ÉPREUVE D'ANALYSE ET PROBABILITÉS
AGRÉGATION EXTERNE 2003

Corrigé

I. ENTROPIE ET VARIATION TOTALE

I.1. Signe de l'entropie.

a. La fonction Ψ est bien convexe sur $\mathbb{R}$: c'est en effet une fonction de classe C^2 sur $]0,+\infty[$ et l'on a $\Psi'(x) = 1 + \ln x$, $\Psi''(x) = 1/x > 0$, d'où la convexité de Ψ (une fonction de classe C^2 sur un intervalle I de $\mathbb{R}$ est convexe sur cet intervalle si et seulement si $\Psi'' \geq 0$ sur I).

b. On a, puisque $m > 0$ et $g > 0$ sur $[0,1]^n$,

$$\int_{[0,1]^n} m(x)g(x)dx > 0$$

(en effet, si cette intégrale était nulle, on aurait d'après la théorie de l'intégration $mg = 0$ presque partout sur $[0,1]^n$) ; on peut donc appliquer l'inégalité de Jensen (rappelée dans les préliminaires) avec pour intervalle $I =]0,+\infty[$, pour fonction f la fonction $f = m$ (à valeurs dans $]0,+\infty[$), telle que les intégrales concernées soient convergentes (car mg et $m \log m\, g$ sont par hypothèses intégrables au sens de Lebesgue) ; la fonction Ψ étant convexe sur I, on a

$$\begin{aligned}\Big(\int_{[0,1]^n} m(x)g(x)dx\Big) \quad &\times \quad \Big(\ln \int_{[0,1]^n} m(x)g(x)dx\Big) = \Psi\Big(\int_{[0,1]^n} m(x)g(x)dx\Big) \\ &\leq \quad \int_{[0,1]^n} \Psi(m(x))g(x)\,dx = \int_{[0,1]^n} m(x)\ln m(x)g(x)dx\end{aligned}$$

d'où $\mathrm{Ent}_g(m) \geq 0$.

I.2. Inégalités auxilaires.

a. On considère la fonction auxilaire

$$\theta \ : u \in]0,\infty] \to J(u) - \frac{(1-u)^2}{2}\,;$$

sur l'intervalle $]0,\infty[$, θ est dérivable, de dérivée

$$u \in]0,\infty[\to \ln u + 1 - u\,;$$

cette dernière fonction est strictement croissante sur $]0,1]$ car elle est dérivable sur $]0,1[$, de dérivée $u \to 1/u - 1 > 0$; comme $\theta'(1) = 0$, on a $\theta' < 0$

sur $]0,1[$, ce qui prouve que θ est strictement décroissante sur $]0,1]$; mais $\theta(1)=0$, d'où il suit $\theta \geq 0$ sur $]0,1]$; comme de plus $\theta(0)=1$, on a bien $\theta \geq 0$ sur $[0,1]$, d'où

$$J(u) \geq \frac{(1-u)^2}{2}$$

pour $u \in [0,1]$.

b. Si l'on pose, pour $u \geq 1$, $u = 1/v$, on a

$$J(u) = -\frac{\ln v}{v} - \frac{1}{v} + 1 = -\frac{1}{v}(\ln v + v - 1) = -\frac{J(v)}{v}\,;$$

on a donc

$$\frac{J(u)}{u} = v - 1 - \log v\ ;$$

pour $v \in]0,1[$, on a

$$\begin{aligned} \ln v = \ln\left(1-(1-v)\right) &= -(1-v) - \frac{(1-v)^2}{2} - \ldots - \frac{(1-v)^n}{n} - \ldots \\ &\leq -(1-v) - \frac{(1-v)^2}{2}\,, \end{aligned}$$

d'où, pour $v \in]0,1[$,

$$\ln v \leq v - 1 - \frac{(1-v)^2}{2}\,,$$

ou encore

$$v - 1 - \ln v \geq \frac{(1-v)^2}{2}\,,$$

inégalité restant valide pour $v \in]0,1]$; finalement, on a, pour $u = 1/v \in [1,+\infty[$,

$$K(u) = \frac{J(u)}{u} \geq \frac{1}{2}(1-v)^2 = \frac{1}{2}\Big(1-\frac{1}{u}\Big)^2.$$

c. On a, pour tout $u \in]0,+\infty[$,

$$J(u) = J(u)\mathbf{1}_{]0,1[}(u) + J(u)\mathbf{1}_{[1,+\infty[}(u)\,; \qquad (\dagger)$$

or, d'après **I.2.a**,

$$J(u)\mathbf{1}_{]0,1[}(u) \geq \frac{1}{2}(1-u)^2 \times \mathbf{1}_{]0,1[}(u) = \frac{1}{2}[1-u]_+^2$$

tandis que, d'après **I.2.b**,

$$J(u)\mathbf{1}_{[1,+\infty[}(u) = uK(u)\mathbf{1}_{[1,+\infty[}(u) \geq \frac{u}{2}\Big(1-\frac{1}{u}\Big)^2 \times \chi_{[1,\infty[}(u) = \frac{u}{2}\Big[1-\frac{1}{u}\Big]_+^2\,;$$

on a donc l'inégalité voulue en ajoutant dans (†).

d. On pose $m = h/g$ sur $[0,1]^n$ (ce qui est licite car $g > 0$) ; comme h est une densité, $h = mg$ est intégrable sur $[0,1]^n$, de même que $mg \ln m = h \ln(h/g)$ par hypothèses ; on peut donc bien définir l'entropie $\mathrm{Ent}_g(h/g)$. Par définition

$$\mathrm{Ent}_g(h/g) = \int_{[0,1]^n} \ln\Big(\frac{h(x)}{g(x)}\Big) h(x)dx - 1 \times \ln 1 = \int_{[0,1]^n} \ln\Big(\frac{h(x)}{g(x)}\Big) h(x)dx$$

puisque

$$\int_{[0,1]^n} \frac{h(x)}{g(x)} g(x)dx = \int_{[0,1]^n} h(x)dx = 1$$

(h étant une densité sur $[0,1]^n$).

On a, par définition de l'entropie

$$\begin{aligned}\mathrm{Ent}_g(tm) &= t\int_{[0,1]^n} m(x)(\ln t + \ln(m(x))g(x)dx \\ &\qquad -t\Big(\int_{[0,1]^n} m(x)g(x)dx\Big)\Big(\ln t + \int_{[0,1]^n} m(x)g(x)dx\Big) \\ &= t\mathrm{Ent}_g(m)\,.\end{aligned}$$

e. Il faut dans cette question que l'entropie $\mathrm{Ent}_r(q/r)$ soit bien définie (ce qui est implicite), donc que la fonction

$$x \to q(x)\ln\Big(\frac{q(x)}{r(x)}\Big)$$

soit intégrable sur $[0,1]^n$, ce que l'on supposera (en plus du fait que q et r sont des densités).

