optimisation sous la contrainte $\mathcal{C}$, $\nabla(x^*)$ appartient à l'orthogonal de $\operatorname{Ker} s$.

Donc $<\nabla(x^*), \delta> = 0$. On a alors $h_N(x^* + \delta) = h_N(x^*) + \dfrac{1}{2}\,q_{x^* + \theta\,\delta}(\delta)$.

En tout point de $\mathcal{O}$ la hessienne de h_N est une matrice diagonale à éléments diagonaux strictement négatifs.

Donc en tout point de $\mathcal{O}$ la hessienne de h_N est une matrice symétrique à valeurs propres strictement négatives.

Alors $\forall y \in \mathcal{O},\ \forall z \in \mathbb{R}^N,\ q_y(z) \leqslant 0$. Mieux $\forall y \in \mathcal{O},\ \forall z \in \mathbb{R}^N - \{0_{\mathbb{R}^N}\},\ q_y(z) < 0$.

Comme $x^* + \theta\,\delta$ appartient à $\mathcal{O}$, $q_{x^* + \theta\,\delta}(\delta) \leqslant 0$. Mieux $q_{x^* + \theta\,\delta}(\delta) < 0$ si $\delta \neq 0_{\mathbb{R}^N}$.

Donc $h_N(x) - h_N(x^*) = h_N(x^* + \delta) - h_N(x) = \dfrac{1}{2}\, q_{x^*+\theta\,\delta}(\delta) \leqslant 0$ et même $h_N(x) - h_N(x^*) < 0$ si $\delta \neq 0_{\mathbb{R}^n}$ donc si $x \neq x^*$.

$h_N(x) \leqslant h_N(x^*)$ avec égalité si et seulement si $x = x^*$.

Finalement : $x^* = \left(\dfrac{1}{N}, \dfrac{1}{N}, \dots, \dfrac{1}{N}\right) \in \mathcal{C}$ et $\forall x \in \mathcal{C},\ h_N(x) \leqslant h_N(x^*)$ avec égalité si et seulement si $x = x^*$. Ainsi :

$$\boxed{h_N \text{ admet en } x^* = \left(\frac{1}{N}, \frac{1}{N}, \dots, \frac{1}{N}\right) \text{ un maximum global strict sous la contrainte } \sum_{k=1}^{N} x_k = 1.}$$

Remarque 1 *Notons que* $h_N(x^*) = -\displaystyle\sum_{k=1}^{N} \left(\frac{1}{N}\ \ln\frac{1}{N}\right) = -N\left(\frac{1}{N}\ \ln\frac{1}{N}\right) = -\ln\frac{1}{N} = \ln N.$

Alors si $x = (x_1, x_2, \dots, x_N)$ *est un élément de* $\big(]0,1[\big)^N$ *tel que* $\sum\limits_{k=1}^{N} x_k = 1$, $h_N(x) \leqslant h_n(x^*)$ *donc* $-\sum\limits_{k=1}^{N} x_k\ \ln x_k \leqslant \ln N$*. Ainsi*

$$\boxed{\forall (x_1, x_2, \dots, x_N) \in \big(]0,1[\big)^N,\ \sum_{k=1}^{N} x_k = 1 \Rightarrow -\sum_{k=1}^{N} x_k\ \ln x_k \leqslant \ln N.}$$

Mieux encore :

$$\boxed{\forall (x_1, x_2, \dots, x_N) \in \big(]0,1[\big)^N,\ \sum_{k=1}^{N} x_k = 1 \text{ et } (x_1, x_2, \dots, x_N) \neq \left(\frac{1}{N}, \frac{1}{N}, \dots, \frac{1}{N}\right) \Rightarrow -\sum_{k=1}^{N} x_k\ \ln x_k < \ln N.}$$

Remarque 2 *On peut obtenir le résultat de c) plus simplement de la manière suivante. Il est facile de voir que* $\psi : x \to -x\ \ln x$ *est concave sur* $]0, +\infty[$ *(sa dérivée seconde est négative). Elle est aussi de classe* $\mathcal{C}^1$ *sur* $]0, +\infty[$*...*

Alors $\forall (x_1, x_2, \dots, x_N) \in]0, +\infty[,\ \psi\left(\dfrac{x_1 + x_2 + \cdots + x_N}{N}\right) \geqslant \dfrac{\psi(x_1) + \psi(x_2) + \cdots + \psi(x_N)}{N}$ *(voir les remarques de* **II Q5***).*

Ainsi $\forall (x_1, x_2, \dots, x_N) \in \mathcal{O},\ \displaystyle\sum_{k=1}^{N} x_k = 1 \Rightarrow \psi\left(\frac{1}{N}\right) \geqslant \frac{\psi(x_1) + \psi(x_2) + \cdots + \psi(x_N)}{N}.$

Soit encore $\forall (x_1, x_2, \dots, x_N) \in \mathcal{O},\ \displaystyle\sum_{k=1}^{N} x_k = 1 \Rightarrow h_N(x^*) = N\,\psi\left(\frac{1}{N}\right) \geqslant \psi(x_1) + \psi(x_2) + \cdots + \psi(x_N) = h_N(x).$

Ce qui donne à h_N *un maximum global en* x^* *sous la contrainte* $\displaystyle\sum_{k=1}^{N} x_k = 1$*. Maximum global nécessairement strict car il n' y a qu'un point critique.*

► *Remarque 3* *Notons que ce nous venons de faire en plus de deux pages peut se faire en quelques lignes par un élève de terminale. Nous saurons nous en souvenir dans Q13...*

En effet, soit $x = (x_1, x_2, \dots, x_N) \in \mathcal{O}$ *tel que* $\displaystyle\sum_{k=1}^{N} x_k = 1$.

Si $k \in [\![1, N]\!]$, $-x_k\ \ln(x_k) = x_k \ln\dfrac{1}{x_k} = x_k \ln\dfrac{1}{N\,x_k} + x_k\ \ln N.$

D'après **Q2.***, si* $k \in [\![1, N]\!]$, $-x_k\ \ln(x_k) \leqslant x_k\left(\dfrac{1}{N\,x_k} - 1\right) + x_k\ \ln N = \dfrac{1}{N} - x_k + x_k\ \ln N$ *avec égalité si et seulement si* $\dfrac{1}{N\,x_k} = 1$ *ou si et seulement si* $x_k = \dfrac{1}{N}$ *car* $x_k > 0$.

Donc $h_N(x) = -\displaystyle\sum_{k=1}^{N} x_k \ln(x_k) \leqslant \sum_{k=1}^{N} \frac{1}{N} - \sum_{k=1}^{N} x_k + \ln N \sum_{k=1}^{N} x_k = 1 - 1 + (\ln N) \times 1 = \ln N = h_N\left(\frac{1}{N}, \frac{1}{N}, \dots, \frac{1}{N}\right)$

avec égalité si et seulement si $(x_1, x_2, \dots, x_N) = \left(\frac{1}{N}, \frac{1}{N}, \dots, \frac{1}{N}\right)$.

Ainsi si $x = (x_1, x_2, \dots, x_N)$ *est un élément de* $\mathcal{O}$ *tel que* $\displaystyle\sum_{k=1}^{N} x_k = 1$ *et distincts de* x^* *:* $h_N(x) < h_N(x^*)$.

d) • Soit U une variable aléatoire qui suit la loi uniforme sur $[\![1, N]\!]$. $\forall k \in [\![1, N]\!],\ P(U = k) = \frac{1}{N}$.

U prend ses valeurs dans $[\![1, N]\!]$. De plus

$H(U) = -\displaystyle\sum_{k=1}^{N} \big(P(U = k) \ln\big(P(U = k)\big)\big) = -\sum_{k=1}^{N} \frac{1}{N} \ln \frac{1}{N} = -N \frac{1}{N} \ln \frac{1}{N} = -\ln \frac{1}{N} = \ln N.$

• Soit X une variable aléatoire qui prend ses valeurs dans $[\![1, N]\!]$ et qui ne suit pas la loi uniforme sur $[\![1, N]\!]$.

Nous allons montrer que $H(X) < H(U)$ en montrant que $H(X) < \ln N$.

Posons $\forall k \in [\![1, N]\!],\ p_k = P(X = k)$. $H(X) = -\sum\limits_{k=1}^{N} t(p_k)$.

Notons que $(x_1, x_2, \dots, x_N) \neq \left(\frac{1}{N}, \frac{1}{N}, \dots, \frac{1}{N}\right)$ car X ne suit pas la loi uniforme sur $[\![1, N]\!]$.

Premier cas : on suppose que $\forall k \in [\![1, N]\!],\ p_k \neq 0$.

