ESSEC II 2019

Un modèle probabiliste d'une expérience aléatoire représente dans un certain sens le désordre qui intervient dans l'expérience et il est donc naturel que des outils soient introduits qui permettent de mesurer l'intensité de ce désordre. C'est le cas de la notion d'entropie qui fait l'objet du présent problème. On considèrera différentes situations et notamment la façon dont on mesure l'information que deux variables aléatoires s'apportent mutuellement.
Dans la première partie on étudie le cas plus simple techniquement de variables dont la loi admet une densité. Les deuxièmes et troisièmes parties sont consacrées au cas discret. Dans la deuxième partie, on introduit les différentes notions d'entropie pour le cas de variables discrètes et dans la troisième partie, on examine comment on peut mesurer l'information apportée mutuellement par deux variables aléatoires.
Toutes les variables aléatoires intervenant dans le problème sont définies sur un espace probabilisé $(\Omega, \mathscr{A}, \mathbb{P})$.
Pour toute variable aléatoire Y, on notera $\mathbb{E}(Y)$ son espérance lorsqu'elle existe.

Première partie : Entropie différentielle d'une variable à densité

1. La fonction logarithme de base 2, notée $\log_2$, est définie sur $\mathbb{R}_+^*$ par : $\log_2(x) = \dfrac{\ln(x)}{\ln(2)}$.

a) Montrer que pour tout (x, y) élément de $\mathbb{R}_+^* \times \mathbb{R}_+^*$, on a : $\log_2(x\,y) = \log_2(x) + \log_2(y)$.

Démonstration.
Soit $(x, y) \in \mathbb{R}_+^* \times \mathbb{R}_+^*$.

$$\begin{array}{rcll} \log_2(x\,y) & = & \dfrac{\ln(x\,y)}{\ln(2)} & \\ & = & \dfrac{\ln(x)+\ln(y)}{\ln(2)} & \textit{(par propriété de la fonction } \ln\textit{)} \\ & = & \dfrac{\ln(x)}{\ln(2)} + \dfrac{\ln(y)}{\ln(2)} & \\ & = & \log_2(x) + \log_2(y) & \end{array}$$

$$\boxed{\forall (x, y) \in \mathbb{R}_+^* \times \mathbb{R}_+^*,\ \log_2(x\,y) = \log_2(x) + \log_2(y)}$$

□

b) Vérifier que pour tout réel α : $\log_2(2^\alpha) = \alpha$.

Démonstration.
Soit $\alpha \in \mathbb{R}$.

$$\log_2(2^\alpha) = \frac{\ln(2^\alpha)}{\ln(2)} = \frac{\alpha \cancel{\ln(2)}}{\cancel{\ln(2)}} = \alpha$$

$$\boxed{\forall \alpha \in \mathbb{R},\ \log_2(2^\alpha) = \alpha}$$

□

c) Montrer que la fonction $\log_2$ est concave sur $\mathbb{R}_+^*$.

Démonstration.
La fonction $\log_2$ est de classe $\mathcal{C}^2$ sur $\mathbb{R}_+^*$ car la fonction ln l'est.
Soit $x \in \mathbb{R}_+^*$.

$$(\log_2)'(x) = \frac{1}{x\ \ln(2)}$$

$$(\log_2)''(x) = -\frac{1}{x^2\ \ln(2)} < 0$$

$$\boxed{\text{On en déduit que la fonction } \log_2 \text{ est concave sur } \mathbb{R}_+^*.}$$

Commentaire

On pouvait répondre à cette question grâce à d'autres caractérisations de la concavité.

1. La fonction $\log_2$ est de classe $\mathcal{C}^1$ sur $\mathbb{R}_+^*$ car la fonction ln l'est.
Soit $x \in \mathbb{R}_+^*$.

$$(\log_2)'(x) = \frac{1}{x\ \ln(2)}$$

La fonction $(\log_2)'$ est donc décroissante sur $\mathbb{R}_+^*$.
On en déduit que la fonction $\log_2$ est concave sur $\mathbb{R}_+^*$.

2.
- Tout d'abord :

$$\text{La fonction } \log_2 \text{ est concave sur } \mathbb{R}_+^*$$
$$\Updownarrow$$
$$\forall \lambda \in [0,1],\ \forall (x,y) \in \mathbb{R}_+^* \times \mathbb{R}_+^*,\ \log_2\big(\lambda\, x + (1-\lambda) y\big) \geqslant \lambda\, \log_2(x) + (1-\lambda)\log_2(y)$$

- Soit $\lambda \in [0,1]$. Soit $(x,y) \in \mathbb{R}_+^* \times \mathbb{R}_+^*$.

$$\begin{array}{rrcl} & \log_2\big(\lambda\, x + (1-\lambda) y\big) & \geqslant & \lambda\, \log_2(x) + (1-\lambda)\log_2(y) \\ \Leftrightarrow & \dfrac{\ln\big(\lambda\, x + (1-\lambda) y\big)}{\ln(2)} & \geqslant & \lambda\, \dfrac{\ln(x)}{\ln(2)} + (1-\lambda)\dfrac{\ln(y)}{\ln(2)} \\ \Leftrightarrow & \ln\big(\lambda\, x + (1-\lambda) y\big) & \geqslant & \lambda\, \ln(x) + (1-\lambda)\ln(y) \qquad \textit{(car } \ln(2) > 0\textit{)} \end{array}$$

Cette dernière inégalité est vraie car la fonction ln est concave sur $\mathbb{R}_+^*$. Par équivalence, la première est vraie.

On en déduit que la fonction $\log_2$ est concave sur $\mathbb{R}_+^*$.

□

2. Soit X une variable aléatoire réelle à densité, et soit f une densité de X. On appelle **support** de f l'ensemble $I = \{x \in \mathbb{R} \mid f(x) > 0\}$, et on suppose que I est un intervalle de $\mathbb{R}$ d'extrémités a et b ($a < b$, a et b finis ou infinis). L'**entropie différentielle** de X est, sous réserve d'existence, le réel :

$$h(X) \ = \ -\int_a^b \ f(x) \ \log_2\big(f(x)\big) \ dx$$

Montrer : $h(X) = -\mathbb{E}\Big(\log_2\big(f(X)\big)\Big)$.

Démonstration.

- Par théorème de transfert, la v.a.r. $\log_2\big(f(X)\big)$ admet une espérance si et seulement si l'intégrale $\displaystyle\int_{-\infty}^{+\infty} \log_2\big(f(x)\big) \ f(x) \ dx$ est absolument convergente.
- De plus, par définition du support, la fonction f est nulle en dehors de l'intervalle I d'extrémités a et b, donc :

$$\int_{-\infty}^{+\infty} \ f(x) \ \log_2\big(f(x)\big) \ dx \ = \ \int_a^b \ f(x) \ \log_2\big(f(x)\big) \ dx$$

$$\boxed{\text{Ainsi, sous réserve de convergence absolue :} \quad h(X) \ = \ -\int_a^b \ f(x) \ \log_2\big(f(x)\big) \ dx \ = \ -\mathbb{E}\Big(\log_2\big(f(x)\big)\Big).}$$

Commentaire

L'énoncé précise que I est un intervalle de $\mathbb{R}$ **d'extrémités** a et b. C'est donc l'un des intervalles suivants : $]a, b[$, $]a, b]$, $[a, b[$ ou $[a, b]$.

□

3. Soit X une variable aléatoire de densité f de support I, intervalle de $\mathbb{R}$ d'extrémités a et b. On suppose que X admet une entropie différentielle.

a) Soit c un réel, et soit Y la variable aléatoire définie par $Y = c + X$.

(i) Déterminer une densité de Y.

Démonstration.

- Déterminons la fonction de répartition de la v.a.r. Y.
  Soit $x \in \mathbb{R}$.

$$F_Y(x) \ = \ \mathbb{P}([Y \leqslant x]) \ = \ \mathbb{P}([c + X \leqslant x]) \ = \ \mathbb{P}([X \leqslant x - c]) \ = \ F_X(x - c)$$

$$\boxed{\text{Finalement : } F_Y : x \mapsto F_X(x - c).}$$

- La fonction F_Y est continue sur $\mathbb{R}$ car elle est la composée $F_X \circ g_1$ de :
  × $g_1 : x \mapsto x - c$ qui :
    - est continue sur $\mathbb{R}$,
    - vérifie : $g_1(\mathbb{R}) \subset \mathbb{R}$.
  × F_X qui est continue sur $\mathbb{R}$, car X est une v.a.r. à densité.
- La fonction F_Y est de classe $\mathcal{C}^1$ sur $\mathbb{R}$ sauf (éventuellement) en un nombre fini de points avec des arguments similaires à ceux de la continuité sur $\mathbb{R}$.

$$\boxed{\text{On en déduit que } Y \text{ est une v.a.r. à densité.}}$$

Commentaire

- Si F_X est de classe $\mathcal{C}^1$ sur $\mathbb{R}$, alors la fonction F_Y est également de classe $\mathcal{C}^1$ sur $\mathbb{R}$ en tant composée $F_X \circ g_1$ (même démonstration que la continuité de F_Y sur $\mathbb{R}$).
- Si F_X n'est pas de classe $\mathcal{C}^1$ sur $\mathbb{R}$, on peut déterminer plus précisément les points où la fonction F_Y n'est pas de classe $\mathcal{C}^1$.
  - × Comme X est une v.a.r. à densité, sa fonction de répartition F_X est de classe $\mathcal{C}^1$ sur $\mathbb{R}$ sauf en un nombre fini de points. Notons $\{x_1, \ldots, x_n\}$ l'ensemble des points où la fonction F_X n'est pas de classe $\mathcal{C}^1$.
  - × Comme $F_Y : x \mapsto F_X(x - c)$, on en déduit que la fonction F_Y est de classe $\mathcal{C}^1$ sur $\mathbb{R}$ sauf en $x_1 + c$, ..., $x_n + c$.

