DS8 /154

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 considérera 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 : $\log_2(x\,y) = \log_2(x) + \log_2(y)$.

- **1 pt**

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

- **1 pt**

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

- **1 pt :** $\log_2$ **de classe** $\mathcal{C}^2$ **sur** $]0, +\infty[$
- **1 pt :** $\log_2'(x) = \dfrac{1}{x\,\ln(2)}$ **et** $\log_2''(x) = -\dfrac{1}{x^2\,\ln(2)} > 0$

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

- **1 pt : par théorème de transfert,** $\log_2\big(f(X)\big)$ **admet une espérance ssi** $\displaystyle\int_{-\infty}^{+\infty} \log_2\big(f(x)\big) f(x)\,dx$ **est absolument convergente.**
- **1 pt :** $\displaystyle\int_{-\infty}^{+\infty} \log_2\big(f(x)\big)\,dx = \int_a^b \log_2\big(f(x)\big)\,dx$ **car** f **nulle en dehors de** I **d'extrémités** a **et** 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.

- **1 pt :** $f_Y : x \mapsto f(x - c)$

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

- **1 pt :** Y **admet une espérance ssi** $\log_2\big(f_Y(Y)\big)$ **admet une espérance**
- **1 pt :** $\log_2\big(f_Y(Y)\big) = \log_2\big(f(X)\big)$
- **1 pt :** X **admet une entropie différentielle donc** Y **aussi et** $h(Y) = h(X)$

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

(i) Déterminer une densité de Z.

- **1 pt :** $f_Z : x \mapsto \dfrac{1}{\alpha}\, f\left(\frac{x}{\alpha}\right)$

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

- **1 pt :** $\log_2\big(f_Z(Z)\big) = \log_2\left(\dfrac{1}{\alpha}\right) + \log_2\big(f(x)\big)$
- **1 pt :** $\log_2\left(\dfrac{1}{\alpha}\right) = -\log_2(\alpha)$
- **1 pt :** X **admet une entropie différentielle donc** Z **aussi et** $h(Z) = h(X) + \log_2(\alpha)$

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.

- **1 pt :** $f : x \mapsto \begin{cases} 0 & \textbf{si } x \in\ ]-\infty, 0[\\ \dfrac{1}{a} & \textbf{si } x \in [0, a] \\ 0 & \textbf{si } x \in\ ]a, +\infty[\end{cases}$

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

- **1 pt :** X **admet une entropie différentielle ssi** $\displaystyle\int_0^a \log_2\big(f(x)\big) f(x)\ dx$ **est convergente**
- **1 pt :** $x \mapsto \log_2\big(f(x)\big)\, f(x)$ **continue par morceaux sur le segment** $[0, a]$ **donc** X **admet une entropie différentielle**
- **1 pt :** $h(X) = \log_2(a)$

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

- **1 pt :** $h(X) > 0 \Leftrightarrow a > 1$
  0 **si la stricte croissance de** $\exp$ **sur** $\mathbb{R}$ **n'est pas citée**

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

- **1 pt :** $\log_2\big(f_Y(x)\big)\, f_Y(x) = -\dfrac{1}{2} \log_2(2\,\pi)\ f_Y(x) - \dfrac{1}{2\ln(2)}\ x^2\, f_Y(x)$
- **1 pt :** $\displaystyle\int_{-\infty}^{+\infty} f_Y(x)\ dx$ **converge car** f_Y **est une densité et :** $\displaystyle\int_{-\infty}^{+\infty} f_Y(x)\ dx = 1$
- **1 pt :** $\displaystyle\int_{-\infty}^{+\infty} x^2\, f_Y(x)\ dx$ **converge car** Y **admet un moment d'ordre** 2
- **1 pt : par formule de Koenig-Huygens** $\mathbb{E}(Y^2) = \mathbb{V}(Y) + \big(\mathbb{E}(Y)\big)^2 = 1$
- **1 pt :** $\dfrac{1}{\ln(2)} = \log_2(\mathbf{e})$

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.