Alors $\forall k \in [\![1, N]\!],\ p_k > 0$ et $\sum\limits_{k=1}^{N} p_k = 1$.

Mieux, comme $N \geqslant 2$, $\forall k \in [\![1, N]\!],\ p_k \in]0, 1[$ et $\sum\limits_{k=1}^{N} p_k = 1$. De plus $(p_1, p_2, \dots, p_N) \neq \left(\frac{1}{N}, \frac{1}{N}, \dots, \frac{1}{N}\right)$.

Dans ces conditions, **c)** donne : $H(X) = -\displaystyle\sum_{k=1}^{N} p_k \ln p_k = h_N(p_1, p_2, \dots, p_N) < \ln N$. Ainsi $H(X) < H(U)$.

Deuxième cas : on suppose qu'il existe au moins un éléments k_0 de $[\![1, N]\!]$ tel que $p_{k_0} = 0$.

Notons que l'on ne peut pas avoir $p_1 = p_2 = \dots = p_n = 0$ car $\displaystyle\sum_{k=1}^{N} p_k = 1$. Alors $K = \{k \in [\![1, N]\!] \mid p_k \neq 0\}$ est une partie non vide de $[\![1, N]\!]$. Notons r son cardinal. $r < N$ et il existe une bijection w de $[\![1, r]\!]$ sur K.

Rappelons que $t(0) = 0$. Alors $H(X) = -\displaystyle\sum_{k=1}^{N} t(p_k) = -\sum_{k \in K} t(p_k) = -\sum_{k=1}^{r} t(p_{w(k)}) = -\sum_{k=1}^{r} p_{w(k)} \ln p_{w(k)}$.

Si $r = 1$, alors $p_{w(1)} = 1$ et $H(X) = -p_{w(1)} \ln p_{w(1)} = 0$; ainsi $H(X) < \ln N = H(U)$. Supposons $r \geqslant 2$.

$\forall k \in [\![1, r]\!],\ p_{w(k)} > 0$ et $\displaystyle\sum_{k=1}^{r} p_{w(k)} = 1$. Alors $\forall k \in [\![1, r]\!],\ p_{w(k)} \in]0, 1[$ et $\displaystyle\sum_{k=1}^{r} p_{w(k)} = 1$ (car $r \geqslant 2$).

Alors **c)** montre que $H(X) = -\displaystyle\sum_{k=1}^{r} p_{w(k)} \ln p_{w(k)} \leqslant \ln r < \ln N = H(U)$.

Donc dans tous les cas $H(X) < \ln N = H(U)$.

Si X est une variable aléatoire à valeurs dans $[\![1, N]\!]$ qui ne suit pas la loi uniforme sur $[\![1, N]\!]$, son entropie est strictement inférieure à l'entropie des variables aléatoires qui suivent la loi uniforme sur $[\![1, N]\!]$.

$\boxed{\text{Parmi les variables aléatoires à valeurs dans } [\![1,N]\!], \text{ seules les variables aléatoires qui suivent la loi uniforme sur } [\![1,N]\!] \text{ ont une entopie maximum.}}$

9. a) La série de terme général $n\,p_n$ converge donc $\lim\limits_{n\to+\infty}(n\,p_n)=0$.

Comme $\lim\limits_{n\to+\infty}\dfrac{1}{n}=0$: $\lim\limits_{n\to+\infty}p_n=\lim\limits_{n\to+\infty}\left(\dfrac{1}{n}\times(n\,p_n)\right)=0\times 0=0$.

Par croissance comparée il vient alors $\lim\limits_{n\to+\infty}(\sqrt{p_n}\,\ln(p_n))=0$. Donc $\lim\limits_{n\to+\infty}(\sqrt{p_n}\,|\ln(p_n)|)=\lim\limits_{n\to+\infty}|\sqrt{p_n}\,\ln(p_n)|=0$.

Alors $\forall\varepsilon\in\mathbb{R}^{+*}$, $\exists n_0\in\mathbb{N}^*$, $\forall n\in\mathbb{N}^*$, $n\geqslant n_0\Rightarrow\sqrt{p_n}\,|\ln(p_n)|<\varepsilon$. Ne reste plus qu'à faire $\varepsilon=1$ pour pouvoir largement dire que :

$\boxed{\text{il existe un élément } n_0 \text{ de } \mathbb{N}^* \text{ tel que pour tout entier supérieur ou égal à } n_0 : \sqrt{p_n}\,|\ln(p_n)|\leqslant 1.}$

b) Soit n un élément de $\mathbb{N}^*$ tel que $n\geqslant n_0$ et $p_n\leqslant\dfrac{1}{n^3}$.

$0\leqslant\sqrt{p_n}\leqslant\sqrt{\dfrac{1}{n^3}}=\dfrac{1}{n^{\frac{3}{2}}}$ et $0\leqslant\sqrt{p_n}\,|\ln(p_n)|\leqslant 1$. Par produit : $\sqrt{p_n}\,\sqrt{p_n}\,|\ln p_n|\leqslant\dfrac{1}{n^{\frac{3}{2}}}$. Donc $p_n\,|\ln p_n|\leqslant\dfrac{1}{n^{\frac{3}{2}}}$.

$\boxed{\text{si } n \text{ est un élément de } \mathbb{N}^* \text{ tel que } n\geqslant n_0 \text{ et } p_n\leqslant\dfrac{1}{n^3} \text{ alors } p_n\,|\ln(p_n)|\leqslant\dfrac{1}{n^{\frac{3}{2}}}.}$

c) Soit n un élément de $\mathbb{N}^*$ tel que $n\geqslant n_0$.

Si $p_n\leqslant\dfrac{1}{n^3}$, $p_n\,|\ln p_n|\leqslant\dfrac{1}{n^{\frac{3}{2}}}\leqslant\mathrm{Max}\left\{\dfrac{1}{n^{\frac{3}{2}}}\,,\,3p_n\,\ln n\right\}$.

Supposons $p_n>\dfrac{1}{n^3}$. Alors $0\geqslant\ln(p_n)>\ln\dfrac{1}{n^3}=-3\ln n$ donc $0\leqslant-\ln(p_n)<3\ln(n)$ ou $|\ln(p_n)|<3\ln(n)$.

Comme $p_n\geqslant 0$: $p_n\,|\ln(p_n)|\leqslant 3p_n\,\ln(n)\leqslant\mathrm{Max}\left\{\dfrac{1}{n^{\frac{3}{2}}}\,,\,3p_n\,\ln n\right\}$.

$\boxed{\text{Si } n \text{ est un élément de } \mathbb{N}^* \text{ supérieur ou égal à } n_0 : p_n\,|\ln(p_n)|\leqslant\mathrm{Max}\left\{\dfrac{1}{n^{\frac{3}{2}}}\,,\,3p_n\,\ln n\right\}.}$

d) Soit n un élément de $[\![n_0,+\infty[\![$. $0\leqslant p_n\,|\ln(p_n)|\leqslant\mathrm{Max}\left\{\dfrac{1}{n^{\frac{3}{2}}}\,,\,3p_n\,\ln n\right\}\leqslant\dfrac{1}{n^{\frac{3}{2}}}+3p_n\,\ln n$.

Or $\ln n\leqslant n-1$ (d'après **I Q2.**). Donc $\ln n\leqslant n$ et $p_n>0$. Ainsi $3p_n\,\ln n\leqslant 3n\,p_n$.

Finalement $0\leqslant p_n\,|\ln(p_n)|\leqslant\dfrac{1}{n^{\frac{3}{2}}}+3n\,p_n$ et ceci pour tout élément n de $[\![n_0,+\infty[\![$.

La série de terme général $\dfrac{1}{n^{\frac{3}{2}}}$ converge car $\dfrac{3}{2}>1$ et la série de terme général $n\,p_n$ converge par hypothèse.

Alors la série de terme général $\dfrac{1}{n^{\frac{3}{2}}}+3n\,p_n$ converge comme combinaison linéaire de deux séries convergentes.

Les règles de comparaison sur les séries à termes positifs montrent alors la convergence de la série de terme général $p_n\,|\ln(p_n)|$.

$\boxed{\text{La série de terme général } p_n\,|\ln(p_n)| \text{ converge.}}$

Si nous reprenons les notations qui précèdent Q9, dire que la série de terme général $n\,p_n$ converge c'est dire que X possède une espérance car $\forall n \in \mathbb{N}^*,\ n\,p_n \geqslant 0$. Et dire que la série de terme général $p_n\,|\ln p_n|$ converge c'est dire que X possède une entropie.

Donc l'existence de $E(X)$ donne l'existence de $H(X)$.

