Dans ce problème, n désignera un entier supérieur ou égal à 2. On considérera l'ensemble

$$\mathscr{A}_n = \{\boldsymbol{p} = (p_1, \dots, p_n) \in \mathbb{R}^n |\ \forall i,\ p_i \geqslant 0, \ \text{ et } \ p_1 + \cdots + p_n = 1\}.$$

On utilisera également la fonction définie par $x \mapsto x \ln x$, pour $x > 0$, et prolongée en une fonction continue en posant $0 \ln 0 = 0$. On définit alors sur $\mathscr{A}_n$ l'entropie

$$H_n(\boldsymbol{p}) = H_n(p_1, \dots, p_n) = -\sum_{k=1}^{n} p_k \ln p_k.$$

Partie I

Dans cette partie, $\varphi : I \to \mathbb{R}$ désigne une fonction de classe $\mathscr{C}^2$ définie sur un intervalle I de $\mathbb{R}$ à valeurs réelles. On suppose que pour tout $x \in I$, $\varphi''(x) \geqslant 0$ (on dira alors que φ est convexe).

1) Montrer que les fonctions $x \mapsto -\ln x$ et $x \mapsto x \ln x$ sont convexes sur $]0, +\infty[$.

2) Montrer que pour tout $x, y \in I$, $\varphi(y) \geqslant \varphi(x) + (y - x)\varphi'(x)$ (on pourra utiliser une formule de Taylor).

3) Soit $n \geqslant 2$ un entier, $x_1, \dots, x_n \in I$ et $(p_1, \dots, p_n) \in \mathscr{A}_n$. En utilisant la question précédente avec $x = \sum_{i=1}^{n} p_i x_i$, montrer que
$$\varphi\left(\sum_{i=1}^{n} p_i x_i\right) \leqslant \sum_{i=1}^{n} p_i \varphi(x_i).$$

4) Soient $a < b$ deux réels, et $\rho : [a, b] \to \mathbb{R}$ une fonctions continue telle que $\rho \geqslant 0$ et $\displaystyle\int_a^b \rho = 1$. Soit $f : [a, b] \to I$ une fonction continue à valeurs dans I. Montrer que
$$\varphi\left(\int_a^b f(x)\rho(x)dx\right) \leqslant \int_a^b \varphi(f(x))\rho(x)dx.$$

Partie II

1) Soit $n \geqslant 2$ et $\boldsymbol{p} = (p_1, \dots p_n) \in \mathscr{A}_n$. Montrer que $0 \leqslant H_n(\boldsymbol{p}) \leqslant \ln n$. (On pourra considérer d'abord le cas où $p_i > 0$ pour tout $i = 1, \dots, n$.)

2) Soient $\boldsymbol{p}, \boldsymbol{q} \in \mathscr{A}_n$ et $\lambda \in [0, 1]$. Vérifier que $\lambda\boldsymbol{p} + (1 - \lambda)\boldsymbol{q} \in \mathscr{A}_n$, et montrer que
$$H_n(\lambda\boldsymbol{p} + (1 - \lambda)\boldsymbol{q}) \geqslant \lambda H_n(\boldsymbol{p}) + (1 - \lambda)H_n(\boldsymbol{q}).$$

3) Soient $n, m \in \mathbb{N}^*$. Pour $\boldsymbol{p} = (p_1, \dots, p_n) \in \mathscr{A}_n$ et $\boldsymbol{q} = (q_1, \dots q_m) \in \mathscr{A}_m$, on note $\boldsymbol{p} \otimes \boldsymbol{q}$ le vecteur de $\mathbb{R}^{nm}$ défini par $\boldsymbol{p} \otimes \boldsymbol{q} = (p_1 q_1, \dots, p_1 q_m, p_2 q_1, \dots p_2 q_m, \dots, p_n q_m)$.
Vérifier que $\boldsymbol{p} \otimes \boldsymbol{q} \in \mathscr{A}_{nm}$ et que $H_{nm}(\boldsymbol{p} \otimes \boldsymbol{q}) = H_n(\boldsymbol{p}) + H_m(\boldsymbol{q})$.

4) Soit $n \geqslant 3$ et $\boldsymbol{p} = (p_1, \dots, p_n) \in \mathscr{A}_n$ tel que $p_1 + p_2 > 0$. Montrer que
$$H_n(p_1, \dots, p_n) = H_{n-1}(p_1 + p_2, p_3, \dots, p_n) + (p_1 + p_2)H_2\left(\frac{p_1}{p_1 + p_2}, \frac{p_2}{p_1 + p_2}\right).$$

5) Pour $r \in \mathbb{R}_+^* \setminus \{1\}$ et $\boldsymbol{p} = (p_1, \dots, p_n) \in \mathscr{A}_n$, on définit $H_n^r(\boldsymbol{p}) = \dfrac{1}{1 - r} \ln\left(\displaystyle\sum_{i=1}^{n} p_i^r\right)$. Montrer que
$$H_n^r(\boldsymbol{p}) \to H_n(\boldsymbol{p}) \quad \text{quand} \quad r \to 1.$$

Partie III

Le but de cette partie est de montrer que les seules fonctionnelles vérifiant les propriétés énoncées dans la Partie **II** sont de la forme cH_n.

Soit donc $J_n : \mathscr{A}_n \to \mathbb{R}_+$ une suite de fonctions, définie pour $n \geqslant 2$ et vérifiant les propriétés suivantes :

(P1) Continuité : la fonction $p \mapsto J_2(p, (1-p))$ est continue.

(P2) Symétrie : pour tout $n \geqslant 2$, et pour toute permutation σ de $\{1, \dots, n\}$, et pour tout $\boldsymbol{p} = (p_1, \dots, p_n) \in \mathscr{A}_n$, $J_n(p_1, \dots, p_n) = J_n(p_{\sigma(1)}, \dots, p_{\sigma(n)})$.