- **1 pt :** $\log_2\big(f_Z(x)\big)\, f_Z(x) = \log_2(\lambda\, f_Z(x) - \dfrac{\lambda}{\ln(2)}\ x\, f_Z(x)$
- **1 pt : comme** f_Z **est une densité et est nulle en dehors de** $[0, +\infty[$**,** $\displaystyle\int_0^{+\infty} f_Z(x)\ dx$ **converge et vaut** 1
- **1 pt :** $\displaystyle\int_0^{+\infty} x\, f_Z(x)\ dx$ **converge car** Z **admet une espérance et** $\displaystyle\int_0^{+\infty} x\, f_Z(x)\ dx = \mathbb{E}(Z) = \dfrac{1}{\lambda}$
- **1 pt :** $h(Z) = \log_2\left(\frac{\mathbf{e}}{\lambda}\right)$

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 $\mathbb{R}$.

- **1 pt :** f **est continue sur** $\mathbb{R}$
- **1 pt :** $\forall x \in \mathbb{R},\ f(x) \geqslant 0$
- **2 pts :** $\displaystyle\int_{-\infty}^{+\infty} f(x)\ dx$ **converge et vaut** 1
  - × **1 pt : convergence de** $\displaystyle\int_0^{+\infty} f(x)\ dx$ **par calcul**
  - × **1 pt : convergence de** $\displaystyle\int_{-\infty}^{0} f(x)\ dx$ **par parité de** f

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

- **2 pts :** $\log_2\big(f(x)\big)\, f(x) = \big(-1 + \log_2(\lambda)\big)\, f(x) - \dfrac{\lambda}{\ln(2)}\ |x|\, f(x)$
- **0 pt :** $\displaystyle\int_{-\infty}^{+\infty} f(x)\ dx$ **converge car** f **est une densité**

- **2 pts :** $\displaystyle\int_{-\infty}^{+\infty} |x|\, f(x)\ dx$ **converge**

  × **1 pt :** $\displaystyle\int_{0}^{+\infty} |x|\, f(x)\ dx$ **converge car** Z **admet une espérance et** $\displaystyle\int_{0}^{+\infty} |x|\, f(x)\ dx = \mathbb{E}(Z) = \frac{1}{\lambda}$

  × **1 pt :** $\displaystyle\int_{-\infty}^{0} |x|\, f(x)\ dx$ **converge par partié de** $h : x \mapsto |x|\, f(x)$

- **1 pt :** $h(W) = 1 + \log_2 \left(\dfrac{\mathbf{e}}{\lambda}\right)$

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.

- **1 pt : en particulier** $1 \cdot X + 0 \cdot Y$ **suit une loi normale centrée**
- **1 pt :** $\mathbb{V}(X) = \sigma^2$ **donc** $X \hookrightarrow \mathcal{N}\left(0, \sigma^2\right)$

b) Calculer $h(X)$.

- **1 pt :** $X \hookrightarrow \mathcal{N}\left(0, \sigma^2\right)$ **donc** X **suit la même loi que** $\sigma\, Z$ **où** $Z \hookrightarrow \mathcal{N}(0, 1)$
- **1 pt : d'après *4.b)* et *3.b)(ii)*,** X **admet une entropie différentielle et** $h(X) = \dfrac{1}{2}\ \log_2(2\,\pi\,\mathbf{e}\,\sigma^2)$

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.