Soit X une variables aléatoire à valeurs dans $\mathbb{N}^*$ telle que $\forall n \in \mathbb{N}^*,\ p_n > 0$.
Si X possède une espérance alors X possède une entropie.

Nous avons montré le résultat attendu, est-ce le résultat demandé ? On parle de variable aléatoire à valeurs dans $\mathbb{N}^*$ sans restriction, non ?

Allons donc un peu plus loin même si cela n'est pas franchement indispensable...

Dans la suite X est toujours une variable aléatoire à valeurs dans $\mathbb{N}^*$. On suppose que X possède une espérance.

On pose encore $\forall n \in \mathbb{N}^*,\ p_n = P(X = n)$. Mais on ne suppose plus que $\forall n \in \mathbb{N}^*,\ p_n > 0$.

Rappelons que $\forall x \in [0, +\infty[,\ t(x) = \begin{cases} x \ln x & \text{si } x > 0 \\ 0 & \text{si } x = 0 \end{cases}$.

Dans le texte $H(X)$ existe si la série de terme général $p_n\,|\ln(p_n)|$ converge ce qui revient à dire que $H(X)$ existe si la série de terme général $\big|p_n \ln(p_n)\big|$ converge.

Donc pour nous $H(X)$ existe si la série de terme général $\big|t(p_n)\big|$ converge.

En cas d'existence $H(X)$ vaut $\displaystyle\sum_{n=1}^{+\infty} \big|t(p_n)\big|$ ou $\displaystyle -\sum_{n=1}^{+\infty} t(p_n)$.

Notons que notre situation contient le cas (inutile !) où X prend un nombre fini de valeurs.

Version 1 On adapte rapidement la preuve précédente.

Esquissons ! On montre de la même manière que $\lim\limits_{n\to+\infty} p_n = 0$.

$\lim\limits_{x\to 0} \sqrt{x}\,|\ln x| = \lim\limits_{x\to 0} |\sqrt{x}\,\ln x| = 0$. Alors il existe $\gamma \in \mathbb{R}^{+\,*}$ tel que $\forall x \in]0, \gamma[,\ \sqrt{x}\,|\ln x| \leqslant 1$.

Or il existe n_0 dans $\mathbb{N}^*$ tel que $\forall n \in [\![n_0, +\infty[\![,\ p_n < \gamma$ car $(p_n)_{n\in\mathbb{N}^*}$ converge vers 0.

Donc si n est dans $[\![n_0, +\infty[\![$ et si $p_n \neq 0$ alors $p_n \in]0, \gamma[$, ce qui donne : $\sqrt{p_n}\,|\ln(p_n)| \leqslant 1$.

Soit n dans $[\![n_0, +\infty[\![$. Si $p_n = 0$ on a $|t(p_n)| = 0 \leqslant \mathrm{Max}\{\frac{1}{n^{\frac{3}{2}}}, 3p_n \ln n\}$.

Si $p_n \neq 0$ on montre, comme plus haut, que $|t(p_n)| = p_n\,|\ln(p_n)| \leqslant \mathrm{Max}\{\frac{1}{n^{\frac{3}{2}}}, 3p_n \ln n\}$ en faisant deux cas : $p_n \leqslant \dfrac{1}{n^3}$ et $p_n > \dfrac{1}{n^3}$.

Finalement $\forall n \in [\![n_0, +\infty[\![,\ |t(p_n)| \leqslant \mathrm{Max}\{\frac{1}{n^{\frac{3}{2}}}, 3p_n \ln n\}$.

Alors $0 \leqslant |t(p_n)| \leqslant \mathrm{Max}\left\{\dfrac{1}{n^{\frac{3}{2}}}\ , 3p_n \ln n\right\} \leqslant \dfrac{1}{n^{\frac{3}{2}}} + 3p_n \ln n$.

Comme nous l'avons vu plus haut, on obtient alors aisément la convergence de la série de terme général $|t(p_n)|$.

Ainsi $H(X)$ existe.

Version 2 On profite de l'occasion pour proposer une seconde preuve, plus simple.

Soit n un élément de $\mathbb{N}^*$ tel que $p_n \neq 0$.

$\ln\left(\dfrac{1}{p_n\,e^n}\right) \leqslant \dfrac{1}{p_n\,e^n} - 1 \leqslant \dfrac{1}{p_n\,e^n}$ et $p_n > 0$. Donc $p_n \ln\left(\dfrac{1}{p_n\,e^n}\right) \leqslant \dfrac{1}{e^n} = \left(\dfrac{1}{e}\right)^n$.

Or $p_n \ln\left(\dfrac{1}{p_n\, e^n}\right) = -p_n \ln(p_n\, e^n) = -p_n \ln p_n - n\, p_n = |p_n \ln p_n| - n\, p_n = |t(p_n)| - n\, p_n.$

Donc $|t(p_n)| - n\, p_n \leqslant \left(\dfrac{1}{e}\right)^n$. Ainsi $0 \leqslant |t(p_n)| \leqslant n\, p_n + \left(\dfrac{1}{e}\right)^n$.

Notons que cet encadrement vaut encore si $p_n = 0$ car $t(0) = 0$.

Donc $\forall n \in \mathbb{N}^*,\ 0 \leqslant |t(p_n)| \leqslant n\, p_n + \left(\dfrac{1}{e}\right)^n$.

X possède une espérance et $\left|\dfrac{1}{e}\right| < 1$ donc les séries de termes généraux $n\, p_n$ et $\left(\dfrac{1}{e}\right)^n$ sont convergentes.

Donc la série de terme général $n\, p_n + \left(\dfrac{1}{e}\right)^n$ converge. Les règles de comparaison sur les séries à termes positifs montrent alors la convergence de la série de terme général $|t(p_n)|$. Ainsi $H(X)$ existe.

$$\boxed{\text{Toute variable aléatoire à valeurs dans } \mathbb{N}^* \text{ qui possède une espérance possède une entropie.}}$$

Remarque *Nous donnerons dans c) une version 3, ou plutôt une version 2'...*

Exercice *Une urne contient une boule rouge et une boule noire. On tire une boule de l'urne. Si elle est rouge on s'arrête sinon on la remet dans l'urne, on ajoute une autre boule noire dans l'urne et on procède à un autre tirage en respectant le même protocole. X est la variable aléatoire égale au nombre de tirages effectués.*

Montrer que X a une entropie mais n'a pas d'espérance.

10. a) X suit la loi géométrique de paramètre θ. Alors $E(X) = \dfrac{1}{\theta}\cdot$

$$\boxed{X \text{ suit la loi géométrique de paramètre } \theta \text{ et } E(X) = \frac{1}{\theta}\cdot}$$

$\forall n \in \mathbb{N}^*,\ P(X=n) \ln\big(P(X=n)\big) = P(X=n) \ln\big(\theta\,(1-\theta)^{n-1}\big) = \big[\ln\theta + (n-1) \ln(1-\theta)\big]\, P(X=n).$

$\forall n \in \mathbb{N}^*,\ P(X=n) \ln\big(P(X=n)\big) = \big[\ln\theta - \ln(1-\theta)\big]\, P(X=n) + \ln(1-\theta)\, n\, P(X=n).$

La série de terme général $P(X=n)$ converge et $\displaystyle\sum_{n=1}^{+\infty} P(X=n) = 1$.

La série de terme général $n\, P(X=n)$ converge et $\displaystyle\sum_{n=1}^{+\infty} n\, P(X=n) = \frac{1}{\theta}$ car X a une espérance qui vaut $\dfrac{1}{\theta}\cdot$

Ainsi la série de terme général $P(X=n) \ln\big(P(X=n)\big)$ converge comme combinaison linéaire de deux séries convergentes.

Or $\forall n \in \mathbb{N}^*,\ P(X=n)\, \big|\ln\big(P(X=n)\big)\big| = -P(X=n) \ln\big(P(X=n)\big)$.

Alors la série de terme général $P(X=n)\, \big|\ln\big(P(X=n)\big)\big|$ converge ce qui assure l'existence de l'entropie de X.

De plus : $H(X) = \displaystyle\sum_{n=1}^{+\infty} \Big(P(X=n)\, \big|\ln\big(P(X=n)\big)\big|\Big) = -\sum_{n=1}^{+\infty} \Big(P(X=n) \ln\big(P(X=n)\big)\Big)$.

$H(X) = -\big[\ln(\theta) - \ln(1-\theta)\big] \displaystyle\sum_{n=1}^{+\infty} P(X=n) - \ln(1-\theta) \sum_{n=1}^{+\infty} n\, P(X=n)$ (toutes les séries convergent).