(P3) Maximalité : pour tous $n \geqslant 2$, $\boldsymbol{p} = (p_1, \dots, p_n) \in \mathscr{A}_n$, $J_n(p_1, \dots, p_n) \leqslant J_n(1/n, \dots, 1/n)$.

(P4) Extensibilité : pour tout $\boldsymbol{p} = (p_1, \dots, p_n) \in \mathscr{A}_n$, $J_{n+1}(p_1, \dots, p_n, 0) = J_n(p_1, \dots, p_n)$.

(P5) Additivité : pour tous $n, m \geqslant 2$, $\boldsymbol{p} \in \mathscr{A}_n$, $\boldsymbol{q} \in \mathscr{A}_m$, $J_{nm}(\boldsymbol{p} \otimes \boldsymbol{q}) = J_n(\boldsymbol{p}) + J_m(\boldsymbol{q})$.

(P6) Récursivité : pour tous $n \geqslant 3$, $\boldsymbol{p} = (p_1, \dots, p_n) \in \mathscr{A}_n$ tel que $p_1 + p_2 > 0$,

$$J_n(p_1, \dots, p_n) = J_{n-1}(p_1 + p_2, p_3, \dots, p_n) + (p_1 + p_2) J_2\left(\frac{p_1}{p_1 + p_2}, \frac{p_2}{p_1 + p_2}\right).$$

1) On pose pour tout $n \geqslant 2$, $\psi(n) = J_n\left(\frac{1}{n}, \dots, \frac{1}{n}\right)$. Montrer que la suite $(\psi(n))_{n \geqslant 2}$ est croissante, et que pour tous $n, m \geqslant 2$, $\psi(nm) = \psi(n) + \psi(m)$.

Dans toute la suite, on pose $c = \psi(2)/\ln 2$.

2) a) Si $c = 0$, montrer que $\psi(n) = 0$ pour tout $n \geqslant 2$.

b) On suppose $c \neq 0$. Soit $n \geqslant 2$. En encadrant, pour tout entier r, n^r entre deux puissances de 2, montrer que $\psi(n) = c \ln n$.

3) Montrer par récurrence sur $n \geqslant 2$, que pour tous $(p_1, p_2) \in \mathscr{A}_2$ tels que $p_1 > 0$ et $p_2 > 0$, et pour tous $(q_1, \dots, q_n)$ et $(r_1, \dots, r_n)$ dans $\mathscr{A}_n$, on a :

$$J_{2n}(p_1 q_1, \dots, p_1 q_n, p_2 r_1, \dots, p_2 r_n) = J_2(p_1, p_2) + p_1 J_n(q_1, \dots, q_n) + p_2 J_n(r_1, \dots, r_n).$$

4) a) Soient k et n deux entiers tels que $2 \leqslant k \leqslant n-2$, montrer que

$$\psi(n) = J_2\left(\frac{k}{n}, 1 - \frac{k}{n}\right) + \left(1 - \frac{k}{n}\right)\psi(n-k) + \frac{k}{n}\psi(k).$$

b) En déduire que pour tout $p \in [0, 1]$, on a

$$J_2(p, 1-p) = cH_2(p, 1-p) = -c\left(p \ln p + (1-p)\ln(1-p)\right).$$

5) Montrer par récurrence sur $n \geqslant 2$ que l'on a $J_n = cH_n$.

Partie IV

Soient $\boldsymbol{p} = (p_1, \dots, p_n)$ et $\boldsymbol{q} = (q_1, \dots, q_n)$ deux éléments distincts de $\mathscr{A}_n$. On suppose de plus que $q_i > 0$ pour tout i, et on définit l'entropie relative de $\boldsymbol{p}$ par rapport à $\boldsymbol{q}$ par

$$D_n(\boldsymbol{p} \| \boldsymbol{q}) = \sum_{i=1}^{n} p_i \ln \frac{p_i}{q_i}.$$

Soit $\mathscr{I} = \{i \in \{1, \dots, n\} | p_i \geqslant q_i\}$. On pose $\tilde{p} = \sum_{i \in \mathscr{I}} p_i$ et $\tilde{q} = \sum_{i \in \mathscr{I}} q_i$.

1) a) Montrer que $\mathscr{I} \neq \varnothing$ et que $\operatorname{Card} \mathscr{I} < n$.

b) Soit $\mathscr{J} \subset \{1, \ldots, n\}$ non vide. En utilisant la convexité de $-\ln$, montrer

$$\sum_{i \in \mathscr{J}} p_i \ln \frac{p_i}{q_i} \geqslant \left(\sum_{i \in \mathscr{J}} p_i \right) \ln \frac{\sum_{i \in \mathscr{J}} p_i}{\sum_{j \in \mathscr{J}} q_j}.$$

c) En déduire que $D_n(\boldsymbol{p}||\boldsymbol{q}) \geqslant 0$ et que $D_n(\boldsymbol{p}||\boldsymbol{q}) \geqslant \tilde{p} \ln \dfrac{\tilde{p}}{\tilde{q}} + (1 - \tilde{p}) \ln \dfrac{1 - \tilde{p}}{1 - \tilde{q}}$.

2) Montrer que $\displaystyle\sum_{i=1}^{n} |p_i - q_i| = 2(\tilde{p} - \tilde{q})$.

3) En utilisant la fonction $t \mapsto \tilde{p} \ln t + (1 - \tilde{p}) \ln(1 - t)$, montrer que

$$D_n(\boldsymbol{p}||\boldsymbol{q}) \geqslant \frac{1}{2} \left(\sum_{i=1}^{n} |p_i - q_i| \right)^2.$$