- **1 pt :** X **et** Y **suivent une loi normale, donc admettent un moment d'ordre** 2

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

- **1 pt :** $\lambda^2\, \mathbb{E}(Y^2) + 2\lambda\, \mathbb{E}(XY) + \mathbb{E}(X^2) = \mathbb{E}\left((\lambda\, Y + X)^2\right)$
- **1 pt :** $\mathbb{E}\left((\lambda\, Y + X)^2\right) \geqslant 0$ **par croissance de l'espérance**

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

- **1 pt :** $g : \lambda \mapsto \lambda^2\, \mathbb{E}(Y^2) + 2\lambda\, \mathbb{E}(XY) + \mathbb{E}(X^2)$ **fonction polynomiale de degré** 2 **de signe constant (positif) donc** $\Delta \leqslant 0$
- **1 pt :** $\Delta \leqslant 0\ \Leftrightarrow\ \big(\mathbb{E}(XY)\big)^2 \leqslant \mathbb{E}(Y^2)\, \mathbb{E}(X^2)$

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

- **1 pt :** $\big(\mathbb{E}(XY)\big)^2 \leqslant \mathbb{E}(Y^2)\ \mathbb{E}(X^2) \Leftrightarrow |\mathbb{E}(XY)| \leqslant \sqrt{\mathbb{E}(Y^2)}\ \sqrt{\mathbb{E}(X^2)}$ **par stricte croissance de** $\sqrt{\cdot}$ **sur** $[0, +\infty[$
- **1 pt :** $\mathbb{E}(X^2) = \mathbb{V}(X) + \big(\mathbb{E}(X)\big)^2 = \sigma^2$
- **1 pt :** $|\mathbb{E}(XY)| \leqslant \sqrt{\mathbb{E}(Y^2)}\ \sqrt{\mathbb{E}(X^2)} \Leftrightarrow |\rho| \leqslant 1$ **car** $\sigma^2 > 0$

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

- **1 pt :** $\mathbb{E}(XY) = \mathbb{E}(X)\,\mathbb{E}(Y)$ **car X et Y sont indépendantes**
- **1 pt :** $\rho = 0$ **car X et Y centrées**

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

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

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

- **1 pt :** $h(X,Y) = 0 \;\Leftrightarrow\; 2\pi\,\mathbf{e}\,\sigma^2\,\sqrt{1-\rho^2} = 1$
- **1 pt :** $2\pi\,\mathbf{e}\,\sigma^2\,\sqrt{1-\rho^2} = 1 \;\Leftrightarrow\; \rho \in \left\{-\sqrt{1-\dfrac{1}{(2\pi\mathbf{e}\sigma^2)^2}}\;,\;\sqrt{1-\dfrac{1}{(2\pi\mathbf{e}\sigma^2)^2}}\right\}$

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

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

Calculer $I(X,Y)$.

- **2 pts :** $I(X,Y) = -\dfrac{1}{2}\,\log_2(1-\rho^2)$ **(avec *5.b)* et *1.a)*)**

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

- **2 pts :** $I(X,Y) \geqslant 0 \;\Leftrightarrow\; \rho^2 \geqslant 0$ **et cette dernière propriété est vraie.**

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

- **1 pt : d'après *5.d)(ii)* :** $I(X,Y) = -\dfrac{\ln(1-\rho^2)}{2\,\ln(2)}$
- **1 pt :** $\displaystyle\lim_{\rho\to 1} I(X,Y) = +\infty$

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

- **1 pt : la v.a.r. $g(X)$ admet une espérance car c'est une v.a.r. finie**
- **1 pt : par théorème de transfert :** $\mathbb{E}\big(g(X)\big) = \displaystyle\sum_{k=0}^{n} \log_2\big(\mathbb{P}([X=k])\big)\;\mathbb{P}([X=k]) = H(X)$

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

- **1 pt :** $\forall k \in [\![0,n]\!]$**,** $0 < \mathbb{P}([X=k]) \leqslant 1$ **car X de support $[\![0,n]\!]$**
- **1 pt :** $0 < \mathbb{P}([X=k]) \leqslant 1$ **donc** $\mathbb{P}([X=k])\;\log_2\big(\mathbb{P}([X=k])\big) \leqslant 0$
- **1 pt :** $H(X) \geqslant 0$ **par sommation**

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, associie $H(X)$.