$H(X) = -\big[\ln(\theta) - \ln(1-\theta)\big] \times 1 - \ln(1-\theta) \times \dfrac{1}{\theta}\cdot$

$H(X) = -\ln\theta + \ln(1-\theta) - \dfrac{\ln(1-\theta)}{\theta} = -\dfrac{\theta \ln\theta + (1-\theta) \ln(1-\theta)}{\theta}.$

Une variable aléatoire qui suit la loi géométrique de paramètre θ possède une entropie qui vaut : $-\dfrac{\theta\,\ln\theta+(1-\theta)\,\ln(1-\theta)}{\theta}$.

b) Il s'agit de simuler une variable aléatoire qui suit la loi géométrique de paramètre θ donc en gros de simuler le temps d'attente de la réalisation d'un événement, qui se produit avec la probabilité θ, associé à une expérienece aléatoire que l'on itère, les itérations étant indépendantes. Rien que du classique, non ?

```
function X(theta:real):real;

var n:integer;

begin

n:=0;
repeat
n:=n+1;
until (random<=theta);
X:=n;

end;
```

c) Notons que $H(Y)$ existe car Y possède une espérance (qui vaut $E(X)$).

$H(Y)-H(X)=\displaystyle\sum_{n=1}^{+\infty} q_n\,\big|\ln(q_n)\big|-\sum_{n=1}^{+\infty} p_n\,\big|\ln(p_n)\big|$.

$H(Y)-H(X)=-\displaystyle\sum_{n=1}^{+\infty} q_n\,\ln(q_n)+\sum_{n=1}^{+\infty} p_n\,\ln(p_n)=\sum_{n=1}^{+\infty}\left(q_n\,\ln\frac{1}{q_n}+p_n\,\ln p_n\right)=\sum_{n=1}^{+\infty}\left(q_n\,\ln\left(\frac{p_n}{q_n}\right)+(p_n-q_n)\,\ln p_n\right)$.

$\forall n\in\mathbb{N}^*,\ (p_n-q_n)\,\ln(p_n)=(p_n-q_n)\,\ln\big(\theta\,(1-\theta)^{n-1}\big)=(p_n-q_n)\,\big[\ln\theta+(n-1)\,\ln(1-\theta)\big]$.

$\forall n\in\mathbb{N}^*,\ (p_n-q_n)\,\ln(p_n)=(p_n-q_n)\,\big[\ln\theta-\ln(1-\theta)+n\,\ln(1-\theta)\big]=\big[\ln\theta-\ln(1-\theta)\big]\,(p_n-q_n)+\big[\ln(1-\theta)\big]\,(n\,p_n-n\,q_n)$.

$\forall n\in\mathbb{N}^*,\ (p_n-q_n)\,\ln(p_n)=\big[\ln\theta-\ln(1-\theta)\big]\,p_n-\big[\ln\theta-\ln(1-\theta)\big]\,q_n+\big[\ln(1-\theta)\big]\,n\,p_n-\big[\ln(1-\theta)\big]\,n\,q_n$.

Rappelons que les séries de terme généraux p_n, q_n, $n\,p_n$ et $n\,q_n$ convergent.

Alors la série de terme général $(p_n-q_n)\,\ln(p_n)$ converge comme combinaison linéaire de quatre séries convergentes.

De plus :

$$\sum_{n=1}^{+\infty}(p_n-q_n)\,\ln(p_n)=\big[\ln\theta-\ln(1-\theta)\big]\sum_{n=1}^{+\infty}p_n-\big[\ln\theta-\ln(1-\theta)\big]\sum_{n=1}^{+\infty}q_n+\big[\ln(1-\theta)\big]\sum_{n=1}^{+\infty}(n\,p_n)-\big[\ln(1-\theta)\big]\sum_{n=1}^{+\infty}(n\,q_n).$$

Donc $\displaystyle\sum_{n=1}^{+\infty}(p_n-q_n)\,\ln(p_n)=\big[\ln\theta-\ln(1-\theta)\big]\times 1-\big[\ln\theta-\ln(1-\theta)\big]\times 1+\big[\ln(1-\theta)\big]\,E(X)-\big[\ln(1-\theta)\big]\,E(Y)=0$ car $E(X)=E(Y)$.

$H(Y)-H(X)=\displaystyle\sum_{n=1}^{+\infty}\left(q_n\,\ln\left(\frac{p_n}{q_n}\right)+(p_n-q_n)\,\ln p_n\right)$ et les séries de termes généraux $q_n\,\ln\left(\dfrac{p_n}{q_n}\right)$ et $(p_n-q_n)\,\ln p_n$ convergent, donc :

$$H(Y)-H(X)=\sum_{n=1}^{+\infty}q_n\,\ln\left(\frac{p_n}{q_n}\right)+\sum_{n=1}^{+\infty}(p_n-q_n)\,\ln p_n=\sum_{n=1}^{+\infty}q_n\,\ln\left(\frac{p_n}{q_n}\right)+0=\sum_{n=1}^{+\infty}q_n\,\ln\left(\frac{p_n}{q_n}\right).$$

Alors

$$\boxed{H(Y)-H(X)=\sum_{n=1}^{+\infty} q_n\,\ln\left(\frac{p_n}{q_n}\right).}$$

d) $\forall n\in\mathbb{N}^*,\ \ln\left(\dfrac{p_n}{q_n}\right)\leqslant \dfrac{p_n}{q_n}-1$ et $q_n>0$. Alors $\forall n\in\mathbb{N}^*,\ q_n\,\ln\left(\dfrac{p_n}{q_n}\right)\leqslant q_n\left(\dfrac{p_n}{q_n}-1\right)=p_n-q_n$.

Ainsi $H(Y)-H(X)=\displaystyle\sum_{n=1}^{+\infty} q_n\,\ln\left(\frac{p_n}{q_n}\right)\leqslant \sum_{n=1}^{+\infty} p_n-\sum_{n=1}^{+\infty} q_n$ car les trois séries convergent.

Or $\displaystyle\sum_{n=1}^{+\infty} p_n-\sum_{n=1}^{+\infty} q_n=1-1=0$ donc $H(Y)-H(X)\leqslant 0$.

$$\boxed{H(Y)-H(X) \text{ ou } H(Y)\leqslant H(X)\,!}$$

Affinons. Supposons que Y ne suivent pas la loi géométrique de paramètre θ. Alors il existe un élément m de $\mathbb{N}^*$ tel que $q_m\neq p_m$.

Donc $\dfrac{p_m}{q_m}\neq 1$ et ainsi $\ln\left(\dfrac{p_m}{q_m}\right)<\dfrac{p_m}{q_m}-1$. Or q_m est strictement positif donc $q_m\,\ln\left(\dfrac{p_m}{q_m}\right)<q_m\,\dfrac{p_m}{q_m}-q_m=p_m-q_m$.
$\forall n\in\mathbb{N}^*,\ \ln\left(\dfrac{p_n}{q_n}\right)\leqslant p_n-q_n$ et il existe un élément m de $\mathbb{N}^*$ tel que $q_m\,\ln\left(\dfrac{p_m}{q_m}\right)<p_m-q_m$.

Alors $H(Y)-H(X)=\displaystyle\sum_{n=1}^{+\infty} q_n\,\ln\left(\frac{p_n}{q_n}\right)<\sum_{n=1}^{+\infty}(p_n-q_n)=\sum_{n=1}^{+\infty} p_n-\sum_{n=1}^{+\infty} q_n=1-1=0.$

Donc $H(Y)<H(X)$

$$\boxed{\text{Si } Y \text{ ne suit pas la loi géométrique de paramètre } \theta : H(Y)<H(X).}$$

Moralité : les variables aléatoire à valeurs dans $\mathbb{N}^*$ et d'espérance α donnée ($\alpha\in]1,+\infty[$) d'entropie maximum sont LES variables aléatoires qui suivent la loi géométrique de d'espérance α. Sauf que nous l'avons pas vraiment montré ! !

▶ *Remarque* *En effet dans ce qui précède :* $\forall n\in\mathbb{N}^*,\ P(Y=n)=q_n>0$. *Offrons nous donc une question supplémentaire.*

e) Pour le fun... Nous allons dans cette question prendre un peu de hauteur.

1) Nous supprimerons l'hypothèse indiquant que la série de terme général $q_n\,\ln\left(\dfrac{p_n}{q_n}\right)$ converge car elle est redondante.

2) Nous ne supposerons plus que $\forall n\in\mathbb{N}^*,\ q_n\neq 0$.

3) Nous ne supposerons qu'à la fin que Y a même espérance que X. Cela nous permettra de retrouver l'existence de $H(Y)$ à partir de l'existence de $E(Y)$ (c'est la version 3 ou 2' annoncée).