On utilise l'inégalité établie au **I.2.c** pour $u = r(x)/q(x) \in]0,+\infty[$ et $x \in [0,1]$; on a donc

$$\Big[1 - \frac{r(x)}{q(x)}\Big]_+^2 + \frac{r(x)}{q(x)}\Big[1 - \frac{q(x)}{r(x)}\Big]_+^2 \leq 2\bigg(\frac{r(x)}{q(x)} \times \ln\Big(\frac{r(x)}{q(x)}\Big) - \frac{r(x)}{q(x)} + 1\bigg)\,,$$

soit

$$\Big[1 - \frac{r(x)}{q(x)}\Big]_+^2 q(x) + \Big[1 - \frac{q(x)}{r(x)}\Big]_+^2 r(x) \leq 2\bigg(r(x) \times \ln\Big(\frac{r(x)}{q(x)}\Big) - r(x) + q(x)\bigg)\,;$$

par symétrie (relativement à q et r) du membre de gauche, on a, pour tout $x \in [0,1]^n$,

$$\Big[1 - \frac{r(x)}{q(x)}\Big]_+^2 q(x) + \Big[1 - \frac{q(x)}{r(x)}\Big]_+^2 r(x) \leq 2\bigg(q(x) \times \ln\Big(\frac{q(x)}{r(x)}\Big) - q(x) + r(x)\bigg)\,;$$

on intègre ensuite cette inégalité entre fonctions mesurables positives sur $[0,1]^n$; on trouve

$$\delta^2(r \mid q) + \delta^2(q \mid r) \leq 2 \int_{[0,1]^n} q(x) \ln \Big(\frac{q(x)}{r(x)}\Big) = 2\,\mathrm{Ent}_r(q/r)$$

d'après le résultat du **I.2.d**.

I.3. Formule pour d_{VT}.

a. On a

$$\begin{aligned} \int_{[0,1]^n} [q(x) - r(x)]_+ \, dx &= \int_{\{x \in [0,1]^n \,;\, r(x) \leq q(x)\}} (q(x) - r(x)) dx \\ &= -\int_{\{x \in [0,1]^n \,;\, r(x) \leq q(x)\}} (q(x) - r(x)) \, dx \\ &= \int_{[0,1]^n} [q(x) - r(x)]_- \, dx \end{aligned}$$

(où $[u]_- := \max(0, -u)$) puisque q et r ont même intégrale (1) sur $[0,1]^n$. Comme

$$|q(x) - r(x)| = [q(x) - r(x)]_+ + [q(x) - r(x)]_-$$

pour tout $x \in [0,1]^n$, on obtient en intégrant sur $[0,1]^n$

$$2 \int_{[0,1]^n} [q(x) - r(x)]_+ \, dx = \int_{[0,1]^n} |q(x) - r(x)| \, dx \,,$$

ce qui est l'inégalité voulue.

b. Soient $E := \{x \in [0,1]^n \,;\, r(x) \leq q(x)\}$; on a

$$\begin{aligned} \int_{[0,1]^n} \mathbf{1}_E(x) q(x) \, dx - \int_{[0,1]^n} \mathbf{1}_E(y) r(y) \, dy &= \int_E (q(x) - r(x)) \, dx \\ &= \int_{[0,1]^n} [q(x) - r(x)]_+ \, dx \end{aligned}$$

et par conséquent

$$\begin{aligned} d_{VT}(q, r) &\geq \Big| \int_{[0,1]^n} \mathbf{1}_E(x) q(x) \, dx - \int_{[0,1]^n} \mathbf{1}_E(y) r(y) \, dy \Big| \\ &\geq \int_{[0,1]^n} [q(x) - r(x)]_+ \, dx \\ &\geq \frac{1}{2} \int_{[0,1]^n} |q(x) - r(x)| \, dx \end{aligned}$$

d'après le **I.3.a**. Si maintenant f est une fonction borélienne sur $[0,1]^n$, on a, en découpant $[0,1]^n = E \cup ([0,1]^n \setminus E)$:

$$\begin{aligned}&\Big|\int_{[0,1]^n} f(x)q(x)\,dx - \int_{[0,1]^n} f(y)r(y)\,dy\Big|\\ &\quad = \Big|\int_{[0,1]^n} f(x)[q(x)-r(x)]_+\,dx - \int_{[0,1]^n} f(y)[r(y)-q(y)]_+\,dy\Big|.\end{aligned}$$

Si A et B sont deux nombres positifs, on a $|A-B| \leq \max(A,B)$; on a donc

$$\begin{aligned}&\Big|\int_{[0,1]^n} f(x)q(x)\,dx - \int_{[0,1]^n} f(y)r(y)\,dy\Big|\\ \leq &\quad \max\Big(\int_{[0,1]^n} f(x)[q(x)-r(x)]_+\,dx\,,\int_{[0,1]^n} f(x)[r(x)-q(x)]_+\,dx\Big)\\ \leq &\quad \frac{1}{2}\int_{[0,1]^n} |q(x)-r(x)|\,dx\,.\end{aligned}$$

En prenant le sup pour toutes les fonctions boréliennes de $[0,1]^n$ dans $[0,1]$, on a aussi la majoration

$$d_{VT}(q,r) \leq \frac{1}{2}\int_{[0,1]^n} |q(x)-r(x)|\,dx\,,$$

ce qui prouve, avec l'autre inégalité obtenue prédémment, l'égalité voulue.

II. Démonstration de l'inégalité principale pour $n = 1$

II.1. Unicité ?

Supposons que Π admette deux couples (k_1, k_2) et (l_1, l_2) de densités marginales ; on a donc, pour toute fonction $f \in \mathbf{B}_1$, pour toute fonction $g \in \mathbf{B}_1$,

$$\int_0^1 (k_1(x) - l_1(x))f(x)\,dx = \int_0^1 (k_2(x) - l_2(x))\,g(x)dx = 0\,.$$

Ceci est en particulier vrai si f est une fonction étagée positive minorant la fonction caractéristique de l'ensemble $\{k_1 > l_1\}$; on déduit donc en utilisant le fait que

$$\int_{\{k_1>l_1\}} dx = \lim_{n\to\infty}\int_0^1 \varphi_n(x)\,dx\,,$$

où $(\varphi_n)_n$ est une suite de telles fonctions étagées (théorie de l'intégration) que

$$\int_{\{k_1>l_1\}} dx = 0\,,$$

soit $k_1 \leq l_1$ presque partout ; en renversant les rôles de k_1 et l_1, on trouve $k_1(x) = l_1(x)$ pour presque tout x. Le même raisonnement avec g montre $k_2(x) = l_2(x)$ pour presque tout x.