- **1 pt :** $H(X) = -(1-p)\ \log_2(1-p) - p\ \log_2(p)$

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

- **1 pt :** ψ **de classe** $\mathcal{C}^2$ **sur** $]0,1[$
- **1 pt :** $\psi'(p) = \log_2(1-p) - \log_2(p)$
- **1 pt :** $\psi''(p) = -\dfrac{1}{\ln(2)}\left(\dfrac{1}{1-p} + \dfrac{1}{p}\right) \leqslant 0$

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

- **1 pt :** $\psi'(p) \geqslant 0 \ \Leftrightarrow\ p \leqslant \dfrac{1}{2}$**, donc** ψ **est maximale en** $p_0 = \dfrac{1}{2}$

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 $H(X)$.

- **2 pts :** $H(X) = \dfrac{7}{4}$

7. On souhaite écrire une fonction en **Python** 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 la liste `P` de **Python** est telke que pour tout k de A, `P[k]` $= \mathbb{P}([X=k])$. Compléter la fonction ci-dessous d'argument `P` qui renvoie l'entropie de X, c'est-à-dire $-\sum\limits_{k=0}^{n} \mathbb{P}([X=k])\ \log_2\big(\mathbb{P}([X=k])\big)$.

```
def Entropie(P) :
    ...
    return h
```

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

- **4 pts :**
  - × **1 pt : initialisations**
  - × **2 pts : boucle** `for`
  - × **1 pt :** `h = -S`

```
import numpy as np
def Entropie(P) :
    n = len(P)
    S = 0
    for i in range(n) :
        S = S + P[k] * (np.log(P[k]) / np.log(2))
    h = -S
    return h
```

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 $\{x_1, x_2, \dots, x_N\}$ 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 : $p_i < 1$.

- **1 pt : $(\,[X = x_k]\,)_{k \in [\![1,N]\!]}$ SCE donc $\sum\limits_{k=1}^{N} \mathbb{P}([X = x_k]) = 1$**
- **1 pt : $\forall k \in [\![1, N]\!]$, $p_k > 0$ car le support de X est $\{x_1, \dots, x_N\}$**
- **1 pt : $p_i = 1 - \sum\limits_{\substack{k=1 \\ k \neq i}}^{N} p_k < 1$**

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 $\mathrm{Card}(A) = N$, on a : $\mathbb{E}\big(\varphi(X)\big) \geqslant \varphi\big(\mathbb{E}(X)\big)$**

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

- **1 pt : X et $\varphi(X)$ admettent des espérances car ce sont des v.a.r. finies**
- **1 pt : par théorème de transfert : $\mathbb{E}\big(\varphi(X)\big) = \lambda\,\varphi(x_1) + (1-\lambda)\,\varphi(x_2)$ où $\lambda = \mathbb{P}([X = x_1])$**
- **1 pt : $\varphi\big(\mathbb{E}(X)\big) = \varphi(\lambda\, x_1 + (1-\lambda)\, x_2)$**
- **1 pt : comme $\lambda \in [0, 1]$, on conclut par convexité de φ**

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 $\forall i \in [\![1, N-1]\!]$, $0 < p'_i < 1$.

- **1 pt : $(\,[X = x_i]\,)_{i \in [\![1,N]\!]}$ est un SCE donc : $\sum\limits_{i=1}^{N} p_i = 1$.**
- **1 pt : $\sum\limits_{i=1}^{N-1} p'_i = \dfrac{1}{1-p_N} \sum\limits_{i=1}^{N-1} p_i = \dfrac{1}{1-p_N}\,(1-p_N) = 1$**
- **1 pt : $\forall k \in [\![1, N]\!]$, $p_k > 0$ (car le support de X est $\{x_1, \dots, x_N\}$), donc $p'_k > 0$**
- **1 pt : $\forall k \in [\![1, N]\!]$, $p'_k < 1$ avec le même raisonnement qu'en question *8.a)***

(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 : $\sum\limits_{i=1}^{N-1} p'_i\,\varphi(x_i) \geqslant \varphi\left(\sum\limits_{i=1}^{N-1} p'_i\, x_i\right)$.