4) Nous montrerons que les variables aléatoires à valeurs dans $\mathbb{N}^*$ et d'espérance α donnée ($\alpha\in]1,+\infty[$) d'entropie maximum sont LES variables aléatoires qui suivent la loi géométrique d'espérance α.

Ici X ne change pas. On part d'une variable aléatoire Y qui prend ses valeurs dans $\mathbb{N}^*$ et qui possède une espérance.

On pose toujours $\forall n\in\mathbb{N}^*,\ q_n=P(Y=n)$.

Montrons que $H(Y)$ existe. Il suffit de montrer que la série de terme général $|t(q_n)|$ converge.

Soit n un élément de $\mathbb{N}^*$ tel que $q_n \neq 0$.

$|t(q_n)| = -q_n \,\ln q_n = q_n \,\ln\left(\dfrac{1}{q_n}\right) = q_n \,\ln\left(\dfrac{p_n}{q_n}\right) - q_n \,\ln p_n$.

Or $\ln\left(\dfrac{p_n}{q_n}\right) \leqslant \dfrac{p_n}{q_n} - 1$, avec égalité si et seulement si $\dfrac{p_n}{q_n} = 1$, et $q_n > 0$ donc

$q_n \,\ln\left(\dfrac{p_n}{q_n}\right) \leqslant q_n \,\left(\dfrac{p_n}{q_n} - 1\right) = p_n - q_n$ avec égalité si et seulement si $p_n = q_n$.

Donc $|t(q_n)| \leqslant p_n - q_n - q_n \,\ln p_n$ avec égalité si et seulement si $p_n = q_n$.

Remarque *Arrêtons nous un instant pour observer que nous sommes sur les traces de la version 2.*

$|t(q_n)| \leqslant p_n - q_n - q_n \,\ln p_n \leqslant p_n - q_n \,\ln p_n$ *vaut finalement pour* p_n *réel strictement positif quelconque.*

Donc pour $p_n = \left(\dfrac{1}{e}\right)^n$ *ce qui donne* $|t(q_n)| \leqslant p_n - q_n \,\ln p_n = \left(\dfrac{1}{e}\right)^n + n\,q_n$*... Fin de la remarque.*

Soit n un élément de $\mathbb{N}^*$ tel que $q_n = 0$.

Alors $p_n \neq q_n$ (car $p_n > 0$), $|t(q_n)| = 0$ et $p_n - q_n - q_n \,\ln p_n = p_n > 0$.

Donc $p_n \neq q_n$ et $|t(q_n)| < p_n - q_n - q_n \,\ln p_n$.

Ainsi pour tout n dans $\mathbb{N}^*$, $\big|t(q_n)\big| \leqslant p_n - q_n - q_n \,\ln p_n$ avec égalité si et seulement si $p_n = q_n$. Or :

$\forall n \in \mathbb{N}^*,\ q_n \,\ln(p_n) = q_n \,\ln\big(\theta\,(1-\theta)^{n-1}\big) = \big[\ln\theta\big]\,q_n + (n-1)\,\big[\ln(1-\theta)\big]\,q_n = \big[\ln\theta - \ln(1-\theta)\big]\,q_n + \big[\ln(1-\theta)\big]\,n\,q_n$.

Alors pour tout n dans $\mathbb{N}^*$, $0 \leqslant |t(q_n)| \leqslant p_n - q_n - \big[\ln\theta - \ln(1-\theta)\big]\,q_n - \big[\ln(1-\theta)\big]\,n\,q_n$ avec égalité si et seulement si $p_n = q_n$.

Les séries de termes généraux p_n, q_n, $n\,q_n$ sont convergentes (Y possède une espérance) donc la séries de terme général $p_n - q_n - \big[\ln\theta - \ln(1-\theta)\big]\,q_n - \big[\ln(1-\theta)\big]\,n\,q_n$ converge comme combinaison linéaire de quatre séries convergentes.

Les règles de comparaison des séries à termes positifs montrent alors que la série de terme général $|t(q_n)|$ converge.

Ainsi $H(Y)$ existe.

Remarque *Nous venons de retrouver pour la troisième fois que si* $E(Y)$ *existe alors* $H(Y)$ *existe !*

De plus $H(Y) = \displaystyle\sum_{n=1}^{+\infty} \big|t(p_n)\big| \leqslant \sum_{n=1}^{+\infty} p_n - \sum_{n=1}^{+\infty} q_n - \big[\ln\theta - \ln(1-\theta)\big] \sum_{n=1}^{+\infty} q_n - \big[\ln(1-\theta)\big] \sum_{n=1}^{+\infty} (n\,q_n)$.

Alors : $H(Y) \leqslant 1 - 1 - \big[\ln\theta - \ln(1-\theta)\big] \times 1 - \big[\ln(1-\theta)\big]\,E(Y) = -\big[\ln\theta - \ln(1-\theta)\big] - \big[\ln(1-\theta)\big]\,E(Y)$.

Donc $H(Y) \leqslant -\big[\ln\theta - \ln(1-\theta)\big] - \big[\ln(1-\theta)\big]\,E(Y)$.

Supposons que Y ne suivent pas la même loi que X. Alors il existe m dans $\mathbb{N}^*$ tel que $p_m \neq q_m$. Donc :

1) $\forall n \in \mathbb{N}^*,\ |t(q_n)| \leqslant p_n - q_n - \big(\ln\theta - \ln(1-\theta)\big)\,q_n - \ln(1-\theta)\,n\,q_n$;

2) $\exists m \in \mathbb{N}^*,\ |t(q_m)| < p_m - q_m - \big(\ln\theta - \ln(1-\theta)\big)\,q_m - \ln(1-\theta)\,m\,q_m$.

Dans ces conditions :

$$H(Y) = \sum_{n=1}^{+\infty} |t(p_n)| < \sum_{n=1}^{+\infty} \Big(p_n - q_n - \big[\ln\theta - \ln(1-\theta)\big]\,q_n - \big[\ln(1-\theta)\big]\,n\,q_n\Big).$$

$$H(Y) < \sum_{n=1}^{+\infty} p_n - \sum_{n=1}^{+\infty} q_n - \big[\ln\theta - \ln(1-\theta)\big] \sum_{n=1}^{+\infty} q_n - \big[\ln(1-\theta)\big] \sum_{n=1}^{+\infty} (n\,q_n) = -\big[\ln\theta - \ln(1-\theta)\big] - \big[\ln(1-\theta)\big]\,E(Y).$$

Finalement :

1) $H(Y) \leqslant -\big[\ln\theta - \ln(1-\theta)\big] - \big[\ln(1-\theta)\big]\,E(Y)$;

2) si Y ne suit pas la même loi que X, $H(Y) < -\big[\ln\theta - \ln(1-\theta)\big] - \big[\ln(1-\theta)\big]\, E(Y)$.

Pour finir supposons maintenant que Y a même espérance que X. Alors $E(Y) = \dfrac{1}{\theta}\cdot$

Ainsi $-\big[\ln\theta - \ln(1-\theta)\big] - \big[\ln(1-\theta)\big]\, E(Y) = -\big[\ln\theta - \ln(1-\theta)\big] - \big[\ln(1-\theta)\big]\, \dfrac{1}{\theta} = -\dfrac{\theta\,\ln\theta - \theta\,\ln(1-\theta) + \ln(1-\theta)}{\theta}\cdot$

Ainsi $-\big[\ln\theta - \ln(1-\theta)\big] - \big[\ln(1-\theta)\big]\, E(Y) = -\dfrac{\theta\,\ln\theta + (1-\theta)\,\ln(1-\theta)}{\theta} = H(X)$.

Par conséquent $H(Y) \leqslant H(X)$ et si Y ne suit pas la même loi que X, $H(Y) < H(X)$. Plus de doute :

Les variables aléatoire à valeurs dans $\mathbb{N}^*$, d'espérance α donnée ($\alpha \in]1,+\infty[$) et d'entropie maximum sont LES variables aléatoires qui suivent la loi géométrique d'espérance α.

Partie IV. Entropie et taux de rendement asymptotique

11. a) Aucune raison de préférer Z_n à Z, alors montrons le lemme suivant.

Soit R une variable aléatoire sur $(\Omega, \mathcal{A}, P)$.

L'application S de Ω dans $\mathbb{R}$ définie par $\forall \omega \in \Omega,\ S(\omega) = e^{R(\omega)}$ est une variable aléatoire sur $(\Omega, \mathcal{A}, P)$.

Soit x un réel. Montrons que $S^{-1}(]-\infty, x])$ appartient à $\mathcal{A}$.

