UC 313 J. 5007

SESSION 2003

---

**Filière MP**

**MATHÉMATIQUES**

---

Épreuve commune aux ENS de Paris et Cachan

---

Durée : 4 heures

---

*L'usage de calculatrices électroniques de poche à alimentation autonome, non imprimantes et sans document d'accompagnement, est autorisé. Cependant, une seule calculatrice à la fois est admise sur la table ou le poste de travail, et aucun échange n'est autorisé entre les candidats.*

Le but de ce problème est l'étude mathématique de la notion d'*entropie*. La partie I est consacrée à des questions préliminaires et les parties suivantes à l'étude des propriétés et des liens entre différentes définitions de l'entropie.

**Notations et rappels :**

On notera $\mathbb{R}$ l'ensemble des nombres réels, et $\mathbb{R}_+$ l'ensemble des nombres réels positifs ou nuls.

On dira qu'une fonction $f$ de $\mathbb{R}$ dans $\mathbb{R}$ est à support compact s'il existe un intervalle fermé borné $[a, b]$ de $\mathbb{R}$ tel que $f(x) = 0$ pour tout $x \notin [a, b]$.

On aura besoin de considérer la fonction $x \mapsto x \ln x$ définie pour $x > 0$, et par continuité, lorsque $x = 0$, on posera $x \ln x = 0$.

D'autre part, on rappelle la valeur de l'intégrale de Gauss:

$$\int_0^{+\infty} e^{-x^2} dx = \frac{\sqrt{\pi}}{2}.$$

## Partie I

Dans cette partie, $\varphi$ désigne une fonction de classe $C^2$ définie sur un intervalle $I$ de $\mathbb{R}$ et à valeurs réelles. On suppose que la dérivée seconde de $\varphi$ est positive sur $I$ (on dira alors que $\varphi$ est convexe).

**I-1.** A l'aide d'une formule de Taylor, montrer que

$$\forall x, y \in I \quad \varphi(y) \geqslant \varphi(x) + (y - x)\varphi'(x).$$

**I-2.** Soit un entier $n \geqslant 2$, soient $x_1, \ldots, x_n$ dans $I$ et soient $a_1, \ldots, a_n$ dans $[0, 1]$ tels que $\sum_{i=1}^n a_i = 1$. En utilisant la question précédente avec $x = \sum_{i=1}^n a_i x_i$, montrer que

$$\varphi(\sum_{i=1}^n a_i x_i) \leqslant \sum_{i=1}^n a_i \varphi(x_i).$$

**I-3.** Soient $f$ et $g$ deux fonctions continues définies sur un intervalle $[a, b]$ de $\mathbb{R}$ et à valeurs réelles. On suppose que: $f$ est positive sur $[a, b]$, $\int_a^b f(x)dx = 1$ et $g$ est à valeurs dans $I$. Montrer que

$$\varphi\left(\int_a^b g(x)f(x)dx\right) \leqslant \int_a^b \varphi(g(x))f(x)dx.$$

## Partie II

Pour tout entier $n \geqslant 2$, on note $\mathcal{A}_n$ l'ensemble des $\mathbf{p} = (p_1, ..., p_n)$ de $\mathbb{R}^n$ tels que pour tout $i$ entier entre 1 et $n$, $p_i \geqslant 0$ et tels que $\sum_{i=1}^n p_i = 1$. On définit alors sur $\mathcal{A}_n$ la fonction *entropie*, notée $H_n$, par

$$H_n(\mathbf{p}) = H_n(p_1, ..., p_n) = -\sum_{i=1}^{n} p_i \ln p_i.$$

**II-1.** Soit un entier $k \geqslant 2$, et soit $\mathbf{p} = (p_1, ..., p_k)$ dans $\mathcal{A}_k$ tel que pour tout $i$ entre 1 et $k$, $p_i > 0$. A l'aide de la convexité de la fonction $x \mapsto -\ln x$, montrer que

$$0 \leqslant H_k(\mathbf{p}) \leqslant \ln k.$$

En déduire que pour tout entier $n \geqslant 2$. et tout $\mathbf{p} \in \mathcal{A}_n$, on a: $0 \leqslant H_n(\mathbf{p}) \leqslant \ln(n)$.