- **1 pt : $\mathbb{E}\big(\varphi(Y)\big) = \sum\limits_{i=1}^{N-1} \varphi(x_i)\, p'_i$ et $\varphi\big(\mathbb{E}(Y)\big) = \varphi\left(\sum\limits_{i=1}^{N-1} x_i\, p'_i\right)$**
- **1 pt : on peut appliquer $\mathcal{P}(N-1)$**
  0 si les hypothèses ne sont pas vérifiées (φ convexe et Y v.a.r. de loi à support $A = \{x_1, \dots, x_{N-1}\} \subset \mathbb{R}_+$ avec $\mathrm{Card}(A) = N-1$)

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

- **1 pt :** $\mathbb{E}\big(\varphi(X)\big) = \sum\limits_{i=1}^{N} \varphi(x_i)\, p_i = (1-p_N) \sum\limits_{i=1}^{N-1} \varphi(x_i)\, p'_i + \varphi(x_N)\, p_N$
- **1 pt : d'après la question précédente :** $\mathbb{E}\big(\varphi(X)\big) \leqslant (1-p_N)\ \varphi\left(\sum\limits_{i=1}^{N-1} x_i\, p'_i\right) + p_N\, \varphi(x_N)$
- **1 pt : par convexité de φ :** $(1-p_N)\, \varphi\left(\sum\limits_{i=1}^{N-1} x_i\, p'_i\right) + p_N\, \varphi(x_n) \ \geqslant\ \varphi\left((1-p_N) \sum\limits_{i=1}^{N-1} x_i\, p'_i + p_N\, x_N\right)$
- **1 pt :** $\varphi\left((1-p_N) \sum\limits_{i=1}^{N-1} x_i\, p'_i + p_N\, x_N\right) = \varphi\big(\mathbb{E}(X)\big)$

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

- **1 pt : si φ concave, alors $-\varphi$ convexe et on peut appliquer l'inégalité de Jensen**

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

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

- **4 pts :** $\sum\limits_{k=0}^{n} p_k \log_2\left(\dfrac{1}{(n+1)\, p_k}\right) \ \leqslant\ \log_2\left(\sum\limits_{k=0}^{n} \dfrac{p_k}{(n+1)\, p_k}\right)$
  - × **1 pt : introduction de la v.a.r.** $Y : \omega \mapsto \begin{cases} \dfrac{1}{(n+1)\, p_0} & \textbf{si } X(\omega) = 0 \\ \quad \vdots & \\ \dfrac{1}{(n+1)\, p_n} & \textbf{si } X(\omega) = n \end{cases}$
  - × **1 pt : on peut appliquer l'inégalité de Jensen à Y et $\log_2$ (qui est concave sur $\mathbb{R}_+^*$ d'après *1.c)*)**
  - × **1 pt :** $\mathbb{E}(Y) = \sum\limits_{k=0}^{n} \dfrac{p_k}{(n+1)\, p_k}$
  - × **1 pt :** $\mathbb{E}\big(\log_2(Y)\big) = \sum\limits_{k=0}^{n} \log_2\left(\dfrac{1}{(n+1)\, p_k}\right) p_k$
- **1 pt :** $\log_2\left(\sum\limits_{k=0}^{n} \dfrac{p_k}{(n+1)\, p_k}\right) = 0$

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

- **2 pts (dont 1 pt pour $(\,[X = k]\,)_{k \in [\![0,n]\!]}$ SCE)**

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

- **1 pt :** $H(X) = \log_2(n+1) + \sum\limits_{k=0}^{n} p_k \ \log_2\left(\dfrac{1}{(n+1)\, p_k}\right)$
- **1 pt :** $\sum\limits_{k=0}^{n} p_k \ \log_2\left(\dfrac{1}{(n+1)\, p_k}\right) \leqslant 0$ **d'après *9.a)***