$S^{-1}(]-\infty, x]) = \{\omega \in \Omega \mid S(\omega) \leqslant x\} = \{\omega \in \Omega \mid e^{R(\omega)} \leqslant x\}$.

Si x est négatif ou nul $S^{-1}(]-\infty, x]) = \{\omega \in \Omega \mid e^{R(\omega)} \leqslant x\} = \emptyset$. Donc $S^{-1}(]-\infty, x])$ appartient à $\mathcal{A}$.

Supposons que x est strictement positif. Comme ln est **strictement** croissante sur $]0, +\infty[$:

$S^{-1}(]-\infty, x]) = \{\omega \in \Omega \mid e^{R(\omega)} \leqslant x\} = \{\omega \in \Omega \mid R(\omega) \leqslant \ln x\} = R^{-1}(]-\infty, \ln x])$.

Or $R^{-1}(]-\infty, \ln x])$ est dans $\mathcal{A}$ car R est une variable aléatoire sur $(\Omega, \mathcal{A}, P)$, donc $S^{-1}(]-\infty, x])$ appartient encore à $\mathcal{A}$.

Finalement $\forall x \in \mathbb{R},\ S^{-1}(]-\infty, x]) \in \mathcal{A}$ donc S est une variable aléatoire sur $(\Omega, \mathcal{A}, P)$.

Ceci achève la preuve du lemme. Plus de doute alors.

Pour tout n dans $\mathbb{N}^*$, Z_n est une variable aléatoire sur $(\Omega, \mathcal{A}, P)$.

Z est une variable aléatoire sur $(\Omega, \mathcal{A}, P)$.

b) Notons F_X la fonction de répartition de X. Soit x un réel strictement positif.

$P(|X| \geqslant x) = 1 - P(|X| < x) = 1 - P(-x < X < x) = 1 - \big(P(X < x) - P(X \leqslant -x)\big) = 1 - P(X < x) + P(X \leqslant -x)$.

L'événement $\{X \leqslant x - 1\}$ est contenu dans l'événement $\{X < x\}$.

Par croissance de la probabilité on obtient $P(X \leqslant x-1) \leqslant P(X < x)$ puis $-P(X < x) \leqslant -P(X \leqslant x - 1)$.

Alors $0 \leqslant P(|X| \geqslant x) = 1 - P(X < x) + P(X \leqslant -x) \leqslant 1 - P(X \leqslant x - 1) + P(X \leqslant -x) = 1 - F_X(x-1) + F_X(-x)$.

Pour tout réel x strictement positif : $0 \leqslant P(|X| \geqslant x) \leqslant 1 - F_X(x-1) + F_X(-x)$.

$\lim\limits_{z \to -\infty} F_X(z) = 0$ et $\lim\limits_{z \to +\infty} F_X(z) = 1$ donc $\lim\limits_{x \to +\infty} \big(1 - F_X(x-1) + F_X(-x)\big) = 1 - 1 + 0 = 0$.

Alors par encadrement on obtient : $\lim\limits_{x \to +\infty} P(|X| \geqslant x) = 0$.

En utilisant la définition de la limite d'une fonction on peut dire que :

$\exists A \in \mathbb{R}^{+*},\ \forall x \in]0, +\infty[,\ x > A \Rightarrow P(|X| \geqslant x) < \alpha$ car $\alpha > 0$.

Posons $s = A + 1$. $s > 0$ et $s > A$ donc $P(|X| \geqslant s) < \alpha$.

$$\boxed{\text{Il existe un réel } s \text{ strictement positif tel que } P(|X| \geqslant s) < \alpha.}$$

c) Soient K_1, K_2 et K_3 trois éléments de $\mathcal{A}$.

$P(K_1 \cup K_2 \cup K_3) = P(K_1 \cup K_2) + P(K_3) - P\big((K_1 \cup K_2) \cap K_3\big) \leqslant P(K_1 \cup K_2) + P(K_3)$.

$P(K_1 \cup K_2 \cup K_3) \leqslant P(K_1) + P(K_2) - P(K_1 \cap K_2) + P(K_3) \leqslant P(K_1) + P(K_2) + P(K_3)$.

$$\boxed{\text{Si } K_1,\ K_2 \text{ et } K_3 \text{ trois éléments de } \mathcal{A} : P(K_1 \cup K_2 \cup K_3) \leqslant P(K_1) + P(K_2) + P(K_3).}$$

Considérons les événements :

$K'_0 = \big\{|Z_n - Z| \geqslant \varepsilon\big\}$, $K'_1 = \big\{|X| \geqslant s\big\}$, $K'_2 = \big\{|X_n - X| \geqslant 1\big\}$ et $K'_3 = \big\{|X_n - X| \geqslant \varepsilon\, e^{-1-s}\big\}$.

Nous devons montrer que $P(K'_0) \leqslant P(K'_1) + P(K'_2) + P(K'_3)$.

D'après ce qui précède il suffit de montrer que $K'_0 \subset K'_1 \cup K'_2 \cup K'_3$, non ?

Soit ω un élément de K'_0. Montrons en raisonnant par l'absurde que $\omega \in K'_1 \cup K'_2 \cup K'_3$.

Supposons donc que ω n'appartient pas à $K'_1 \cup K'_2 \cup K'_3$. Alors ω n'appartient ni à K'_1, ni à K'_2, ni à K'_3.

Donc nous avons $|Z_n(\omega) - Z(\omega)| \geqslant \varepsilon$, $|X(\omega)| < s$, $|X_n(\omega) - X(\omega)| < 1$ et $|X_n(\omega) - X(\omega)| < \varepsilon\, e^{-1-s}$.

Établissons un résultat préliminaire. La fonction $v : x \to e^x$ est de classe $\mathcal{C}^1$ sur $\mathbb{R}$ et $v = v'$.

L'inégalité de Taylor-lagrange (!) appliquée à v à l'ordre 1 donne $\forall (x,y) \in \mathbb{R}^2,\ |v(x) - v(y)| \leqslant |x - y| \displaystyle\operatorname*{Max}_{z \in [x,y]} |v''(z)|$.

$\forall (x,y) \in \mathbb{R}^2,\ |e^x - e^y| \leqslant |x - y| \displaystyle\operatorname*{Max}_{z \in [x,y]} e^z = e^{\mathrm{Max}(x,y)}\, |x - y|$.

Alors $\varepsilon \leqslant |Z_n(\omega) - Z(\omega)| = |e^{X_n(\omega)} - e^{X(\omega)}| \leqslant e^{\mathrm{Max}\left(X_n(\omega), X_n(\omega)\right)}\, |X_n(\omega) - X(\omega)| < e^{\mathrm{Max}\left(X_n(\omega), X_n(\omega)\right)}\, \left(\varepsilon\, e^{-1-s}\right)$.

Donc $\varepsilon \leqslant |Z_n(\omega) - Z(\omega)| < \varepsilon\, e^{\mathrm{Max}\left(X_n(\omega), X_n(\omega)\right) - 1 - s}$.

$|X(\omega)| < s$, $|X_n(\omega) - X(\omega)| < 1$ (car ω n'appartient ni à K'_1 ni à K'_2). Donc $-s < X(\omega) < s$ et $-1 < X_n(\omega) - X_(\omega) < 1$.

Alors $X(\omega) < s < 1 + s$ et $X_n(\omega) < X(\omega) + 1 < 1 + s$.

Donc $\mathrm{Max}\left(X_n(\omega), X_n(\omega)\right) < 1 + s$ ou $\mathrm{Max}\left(X_n(\omega), X_n(\omega)\right) - 1 - s < 0$. Ce qui donne $e^{\mathrm{Max}\left(X_n(\omega), X_n(\omega)\right) - 1 - s} < 1$.

Ainsi $\varepsilon \leqslant |Z_n(\omega) - Z(\omega)| < \varepsilon\, e^{\mathrm{Max}\left(X_n(\omega), X_n(\omega)\right) - 1 - s} < \varepsilon$ (car ε est strictement positif) ! !

Cette légère contradiction montre que si ω appartient à K'_0, il appartient à $K'_1 \cup K'_2 \cup K'_3$ et achève de montrer que $K'_0 \subset K'_1 \cup K'_2 \cup K'_3$.

La croissance de P et le premier résultat du c) donnent : $P(K'_0) \leqslant P(K'_1 \cup K'_2 \cup K'_3) \leqslant P(K'_1) + P(K'_2) + P(K'_3)$.

Ainsi :

$$\boxed{P(|Z_n - Z| \geqslant \varepsilon) \leqslant P(|X| \geqslant s) + P(|X_n - X| \geqslant 1) + P(|X_n - X| \geqslant \varepsilon\, e^{-1-s}).}$$

d) On a encore $P(|Z_n - Z| \geqslant \varepsilon) < \alpha + P(|X_n - X| \geqslant 1) + P(|X_n - X| \geqslant \varepsilon\, e^{-1-s})$ grace à b).