**II-2.** Soient $\mathbf{p}$ et $\mathbf{q}$ dans $\mathcal{A}_n$, et $\lambda \in [0,1]$. Vérifier que le vecteur $\lambda\mathbf{p} + (1-\lambda)\mathbf{q}$ appartient aussi à $\mathcal{A}_n$, et montrer (en utilisant la convexité de la fonction $x \mapsto x \ln x$) que

$$H_n(\lambda\mathbf{p} + (1-\lambda)\mathbf{q}) \geqslant \lambda H_n(\mathbf{p}) + (1-\lambda) H_n(\mathbf{q}).$$

**II-3.a.** Pour $\mathbf{p} = (p_1, ..., p_n)$ dans $\mathcal{A}_n$ et $\mathbf{q} = (q_1, ..., q_m)$ dans $\mathcal{A}_m$, on note $\mathbf{p.q}$ le vecteur de $\mathbb{R}^{nm}$ défini par $\mathbf{p.q} = (p_1q_1, \ldots, p_1q_m, p_2q_1, \ldots, p_2q_m, \ldots, p_nq_1, \ldots, p_nq_m)$. Vérifier que $\mathbf{p.q}$ appartient à $\mathcal{A}_{nm}$ et que

$$H_{nm}(\mathbf{p.q}) = H_n(\mathbf{p}) + H_m(\mathbf{q}).$$

**II-3.b.** Soit $n \geqslant 3$ et $\mathbf{p} = (p_1, ..., p_n)$ dans $\mathcal{A}_n$ tel que $p_1 + p_2 > 0$. Vérifier que

$$H_n(p_1, ..., p_n) = H_{n-1}(p_1 + p_2, p_3, ..., p_n) + (p_1 + p_2) H_2\left(\frac{p_1}{p_1 + p_2}, \frac{p_2}{p_1 + p_2}\right).$$

**II-4.** Soient $\mathbf{p} = (p_1, ..., p_n)$ et $\mathbf{q} = (q_1, ..., q_n)$ deux éléments distincts de $\mathcal{A}_n$. On suppose de plus que pour tout $i$, $q_i > 0$, et on définit alors l'*entropie relative* de $\mathbf{p}$ et $\mathbf{q}$ par

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

Soit $\mathcal{I}$ l'ensemble des indices $i$ entre 1 et $n$ tels que $p_i \geqslant q_i$. On pose alors $\tilde{p} = \sum_{i \in \mathcal{I}} p_i$ et $\tilde{q} = \sum_{i \in \mathcal{I}} q_i$.

**II-4.a.** Montrer que $\mathcal{I}$ est non vide et de cardinal strictement inférieur à $n$. Puis montrer, à l'aide de la convexité de la fonction $x \mapsto -\ln x$, que pour tout sous-ensemble $\mathcal{J}$ non vide de $\{1, \ldots, n\}$, on a:

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

En déduire alors que

$$D_n(\mathbf{p}, \mathbf{q}) \geqslant \tilde{p} \ln \frac{\tilde{p}}{\tilde{q}} + (1-\tilde{p}) \ln \frac{1-\tilde{p}}{1-\tilde{q}}.$$

**II-4.b.** Montrer que

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

**II-4.c.** A l'aide d'une étude de la fonction $t \mapsto \tilde{p} \ln t + (1-\tilde{p}) \ln(1-t)$, montrer que

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

**II-5.** Soit une suite de fonctions $J_n : \mathcal{A}_n \to \mathbb{R}_+$, définie pour $n \geqslant 2$ , et vérifiant les propriétés suivantes:
(P1) Continuité: la fonction $p \mapsto J_2(p, 1-p)$ est continue sur $[0, 1]$.
(P2) Symétrie: $\forall n \geqslant 2$, $\forall \sigma$ permutation de $\{1, \ldots, n\}$, et $\forall \mathbf{p} = (p_1, \ldots, p_n) \in \mathcal{A}_n$,