II.2. Étude d'une classe particulière.

a. Supposons tout d'abord que $\Lambda \in \mathcal{L}_{1\times 1}$. On doit avoir à la fois la positivité de Λ et le fait que $\Lambda(\mathbf{1}) = 1$.

- la dernière condition équivaut à
$$\int_0^1 \phi(x)dx + \int_0^1 \int_0^1 \psi(x,y)\,dx\,dy = 1\,; \qquad (\dagger\dagger)$$
- la fonction ψ doit être positive ou nulle presque partout sur $\{(x,y) \in [0,1]\times[0,1]\,;\, x \neq y\}$, donc presque partout sur $[0,1]\times[0,1]$ au sens de la mesure de Lebesgue 2-dimensionnelle (car la diagonale du carré est de mesure nulle au sens de la mesure de Lebesgue 2-dimensionnelle) : en effet, si l'ensemble $\{(x,y) \in [0,1]\times[0,1]\,;\, x \neq y\,,\, \psi(x) < 0\}$ était de mesure non nulle, il existerait un borélien C_0 de $\{(x,y) \in [0,1]\times[0,1]\,;\, x \neq y\}$ de mesure de Lebesgue 2-dimensionnelle m_2 strictement positive et sur lequel ψ serait majorée par une constante strictement négative $-1/n_0$ et l'on aurait une contradiction en écrivant
$$\Lambda(\mathbf{1}_{C_0}) = \int_0^1 \int_0^1 \psi(x,y)\mathbf{1}_{C_0}(x,y)dxdy \geq 0$$
et
$$\int_0^1 \int_0^1 \psi(y)\mathbf{1}_{C_0}(x,y)dxdy \leq -\frac{m_2(C_0)}{n_0} < 0\,;$$
- la fonction ϕ est positive ou nulle presque partout sur $[0,1]$: si ceci n'était pas vrai,il existerait un borélien B_0 de $[0,1]$ de mesure de Lebesgue 1-dimensionnelle m_1 strictement positive sur lequel ϕ serait majorée par une constante strictement négative $-1/n_0$; l'ensemble $C_0 := \{(x,y) \in [0,1]\times[0,1]\,;\, x = y\,,\, x \in B_0\}$ est un borélien de $[0,1]\times[0,1]$ de mesure de Lebesgue nulle (au sens de la mesure de Lebesgue sur $[0,1]\times[0,1]$) puisque qu'inclus dans $\{(x,y) \in [0,1]\times[0,1]\,;\, x = y\}$ qui est de mesure de Lebesgue 2-dimensionnelle nulle ; on aurait donc à la fois, du fait de la positivité de Λ,
$$\Lambda(\mathbf{1}_{C_0}) = \int_{B_0} \varphi(x)dx \geq 0$$
et
$$\int_{B_0} \varphi(x)dx \leq -m_1(B_0)/n_0 < 0\,,$$
ce qui conduit à une contradiction.

Si $\phi \geq 0$ presque partout sur $[0,1]$ (au sens de la mesure de Lebesgue m_1), $\psi \geq 0$ presque partout sur $[0,1] \times [0,1]$ (au sens de la mesure de Lebesgue m_2), avec de plus la condition ($\dagger\dagger$), alors $\Lambda \in \mathcal{L}_{1\times 1}$, ce qui prouve donc que ces trois conditions sont bien les conditions nécessaires et suffisantes.

b. Si f est une fonction borélienne bornée de $[0,1]$ dans $\mathbb{R}$, et si $h(x,y) := f(x)$,

$$\begin{aligned}\Lambda_1(f) &= \int_0^1 \phi(x)f(x)dx + \int_0^1\int_0^1 \psi(x,y)f(x)dxdy \\ &= \int_0^1 \phi(x)f(x)dx + \int_0^1 \Big[\int_0^1 \psi(x,y)dy\Big]f(x)dx\end{aligned}$$

d'après le théorème de Fubini (f est bornée et ψ est intégrable, donc $(x,y) \to f(x)\psi(x,y)$ est bien intégrable sur $[0,1] \times [0,1]$ et le théorème de Fubini s'applique). Il y a donc une densité marginale :

$$x \to l_1(x) := \phi(x) + \int_0^1 \psi(x,y)dy$$

(c'est une densité car c'est une fonction positive et d'intégrale 1 sur $[0,1]$ en vertu de la condition ($\dagger\dagger$)) ; de la même manière, on a, pour toute fonction borélienne bornée sur $[0,1]$,

$$\begin{aligned}\Lambda_2(g) &= \int_0^1 \phi(y)g(y)dy + \int_0^1\int_0^1 \psi(x,y)g(y)dxdy \\ &= \int_0^1 \phi(y)f(y)dy + \int_0^1 \Big[\int_0^1 \psi(x,y)dx\Big]g(y)dy\,,\end{aligned}$$

ce qui prouve l'existence d'une seconde densité marginale :

$$x \to l_2(x) := \phi(x) + \int_0^1 \psi(y,x)dy\,.$$

II.3. Interprétation variationnelle de d_{VT}.

a. Soit f une fonction borélienne de $[0,1]$ dans $[0,1]$; on a

$$\int_0^1 f(x)q(x)dx = \Pi(h_1)\,,$$

où $h_1(x,y) := f(x)$. De même

$$\int_0^1 f(y)r(y)dy = \Pi(h_2)\,,$$

où $h_2(x,y) := f(y)$; on a donc

$$\Big|\int_0^1 f(x)q(x)dx - \int_0^1 f(y)r(y)dy\Big| = \Big|\Pi(h_1 - h_2)\Big| \leq \Pi\Big(|h_1 - h_2|\Big)$$