- Pour déterminer une densité f_Y de Y, on dérive la fonction F_Y en les points où elle est de classe $\mathcal{C}^1$. Soit $x \in \mathbb{R}$.
  - × Si F_X est de classe $\mathcal{C}^1$ en x.
  $$f_Y(x) = F'_Y(x) = (F_X \circ g_1)'(x) = (F'_X \circ g_1)(x) \times g'_1(x) = F'_X\big(g_1(x)\big) \times 1 = F'_X(x-c) = f(x-c)$$
  - × Si F_X n'est pas de classe $\mathcal{C}^1$ en x, on choisit : $f_Y(x) = f(x - c)$.

$$\boxed{\text{Finalement : } f_Y : x \mapsto f(x - c).}$$

Commentaire

Afin de permettre une bonne compréhension du calcul, on a rédigé ici en détails la dérivation de la composée. Cependant, fournir le résultat correct permet sans aucun doute d'obtenir la totalité des points alloués à ce calcul.

□

(ii) Justifier l'existence de l'entropie différentielle $h(Y)$, et la déterminer en fonction de $h(X)$.

Démonstration.

- D'après la question ***2.***, la v.a.r. Y admet une entropie différentielle si et seulement si la v.a.r. $\log_2\big(f_Y(Y)\big)$ admet une espérance.
- On remarque de plus :

$$\begin{aligned} \log_2\big(f_Y(Y)\big) &= \log_2\big(f(Y-c)\big) && \textit{(d'après la question précédente)} \\ &= \log_2\Big(f\big((X+\cancel{c}) - \cancel{c}\big)\Big) && \textit{(par définition de la v.a.r. } Y\textit{)} \\ &= \log_2\big(f(X)\big) \end{aligned}$$

- Or, d'après l'énoncé, la v.a.r. X admet une entropie différentielle. Donc la v.a.r. $\log_2\big(f(X)\big)$ admet une espérance.
  On en déduit que la v.a.r. $\log_2\big(f_Y(Y)\big)$ admet une espérance.

$$\boxed{\text{Ainsi, la v.a.r. } Y \text{ admet une entropie différentielle.}}$$

- De plus, toujours d'après la question ***2.*** :

$$h(Y) = -\mathbb{E}\Big(\log_2\big(f_Y(Y)\big)\Big) = -\mathbb{E}\Big(\log_2\big(f(X)\big)\Big) = h(X)$$

$$\boxed{h(Y) = h(X)}$$

□

b) Soit α un réel strictement positif, et soit Z la variable aléatoire définie par $Z = \alpha\, X$.

(i) Déterminer une densité de Z.

Démonstration.

- Déterminons la fonction de répartition de la v.a.r. Z.
  Soit $x \in \mathbb{R}$.

$$\begin{aligned} F_Z(x) &= \mathbb{P}([Z \leqslant x]) \ = \ \mathbb{P}([\alpha\, X \leqslant x]) \\ &= \mathbb{P}\left(\left[X \leqslant \frac{x}{\alpha}\right]\right) && \textit{(car } \alpha > 0\textit{)} \\ &= F_X\left(\frac{x}{\alpha}\right) \end{aligned}$$

$$\boxed{\text{Finalement : } F_Z : x \mapsto F_X\left(\frac{x}{\alpha}\right).}$$

- La fonction F_Z est continue sur $\mathbb{R}$ car elle est la composée $F_X \circ g_2$ de :
  × $g_2 : x \mapsto \dfrac{x}{\alpha}$ qui :
    - est continue sur $\mathbb{R}$,
    - vérifie : $g_2(\mathbb{R}) \subset \mathbb{R}$.

  × F_X qui est continue sur $\mathbb{R}$, car X est une v.a.r. à densité.
- La fonction F_Z est de classe $\mathcal{C}^1$ sur $\mathbb{R}$ sauf (éventuellement) en un nombre fini de points avec des arguments similaires à ceux de la continuité sur $\mathbb{R}$.

Commentaire

- Si F_X est de classe $\mathcal{C}^1$ sur $\mathbb{R}$, alors la fonction F_Z est également de classe $\mathcal{C}^1$ sur $\mathbb{R}$ en tant que composée $F_X \circ g_2$ (même démonstration que la continuité de F_Z sur $\mathbb{R}$).
- Si F_X n'est pas de classe $\mathcal{C}^1$ sur $\mathbb{R}$,
  × on note comme précédemment $\{x_1, \ldots, x_n\}$ l'ensemble des points où la fonction F_X n'est pas de classe $\mathcal{C}^1$,
  × comme $F_Z : x \mapsto F_X\left(\dfrac{x}{\alpha}\right)$, on en déduit que la fonction F_Z est de classe $\mathcal{C}^1$ sur $\mathbb{R}$ sauf en $\alpha\, x_1$, ..., $\alpha\, x_n$.

- Pour déterminer une densité f_Z de Z, on dérive la fonction F_Z en les points où elle est de classe $\mathcal{C}^1$. Soit $x \in \mathbb{R}$.
  × Si F_X est de classe $\mathcal{C}^1$ en x.

$$f_Z(x) = F_Z'(x) = (F_X \circ g_2)'(x) = (F_X' \circ g_2)(x) \times g_2'(x) = F_X'\big(g_2(x)\big) \times \frac{1}{\alpha} = \frac{1}{\alpha}\, F_X'\left(\frac{x}{\alpha}\right) = \frac{1}{\alpha}\, f\left(\frac{x}{\alpha}\right)$$

  × Si F_X n'est pas de classe $\mathcal{C}^1$ en x, on choisit : $f_Z(x) = \dfrac{1}{\alpha}\, f\left(\dfrac{x}{\alpha}\right)$.

$$\boxed{\text{Finalement : } f_Z : x \mapsto \frac{1}{\alpha}\, f\left(\frac{x}{\alpha}\right).}$$

□

(ii) Justifier l'existence de l'entropie différentielle $h(Z)$, et la déterminer en fonction de $h(X)$.

Démonstration.

- D'après la question ***2.***, la v.a.r. Y admet une entropie différentielle si et seulement si la v.a.r. $\log_2\big(f_Z(Z)\big)$ admet une espérance.
- On remarque de plus :

$$\begin{array}{rcll} \log_2\big(f_Z(Z)\big) & = & \log_2\left(\dfrac{1}{\alpha}\, f\left(\dfrac{X}{\alpha}\right)\right) & \textit{(d'après la question précédente)} \\ & = & \log_2\left(\dfrac{1}{\alpha}\, f\left(\dfrac{\cancel{\alpha}\, X}{\cancel{\alpha}}\right)\right) & \textit{(par définition de la v.a.r. } Y\textit{)} \\ & = & \log_2\left(\dfrac{1}{\alpha}\, f(X)\right) & \\ & = & \log_2\left(\dfrac{1}{\alpha}\right) + \log_2\big(f(X)\big) & \textit{(d'après la question } \textbf{1.a)}\textit{)} \end{array}$$

  Enfin :

$$\log_2\left(\dfrac{1}{\alpha}\right) = \dfrac{\ln\left(\frac{1}{\alpha}\right)}{\ln(2)} = \dfrac{-\ln(\alpha)}{\ln(2)} = -\log_2(\alpha)$$

$$\boxed{\text{Ainsi : } \log_2\big(f_Z(Z)\big) = \log_2\big(f(X)\big) - \log_2(\alpha).}$$

- Or, d'après l'énoncé, la v.a.r. X admet une entropie différentielle. Donc la v.a.r. $\log_2\big(f(X)\big)$ admet une espérance.
  On en déduit que la v.a.r. $\log_2\big(f_Z(Z)\big)$ admet une espérance en tant que transformée affine d'une v.a.r. qui en admet une.

$$\boxed{\text{Ainsi, la v.a.r. } Z \text{ admet une entropie différentielle.}}$$

- De plus, toujours d'après la question ***2.*** :

$$\begin{array}{rcll} h(Z) & = & -\mathbb{E}\Big(\log_2\big(f_Z(Z)\big)\Big) & \\ & = & -\mathbb{E}\Big(\log_2\big(f(X)\big) - \log_2(\alpha)\Big) & \\ & = & -\Big(\mathbb{E}\big(\log_2\big(f(X)\big)\big) - \log_2(\alpha)\Big) & \textit{(par linéarité de l'espérance)} \\ & = & h(X) + \log_2(\alpha) & \end{array}$$

$$\boxed{h(Z) = h(X) + \log_2(\alpha)}$$

□

4. On détermine dans cette question l'entropie différentielle de quelques variables aléatoires suivant des lois classiques.

a) Soit $a > 0$. On considère X une variable aléatoire de loi uniforme sur $[0, a]$.

(i) Donner une densité de X.