La suite $(X_n)_{n \in \mathbb{N}^*}$ converge en probabilité vers X donc, pour tout réel γ strictement positif $\displaystyle\lim_{n \to +\infty} P(|X_n - X| \geqslant \gamma) = 0$.

Comme $1 > 0$ et $\varepsilon\, e^{-1-s} > 0$, $\displaystyle\lim_{n \to +\infty} P(|X_n - X| \geqslant 1) = \lim_{n \to +\infty} P(|X_n - X| \geqslant \varepsilon\, e^{-1-s}) = 0$.

Donc $\lim_{n\to+\infty}\Big(P(|X_n-X|\geqslant 1)+P(|X_n-X|\geqslant\varepsilon\,e^{-1-s})\Big)=0$.

La définition de la limite d'une suite permet de dire qu'il existe n_1 dans $\mathbb{N}^*$ tel que :

$\forall n\in\mathbb{N}^*,\ n\geqslant n_1\Rightarrow P(|X_n-X|\geqslant 1)+P(|X_n-X|\geqslant\varepsilon\,e^{-1-s})<\alpha$.

Donc $\forall n\in\mathbb{N}^*,\ n\geqslant n_1\Rightarrow P(|Z_n-Z|\geqslant\varepsilon)<\alpha+P(|X_n-X|\geqslant 1)+P(|X_n-X|\geqslant\varepsilon\,e^{-1-s})<2\,\alpha$.

Nous avons donc montré que : $\forall\varepsilon\in\mathbb{R}^{+*},\ \forall\alpha\in\mathbb{R}^{+*},\ \exists n_1\in\mathbb{N}^*,\ \forall n\in\mathbb{N}^*,\ n\geqslant n_1\Rightarrow P(|Z_n-Z|\geqslant\varepsilon)<2\,\alpha$.

Ceci suffit pour dire que $\forall\varepsilon\in\mathbb{R}^{+*},\ \lim_{n\to+\infty}P(|Z_n-Z|\geqslant\varepsilon)=0$ (car $\alpha\to 2\,\alpha$ définit une bijection de $\mathbb{R}^{+*}$ sur $\mathbb{R}^{+*}$!). Ainsi :

$$\boxed{\text{La suite } (Z_n)_{n\in\mathbb{N}^*} \text{ converge en probabilité ver } Z.}$$

Dans $(\Omega,\mathcal{A},P)$, si une suite $(X_n)_{n\in\mathbb{N}^*}$ de variables aléatoires réelles converge en probabilité vers la variable aléatoire X, $\left(e^{X_n}\right)_{n\in\mathbb{N}^*}$ est une suite de variables aléatoires qui converge vers la variable aléatoire e^X.

12. a) Notons qu'une récurrrence très simple montre que pour tout élément n de $\mathbb{N}^*$, R_n prend des valeurs strictement positives.

Soit n un élément de $[\![2,+\infty[\![$. Supposons que R_{n-1} ait pris la valeur a ($a>0$).

Pour tout k dans $[\![1,N]\!]$, le parieur mise une somme égale à $f_k\,a$ sur le cheval numéro k dans la $n^{\text{ème}}$ course.

Pour tout k dans $[\![1,N]\!]$, si G_n prend la valeur k, le parieur aura à l'issue de la $n^{\text{ème}}$ course une somme égale à $(f_k\,a)\,c_k$ et M_n prendra alors la valeur $f_k\,c_k$.

Nous pouvons alors dire que que $M_n=\displaystyle\sum_{k=1}^{N}\Big(f_k\,c_k\,\mathbb{1}_{\{G_n=k\}}\Big)$ ou que plus simplement $M_n=f_{G_n}\,c_{G_n}$.

Montrons que ceci vaut encore pour $n=1$.

Pour tout k dans $[\![1,N]\!]$, si G_1 prend la valeur k, le parieur aura à l'issue de la première course une somme égale à $(f_k\,r_0)\,c_k$ et M_1 prendra alors la valeur $f_k\,c_k$.

Donc, ici encore, $M_1=f_{G_1}\,c_{G_1}$ et $M_1=\displaystyle\sum_{k=1}^{N}\Big(f_k\,c_k\,\mathbb{1}_{\{G_1=k\}}\Big)$.

Finalement $\forall n\in\mathbb{N}^*,\ M_n=f_{G_n}\,c_{G_n}$ et $M_n=\displaystyle\sum_{k=1}^{N}\Big(f_k\,c_k\,\mathbb{1}_{\{G_n=k\}}\Big)$.

Pour tout n dans $\mathbb{N}^*$, M_n est une fonction de G_n. Comme $(G_n)_{n\in\mathbb{N}^*}$ est une suite de variables aléatoires indépendantes sur $(\Omega,\mathcal{A},P)$, $(M_n)_{n\in\mathbb{N}^*}$ est une suite de variables aléatoires indépendantes sur $(\Omega,\mathcal{A},P)$.

$\forall k\in[\![1,N]\!],\ f_k>0$ et $c_k>0$ donc, pour tout n dans $\mathbb{N}^*$, M_n prend des valeurs strictement positives.

Les variables aléatoires de la suite $(G_n)_{n\in\mathbb{N}^*}$ ayant même loi il en est de même pour les variables aléatoires de la suite $(M_n)_{n\in\mathbb{N}^*}$.

$$\boxed{(M_n)_{n\in\mathbb{N}^*} \text{ est une suites de variables aléatoires indépendantes, à valeurs strictement positives et de même loi.}}$$

b) $\forall n\in\mathbb{N}^*,\ \ln(T_n+1)=\ln\left(\dfrac{R_n}{r_0}\right)^{\frac{1}{n}}=\ln\left(\displaystyle\prod_{i=1}^{n}M_i\right)^{\frac{1}{n}}=\dfrac{1}{n}\displaystyle\sum_{i=1}^{n}\ln(M_i)$.

$(M_n)_{n\in\mathbb{N}^*}$ est une suite de variables aléatoires indépendantes sur $(\Omega,\mathcal{A},P)$ ayant même loi donc $\big(\ln(M_n)\big)_{n\in\mathbb{N}^*}$ est une suite de variables aléatoires indépendantes sur $(\Omega,\mathcal{A},P)$ ayant même loi.

De plus $\ln(M_1)$ possède une espérance et une variance. Alors les variables aléatoires de la suite $\big(\ln(M_n)\big)_{n\in\mathbb{N}^*}$ ont une espérance commune égale à $E\big(\ln(M_1)\big)$ et une variance commune.

La loi faible des grands nombres montre alors que la suite $\left(\dfrac{1}{n}\displaystyle\sum_{i=1}^{n}\ln M_i\right)_{n\in\mathbb{N}^*}$ converge en probabilité vers la variable certaine $E\big(\ln(M_1)\big)$.

Or $\forall n\in\mathbb{N}^*,\ \ln(T_n+1)=\dfrac{1}{n}\displaystyle\sum_{i=1}^{n}\ln M_i$ donc la suite $\big(\ln(T_n+1)\big)_{n\in\mathbb{N}^*}$ converge en probabilité vers la variable certaine $E\big(\ln(M_1)\big)$.

Alors **Q11.** montre que la suite $\big(e^{\ln(T_n+1)}\big)_{n\in\mathbb{N}^*}$ converge en probabilité vers la variable certaine $e^{E(\ln(M_1))}$.

Ainsi la suite $\big(T_n+1\big)_{n\in\mathbb{N}^*}$ converge en probabilité vers la variable certaine $e^{E(\ln(M_1))}$.

Alors $\forall\varepsilon\in\mathbb{R}^{+*},\ \displaystyle\lim_{n\to+\infty}P\left(\left|(T_n+1)-e^{E(\ln(M_1))}\right|\geqslant\varepsilon\right)=0$ ou $\forall\varepsilon\in\mathbb{R}^{+*},\ \displaystyle\lim_{n\to+\infty}P\left(\left|T_n-\big(e^{E(\ln(M_1))}-1\big)\right|\geqslant\varepsilon\right)=0$.

Plus de doute :

$$\boxed{\text{la suite de variables aléatoires } (T_n)_{n\in\mathbb{N}^*} \text{ converge en probabilité vers la variable aléatoire certaine égale à } \tau \text{ où } \tau=e^{E(\ln(M_1))}-1.}$$

Remarque Notons que $\ln(M_1)$ *ne prend qu'un nombre fini de valeurs donc* $\ln(M_1)$ *possède une espérance et une variance !*

13. a) Rappelons que $M_1=\displaystyle\sum_{k=1}^{N}\Big(f_k\,c_k\,\mathbb{1}_{\{G_1=k\}}\Big)$ donc $\ln(M_1)=\displaystyle\sum_{k=1}^{N}\Big(\ln(f_k\,c_k)\,\mathbb{1}_{\{G_1=k\}}\Big)$, non ?