(d'après la positivité de Π) ; mais on a, pour tout $(x,y) \in [0,1] \times [0,1]$,

$$|h_1(x,y) - h_2(x,y)| \leq \mathbf{1}_{x \neq y}(x,y)$$

pour tout $(x,y) \in [0,1] \times [0,1]$; ceci est en effet vrai si $x = y$ (le premier membre est nul) et si $x \neq y$ (le premier membre est inférieur à 1 car $f(x)$ et $f(y)$ sont dans $[0,1]$). Par positivité de Π, on a donc

$$\Big|\int_0^1 f(x)q(x)dx - \int_0^1 f(y)r(y)dy\Big| \leq \Pi(\mathbf{1}_{x \neq y})\,,$$

d'où, par définition de $d_{VT}(q,r)$ comme le sup des expressions figurant au membre de gauche,

$$d_{VT}(q,r) \leq \Pi(\mathbf{1}_{x \neq y})\,.$$

b. La forme Λ_0 est bien positive car ϕ_0 et ψ_0 sont des fonctions positives ; de plus, grâce au théorème de Fubini,

$$\int_0^1 \int_0^1 \psi_0(x,y)dxdy = \frac{1}{d_{VT}(q,r)}\Big(\int_0^1 [q(x) - r(x)]_+ dx\Big)\Big(\int_0^1 [r(y) - q(y)]_+ dy\Big)\,;$$

or on a vu (au **I.3.b**) que

$$d_{VT}(q,r) = \int_0^1 [q(x) - r(x)]_+ dx = \int_0^1 [r(y) - q(y)]_+ dy\,.$$

On a donc bien

$$\int_0^1 \int_0^1 \psi_0(x,y)dxdy = \int_0^1 [q(x) - r(x)]_+ dx = \int_0^1 [r(y) - q(y)]_+ dy\,.$$

On a

$$\int_0^1 \phi_0(x)dx = \int_{q \leq r} q(x)dx + \int_{q > r} r(x)dx\,;$$

de plus

$$\int_0^1 \int_0^1 \psi_0(x,y)dxdy = \int_{q > r} (q(x) - r(x))dx\,;$$

on a donc

$$\begin{aligned}\int_0^1 \phi_0(x)dx + \int_0^1 \int_0^1 \psi_0(x,y)dxdy &= \int_{q \leq r} q(x)dx \\ &\quad + \int_{q > r} (q(x) - r(x) + r(x))dx \\ &= \int_0^1 q(x)dx = 1\,.\end{aligned}$$

La forme linéaire définie par (11) est bien dans $\mathcal{L}_{1\times 1}$. Les densités marginales se calculent grâce aux formules établies au **II.2.b** et l'on a

$$\begin{aligned} l_1(x) &= \phi_0(x) + \int_0^1 \psi_0(x,y)dy = \min(q(x), r(x)) + [q(x) - r(x)]_+ = q(x) \\ l_2(y) &= \phi_0(y) + \int_0^1 \psi_0(x,y)dx = \min(q(y), r(y)) + [r(y) - q(y)]_+ = r(y)\,, \end{aligned}$$

d'où $\Lambda_0 \in \mathcal{L}(q,r)$.
On a

$$\Lambda_0(\mathbf{1}_{x\neq y}) = \int_0^1 \int_0^1 \psi_0(x,y)\mathbf{1}_{x\neq y}\, dxdy = \int_0^1 \int_0^1 \psi_0(x,y)dxdy = d_{VT}(q,r)\,.$$

c. D'après **II.3.a**, on a

$$\inf_{\Pi\in\mathcal{L}(q,r)} \Pi(\mathbf{1}_{x\neq y}) \geq d_{VT}(q,r)\,;$$

si $d_{VT}(q,r) > 0$, on a vu au **II.3.b** que la valeur $d_{VT}(q,r)$ était atteinte pour $\Pi = \Lambda_0$; on a donc dans ce cas

$$\inf_{\Pi\in\mathcal{L}(q,r)} \Pi(\mathbf{1}_{x\neq y}) = d_{VT}(q,r)\,.$$

Si $d_{VT}(q,r) = 0$, on a d'après la formule (9) établie au **I.3.a**, $q = r$ presque partout. Si l'on pose

$$\Lambda(h) = \int_0^1 q(x)h(x,x)\, dx$$

on définit bien un élément de $\mathcal{L}(q,q)$; on a $\Lambda(\mathbf{1}_{x\neq y}) = 0$, donc dans ce cas aussi

$$\inf_{\Pi\in\mathcal{L}(q,q)} \Pi(\mathbf{1}_{x\neq y}) = d_{VT}(q,q) = 0\,.$$

II.4. Inégalité principale dans le cas $n = 1$.

On suppose dans un premier temps $d_{VT}(q,r) > 0$, c'est-à-dire que q et r ne sont pas égales presque partout. On a

$$\begin{aligned} \Lambda_0(h_{2,\alpha}) &= \int_0^1 \int_0^1 \psi_0(x,y)\alpha(y)dxdy = \int_0^1 [r(y) - q(y)]_+ \alpha(y)dy \\ &= \int_0^1 \Big[1 - \frac{q(y)}{r(y)}\Big]_+ \alpha(y)r(y)dy \\ &\leq \left(\int_0^1 \Big[1 - \frac{q(y)}{r(y)}\Big]_+^2 r(y)dy\right)^{1/2} \times \left(\int_0^1 \alpha^2 r(y)dy\right)^{1/2} \\ &\leq \delta(q\,|\,r) \leq \sqrt{\delta^2(q\,|\,r) + \delta^2(r\,|\,q)} \leq \sqrt{2\mathrm{Ent}_r(q/r)} \end{aligned}$$

d'après l'inégalité de Cauchy-Schwarz (dans l'espace $L^2([0,1], r(y)dy)$) et l'inégalité (7) établie au **I.2.e**. On a la même majoration pour $\Lambda_0(h_{1,\beta})$ si $\int_0^1 \beta^2(x)q(x)dx \le 1$. Ceci implique donc, si $d_{VT}(q,r) > 0$,

$$d_2(q\,|\,r) := \inf_{\Pi \in \mathcal{L}(q,r)} \sup_\alpha \Pi(h_{2,\alpha}) \le \sqrt{2\mathrm{Ent}_r(q/r)}$$

ainsi que

$$d_2(r\,|\,q) := \inf_{\Pi \in \mathcal{L}(q,r)} \sup_\alpha \Pi(h_{1,\beta}) \le \sqrt{2\mathrm{Ent}_r(q/r)}\,.$$

Si $d_{VT}(q,r) = 0$, on travaille avec la forme Λ introduite au **II.3.c** et on voit que $\Lambda(\Pi_{2,\alpha}) = \Lambda(\Pi_{1,\beta}) = 0$ pour tout choix de α et β comme indiqué ; mais dans ce cas $\mathrm{Ent}_r(q/r) = E_r(1) = 0$ comme on le vérifie avec la définition (1) de l'entropie ; les inégalités sont encore valides dans ce cas car $d_2(q\,|\,r) = d_2(r\,|\,q) = 0 = \mathrm{Ent}_r(q/r) = 0$ si $q = r$ presque partout.