Démonstration.

$$\boxed{\text{Une densité } f \text{ de } X \text{ est : } f : x \mapsto \begin{cases} 0 & \text{si } x \in \,]-\infty, 0[\\ \dfrac{1}{a} & \text{si } x \in [0, a] \\ 0 & \text{si } x \in \,]a, +\infty[\end{cases}}$$

□

(ii) Justifier l'existence de l'entropie différentielle $h(X)$, et la déterminer.

Démonstration.

- On commence par remarquer que, par définition de f, son support I est l'intervalle $[0, a]$. Alors, d'après la définition de l'entropie, X admet une entropie différentielle si et seulement si l'intégrale $\displaystyle\int_0^a f(x) \log_2 \big(f(x)\big)\, dx$ est convergente.
- Or la fonction $x \mapsto f(x) \log_2 \big(f(x)\big)$ est continue par morceaux sur le segment $[0, a]$. On en déduit que l'intégrale $\displaystyle\int_0^a f(x) \log_2 \big(f(x)\big)\, dx$ converge.

$$\boxed{\text{Ainsi, la v.a.r. } X \text{ admet une entropie différentielle.}}$$

- De plus :

$$\begin{aligned}
h(X) &= -\int_0^a f(x) \log_2 \big(f(x)\big)\, dx \\
&= -\int_0^a \frac{1}{a} \log_2\left(\frac{1}{a}\right) dx && \textit{(par définition de } f \textit{ sur } [0,a]\textit{)} \\
&= \frac{1}{a} \log_2(a) \int_0^a dx \\
&= \frac{1}{a} \log_2(a)\ \big[\ x\ \big]_0^a \\
&= \frac{1}{a} \log_2(a)\, (a-0) \\
&= \log_2(a)
\end{aligned}$$

$$\boxed{h(X) = \log_2(a)}$$

□

(iii) Déterminer une condition nécessaire et suffisante sur a pour que $h(X) > 0$.

Démonstration.
D'après la question précédente :

$$\begin{aligned}
h(X) > 0 &\Leftrightarrow \log_2(a) > 0 \\
&\Leftrightarrow \frac{\ln(a)}{\ln(2)} > 0 \\
&\Leftrightarrow \ln(a) > 0 && \textit{(car } \ln(2) > 0\textit{)} \\
&\Leftrightarrow a > 1 && \textit{(par stricte croissance de la fonction } \exp \textit{ sur } \mathbb{R}\textit{)}
\end{aligned}$$

$$\boxed{h(X) > 0 \ \Leftrightarrow\ a > 1}$$

□

b) On considère Y une variable aléatoire de loi normale centrée réduite. Montrer que Y admet une entropie différentielle et : $h(Y) = \dfrac{1}{2}\ \log_2(2\pi e)$.

Démonstration.

- Une densité de Y est : $f_Y : x \mapsto \dfrac{1}{\sqrt{2\pi}}\ e^{-\frac{x^2}{2}}$.
  On en déduit que le support de f_Y est l'intervalle $I =]-\infty, +\infty[$.
  Alors, la v.a.r. Y admet une entropie différentielle si et seulement si l'intégrale $\displaystyle\int_{-\infty}^{+\infty} f_Y(x)\ \log_2\big(f_Y(x)\big)\ dx$ est convergente.
- Soit $x \in \mathbb{R}$.

$$\begin{array}{rcll}
\log_2\big(f_Y(x)\big) & = & \log_2\left(\dfrac{1}{\sqrt{2\pi}}\ e^{-\frac{x^2}{2}}\right) & \\[2mm]
& = & \log_2\left(\dfrac{1}{\sqrt{2\pi}}\right) + \log_2\left(e^{-\frac{x^2}{2}}\right) & \textit{(d'après la question } \textbf{\textit{1.a)}}\textit{)} \\[2mm]
& = & \dfrac{\ln\left(\frac{1}{\sqrt{2\pi}}\right)}{\ln(2)} + \dfrac{\ln\left(e^{-\frac{x^2}{2}}\right)}{\ln(2)} & \\[2mm]
& = & \dfrac{-\frac{1}{2}\ \ln(2\pi)}{\ln(2)} + \dfrac{-\frac{x^2}{2}}{\ln(2)} & \\[2mm]
& = & -\dfrac{1}{2}\ \dfrac{\ln(2\pi)}{\ln(2)} - \dfrac{1}{\ln(2)}\ \dfrac{x^2}{2} & \\[2mm]
& = & -\dfrac{1}{2}\ \log_2(2\pi) - \dfrac{1}{2\ln(2)}\ x^2 &
\end{array}$$

  Ainsi :

$$\begin{array}{rcl}
f_Y(x)\ \log_2\big(f_Y(x)\big) & = & f_Y(x)\left(-\dfrac{1}{2}\ \log_2(2\pi) - \dfrac{1}{2\ln(2)}\ x^2\right) \\[2mm]
& = & -\dfrac{1}{2}\ \log_2(2\pi)\ f_Y(x) - \dfrac{1}{2\ln(2)}\ x^2\ f_Y(x)
\end{array}$$

- Or,
  × l'intégrale $\displaystyle\int_{-\infty}^{+\infty} f_Y(x)\ dx$ est convergente car f_Y est une densité. De plus :
  $$\int_{-\infty}^{+\infty} f_Y(x)\ dx \ = \ 1$$
  × l'intégrale $\displaystyle\int_{-\infty}^{+\infty} x^2\ f_Y(x)\ dx$ est convergente car la v.a.r. Y admet un moment d'ordre 2. De plus :
  $$\int_{-\infty}^{+\infty} x^2\ f_Y(x)\ dx \ = \ \mathbb{E}(Y^2)$$
  D'après la formule de Koenig-Huygens : $\mathbb{V}(Y) \ = \ \mathbb{E}(Y^2) - \big(\mathbb{E}(Y)\big)^2$. Ainsi :
  $$\int_{-\infty}^{+\infty} x^2\ f_Y(x)\ dx \ = \ \mathbb{E}(Y^2) \ = \ \mathbb{V}(X) + \big(\mathbb{E}(Y)\big)^2 \ = \ 1 - 0 \ = \ 1$$

× On en déduit que l'intégrale $\displaystyle\int_{-\infty}^{+\infty} f_Y(x)\ \log_2\big(f_Y(x)\big)\ dx$ est convergente en tant que combinaison linéaire d'intégrales convergentes.

$\boxed{\text{Ainsi, la v.a.r. } Y \text{ admet une entropie différentielle.}}$

× De plus :

$$\begin{aligned}
& h(Y) \\
= \ & -\int_{-\infty}^{+\infty} f_Y(x)\ \log_2\big(f_Y(x)\big)\ dx \\
= \ & -\int_{-\infty}^{+\infty} -\frac{1}{2}\ \log_2(2\pi)\, f_Y(x) - \frac{1}{2\ln(2)} x^2\ f_Y(x)\ dx && \textit{(d'après les calculs précédents)} \\
= \ & \frac{1}{2}\ \log_2(2\pi) \int_{-\infty}^{+\infty} f_Y(x)\ dx + \frac{1}{2\ln(2)} \int_{-\infty}^{+\infty} x^2\ f_Y(x)\ dx && \textit{(par linéarité de l'intégrale, car les intégrales en présence sont convergentes)} \\
= \ & \frac{1}{2}\ \log_2(2\pi)\times 1 + \frac{1}{2\ln(2)}\times 1 \\
= \ & \frac{1}{2}\left(\log_2(2\pi) + \frac{1}{\ln(2)}\right) \\
= \ & \frac{1}{2}\left(\log_2(2\pi) + \frac{\ln(\mathrm{e})}{\ln(2)}\right) \\
= \ & \frac{1}{2}\big(\log_2(2\pi) + \log_2(\mathrm{e})\big)
\end{aligned}$$

$\boxed{\text{D'après la question } \textbf{\textit{1.a)}}\text{, on obtient : } h(Y) = \frac{1}{2}\ \log_2(2\pi\mathrm{e}).}$

Commentaire

On a en particulier démontré dans nos calculs l'égalité suivante :

$$\frac{1}{\ln(\mathrm{e})} = \log_2(\mathrm{e})$$

□

c) On considère Z une variable aléatoire de loi exponentielle de paramètre λ ($\lambda > 0$). Justifier l'existence de l'entropie différentielle $h(Z)$ et la déterminer.

Démonstration.

- Une densité de Z est :

$$f_Z : x \mapsto \begin{cases} 0 & \text{si } x \in\ ]-\infty, 0[\\ \lambda\,\mathrm{e}^{-\lambda x} & \text{si } x \in [0, +\infty[\end{cases}$$

  On en déduit que le support de f_Z est l'intervalle $I = [0, +\infty[$.
  Alors, la v.a.r. Z admet une entropie différentielle si et seulement si l'intégrale $\displaystyle\int_0^{+\infty} f_Z(x)\ \log_2\big(f_Z(x)\big)\ dx$ est convergente.