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

- **1 pt :** $H(X) = \log_2(N+1)$

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

- **1 pt : FPT sur le SCE** $(\,[X = k]\,)_{k\in[\![0,n]\!]}$
- **1 pt : indépendance de** X **et** Y
- **1 pt :** X **et** Y **ont même loi**

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

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

- **1 pt :** $\varphi : x \mapsto 2^{-x}$ **convexe sur** $\mathbb{R}_+$
- **1 pt :** $Z = -\log_2\big(v(X)\big)$ **v.a.r. finie et de loi à support** $A \subset \mathbb{R}_+$
- **1 pt : par inégalité de Jensen :** $\mathbb{E}\left(2^{\log_2\left(v(X)\right)}\right) \geqslant \mathbb{E}\left(2^{\log_2\left(v(X)\right)}\right)$
- **1 pt :** $2^{\log_2\left(v(X)\right)} = v(X)$

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

- **1 pt : d'après *6.a)* :** $H(X) = -\mathbb{E}\Big(\log_2\big(v(X)\big)\Big) = 2^{-H(X)}$
- **1 pt :** $\mathbb{E}\big(v(X)\big) = \mathbb{P}([X = Y])$

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

- **1 pt : on considère** X **et** Y **indépendantes de loi** $\mathcal{U}([\![0, n]\!])$
- **1 pt : d'après *9.d)*,** $2^{-H(X)} = \dfrac{1}{n+1}$
- **1 pt : d'après *10.a)*,** $\mathbb{P}([X = Y]) = \dfrac{1}{n+1}$

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 $g : \{0, 1, \dots, n\}^2 \to \mathbb{R} \cup \{-\infty\}$ 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 : $H(X, Y) = -\mathbb{E}\big(g(X, Y)\big)$.

- **1 pt :** $g(X, Y)$ **v.a.r. finie donc admet une espérance**
- **1 pt : par théorème de transfert,** $\mathbb{E}\big(g(X, Y)\big) = -H(X, Y)$

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

- **1 pt**

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

- **1 pt :** $H(Y \mid X = k) = -\dfrac{1}{\mathbb{P}([X=k])} \left(\sum_{j=0}^{n} \mathbb{P}([X=k] \cap [Y=j]) \ \log_2 \big(\mathbb{P}([X=k] \cap [Y=j])\big) \log_2 \big(\mathbb{P}([X = \right.$
- **1 pt : par FPT** $\mathbb{P}([X=k]) = \sum_{j=0}^{n} \mathbb{P}([X=k] \cap [Y=j])$
- **1 pt :** $H(Y \mid X) = H(X,Y) - H(X)$

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

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

- **1 pt :** $H(X,Y) = H(Y,X)$ **d'après *11.b)***
- **1 pt : d'après *11.c)*,** $H(X) + H(Y \mid X) = H(Y) + H(X \mid Y)$

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 : $H(X) = \dfrac{7}{4}$.

- **2 pts : par FPT sur le SCE** $\big(\, [Y=j] \,\big)_{j \in [\![0,3]\!]}$ **:**

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

- **1 pt : d'après *6.d)*,** $H(X) = \dfrac{7}{4}$

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

- **1 pt : par FPT sur le SCE** $(\,[X=k]\,)_{k\in[\![0,3]\!]}$**,** $Y \hookrightarrow \mathcal{U}([\![0,3]\!])$
- **1 pt : d'après *9.d)*,** $H(Y)=2$

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

- **1 pt : on a les valeurs** $\mathbb{P}_{[Y=j]}([X=k])$ **suivantes :**

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

- **1 pt :** $H(X\,|\,Y=0) = H(X\,|\,Y=1) = \dfrac{7}{4}$
- **1 pt :** $H(X\,|\,Y=2) = 2$
- **1 pt :** $H(X\,|\,Y=3) = 0$
- **1 pt :** $H(X\,|\,Y) = \dfrac{11}{8}$

d) Que vaut $H(Y\,|\,X)$?