III. Une première inégalité de concentration

III.1. Une inégalité pour des sommes.

a. On a, pour tout $\lambda \ge 0$,

$$\mathrm{ch}\,\lambda = \sum_{k=0}^{\infty} \frac{\lambda^{2k}}{(2k)!} \le \sum_{k=0}^{\infty} \frac{\lambda^{2k}}{2^k\,k!} = \exp\Big(\frac{\lambda^2}{2}\Big)$$

car $(2k)! \ge 2 \times 4 \times \cdots \times 2k$ pour tout $k \ge 1$ (c'est aussi vrai par convention si $k = 0$ et $0! = 1$).

b. Si $\lambda \ge 0$, la fonction $f\ : x \to e^{\lambda x}$ est convexe sur $\mathbb{R}$, donc sur l'intervalle $[-1,1]$; elle est donc majorée sur $[-1,1]$ par l'unique fonction affine $x \to \alpha x + \beta$ telle que $-\alpha + \beta = e^{-\lambda} = f(-1)$ et $\alpha + \beta = e^{\lambda} = f(1)$ (le graphe d'une fonction convexe est sous la corde joignant deux points de ce graphe) ; le calcul de α et β donne $\alpha = \mathrm{sh}\,\lambda$ et $\beta = \mathrm{ch}\,\lambda$. On a donc

$$\forall x \in [-1,1]\,,\ e^{\lambda x} \le \mathrm{ch}\,\lambda + x\mathrm{sh}\,\lambda \le \exp\Big(\frac{\lambda^2}{2}\Big) + x\,\mathrm{sh}\,\lambda$$

si l'on utilise l'inégalité (17) établie au **III.1.a**.

c. Comme X est une variable aléatoire (définie sur un espace probabilisé $(\Omega, \mathcal{T}, P)$) à valeurs dans $[-1,1]$, X est bornée, donc intégrable relativement à la mesure positive P ; il en est de même pour les variables aléatoires $e^{\lambda X}$ et $e^{-\lambda X}$. Comme X prend ses valeurs dans $[-1,1]$, on a, d'après l'inégalité (18) établie au **III.1.b**

$$\forall \omega \in \Omega\,,\ e^{\lambda X(\omega)} \le \exp\Big(\frac{\lambda^2}{2}\Big) + \mathrm{sh}\,\lambda \times X(\omega).$$

On a donc, du fait de la monotonie de l'opération de prise d'espérance :

$$\begin{aligned} E(e^{\lambda X}) \le \exp(\lambda^2)\int_\Omega dP + \operatorname{sh}\lambda \int_\Omega X(\omega)dP &= \exp\Big(\frac{\lambda^2}{2}\Big) + \operatorname{sh}\lambda \times E(X) \\ &= \exp\Big(\frac{\lambda^2}{2}\Big) \end{aligned}$$

puisque X est d'espérance nulle ; le même résultat vaut si l'on remplace X par $-X$ qui a les mêmes propriétés.

d. On utilise l'inégalité de Chernoff rappelée dans l'en-tête de la section III ; on a, pour tout $\lambda \ge 0$:

$$P(X \ge a) \le e^{-\lambda a} E(e^{\lambda X}) \le e^{-\lambda a} e^{\lambda^2/2}\,;$$

en particulier, pour $\lambda = a$,

$$P(X \ge a) \le e^{-a^2/2}\,.$$

On a

$$P(X \le -a) = P(-X \ge a) \le e^{-a^2/2}$$

pour les mêmes raisons (X et $-X$ ont les mêmes propriétés) ; on a donc bien

$$P(|X| \ge a) \le P(X \ge a) + P(X \le -a) \le 2\,\exp\Big(-\frac{a^2}{2}\Big)$$

pour tout $a \ge 0$.

e. D'après l'inégalité de Chernoff, on a, pour tout $a \ge 0$, pour tout $n \ge 1$, pour tout $\lambda \ge 0$,

$$P(X_1 + \cdots + X_n \ge a\sqrt{n}) \le e^{-\lambda a\sqrt{n}} E(e^{\lambda(X_1+\cdots+X_n)})\,;$$

mais l'indépendance des X_i implique celle des $e^{\lambda X_i}$ et l'on a donc

$$E(e^{\lambda(X_1+\cdots+X_n)}) = \prod_{i=1}^n E(e^{\lambda X_i}) \le \exp\Big(\frac{n\lambda^2}{2}\Big)\,;$$

on a donc, finalement :

$$P(X_1 + \cdots + X_n \ge a\sqrt{n}) \le e^{-\lambda a\sqrt{n}} \exp\Big(\frac{n\lambda^2}{2}\Big)\,;$$

si l'on prend $\lambda = a/\sqrt{n}$, on trouve

$$P(X_1 + \cdots + X_n \ge a\sqrt{n}) \le \exp\Big(-\frac{a^2}{2}\Big)\,;$$

de même, en remplacant chaque X_i par $-X_i$

$$P(X_1 + \cdots + X_n \leq -a\sqrt{n}) \leq \exp\left(-\frac{a^2}{2}\right),$$

donc, au final

$$P(|X_1 + \cdots + X_n| \geq a\sqrt{n}) \leq 2\exp\left(-\frac{a^2}{2}\right)$$

pour tout $a \geq 0$, pour tout $n \geq 1$.

III.2. Optimalité ?

a. C'est une équation différentielle linéaire du premier ordre résoluble en y' ; les solutions de cette équation différentielle sur $\mathbb{R}$ sont les fonctions

$$x \to C\exp(-x^2/2),$$

où $C \in \mathbb{R}$ (ce sont des solutions maximales car définies sur $\mathbb{R}$ et l'on peut ensuite invoquer le théorème d'existence et d'unicité globale de Cauchy-Lipschitz).

b. On a

$$\int_{-\infty}^{+\infty} \varphi(t)dt = 1$$

car φ est une densité. Donc, pour $x >> 0$,

$$1 - \Phi(x) = \int_x^{+\infty} \varphi(t)dt = -\int_x^{+\infty} \frac{\varphi'(t)}{t}dt = -\left[\frac{\varphi(t)}{t}\right]_x^{+\infty} - \int_x^{+\infty} \frac{\varphi(t)}{t^2}dt$$

en intégrant par parties ; on peut d'ailleurs recommencer en écrivant

$$\int_x^{+\infty} \frac{\varphi(t)}{t^2}dt = -\int_x^{+\infty} \frac{\varphi'(t)}{t^3}dt = -\left[\frac{\varphi(t)}{t^3}\right]_x^{+\infty} - 3\int_x^{+\infty} \frac{\varphi(t)}{t^4}dt\,.$$