- Soit $x \in [0, +\infty[$.

$$\begin{array}{rcll} \log_2\big(f_Z(x)\big) & = & \log_2\big(\lambda\, \mathrm{e}^{-\lambda\, x}\big) & \\ & = & \log_2(\lambda) + \log_2\big(\mathrm{e}^{-\lambda\, x}\big) & \textit{(d'après la question } \textbf{\textit{1.a)}}\textit{)} \\ & = & \log_2(\lambda) + \dfrac{\ln\big(\mathrm{e}^{-\lambda\, x}\big)}{\ln(2)} & \\ & = & \log_2(\lambda) + \dfrac{-\lambda\, x}{\ln(2)} & \\ & = & \log_2(\lambda) - \dfrac{\lambda}{\ln(2)}\ x & \end{array}$$

  Ainsi :

$$\begin{array}{rcl} f_Z(x)\ \log_2\big(f_Z(x)\big) & = & f_Y(x)\left(\log_2(\lambda) - \dfrac{\lambda}{\ln(2)}\ x\right) \\ & = & \log_2(\lambda)\ f_Z(x) - \dfrac{\lambda}{\ln(2)}\ x\, f_Z(x) \end{array}$$

- Or,
  - × l'intégrale $\displaystyle\int_0^{+\infty} f_Z(x)\ dx$ est convergente car f_Z est une densité. De plus, comme f_Z est nulle en dehors de $[0, +\infty[$:

$$\int_0^{+\infty} f_Z(x)\ dx \ = \ \int_{-\infty}^{+\infty} f_Z(x)\ dx \ = \ 1$$

  - × l'intégrale $\displaystyle\int_0^{+\infty} x\, f_Z(x)\ dx$ est convergente car la v.a.r. Z admet une espérance. De plus, toujours comme f_Z est nulle en dehors de $[0, +\infty[$:

$$\int_0^{+\infty} x\, f_Z(x)\ dx \ = \ \int_{-\infty}^{+\infty} x\, f_Z(x)\ dx \ = \ \mathbb{E}(Z) \ = \ \frac{1}{\lambda}$$

- On en déduit que l'intégrale $\displaystyle\int_0^{+\infty} f_Z(x)\ \log_2\big(f_Z(x)\big)\ dx$ est convergente en tant que combinaison linéaire d'intégrales convergentes.

$$\boxed{\text{Ainsi, la v.a.r. } Z \text{ admet une entropie différentielle.}}$$

- De plus :

$$\begin{aligned}
& h(Z) \\
= \ & -\int_0^{+\infty} f_Z(x)\, \log_2\big(f_Z(x)\big)\, dx \\
= \ & -\int_0^{+\infty} \log_2(\lambda)\, f_Z(x) - \frac{\lambda}{\ln(2)}\, x\, f_Z(x)\, dx && \textit{(d'après les calculs précédents)} \\
= \ & -\log_2(\lambda) \int_0^{+\infty} f_Z(x)\, dx + \frac{\lambda}{\ln(2)} \int_0^{+\infty} x\, f_Z(x)\, dx && \textit{(par linéarité de l'intégrale, car les intégrales en présence sont convergentes)} \\
= \ & -\log_2(\lambda) \times 1 + \frac{\cancel{\lambda}}{\ln(2)} \times \frac{1}{\cancel{\lambda}} \\
= \ & -\log_2(\lambda) + \log_2(\mathrm{e})
\end{aligned}$$

$$\boxed{\text{D'après la question } \textit{\textbf{1.a)}}, \text{ on obtient : } h(Z) \ = \ \log_2\left(\frac{\mathrm{e}}{\lambda}\right).}$$

□

d) Soit f la fonction définie sur $\mathbb{R}$ par $f(x) = \dfrac{1}{2}\,\lambda\,\mathrm{e}^{-\lambda\,|x|}$ $(\lambda > 0)$.

(i) Montrer que f est une densité de probabilité sur $\mathbb{R}$.

Démonstration.

- La fonction $x \mapsto \mathrm{e}^{-\lambda|x|}$ est continue sur $\mathbb{R}$ car elle est la composée $\exp \circ g_3$ de :
  - × $g_3 : x \mapsto -\lambda|x|$ qui :
    - est continue sur $\mathbb{R}$ en tant que produit de fonctions continues sur $\mathbb{R}$,
    - vérifie : $g_3(\mathbb{R}) \subset \mathbb{R}$ (en fait on a même : $g_3(\mathbb{R}) \subset \,]-\infty, 0]$).
  - × exp qui est continue sur $\mathbb{R}$.

$$\boxed{\text{On en déduit que la fonction } f \text{ est continue sur } \mathbb{R}.}$$

Commentaire

Afin de permettre une bonne compréhension de ce point, on a rédigé ici en détails la continuiré de la composée. Cependant, à ce stade du sujet, une phrase plus succinte du type « la fonction f est continue sur $\mathbb{R}$ en tant que composée et produit de fonctions continues sur $\mathbb{R}$ » permettait sans doute d'obtenir la totalité des points alloués à cette question.

- Soit $x \in \mathbb{R}$.
  Comme $\lambda > 0$ et $\mathrm{e}^{-\lambda|x|} > 0$, on obtient : $f(x) > 0$.

$$\boxed{\forall x \in \mathbb{R},\ f(x) \geqslant 0}$$

- Montrons que l'intégrale $\displaystyle\int_{-\infty}^{+\infty} f(x)\ dx$ converge et vaut 1.

  × Soit $A \in [0, +\infty[$.

$$\begin{aligned} \int_0^A f(x)\ dx &= \int_0^A \frac{1}{2}\ \lambda\, \mathrm{e}^{-\lambda|x|}\ dx \\ &= \frac{1}{2} \int_0^A \lambda\, \mathrm{e}^{-\lambda\, x}\ dx \\ &= \frac{1}{2}\ \Big[-\mathrm{e}^{-\lambda x} \Big]_0^A \\ &= \frac{1}{2} \left(-\mathrm{e}^{-\lambda A} + 1\right) \end{aligned}$$

  Or, comme $\lambda > 0$: $\displaystyle\lim_{A \to +\infty} \mathrm{e}^{-\lambda A} = 0$.

  $\boxed{\text{On en déduit que l'intégrale } \displaystyle\int_0^{+\infty} f(x)\ dx \text{ est convergente et vaut } \frac{1}{2}.}$

  × Soit $A \in]-\infty, 0]$.

$$\begin{aligned} \int_A^0 f(x)\ dx &= \int_A^0 \frac{1}{2}\ \lambda\, \mathrm{e}^{-\lambda|x|}\ dx \\ &= \frac{1}{2} \int_A^0 \lambda\, \mathrm{e}^{\lambda\, x}\ dx \\ &= \frac{1}{2}\ \Big[\mathrm{e}^{\lambda x} \Big]_0^A \\ &= \frac{1}{2} \left(1 - \mathrm{e}^{\lambda A}\right) \end{aligned}$$

  Or, comme $\lambda > 0$: $\displaystyle\lim_{A \to -\infty} \mathrm{e}^{\lambda A} = 0$.

  $\boxed{\text{On en déduit que l'intégrale } \displaystyle\int_{-\infty}^{0} f(x)\ dx \text{ est convergente et vaut } \frac{1}{2}.}$

  $\boxed{\text{Finalement, l'intégrale } \displaystyle\int_{-\infty}^{+\infty} f(x)\ dx \text{ est convergente.}}$

  × De plus :

$$\int_{-\infty}^{+\infty} f(x)\ dx \ = \ \int_{-\infty}^{0} f(x)\ dx + \int_0^{+\infty} f(x)\ dx \ = \ \frac{1}{2} + \frac{1}{2} \ = \ 1$$

$$\boxed{\int_{-\infty}^{+\infty} f(x)\ dx \ = \ 1}$$

  $\boxed{\text{On en déduit que la fonction } f \text{ est une densité de probabilité.}}$

Commentaire

L'intégrale $\displaystyle\int_{-\infty}^{+\infty} f(x)\ dx$ est impropre en ses deux bornes. On rappelle que pour démontrer que de telles intégrales sont convergentes, on démontre :

1. que l'intégrale $\displaystyle\int_{c}^{+\infty} f(x)\ dx$ est convergente,

2. que l'intégrale $\displaystyle\int_{-\infty}^{c} f(x)\ dx$ est convergente.

où c est un réel fixé. Fréquemment, on choisit $c = 0$, comme c'est le cas pour cette question.

□

(ii) Soit W une variable aléatoire de densité f. Justifier l'existence de l'entropie différentielle $h(W)$ et la déterminer.

Démonstration.

- Par définition de f, son support est $]-\infty, +\infty[$.
  Ainsi, la v.a.r. W admet une entropie différentielle si et seulement si l'intégrale $\displaystyle\int_{-\infty}^{+\infty} f(x) \log_2\big(f(x)\big)\ dx$ est convergente.
- Soit $x \in \mathbb{R}$.

$$\begin{array}{rcll}
\log_2\big(f(x)\big) & = & \log_2\left(\dfrac{1}{2}\,\lambda\,\mathrm{e}^{-\lambda\,|x|}\right) & \\[2ex]
& = & -\log_2(2) + \log_2(\lambda) + \log_2\left(\mathrm{e}^{-\lambda\,|x|}\right) & \textit{(d'après la question \textbf{1.a)})} \\[2ex]
& = & -1 + \log_2(\lambda) + \dfrac{\ln\left(\mathrm{e}^{-\lambda\,|x|}\right)}{\ln(2)} & \textit{(car } \log_2(2^1) = 1 \textit{ d'après \textbf{1.b)})} \\[2ex]
& = & -1 + \log_2(\lambda) + \dfrac{-\lambda\,|x|}{\ln(2)} & \\[2ex]
& = & -1 + \log_2(\lambda) - \dfrac{\lambda}{\ln(2)}\ |x| &
\end{array}$$

  Ainsi :

$$\begin{array}{rcl}
f(x) \log_2\big(f(x)\big) & = & f(x)\left(-1 + \log_2(\lambda) - \dfrac{\lambda}{\ln(2)}\ |x|\right) \\[2ex]
& = & \big(-1 + \log_2(\lambda)\big)\, f(x) - \dfrac{\lambda}{\ln(2)}\ |x|\, f(x)
\end{array}$$

- Or l'intégrale $\displaystyle\int_{-\infty}^{+\infty} f(x)\ dx$ est convergente car f est une densité. De plus :

$$\int_{-\infty}^{+\infty} f(x)\ dx \;=\; 1$$

- Démontrons maintenant que l'intégrale $\displaystyle\int_{-\infty}^{+\infty} |x|\, f(x)\ dx$ est convergente.
  × La fonction $x \mapsto |x|\, f(x)$ est continue sur $\mathbb{R}$ en tant que produit de fonctions continues sur $\mathbb{R}$.