$$J_n(p_1, \ldots, p_n) = J_n(p_{\sigma(1)}, \ldots, p_{\sigma(n)}).$$

(P3) Maximalité: $\forall n \geqslant 2$, $\forall \mathbf{p} = (p_1, \ldots, p_n) \in \mathcal{A}_n$, $J_n(p_1, \ldots, p_n) \leqslant J_n(1/n, \ldots, 1/n)$.
(P4) Extensibilité: $\forall n \geqslant 2$, $\forall \mathbf{p} = (p_1, \ldots, p_n) \in \mathcal{A}_n$, $J_{n+1}(p_1, \ldots, p_n, 0) = J_n(p_1, \ldots, p_n)$.
(P5) Additivité: $\forall n, m \geqslant 2$, $\forall \mathbf{p} \in \mathcal{A}_n$, $\forall \mathbf{q} \in \mathcal{A}_m$, $J_{nm}(\mathbf{p}.\mathbf{q}) = J_n(\mathbf{p}) + J_m(\mathbf{q})$.
(P6) Récursivité: $\forall n \geqslant 3$, $\forall \mathbf{p} = (p_1, \ldots, p_n) \in \mathcal{A}_n$ tel que $p_1 + p_2 > 0$,

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

**II-5.a.** On pose pour tout $n \geqslant 2$, $\psi(n) = J_n(1/n, \ldots, 1/n)$. Montrer que la suite $(\psi(n))_{n \geqslant 2}$ est croissante, et que pour tous les entiers $n, m \geqslant 2$, on a $\psi(nm) = \psi(n) + \psi(m)$.

**II-5.b.** On pose dans toute la suite $c = \psi(2)/\ln 2$. Montrer tout d'abord que si $c = 0$ alors $\psi(n) = 0$ pour tout $n \geqslant 2$. On suppose maintenant $c \neq 0$. Soit un entier $n > 2$. En encadrant, pour tout entier $r$ non nul, $n^r$ entre des puissances de 2, montrer que $\psi(n) = c \ln(n)$.

**II-5.c.** Montrer par récurrence sur $n \geqslant 2$ que pour tout $(p_1, p_2) \in \mathcal{A}_2$ tel que $p_1 > 0$ et $p_2 > 0$, et tous $(q_1, \ldots, q_n)$ et $(r_1, \ldots, r_n)$ dans $\mathcal{A}_n$, on a:

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

**II-5.d.** Soient $k$ et $n$ deux entiers tels que $2 \leqslant k \leqslant n-2$, montrer que

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

Puis en déduire que pour tout $p \in [0, 1]$, on a

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

**II-5.e.** Montrer par récurrence sur $n$ que pour tout $n \geqslant 2$, on a $J_n = cH_n$.

## Partie III

Dans cette partie on notera $\mathcal{F}$ l'ensemble des fonctions $f$ de $\mathbb{R}$ dans $\mathbb{R}_+$, continues et telles que la fonction $x \mapsto x^2 f(x)$ soit bornée sur $\mathbb{R}$, l'intégrale $\int_{-\infty}^{+\infty} x^2 f(x)dx$ converge et $\int_{-\infty}^{+\infty} f(x)dx = 1$.
On notera de plus $\mathcal{F}_0$ l'ensemble des fonctions de $\mathcal{F}$ qui sont à support compact.

**III-1.** Pour tout $f$ appartenant à $\mathcal{F}$, on définit l'*entropie différentielle* de $f$ par:

$$H(f) = -\int_{-\infty}^{+\infty} f(x) \ln f(x) dx.$$

Montrer que cette intégrale est bien définie.

**III-2.** Soient $\mu$ un réel et $\sigma$ un réel strictement positif. On définit sur $\mathbb{R}$ la fonction $g_{\mu,\sigma}$ par

$$g_{\mu,\sigma}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}.$$

Montrer que $g_{\mu,\sigma}$ appartient à $\mathcal{F}$, et que

$$\int_{-\infty}^{+\infty} x g_{\mu,\sigma}(x)dx = \mu \quad \text{et} \quad \int_{-\infty}^{+\infty} (x-\mu)^2 g_{\mu,\sigma}(x)dx = \sigma^2.$$

En déduire la valeur de $H(g_{\mu,\sigma})$.

**III-3.a.** Soit $f \in \mathcal{F}$ telle que $\forall x \in \mathbb{R}$, $f(x) > 0$. Pour tout $\mu$ réel et tout $\sigma$ réel strictement positif, on pose

$$K_{\mu,\sigma}(f) = \int_{-\infty}^{+\infty} f(x) \ln \frac{f(x)}{g_{\mu,\sigma}(x)} dx.$$

Montrer que cette quantité est bien définie, et montrer de plus, à l'aide de la convexité de la fonction $t \mapsto -\ln t$, que $K_{\mu,\sigma}(f) \geqslant 0$.