On voit ainsi que

$$1 - \Phi(x) = \frac{\varphi(x)}{x} - \frac{\varphi(x)}{x^3} + 3\int_x^{+\infty} \frac{\varphi(t)}{t^4}dt\,;$$

comme la dernière intégrale est convergente, on a bien

$$1 - \Phi(x) \simeq \frac{\varphi(x)}{x}$$

lorsque x tend vers $+\infty$.

c. Si $(X_n)_n$ est une suite de variables aléatoires centrées, indépendantes et de même loi, de moyenne nulle et de variance 1 (chaque X_n ne prenant que les valeurs 1 et -1 avec la probabilité $1/2$), il résulte du théorème limite-centrale que la suite de variables

$$Y_n := \frac{X_1 + \cdots + X_n}{\sqrt{n}}$$

converge en loi vers une loi normale. Ceci implique, pour $x > 0$:

$$\lim_{n\to\infty} P(|Y_n| \geq x) = \int_{|t|\geq x} \varphi(t)dt = 2(1 - \Phi(x)) \,.$$

Supposons l'existence de $A > 0$ et $\kappa > 1/2$ telles que

$$P(|Y_n| \geq a) \leq A \exp(-\kappa a^2)$$

pour tout $n \geq 1$, pour tout $a \geq 0$; en faisant tendre n vers l'infini, on trouverait, pour tout $a \geq 0$,

$$2(1 - \Phi(a)) \leq Ae^{-\kappa a^2} \,;$$

en faisant tendre maintenant a vers $+\infty$, il viendrait (d'après la conclusion (22) de **III.2.b**)

$$2\frac{\varphi(a)}{a} = 2\frac{\exp(-a^2/2)}{a} \leq 2Ae^{-\kappa a^2}$$

pour $a >> 0$; ceci impliquerait

$$e^{(\kappa-1/2)a^2} \leq Aa$$

pour a assez grand, ce qui est exclus car $\kappa - 1/2 > 0$. L'existence de A et κ est donc impossible.

IV. Premières applications de l'inégalité principale

IV.1. Inégalités pour des fonctions convexes sur $[0,1]^n$.

a. Supposons $x > y$; par la formule des accroissements finis, il existe $\xi \in]x,y[$ tel que $f(x) - f(y) = (x-y)f'(\xi)$; mais la fonction f' est une fonction croissante (puisque f est convexe) et l'on a donc

$$f(x) - f(y) \leq (x-y)f'(\xi) \leq (x-y)f'(x) \leq |f'(x)|\, \mathbf{1}_{x\neq y}(x,y)$$

puisque $0 < x - y < 1$. Si maintenant $x < y$, on a

$$f(x) - f(y) = (x-y)f'(\xi)$$

avec $\xi \in]x, y[$ et, par convexité de f, $f'(\xi) \geq f'(x)$; on a donc dans ce cas

$$f(x) - f(y) \leq (x-y)f'(x) \leq |x-y|\,|f'(x)| \leq |f'(x)|\,\mathbf{1}_{x\neq y}(x,y)$$

dans ce cas encore. Enfin, si $x = y$, on a $f(x) - f(y) = 0 \leq \mathbf{1}_{x\neq y}(x,y)$, ce qui montre que pour tout $x, y \in]0,1[$,

$$f(x) - f(y) \leq |f'(x)|\,\mathbf{1}_{x\neq y}(x,y)\,.$$

b. On utilise, si x et y sont des points de $[0,1]^n$, la fonction

$$f_{x,y} \ : t \in [0,1] \to f((1-t)x + ty)$$

qui est une fonction convexe d'une variable t ; on a

$$f_{x,y}(1) - f_{x,y}(0) = f'_{x,y}(\xi) = \sum_{i=1}^{n}(x_i - y_i)\partial_i f((1-\xi)x + \xi y)$$

pour un certain $\xi \in]0,1[$. Ensuite, on raisonne comme au **IV.1.a** pour affirmer que, pour tout $i = 1, ..., n$,

$$(x_i - y_i)\partial_i f((1-\xi)x + \xi y) \leq |x_i - y_i|\,|\partial_i f(x)|\,\mathbf{1}_{x_i\neq y_i}(x,y) \qquad (*)$$

en utilisant le fait que si $x_i > y_i$, on a

$$\partial_i f((1-\xi)x + \xi y) \geq \partial_i f(x)$$

car le graphe de f est au dessus de son plan tangent en tout point tandis que, si $y_i > x_i$, on a

$$\partial_i f(\xi y + (1-\xi)x) \leq \partial_i f(x)$$

pour les mêmes raisons. On a l'inégalité (25) voulue en additionnant les inégalités (*).

IV.2. Résultats itermédiaires.

a. On suppose que q est une densité strictement positive ; si $I_q(f) = 0$, on a $\nabla f \equiv 0$ dans $]0,1[^n$, ce qui implique que f est constante dans $[0,1]^n$. Le membre de gauche de la formule (29) vaut donc 0 puisque q et r sont toutes les deux des densités sur $[0,1]^n$; on a bien l'égalité avec le membre de droite. Supposons donc $I_q(f) > 0$ et prenons $\Pi \in \mathcal{L}(q,r)$. On a

$$\int_{[0,1]^n} f(x)q(x)dx - \int_{[0,1]^n} f(y)r(y)dy \;\; = \;\; \Pi((x,y) \to f(x) - f(y))$$

$$
\begin{aligned}
&\leq \Pi\Big(\sum_{i=1}^{n} |\partial_i f(x)|\, \mathbf{1}_{x_i \neq y_i}\Big) \\
&\leq \sqrt{I_q(f)}\,\Pi\left(\sum_{i=1}^{n} \frac{|\partial_i f(x)|}{\sqrt{I_q(f)}}\, \mathbf{1}_{x_i \neq y_i}\right) \\
&\leq \sqrt{I_q(f)}\, \sup_{\beta} \Pi(h_{1,\beta})\,.
\end{aligned}
$$

En prenant l'inf sur tous les $\Pi \in \mathcal{L}(q,r)$, on trouve

$$
\int_{[0,1]^n} f(x)q(x)dx - \int_{[0,1]^n} f(y)r(y)dy \leq d_2(r \mid q)\,\sqrt{I_q(f)} \leq \sqrt{2 I_q(f) \mathrm{Ent}_r(q/r)}
$$

d'après l'inégalité admise (28).