× Soit $A \in [0, +\infty[$.

$$\begin{aligned} \int_0^A |x|\, f(x)\ dx &= \frac{1}{2} \int_0^A x\ \lambda\, \mathrm{e}^{-\lambda x}\ dx \\ &= \frac{1}{2} \int_0^A x\, f_Z(x)\ dx \qquad \textit{(car } Z \hookrightarrow \mathcal{E}(\lambda)\textit{)} \end{aligned}$$

Or la v.a.r. Z admet une espérance. Ainsi, avec le même raisonnement qu'en question ***4.c)***, on en déduit que l'intégrale $\displaystyle\int_0^{+\infty} x\, f_Z(x)\ dx$ est convergente et :

$$\int_0^{+\infty} x\, f_Z(x)\ dx \ = \ \mathbb{E}(Z) \ = \ \frac{1}{\lambda}$$

$$\boxed{\begin{array}{c} \text{Finalement, l'intégrale } \displaystyle\int_0^{+\infty} |x| f(x)\ dx \text{ est convergente et :} \\ \displaystyle\int_0^{+\infty} |x|\, f(x)\ dx \ = \ \frac{1}{2} \times \frac{1}{\lambda}. \end{array}}$$

× Soit $A \in\ ]-\infty, 0]$.

$$\begin{aligned} \int_A^0 |x|\, f(x)\ dx &= \frac{1}{2} \int_A^0 (-x)\, \lambda \mathrm{e}^{-\lambda\,(-x)}\ dx \\ &= -\frac{1}{2} \int_A^0 x\, \lambda \mathrm{e}^{\lambda x}\ dx \end{aligned}$$

Pour calculer l'intégrale $\displaystyle\int_A^0 x\, \lambda \mathrm{e}^{\lambda x}\ dx$, on procède par intégration par parties (IPP).

$$\left|\begin{array}{rclrcl} u(x) &=& x & u'(x) &=& 1 \\ v'(x) &=& \lambda \mathrm{e}^{\lambda x} & v(x) &=& \mathrm{e}^{\lambda x} \end{array}\right.$$

Cette IPP est valide car les fonctions u et v sont de classe $\mathcal{C}^1$ sur $[A, 0]$. On obtient :

$$\begin{aligned} \int_A^0 x\, \lambda \mathrm{e}^{\lambda x}\ dx &= \Big[\ x\ \lambda \mathrm{e}^{\lambda x}\ \Big]_A^0 - \int_A^0 \mathrm{e}^{\lambda x}\ dx \\ &= \left(0 - \lambda\, A\, \mathrm{e}^{\lambda\, A}\right) - \left[\ \frac{1}{\lambda}\, \mathrm{e}^{\lambda x}\ \right]_A^0 \\ &= -\lambda\, A\, \mathrm{e}^{\lambda\, A} - \frac{1}{\lambda}\, (1 - \mathrm{e}^{\lambda\, A}) \\ &= -\frac{1}{\lambda} + \frac{1}{\lambda}\, \mathrm{e}^{\lambda A} - \lambda\, A\, \mathrm{e}^{\lambda\, A} \end{aligned}$$

Or, comme $\lambda > 0$:

- $\lim\limits_{A \to -\infty} \mathrm{e}^{\lambda A} \ = \ 0$,
- $\lim\limits_{A \to -\infty} A\,\mathrm{e}^{\lambda A} \ = \ 0$, par croissances comparées.

On en déduit que l'intégrale $\displaystyle\int_{-\infty}^{0} x\,\lambda \mathrm{e}^{\lambda x}\,dx$ est convergente et vaut $-\dfrac{1}{\lambda}$.

$$\boxed{\begin{array}{c}\text{Ainsi, l'intégrale } \displaystyle\int_{-\infty}^{0} |x|\,f(x)\,dx \text{ est convergente et :}\\ \displaystyle\int_{-\infty}^{0} |x|\,f(x)\,dx \ = \ -\frac{1}{2} \times \left(-\frac{1}{\lambda}\right) \ = \ \frac{1}{2}\,\frac{1}{\lambda}.\end{array}}$$

$$\boxed{\text{On en déduit que l'intégrale } \displaystyle\int_{-\infty}^{+\infty} |x|\,f(x)\,dx \text{ est convergente.}}$$

De plus :

$$\int_{-\infty}^{+\infty} |x|\,f(x)\,dx \ = \ \int_{-\infty}^{0} |x|\,f(x)\,dx + \int_{0}^{+\infty} |x|\,f(x)\,dx \ = \ \frac{1}{2}\,\frac{1}{\lambda} + \frac{1}{2}\,\frac{1}{\lambda} \ = \ \frac{1}{\lambda}$$

$$\boxed{\int_{-\infty}^{+\infty} |x|\,f(x)\,dx \ = \ \frac{1}{\lambda}}$$

- On en déduit que l'intégrale $\displaystyle\int_{-\infty}^{+\infty} f(x)\,\log_2\big(f(x)\big)\,dx$ est convergente en tant que combinaison linéaire d'intégrales convergentes.

$$\boxed{\text{Ainsi, la v.a.r. } W \text{ admet une entropie différentielle.}}$$

- De plus :

$$\begin{array}{rll}
& h(W) & \\
= & -\displaystyle\int_{-\infty}^{+\infty} f(x)\,\log_2\big(f(x)\big)\,dx & \\
= & -\displaystyle\int_{-\infty}^{+\infty} \big(-1+\log_2(\lambda)\big)\,f(x) - \frac{\lambda}{\ln(2)}\,f(x)\,dx & \textit{(d'après les calculs précédents)} \\
= & \big(1-\log_2(\lambda)\big)\displaystyle\int_{-\infty}^{+\infty} f(x)\,dx - \frac{\lambda}{\ln(2)}\int_{-\infty}^{+\infty} |x|\,f(x)\,dx & \textit{(par linéarité de l'intégrale, car les intégrales en présence sont convergentes)} \\
= & \big(1-\log_2(\lambda)\big) \times 1 - \dfrac{\cancel{\lambda}}{\ln(2)} \times \dfrac{1}{\cancel{\lambda}} & \\
= & 1 - \log_2(\lambda) + \log_2(\mathrm{e}) &
\end{array}$$

$$\boxed{\text{D'après la question } \textbf{\textit{1.a)}}\text{, on obtient : } h(W) = 1 + \log_2\left(\frac{\mathrm{e}}{\lambda}\right).}$$

□

Commentaire

On dit que la v.a.r. W suit la *loi de Laplace de paramètre* $(0, 1)$. L'étude de cette loi est classique aux concours. On pourra par exemple regarder l'épreuve ESSEC I 2017 pour une étude plus détaillée de cette loi.

5. On dit qu'un couple (X, Y) de variables aléatoires est un couple gaussien centré si, pour tout $(\alpha, \beta) \in \mathbb{R}^2$, $\alpha\, X + \beta\, Y$ est une variable de loi normale centrée, c'est-à-dire qu'il existe $\gamma \in \mathbb{R}$ et une variable aléatoire Z de loi normale centrée réduite tels que $\alpha\, X + \beta\, Y$ a même loi que $\gamma\, Z$. On considère un tel couple (X, Y) et on noté σ^2 la variance de X. On suppose : $\sigma^2 > 0$.

a) Montrer que X suit une loi normale centrée.

Démonstration.
Comme (X, Y) est un couple gaussien centré, alors **pour tout** $(\alpha, \beta) \in \mathbb{R}^2$, $\alpha\, X + \beta\, Y$ suit une loi normale centrée. **En particulier**, la v.a.r. $1 \cdot X + 0 \cdot Y = X$ suit une loi normale centrée.
De plus, d'après l'énoncé : $\mathbb{V}(X) = \sigma^2$.

$$\boxed{\text{Finalement : } X \hookrightarrow \mathcal{N}\left(0, \sigma^2\right).}$$

□

b) Calculer $h(X)$.

Démonstration.