**III-3.b.** En déduire que si on définit les réels $m$ et $s$ par $m = \int_{-\infty}^{+\infty} x f(x)dx$ et $s = \sqrt{\int_{-\infty}^{+\infty} (x-m)^2 f(x)dx}$, alors

$$H(f) \leqslant H(g_{m,s}).$$

## Partie IV

On reprend dans cette partie les notations des parties précédentes.

**IV-1.a.** Soit $f \in \mathcal{F}$. Pour tout réel $r \in ]1/2, +\infty[$, et $r \neq 1$, on définit l'*entropie de Renyi d'ordre* $r$ de $f$ par

$$h_r(f) = \frac{1}{1-r} \ln \left( \int_{-\infty}^{+\infty} f^r(x)\, dx \right).$$

Montrer que cette quantité est bien définie. La fonction $f$ étant fixée, on considère la fonction $F$ définie sur $]1/2, +\infty[$ par $r \mapsto F(r) = \int_{-\infty}^{+\infty} f^r(x)\, dx$. Montrer que $F$ est de classe $C^1$, et calculer $F'(1)$.

**IV-1.b.** En déduire que $h_r(f)$ tend vers $H(f)$ lorsque $r$ tend vers 1.

**IV-2.a.** Soient $f$ et $g$ dans $\mathcal{F}$. On appelle alors *produit de convolution* de $f$ et de $g$, que l'on note $f * g$, la fonction définie sur $\mathbb{R}$ par:

$$x \mapsto (f * g)(x) = \int_{-\infty}^{+\infty} f(x-y)g(y)\,dy.$$

Justifier la convergence de cette intégrale pour tout $x$ réel.

**IV-2.b.** Montrer que $f$ est uniformément continue sur $\mathbb{R}$. En déduire que $f * g$ est continue sur $\mathbb{R}$.

**IV-3.a.** On suppose dans la suite que $f$ et $g$ sont dans $\mathcal{F}_0$. Montrer alors que $f*g$ appartient aussi à $\mathcal{F}_0$.

**IV-3.b.** En utilisant la convexité de la fonction $t \mapsto t \ln t$, montrer que

$$H(f * g) \geqslant H(f).$$

**IV-4.a.** Soient $f$ et $g$ dans $\mathcal{F}_0$. On admet dans la suite le résultat suivant (inégalité de Young forte): pour tous $p$, $q$, $r$ réels strictement supérieurs à 1, tels que $\frac{1}{r} = \frac{1}{p} + \frac{1}{q} - 1$, alors

$$\left(\int_{-\infty}^{+\infty} (f * g)^r(x)dx\right)^{1/r} \leqslant \left(\frac{C(p)\,C(q)}{C(r)}\right)^{1/2} \left(\int_{-\infty}^{+\infty} f^p(x)dx\right)^{1/p} \left(\int_{-\infty}^{+\infty} g^q(x)dx\right)^{1/q},$$

où pour tout $x > 1$, on a posé $C(x) = \exp(\frac{1}{x}\ln x + \frac{x-1}{x}\ln\frac{x-1}{x})$ .

Soit $r \in ]1, +\infty[$ et $\lambda \in ]0, 1[$. On pose $p = \frac{r}{r+\lambda(1-r)}$ et $q = \frac{r}{r+(1-\lambda)(1-r)}$; montrer que

$$h_r(f * g) \geqslant \lambda h_p(f) + (1-\lambda)h_q(g) + \frac{1}{2}H_2(\lambda, 1-\lambda) + \frac{r}{2(r-1)}\left(\frac{1}{r}\ln r - \frac{1}{p}\ln p - \frac{1}{q}\ln q\right).$$

**IV-4.b.** On fixe $\lambda \in ]0, 1[$. En faisant tendre $r$ vers 1 dans l'inégalité ci-dessus, montrer que

$$H(f * g) \geqslant \lambda H(f) + (1-\lambda)H(g) + \frac{1}{2}H_2(\lambda, 1-\lambda).$$

**IV-4.c.** Utiliser le résultat de la question précédente avec $\lambda$ choisi de façon optimale pour en déduire que

$$\exp(2H(f * g)) \geqslant \exp(2H(f)) + \exp(2H(g)).$$