b. Si la fonction f est concave, la fonction $-f$ est convexe et l'on a donc, dans $]0,1[^n$,

$$
(-f)(y) - (-f)(x) = f(x) - f(y) \leq \sum_{i=1}^{n} \partial_i f(y) \mathbf{1}_{x_i \neq y_i}\,.
$$

Si $I_r(f) = 0$, les deux membres de la formule (30) sont nuls ; si $I_r(f) > 0$, on a, pour tout $\Pi \in \mathcal{L}(q,r)$:

$$
\begin{aligned}
\int_{[0,1]^n} f(x)q(x)dx - \int_{[0,1]^n} f(y)r(y)dy &= \Pi((x,y) \to f(x) - f(y)) \\
&\leq \Pi\Big(\sum_{i=1}^{n} |\partial_i f(y)|\, \mathbf{1}_{x_i \neq y_i}\Big) \\
&\leq \sqrt{I_r(f)}\,\Pi\left(\sum_{i=1}^{n} \frac{|\partial_i f(y)|}{\sqrt{I_r(f)}}\, \mathbf{1}_{x_i \neq y_i}\right) \\
&\leq \sqrt{I_r(f)}\, \sup_{\beta} \Pi(h_{2,\alpha})\,.
\end{aligned}
$$

IV.3. Deux applications.

a. Les fonctions $m_f g$ et $m_f \ln m_f\, g$ puisque l'entropie $\mathrm{Ent}_r(e^f)$ est définie. On a, d'après le résultat établi au **I.2.b** :

$$
\begin{aligned}
&\mathrm{Ent}_r(m^f) = \frac{1}{R(e^f)} \mathrm{Ent}_r(e^f) \\
= \;& \frac{1}{R(e^f)}\bigg(\int_{[0,1]^n} e^{f(x)} f(x) r(x) dx \\
&\quad - \Big(\int_{[0,1]^n} e^{f(x)} r(x) dx\Big)\Big(\ln\Big[\int_{[0,1]^n} e^{f(x)} r(x) dx\Big]\Big)\bigg)
\end{aligned}
$$

$$
\begin{aligned}
&= \int_{[0,1]^n} f(x)r^f(x)dx - \frac{1}{R(e^f)}\Big(\int_{[0,1]^n} e^{f(x)}r(x)dx\Big)\Big(\ln\Big[\int_{[0,1]^n} e^{f(x)}r(x)dx\Big]\Big) \\
&= \int_{[0,1]^n} f(x)r^f(x)dx - \ln\Big[\int_{[0,1]^n} e^{f(x)}r(x)dx\Big] .
\end{aligned}
$$

Or l'inégalité de Jensen (appliquée avec la fonction convexe exp sur l'intervalle $[0,+\infty[$, la fonction f et la densité r) implique

$$
\exp\Big[\int_{[0,1]^n} f(y)r(y)dy\Big] \leq \int_{[0,1]^n} e^{f(x)}r(x)dx\,;
$$

on a donc, en passant aux logarithmes

$$
-\ln\Big[\int_{[0,1]^n} e^{f(x)}r(x)dx\Big] \leq -\int_{[0,1]^n} f(y)r(y)dy\,,
$$

inégalité que l'on reporte dans la formule précédente pour obtenir l'inégalité (31).

b. Si l'on utilise **IV.2.a** et l'inégalité (31) établie au IV.3.a, on trouve

$$
\mathrm{Ent}_r[m^f] \leq \sqrt{2I_{r_f}(f)\mathrm{Ent}_r(r_f/r)} = \sqrt{2I_{r_f}(f)\mathrm{Ent}_r[m_f]}\,;
$$

on a donc

$$
\mathrm{Ent}_r[m^f] \leq 2\int_{[0,1]^n} \|\nabla f(x)\|^2 \frac{e^{f(x)}r(x)dx}{R(e^f)}\,,
$$

d'où, en multipliant par $R(e^f)$:

$$
\mathrm{Ent}_r(e^f) \leq 2\int_{[0,1]^n} \|\nabla f(x)\|^2 e^{f(x)}r(x)dx\,.
$$

c. On utilise l'inégalité (31) ainsi que l'inégalité (30) établie au **IV.2.b** pour une fonction concave. On obtient

$$
\mathrm{Ent}_r[m^f] \leq \sqrt{2I_r(f)\mathrm{Ent}_r(m_f)}\,,
$$

d'où

$$
\mathrm{Ent}_r[m^f] \leq 2I_r(f)
$$

et par conséquent

$$
\mathrm{Ent}_r[e^f] \leq 2R(e^f)I_r(f) = 2R(e^f) \times \int_{[0,1]^n} \|\nabla f(x)\|^2 e^{f(x)}r(x)dx\,.
$$

IV.4. Une inégalité de Poincaré.

On suppose dans un premier temps f convexe.

Ni le membre de gauche, ni le membre de droite de l'inégalité (34) ne changent lorsque l'on remplace f par $f+K$, où K est une constante (c'est évident en ce qui concerne le membre de droite car $\nabla(f+K)=\nabla(f)$, cela résulte du fait que r est une densité en ce qui concerne le membre de gauche). On peut donc supposer, quitte à retrancher à f la constante

$$K := \int_{[0,1]^n} f(x)r(x)dx\,,$$

que $\int_{[0,1]^n} f(x)r(x)dx = 0$, auquel cas l'inégalité à établir est

$$\int_{[0,1]^n} f^2(x)r(x)dx \leq CI_r(f)\,.$$

Comme f est continue sur $[0,1]^n$, elle est bornée et l'entropie de ϵf pour $\epsilon \in \mathbb{R}$ est bien définie. En appliquant l'inégalité (32) établie au **IV.3.b** avec ϵf à la place de f, on trouve,

$$\mathrm{Ent}_r(e^{\epsilon f}) \leq 2\epsilon^2 \int_{[0,1]^n} \|\nabla f(x)\|^2 e^{\epsilon f(x)} r(x)dx\,.$$

Or

$$\begin{aligned}\mathrm{Ent}_r(e^{\epsilon f}) &= \epsilon \int_{[0,1]^n} e^{\epsilon f(x)} f(x)r(x)dx \\ &\quad - \Big(\int_{[0,1]^n} e^{\epsilon f(x)} r(x)dx \Big) \times \ln \Big[\int_{[0,1]^n} e^{\epsilon f(x)} r(x)dx \Big]\,.\end{aligned}$$