- D'après la question précédente : $X \hookrightarrow \mathcal{N}\left(0, \sigma^2\right)$.

$$\boxed{\text{On en déduit que } X \text{ suit la même loi que la v.a.r. } \sigma\, Z, \text{ où : } Z \hookrightarrow \mathcal{N}(0, 1).}$$

Commentaire

- On rappelle qu'on a le résultat suivant :

$$X \hookrightarrow \mathcal{N}\left(m, \sigma^2\right) \quad \Leftrightarrow \quad a\, X + b \hookrightarrow \mathcal{N}\left(am + b, a^2\, \sigma^2\right)$$

- En particulier :

$$Z \hookrightarrow \mathcal{N}(0, 1) \quad \Leftrightarrow \quad \sigma\, Z \hookrightarrow \mathcal{N}\left(0, \sigma^2\right)$$

- De plus, d'après la question ***4.b)***, la v.a.r. Z admet une entropie différentielle et :

$$h(Z) \;=\; \frac{1}{2}\, \log_2(2\pi \mathrm{e})$$

- Enfin, d'après la question ***3.b)(ii)*** :
  - × la v.a.r. X admet une entropie différentielle,
  - × on obtient de plus :

$$h(X) \;=\; h(Z) + \log_2(\sigma) \;=\; \frac{1}{2}\, \log(2\pi \mathrm{e}) + \frac{1}{2}\, \log_2(\sigma^2) \;=\; \frac{1}{2}\, \log_2(2\pi\, \mathrm{e}\, \sigma^2)$$

$$\boxed{h(X) \;=\; \frac{1}{2}\, \log_2(2\pi\, \mathrm{e}\, \sigma^2)}$$

□

c) On suppose désormais que X et Y suivent la même loi normale centrée de variance σ^2 et on admet que les propriétés de l'espérance des variables discrètes se généralisent aux variables aléatoires quelconques.

(i) Montrer que $\mathbb{E}(XY)$ existe.

Démonstration.
Les v.a.r. X et Y suivent une loi normale.
Elles admettent donc chacune un moment d'ordre 2.

$$\boxed{\text{On en déduit que la v.a.r. } XY \text{ admet une espérance.}}$$

Commentaire

- On rappelle que, dans le cas général, la v.a.r. produit XY admet une espérance si les v.a.r. X et Y admettent **un moment d'ordre** 2.
- On peut se demander d'où provient cette hypothèse liée aux moments d'ordre 2. Elle est issue d'un théorème de domination. Détaillons ce point.
  Remarquons tout d'abord : $(X-Y)^2 \geqslant 0$.
  On en déduit : $X^2 - 2\,XY + Y^2 \geqslant 0$.
  Et, en réordonnant : $XY \leqslant \dfrac{1}{2}\,X^2 + \dfrac{1}{2}\,Y^2$. Ou encore :
  $$0 \ \leqslant\ |XY| \ \leqslant\ \frac{1}{2}\,X^2 + \frac{1}{2}\,Y^2$$
  Comme X et Y admettent un moment d'ordre 2, la v.a.r. $\frac{1}{2}\,X^2 + \frac{1}{2}\,Y^2$ admet une espérance comme combinaison linéaire de v.a.r. qui en admettent une.
  Ainsi, par théorème de domination (présenté seulement dans le programme ECS), la v.a.r. $|XY|$ admet une espérance. Il en est de même de la v.a.r. XY.

□

(ii) Montrer de plus, pour tout réel λ : $\lambda^2\,\mathbb{E}\left(Y^2\right) + 2\lambda\,\mathbb{E}(XY) + \mathbb{E}\left(X^2\right) \ \geqslant\ 0$.

Démonstration.
Soit $\lambda \in \mathbb{R}$.

$$\begin{aligned} \lambda^2\,\mathbb{E}(Y^2) + 2\lambda\,\mathbb{E}(XY) + \mathbb{E}(X^2) &= \mathbb{E}\left(\lambda^2\,Y^2 + 2\lambda\,XY + X^2\right) && \text{(par linéarité de l'espérance)} \\ &= \mathbb{E}\Big((\lambda\,Y + X)^2\Big) \end{aligned}$$

Or : $(\lambda\,Y + X)^2 \geqslant 0$.
Ainsi, par croissance de l'espérance : $\mathbb{E}\Big((\lambda Y + X)^2\Big) \geqslant 0$.

$$\boxed{\text{On en déduit : } \lambda^2\,\mathbb{E}(Y^2) + 2\lambda\,\mathbb{E}(XY) + \mathbb{E}(X^2) \geqslant 0.}$$

Commentaire

On prêtera attention à la différence de nature des deux inégalités citées plus haut :
- × l'inégalité $(\lambda\,Y + X)^2 = 0$ est une inégalité entre **variables aléatoires**. Ainsi, « 0 » est ici la variable aléatoire constante égale à 0.
- × l'inégalité $\mathbb{E}\Big((\lambda\,Y + X)^2\Big) \geqslant 0$ est une inégalité entre **réels**. Ainsi, « 0 » ici tout simplement le réel 0.

□

(iii) En déduire : $\big(\mathbb{E}(XY)\big)^2 \ \leqslant\ \mathbb{E}\left(X^2\right)\,\mathbb{E}\left(Y^2\right)$.

Démonstration.
- La fonction g définie par :
  $$\forall \lambda \in \mathbb{R}, \quad P(\lambda) = \lambda^2\,\mathbb{E}(Y^2) + 2\lambda\,\mathbb{E}(XY) + \mathbb{E}(X^2)$$
  est une fonction polynomiale de degré 2 en λ.

- De plus, d'après la question précédente : $\forall \lambda \in \mathbb{R},\ g(\lambda) \geqslant 0$.
  Cette fonction polynomiale étant de signe constant, on en déduit que le discriminant du polynôme associé est négatif. Or :

$$\Delta \ = \ \big(2\,\mathbb{E}(XY)\big)^2 - 4\,\mathbb{E}(Y^2)\,\mathbb{E}(X^2) \ = \ 4\Big(\big(\mathbb{E}(XY)\big)^2 - \mathbb{E}(Y^2)\,\mathbb{E}(X^2)\Big)$$

  On obtient donc :

$$\Delta \leqslant 0 \quad \Leftrightarrow \quad \big(\mathbb{E}(XY)\big)^2 \ \leqslant \ \mathbb{E}(Y^2)\,\mathbb{E}(X^2)$$

$$\boxed{\text{Ainsi : } \big(\mathbb{E}(XY)\big)^2 \ \leqslant \ \mathbb{E}(Y^2)\,\mathbb{E}(X^2).}$$

Commentaire

On reconnaît ici exactement le principe de démonstration de l'inégalité de Cauchy-Schwarz :

$$\big(\operatorname{Cov}(X,Y)\big)^2 \ \leqslant \ \mathbb{V}(X)\,\mathbb{V}(Y)$$

□

(iv) On pose $\rho = \dfrac{\mathbb{E}(XY)}{\sigma^2}$. Montrer : $\rho \in [-1,1]$.

Démonstration.

- D'après la question précédente :

$$\begin{aligned} \big(\mathbb{E}(XY)\big)^2 \ \leqslant \ \mathbb{E}(Y^2)\,\mathbb{E}(X^2) \quad &\Leftrightarrow \quad \sqrt{\big(\mathbb{E}(XY)\big)^2} \ \leqslant \ \sqrt{\mathbb{E}(Y^2)\,\mathbb{E}(X^2)} && \textit{(par stricte croissance de } \sqrt{\cdot} \textit{ sur } [0,+\infty[\textit{)} \\ &\Leftrightarrow \quad |\mathbb{E}(XY)| \ \leqslant \ \sqrt{\mathbb{E}(Y^2)}\,\sqrt{\mathbb{E}(X^2)} \end{aligned}$$

- De plus, par fomule de Koenig-Huygens :

$$\begin{array}{ccccc} \mathbb{V}(X) & = & \mathbb{E}(X^2) & - & \big(\mathbb{E}(X)\big)^2 \\ \shortparallel & & & & \shortparallel \\ \sigma^2 & & & & 0 \end{array} \qquad \textit{(car } X \textit{ est une v.a.r. centrée)}$$

  Ainsi : $\mathbb{E}(X^2) = \sigma^2$. De même : $\mathbb{E}(Y^2) = \sigma^2$.

- On en déduit :

$$\begin{aligned} |\mathbb{E}(XY)| \ \leqslant \ \sqrt{\mathbb{E}(Y^2)}\,\sqrt{\mathbb{E}(X^2)} \quad &\Leftrightarrow \quad |\mathbb{E}(XY)| \ \leqslant \ \sqrt{\sigma^2}\,\sqrt{\sigma^2} \\ &\Leftrightarrow \quad |\mathbb{E}(XY)| \ \leqslant \ \sigma^2 \\ &\Leftrightarrow \quad \frac{|\mathbb{E}(XY)|}{\sigma^2} \ \leqslant \ 1 \\ &\Leftrightarrow \quad \left|\frac{\mathbb{E}(XY)}{\sigma^2}\right| \ \leqslant \ 1 && \textit{(car } \sigma^2 \geqslant 0\textit{)} \\ &\Leftrightarrow \quad |\rho| \ \leqslant \ 1 \end{aligned}$$

$$\boxed{\text{Ainsi : } -1 \ \leqslant \ \rho \ \leqslant \ 1.}$$

Commentaire

- On reconnaît ici le *coefficient de corrélation linéaire* défini par :

$$\rho \ = \ \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathbb{V}(X)}\,\sqrt{\mathbb{V}(Y)}}$$

dans le cas où X et Y sont deux v.a.r. :
  - × centrées (dans ce cas, par formule de Koenig-Huygens : $\mathrm{Cov}(X,Y) = \mathbb{E}(XY)$),
  - × de variance σ^2.
- Le programme d'ECE ne définit ce coefficient que pour des v.a.r. discrètes. On peut cependant généraliser cette définition au cas de v.a.r. quelconques. Cette question et la suivante montrent que l'on conserve les grandes propriétés du coefficient de corrélation linéaire défini en cours.