- **1 pt : d'après *11.d)*,** $H(Y\,|\,X) = \dfrac{13}{8}$

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

- **1 pt : d'après *11.c)*,** $H(X,Y) = \dfrac{27}{8}$

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

- **1 pt**

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

- **1 pt :** $\log_2\left(\dfrac{\mathbb{P}([X=k]\cap[Y=j])}{\mathbb{P}([X=k])\ \mathbb{P}([Y=j])}\right) = \log_2\left(\mathbb{P}_{[Y=j]}([X=k])\right) - \log_2\left(\mathbb{P}([X=k])\right)$
- **1 pt :** $\sum\limits_{k=0}^{n}\left(\sum\limits_{j=0}^{n}\mathbb{P}([X=k]\cap[Y=j])\ \log_2\left(\mathbb{P}_{[Y=j]}([X=k])\right)\right) = -H(X\mid Y)$
- **1 pt :** $-\sum\limits_{k=0}^{n}\left(\log_2\left(\mathbb{P}([X=k])\right)\ \sum\limits_{j=0}^{n}\mathbb{P}([X=k]\cap[Y=j])\right) = H(X)$ **(car** $(\,[Y=j]\,)_{j\in[\![0,n]\!]}$ **est un SCE)**

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

- **1 pt :** $\sum\limits_{\substack{j=0\\ j\neq k}}^{n}\mathbb{P}([X=k]\cap[X=j])\ \log_2\left(\dfrac{\mathbb{P}([X=k]\cap[X=j])}{\mathbb{P}([X=k])\,\mathbb{P}([X=j])}\right) = 0$
- **1 pt :** $\sum\limits_{\substack{j=0\\ j=k}}^{n}\mathbb{P}([X=k]\cap[X=j])\ \log_2\left(\dfrac{\mathbb{P}([X=k]\cap[X=j])}{\mathbb{P}([X=k])\,\mathbb{P}([X=j])}\right) = -\mathbb{P}([X=k])\ \log_2\left(\mathbb{P}([X=k])\right)$
- **1 pt :** $I(X,X) = H(X)$

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

- **1 pt : si X et Y indépendantes,** $\log_2\left(\dfrac{\mathbb{P}([X=k]\cap[Y=k])}{\mathbb{P}([X=k])\ \mathbb{P}([Y=j])}\right) = 0$
- **1 pt :** $I(X,Y) = 0$

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

- **1 pt**

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 :

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

- **1 pt :** $Z(\Omega) \subset \mathbb{R}_+$
- **1 pt : hypothèses inégalité de de Jensen pour obtenir** $\mathbb{E}\big(\log_2(Z_k)\big) \leqslant \log_2\left(\mathbb{E}(Z_k)\right)$
- **1 pt :** $\mathbb{E}(Z_k) = 1$

c) En déduire : $I(X,Y) \geqslant 0$.

- **1 pt :** $\mathbb{E}\big(\log_2(Z_k)\big) = -\dfrac{1}{\mathbb{P}([X=k])}\ \sum\limits_{j=0}^{n}\mathbb{P}([X=k]\cap[Y=j])\ \log_2\left(\dfrac{\mathbb{P}([X=k]\cap[Y=j])}{\mathbb{P}([X=k])\ \mathbb{P}([Y=j])}\right)$
- **1 pt : comme $\mathbb{P}([X=k]) \geqslant 0$:** $\sum\limits_{j=0}^{n}\mathbb{P}([X=k]\cap[Y=j])\ \log_2\left(\dfrac{\mathbb{P}([X=k]\cap[Y=j])}{\mathbb{P}([X=k])\ \mathbb{P}([Y=j])}\right) \geqslant 0$