Pour tout k dans $[\![1,N]\!]$, $\mathbb{1}_{\{G_1=k\}}$ possède une espérance qui vaut $P(G_1=k)$ donc p_k.

Il se confirme alors que $\ln(M_1)$ possède une espérance qui vaut $\displaystyle\sum_{k=1}^{N}\Big(\ln(f_k\,c_k)\,E\Big(\mathbb{1}_{\{G_1=k\}}\Big)\Big)$.

Donc $E\big(\ln(M_1)\big)=\displaystyle\sum_{k=1}^{N}\big(\ln(f_k\,c_k)\big)\,p_k=\sum_{k=1}^{N}p_k\,\ln(f_k\,c_k)$. Ainsi :

$$\boxed{\tau=e^{\sum_{k=1}^{N}p_k\,\ln(f_k\,c_k)}-1=\exp\left(\sum_{k=1}^{N}p_k\,\ln(f_k\,c_k)\right)-1.}$$

b) Il s'agit donc de trouver le maximum de $(f_1,f_2,\ldots,f_n)\to\exp\left(\displaystyle\sum_{k=1}^{N}p_k\,\ln(f_k\,c_k)\right)-1$ sous la contrainte $\displaystyle\sum_{k=1}^{N}f_k=1$... avec $(f_1,f_2,\ldots,f_N)$ dans $\big(]0,1[\big)^N$.

Utiliser la même méthode que celle de **Q8.** est trop long. Prenons un raccourci et henni soit qui mal y pense (hum, désolé pour cette saillie un peu cavalière, la fatigue sans doute).

$x \to e^x - 1$ est strictement croissante sur $\mathbb{R}$ donc trouver le maximum de $(f_1, f_2, \ldots, f_n) \to \exp\left(\sum_{k=1}^{N} p_k \ln(f_k\, c_k)\right) - 1$ sous la contrainte $\sum_{k=1}^{N} f_k = 1$ revient à trouver le maximum de $(f_1, f_2, \ldots, f_n) \to \sum_{k=1}^{N} p_k \ln(f_k\, c_k)$ sous la contrainte $\sum_{k=1}^{N} f_k = 1$.

Or $\sum_{k=1}^{N} p_k \ln(f_k\, c_k) = \sum_{k=1}^{N} p_k \ln(f_k) + \sum_{k=1}^{N} p_k \ln(c_k)$

Alors trouver le maximum de $(f_1, f_2, \ldots, f_n) \to \exp\left(\sum_{k=1}^{N} p_k \ln(f_k\, c_k)\right) - 1$ sous la contrainte $\sum_{k=1}^{N} f_k = 1$ revient à trouver le maximum de $(f_1, f_2, \ldots, f_n) \to \sum_{k=1}^{N} p_k \ln(f_k)$ sous la contrainte $\sum_{k=1}^{N} f_k = 1$.

Notons que l'ensemble des points qui réalisent le maximum (s'il existe !) de $(f_1, f_2, \ldots, f_n) \to \exp\left(\sum_{k=1}^{N} p_k \ln(f_k\, c_k)\right) - 1$ sous la contrainte $\sum_{k=1}^{N} f_k = 1$ coïncide avec l'ensemble des points qui réalisent le maximum de $(f_1, f_2, \ldots, f_n) \to \sum_{k=1}^{N} p_k \ln(f_k)$ sous la contrainte $\sum_{k=1}^{N} f_k = 1$.

Soit $(f_1, f_2, \ldots, f_N)$ un élément de $\big(]0,1[\big)^N$ tel que $\sum_{k=1}^{N} f_k = 1$.

Pour tout k dans $[\![1, N]\!]$, $p_k \ln(f_k) = p_k \ln\left(\frac{f_k}{p_k}\right) + p_k \ln(p_k) \leqslant p_k \left(\frac{f_k}{p_k} - 1\right) + p_k \ln p_k = f_k - p_k + p_k \ln p_k$ avec égalité si et seulement si $\frac{f_k}{p_k} = 1$ ou si et seulement si $f_k = p_k$ d'après **Q2.** car $p_k > 0$. Alors

$\sum_{k=1}^{N} p_k \ln(f_k) \leqslant \sum_{k=1}^{N} f_k - \sum_{k=1}^{N} p_k + \sum_{k=1}^{N} p_k \ln(p_k) = \sum_{k=1}^{N} p_k \ln(p_k)$ avec égalité si et seulement si $(f_1, f_2, \ldots, f_N) = (p_1, p_2, \ldots, p_N)$.

Rappelons que $\forall k \in [\![1, N]\!],\ p_k \in]0,1[$ et $\sum_{k=1}^{N} p_k = 1$.

Posons $\forall (f_1, f_2, \ldots, f_N) \in \big(]0,1[\big)^N,\ \mu(f_1, f_2, \ldots, f_N) = \sum_{k=1}^{N} p_k \ln(f_k)$ et notons que $\mu(p_1, p_2, \ldots, p_N) = \sum_{k=1}^{N} p_k \ln(p_k)$.

Ce qui précède montre que μ admet un maximum sous la contrainte $\sum_{k=1}^{N} f_k = 1$ et que $(p_1, p_2, \ldots, p_N)$ est le seul point qui le réalise.

Finalement il existe un élément $(f_1, f_2, \ldots, f_N)$ de $\big(]0,1[\big)^N$ et un seul vérifiant $\sum_{k=1}^{N} f_k = 1$ et rendant τ maximum. Cet élément est $(p_1, p_2, \ldots, p_N)$.

La stratégie optimale pour le parieur est de prendre f_k égal à p_k pour tout k dans $[\![1, N]\!]$.

La valeur optimale de τ est alors $\exp\left(\sum_{k=1}^{N} p_k \ln(p_k\, c_k)\right) - 1$.

c) Ici $\displaystyle\sum_{k=1}^{N} \frac{1}{c_k} = 1$.

Si $k \in [\![1, N]\!]$, $-p_k \ \ln(p_k\, c_k) = p_k \ \ln \dfrac{1}{p_k\, c_k} \leqslant p_k \left(\dfrac{1}{p_k\, c_k} - 1\right) = \dfrac{1}{c_k} - p_k$ avec égalité si et seulement si $\dfrac{1}{p_k\, c_k} = 1$ car $p_k > 0$.

Alors $\displaystyle -\sum_{k=1}^{N} p_k \ \ln(p_k\, c_k) \leqslant \sum_{k=1}^{N} \frac{1}{c_k} - \sum_{k=1}^{N} p_k = 1 - 1 = 0$ avec égalité si et seulement si $\forall k \in [\![1, N]\!] \ \ p_k = \dfrac{1}{c_k}\cdot$ Donc :

$\displaystyle\sum_{k=1}^{N} p_k \ \ln(p_k\, c_k) \geqslant 0$ avec égalité si et seulement si $\forall k \in [\![1, N]\!] \ \ p_k = \dfrac{1}{c_k}$ (ce qui est possible car $\forall k \in [\![1, N]\!]$, $c_k > 1$ et $\displaystyle\sum_{k=1}^{N} \frac{1}{c_k} = 1$).

$$\boxed{\sum_{k=1}^{N} p_k \ \ln(p_k\, c_k) \geqslant 0. \text{ De plus } \sum_{k=1}^{N} p_k \ \ln(p_k\, c_k) = 0 \text{ si et seulement si } \forall k \in [\![1, N]\!] \ \ p_k = \frac{1}{c_k}\cdot}$$

Rappelons que la valeur maximale de τ est $= e^{\sum_{k=1}^{N} p_k \ \ln(p_k\, c_k)} - 1$.

Cette valeur est nulle si $(p_1, p_2, \ldots, p_N) = \left(\dfrac{1}{c_1}, \dfrac{1}{c_2}, \ldots, \dfrac{1}{c_N}\right)$ et strictement positive dans le cas contraire.

$$\boxed{\text{Si } \sum_{k=1}^{N} \frac{1}{c_k} = 1, \text{ le parieur ne dispose d'aucune stratégie lui permettant de s'assurer un taux de rendement asymptotique optimal strictement positif si et seulement si } (p_1, p_2, \ldots, p_N) = \left(\frac{1}{c_1}, \frac{1}{c_2}, \ldots, \frac{1}{c_N}\right).}$$