□

(v) Que vaut ρ si X et Y sont indépendantes ?

Démonstration.
Si les v.a.r. X et Y sont indépendantes, alors :

$$\begin{array}{rcll} \mathbb{E}(XY) & = & \mathbb{E}(X)\,\mathbb{E}(Y) & \\ & = & 0 & \textit{(car X et Y sont des v.a.r. centrées)} \end{array}$$

$$\boxed{\text{Ainsi, si les v.a.r. } X \text{ et } Y \text{ sont indépendantes, alors : } \rho \ = \ \frac{\mathbb{E}(XY)}{\sigma^2} \ = \ 0.}$$

Commentaire

Remarquons bien qu'on démontre ici **l'implication** :

$$X \text{ et } Y \text{ indépendantes} \quad \Rightarrow \quad \rho = 0$$

Ce n'est en rien une équivalence.

□

d) On suppose $|\rho| < 1$. On appelle **entropie jointe** du couple (X,Y) le réel :

$$h(X,Y) \ = \ \log_2\left(2\pi\,\mathrm{e}\,\sigma^2\,\sqrt{1-\rho^2}\right)$$

(i) À quelle condition $h(X,Y)$ est-elle nulle ?

Démonstration.
On a les équivalences suivantes :

$$\begin{array}{rcll} h(X,Y)=0 & \Leftrightarrow & \log_2\left(2\pi\,\mathrm{e}\,\sigma^2\,\sqrt{1-\rho^2}\right) \ = \ 0 & \\ & \Leftrightarrow & \dfrac{\ln\left(2\pi\,\mathrm{e}\,\sigma^2\,\sqrt{1-\rho^2}\right)}{\ln(2)} \ = \ 0 & \\ & \Leftrightarrow & \ln\left(2\pi\,\mathrm{e}\,\sigma^2\,\sqrt{1-\rho^2}\right) \ = \ 0 & \\ & \Leftrightarrow & 2\pi\,\mathrm{e}\,\sigma^2\,\sqrt{1-\rho^2} \ = \ 1 & \textit{(par bijectivité de la fonction $\exp$ sur $\mathbb{R}$)} \end{array}$$

De plus :

$$\begin{aligned} 2\pi\,\mathrm{e}\,\sigma^2\sqrt{1-\rho^2} = 1 \quad &\Leftrightarrow \quad \sqrt{1-\rho^2} = \frac{1}{2\pi\mathrm{e}\sigma^2} \\ &\Leftrightarrow \quad 1-\rho^2 = \left(\frac{1}{2\pi\mathrm{e}\sigma^2}\right)^2 && \textit{(par bijectivité de la fonction } x \mapsto x^2 \textit{ sur } [0,+\infty[\textit{)} \\ &\Leftrightarrow \quad \rho^2 = 1 - \frac{1}{(2\pi\mathrm{e}\sigma^2)^2} \end{aligned}$$

$$\boxed{\text{Finalement : } h(X,Y) = 0 \;\Leftrightarrow\; \rho \in \left\{ -\sqrt{1-\frac{1}{(2\pi\mathrm{e}\sigma^2)^2}} \,,\, \sqrt{1-\frac{1}{(2\pi\mathrm{e}\sigma^2)^2}} \right\}.}$$

Commentaire

Cette question est un peu vague :
- × demande-t-on une condition nécessaire ? suffisante ? nécessaire et suffisante ?
- × demande-t-on une condition sur ρ ou sur un autre paramètre ?

Sans plus de précisions, on considèrera que les termes « nécessaire et suffisante » sont sous-entendus. On choisit par ailleurs dans ce corrigé de considérer qu'on demande bien une condition sur le paramètre ρ.

□

(ii) L'**information mutuelle** de X et Y est définie par :

$$I(X,Y) = h(X) + h(Y) - h(X,Y)$$

Calculer $I(X,Y)$.

Démonstration.
On calcule :

$$\begin{aligned} I(X,Y) &= h(X) + h(Y) - h(X,Y) \\ &= \frac{1}{2}\log_2(2\pi\,\mathrm{e}\,\sigma^2) + \frac{1}{2}\log_2(2\pi\,\mathrm{e}\,\sigma^2) - \log_2\left(2\pi\,\mathrm{e}\,\sigma^2\sqrt{1-\rho^2}\right) && \textit{(d'après la question } \textbf{\textit{5.b)}}\textit{)} \\ &= \log_2(2\pi\,\mathrm{e}\,\sigma^2) - \log_2\left(2\pi\,\mathrm{e}\,\sigma^2\sqrt{1-\rho^2}\right) \\ &= -\log_2\left(\frac{\cancel{2\pi\,\mathrm{e}\,\sigma^2}\sqrt{1-\rho^2}}{\cancel{2\pi\,\mathrm{e}\,\sigma^2}}\right) && \textit{(d'après la question } \textbf{\textit{1.a)}}\textit{)} \\ &= -\log_2\left(\sqrt{1-\rho^2}\right) \end{aligned}$$

$$\boxed{\text{Finalement : } I(X,Y) = -\frac{1}{2}\log_2(1-\rho^2)}$$

□

(iii) Montrer : $I(X,Y) \geqslant 0$.

Démonstration.
On a la succession d'équivalences suivante :

$$\begin{array}{rcll} I(X,Y) \geqslant 0 & \Leftrightarrow & -\dfrac{1}{2}\log_2(1-\rho^2) \geqslant 0 & \textit{(d'après la question précédente)} \\ & \Leftrightarrow & \log_2(1-\rho^2) \leqslant 0 & \\ & \Leftrightarrow & \dfrac{\ln(1-\rho^2)}{\ln(2)} \leqslant 0 & \\ & \Leftrightarrow & \ln(1-\rho^2) \leqslant 0 & \textit{(car } \ln(2) > 0\textit{)} \\ & \Leftrightarrow & 1-\rho^2 \leqslant 1 & \textit{(par stricte croissance de la fonction } \exp \textit{ sur } \mathbb{R}\textit{)} \\ & \Leftrightarrow & 0 \leqslant \rho^2 & \end{array}$$

La dernière inégalité est vraie. Ainsi, par équivalence, la première inégalité aussi.

$$\boxed{\text{Finalement : } I(X,Y) \geqslant 0.}$$

□

(iv) Quelle est la limite de $I(X,Y)$ quand ρ tend vers 1 ?

Démonstration.

- D'après la question ***5.d)(ii)*** :

$$I(X,Y) = -\frac{1}{2}\log_2(1-\rho^2) = -\frac{1}{2}\,\frac{\ln(1-\rho^2)}{\ln(2)}$$

- Tout d'abord, comme $\lim\limits_{\rho\to 1} 1-\rho^2 = 0$, alors : $\lim\limits_{\rho\to 1} \ln(1-\rho^2) = -\infty$.

$$\boxed{\text{Comme de plus } -\frac{1}{2\ln(2)} < 0, \text{ on obtient : } \lim_{\rho\to 1} I(X,Y) = +\infty.}$$

□

Deuxième partie : Généralités sur l'entropie des variables discrètes

Soit A un ensemble fini non vide. On dit que X est une variable aléatoire dont la loi est à support A, si X est à valeurs dans A et si pour tout $x \in A$: $\mathbb{P}([X = x]) > 0$.

6. Soit X une variable aléatoire de loi à support $\{0, 1, 2, \ldots, n\}$ où n est un entier naturel. On appelle **entropie** de X le réel :

$$H(X) \ = \ -\sum_{k=0}^{n} \mathbb{P}([X = k]) \log_2 \big(\mathbb{P}([X = k])\big)$$

a) On définit la fonction $g : \{0, \ldots, n\} \to \mathbb{R}$ en posant $g(k) = \log_2 \big(\mathbb{P}([X = k])\big)$ pour k élément de $\{0, 1, \ldots, n\}$. Montrer : $H(X) = -\mathbb{E}\big(g(X)\big)$.

b) Montrer : $H(X) \geqslant 0$.

c) Soit p un réel tel que $0 < p < 1$.
On suppose dans cette question que X suit la loi de Bernoulli $\mathcal{B}(p)$.

(i) Calculer $H(X)$ en fonction de p. On note ψ la fonction qui, à p, associe $H(X)$.

(ii) Montrer que ψ est concave sur $]0, 1[$.

(iii) Déterminer la valeur p_0 où ψ est maximale.

d) On suppose dans cette question que la loi de X est à support $\{0, 1, 2, 3\}$ avec les probabilités :

$$\mathbb{P}([X = 0]) = \frac{1}{2} \quad ; \quad \mathbb{P}([X = 1]) = \frac{1}{4} \quad ; \quad \mathbb{P}([X = 2]) = \mathbb{P}([X = 3]) = \frac{1}{8}$$

Calculer $H(X)$.

7. On souhaite écrire une fonction en **Scilab** pour calculer l'entropie d'une variable aléatoire X dont le support de la loi est de la forme $A = \{0, 1, \ldots, n\}$ où n est un entier naturel. On suppose que le vecteur `P` de **Scilab** est tel que pour tout k de A, `P(k + 1)` $= \mathbb{P}([X = k])$. Compléter la fonction ci-dessous d'argument `P` qui renvoie l'entropie de X, c'est-à-dire $-\sum\limits_{k=0}^{n} \mathbb{P}([X = k]) \log_2 \big(\mathbb{P}([X = k])\big)$.

```
function h = Entropie(P)
...
endfunction
```

Si nécessaire, on pourra utiliser l'instruction `length(P)` qui donne le nombre d'éléments de `P`.
On souhaite maintenant démontrer quelques inégalités concernant l'entropie.

8. On commence par une inégalité générale, appelée **Inégalité de Jensen**.

a) Soit $N \geqslant 2$. Soit X une variable aléatoire de loi à support $\{x_1, x_2, \ldots, x_N\}$ où les x_i sont des éléments distincts de $\mathbb{R}_+$. On pose $\mathbb{P}([X = x_i]) = p_i$.
Montrer que, pour tout $1 \leqslant i \leqslant N$, on a : $p_i < 1$.

On désire démontrer par récurrence la propriété suivante :

$\mathcal{P}(N)$: **Pour toute fonction φ convexe sur $\mathbb{R}_+$, si X est une variable aléatoire de loi à support $A \subset \mathbb{R}_+$ avec $\text{Card}(A) = N$, on a :** $\mathbb{E}\big(\varphi(X)\big) \geqslant \varphi\big(\mathbb{E}(X)\big)$

b) Montrer que $\mathcal{P}(2)$ est vraie.

c) Soit $N \geqslant 3$. On suppose que $\mathcal{P}(N-1)$ est vérifiée. Soit X une variable aléatoire de loi à support $A = \{x_1, x_2, \dots, x_N\}$ où les x_i sont des éléments distincts de $\mathbb{R}_+$. On pose : $\mathbb{P}([X = x_i]) = p_i$. Pour i tel que $1 \leqslant i \leqslant N-1$, on pose : $p'_i = \dfrac{p_i}{1-p_N}$.

(i) Montrer : $\sum\limits_{i=1}^{N-1} p'_i = 1$ et $\forall i \in [\![1, N-1]\!]$, $0 < p'_i < 1$.

(ii) Soit Y une variable aléatoire de loi à support $\{x_1, \dots, x_{N-1}\}$ telle que $\mathbb{P}([Y = x_i]) = p'_i$ pour $1 \leqslant i \leqslant N-1$. Montrer : $\sum\limits_{i=1}^{N-1} p'_i\, \varphi(x_i) \;\geqslant\; \varphi\left(\sum\limits_{i=1}^{N-1} p'_i\, x_i\right)$.

(iii) Montrer : $\mathbb{E}\big(\varphi(X)\big) \;\geqslant\; \varphi\big(\mathbb{E}(X)\big)$.

d) Montrer que, si φ est *concave* sur $\mathbb{R}_+$, on a : $\mathbb{E}\big(\varphi(X)\big) \;\leqslant\; \varphi\big(\mathbb{E}(X)\big)$.

9. Soit X une variable aléatoire de loi à support $\{0, 1, \dots, n\}$. On pose, pour k tel que $0 \leqslant k \leqslant n$, $p_k = \mathbb{P}([X = k])$.

a) Montrer : $\sum\limits_{k=0}^{n} p_k \log_2\left(\dfrac{1}{(n+1)\, p_k}\right) \;\leqslant\; \log_2\left(\sum\limits_{k=0}^{n} \dfrac{p_k}{(n+1)\, p_k}\right) \;=\; 0$.

b) Montrer : $\sum\limits_{k=0}^{n} p_k \log_2\big((n+1)\, p_k\big) \;=\; \log_2(n+1) - H(X)$.

c) Montrer : $H(X) \leqslant \log_2(n+1)$.

d) On suppose que X suit la loi uniforme sur $\{0, 1, \dots, N\}$. Calculer $H(X)$.

10. Soient X et Y deux variables aléatoires *de même loi* à support $\{0, 1, \dots, n\}$. On suppose en outre X et Y indépendantes.

a) Montrer : $\mathbb{P}([X = Y]) \;=\; \sum\limits_{k=0}^{n} \big(\mathbb{P}([X = k])\big)^2$.

b) On pose $v(k) = \mathbb{P}([X = k])$ pour tout k élément de $\{0, 1, \dots, n\}$. Montrer :

$$2^{\mathbb{E}\left(\log_2\left(v(X)\right)\right)} \;\leqslant\; \mathbb{E}\left(2^{\log_2\left(v(X)\right)}\right) \;=\; \mathbb{E}\big(v(X)\big)$$

c) En déduire : $2^{-H(X)} \;\leqslant\; \mathbb{P}([X = Y])$.

d) Donner un exemple de loi où l'inégalité précédente est une égalité.

Troisième partie : Entropie jointe et information mutuelle de deux variables discrètes

Soient X et Y deux variables aléatoires de lois à support $\{0, 1, \dots, n\}$. On appelle **entropie jointe** de X et Y le réel :

$$H(X,Y) \;=\; -\sum_{k=0}^{n} \sum_{j=0}^{n} \mathbb{P}([X = k] \cap [Y = j]) \log_2\big(\mathbb{P}([X = k] \cap [Y = j])\big)$$

avec la convention : $0 \times \log_2(0) = 0$.

11. ***a)*** On définit la fonction $g : \{0, 1, \dots, n\}^2 \to \mathbb{R}^{\{} - \infty\}$ en posant pour $(k, j) \in \{0, 1, \dots, n\}^2$:

$$g(k,j) \;=\; \log_2\big(\mathbb{P}([X = k] \cap [Y = j])\big)$$

Montrer : $H(X,Y) = -\mathbb{E}\big(g(X,Y)\big)$.

b) Montrer : $H(X,Y) = H(Y,X)$.

c) Pour tout k tel que $0 \leqslant k \leqslant n$, on pose :

$$H(Y \mid X = k) \;=\; -\sum_{j=0}^{n} \mathbb{P}_{[X=k]}([Y=j]) \log_2 \big(\mathbb{P}_{[X=k]}([Y=j])\big)$$

On appelle **entropie conditionnelle** de Y sachant X le réel :

$$H(Y \mid X) \;=\; \sum_{k=0}^{n} \mathbb{P}([X=k])\, H(Y \mid X=k)$$

Montrer : $H(X,Y) \;=\; H(X) + H(Y \mid X)$.

d) Montrer que pour tout couple de variables aléatoires X et Y de lois à support $\{0,1,\dots,n\}$, on a :

$$H(X) - H(X \mid Y) \;=\; H(Y) - H(Y \mid X)$$

12. On considère dans cette question deux variables aléatoires de lois à support $\{0,1,2,3\}$. On suppose que la loi conjointe de (X,Y) est donnée par le tableau suivant :

$j \backslash k$	0	1	2	3
0	$\frac{1}{8}$	$\frac{1}{16}$	$\frac{1}{32}$	$\frac{1}{32}$
1	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{32}$	$\frac{1}{32}$
2	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$
3	$\frac{1}{4}$	0	0	0

(on lit dans la $k^{\text{ème}}$ colonne et la $j^{\text{ème}}$ ligne la valeur de $\mathbb{P}([X=k] \cap [Y=j])$)

a) Déterminer la loi de X et montrer : $H(X) = \dfrac{7}{4}$.

b) Déterminer la loi de Y et calculer $H(Y)$.

c) Montrer : $H(X \mid Y) = \dfrac{11}{8}$.

d) Que vaut $H(Y \mid X)$?

e) Calculer $H(X,Y)$.

13. Soient X et Y deux variables aléatoires de lois à support $\{0,1,\dots,n\}$. On appelle **information mutuelle** de X et de Y le réel :

$$I(X,Y) \;=\; \sum_{k=0}^{n} \sum_{j=0}^{n} \mathbb{P}([X=k] \cap [Y=j]) \log_2 \left(\frac{\mathbb{P}([X=k] \cap [Y=j])}{\mathbb{P}([X=k])\, \mathbb{P}([Y=j])} \right)$$

a) Montrer : $I(X,Y) = I(Y,X)$.

b) Montrer : $I(X,Y) = H(X) - H(X \mid Y)$.

c) Montrer : $I(X,X) = H(X)$.

d) Que vaut $I(X,Y)$ si X et Y sont indépendantes ?

14. Soient X et Y deux variables aléatoires de lois à support $\{0, 1, \dots, n\}$. On fixe $0 \leqslant k \leqslant n$. Pour $0 \leqslant j \leqslant n$, on pose : $p_j = \dfrac{\mathbb{P}([X = k] \cap [Y = j])}{\mathbb{P}([X = k])}$. On suppose que $p_j > 0$ pour tout $0 \leqslant j \leqslant n$ et on pose : $x_j = \dfrac{\mathbb{P}([X = k])\,\mathbb{P}([Y = j])}{\mathbb{P}([X = k] \cap [Y = j])}$.

a) Montrer : $\sum\limits_{j=0}^{n} p_j = 1$.

b) Soit Z_k une variable aléatoire de loi à support $\{x_0, \dots, x_n\}$ dont la loi est donnée par $\mathbb{P}([Z_k = x_j]) = p_j$ pour $0 \leqslant j \leqslant n$. Montrer :

$$\mathbb{E}\big(\log_2(Z_k)\big) \leqslant 0$$

c) En déduire : $I(X, Y) \geqslant 0$.