Considérons la fonction

$$\epsilon \to \int_{[0,1]^n} e^{\epsilon f(x)} f(x)r(x)dx = \int_{[0,1]^n} \left(\sum_{k=0}^{\infty} \frac{\epsilon^k (f(x))^k}{k!} \right) f(x)r(x)dx\,;$$

comme

$$\int_{[0,1]^n} \left(\sum_{k=0}^{\infty} \frac{\epsilon^k |f(x)|^k}{k!} \right) |f(x)|r(x)dx = \int_{[0,1]^n} e^{\epsilon |f(x)|} |f(x)|\, r(x)dx < +\infty$$

car f est bornée sur $[0,1]^n$, on peut utiliser le théorème de Fubini et écrire :

$$\int_{[0,1]^n} e^{\epsilon f(x)} f(x)r(x)dx = \sum_{k=1}^{\infty} \Big(\int_{[0,1]^n} [f(x)]^{k+1} r(x)dx \Big) \frac{\epsilon^k}{k!}\,.$$

On a de même, pour tout $\epsilon \in \mathbb{R}$,

$$\begin{aligned}\int_{[0,1]^n} e^{\epsilon f(x)} r(x)dx &= \sum_{k=0}^{\infty}\Big(\int_{[0,1]^n}[f(x)]^k r(x)dx\Big)\frac{\epsilon^k}{k!}\\ &= 1+\sum_{k=2}^{\infty}\Big(\int_{[0,1]^n}[f(x)]^k r(x)dx\Big)\frac{\epsilon^k}{k!}\,.\end{aligned}$$

Au voisinage de $\epsilon = 0$, on peut donc écrire (car $\log(1+u) \simeq u$ au voisinage de 0)

$$\mathrm{Ent}_r(e^{\epsilon f}) = \frac{\epsilon^2}{2}\int_{[0,1]^n} f^2(x)r(x)dx - \mathbf{o}(\epsilon^2)\,;$$

on a donc, au voisinage de $\epsilon = 0$,

$$\frac{\epsilon^2}{2}\int_{[0,1]^n} f^2(x)r(x)dx + \mathbf{o}(\epsilon^2) \le 2\epsilon^2\int_{[0,1]^n}\|\nabla f(x)\|^2 e^{\epsilon f(x)} r(x)dx\,,$$

d'où en divisant par ϵ^2, en faisant tendre ϵ vers 0 et en appliquant le théorème de convergence dominée de Lebesgue ($e^{\epsilon f} \le M$ sur $[0,1]^n$ pour tout ϵ dans $[-1,1]$) :

$$\begin{aligned}\int_{[0,1]^n} f^2(x)r(x)dx &\le 4\lim_{\epsilon\to 0}\int_{[0,1]^n}\|\nabla f(x)\|^2 e^{\epsilon f(x)} r(x)dx\\ &= 4\int_{[0,1]^n}\|\nabla f(x)\|^2 r(x)dx\end{aligned}$$

qui est l'inégalité de Poincaré voulue.

Dans le cas où f est concave, on reprend exactement le même schéma de raisonnement, excepté que l'on utilise l'inégalité (33) établie au **IV.3.c** au lieu de l'inégalité (32) établie au **IV.3.b** ; le développement de

$$\epsilon \to \int_{[0,1]^n} e^{f(x)} r(x)dx$$

(toujours sous l'hypothèse $R(e^f) = 0$) s'écrit en effet $1 + \mathbf{o}(\epsilon^2)$ et ne perturbe pas les calculs en deça de l'ordre 2 dans les développements limités.

Deuxième application : inégalités de concentration

V.1. Cas concave de classe C^1.

a. D'après (30), on a (puisque f, donc $f - R(f)$ est concave)

$$\int_{[0,1]^n}(f - R(f))q(x)dx \le \sqrt{2\tau^2\lambda^2\mathrm{Ent}_r(m)} = \lambda\tau\sqrt{2\mathrm{Ent}_r(m)}\,.$$

On a donc

$$\int_{[0,1]^n} \Big(\lambda(f(x)-R(f)) - \frac{\tau^2\lambda^2}{2}\Big) q(x)dx \le \lambda\tau\sqrt{2\mathrm{Ent}_r(m)} - \frac{\tau^2\lambda^2}{2}\,.$$

La fonction

$$\xi \in [0,+\infty[\to \xi\,\sqrt{2\mathrm{Ent}_r(m)} - \xi^2/2$$

atteint son maximum en $\xi = \sqrt{2\mathrm{Ent}_r(m)}$ et ce maximum vaut exactement $\mathrm{Ent}_r(m)$.

b. Si $m = e^l/R(e^l)$, l'entropie de m s'exprime par

$$\begin{aligned}
\mathrm{Ent}_r(m) &= \frac{1}{R(e^l)} \int_{[0,1]^n} e^{l(x)} l(x) r(x)dx - \ln\Big[\int_{[0,1]^n} e^{l(x)} r(x)dx\Big] \\
&= \int_{[0,1]^n} m(x)l(x)r(x)dx - \ln\Big[\int_{[0,1]^n} e^{l(x)} r(x)dx\Big]\,.
\end{aligned}$$

On pose

$$l(x) = \lambda(f(x)-R(f)) - \tau^2\lambda^2/2\,.$$

L'inégalité établie au **V.1.a** se lit alors

$$\int_{[0,1]^n} l(x)m(x)r(x)dx \le \int_{[0,1]^n} m(x)l(x)r(x)dx - \ln\Big[\int_{[0,1]^n} e^{l(x)} r(x)dx\Big]\,,$$

soit

$$\ln\Big[\int_{[0,1]^n} e^{l(x)} r(x)dx\Big] \le 0\,.$$

En prenant les exponentielles, on trouve exactement l'inégalité (36) voulue.

c. On applique l'inégalité de Chernoff pour la variable aléatoire

$$X := f(X_1,...,X_n) - E(f(X_1,...,X_n))$$

en remarquant que, comme la loi de $(X_1,...,X_n)$ est la loi de densité $r(y) := r_1(y_1)\cdots r_n(y_n)$ (r_i étant la densité de la loi de X_i), on a

$$E[f(X_1,...,X_n)] = \int_{[0,1]^n} f(x)r(x)dx = R(f)$$

(avec les notations de cette partie **V**). D'autre part, si $\lambda > 0$,

$$E(e^{\lambda X}) = \int_{[0,1]^n} e^{\lambda(f(x)-R(f))} r(x)dx \le \exp(\tau^2\lambda^2/2)$$

d'après le **V.1.b**. L'inégalité de Chernoff donne donc que pour tout $\lambda > 0$,

$$P(X \ge a) \le \exp\Big(-\lambda a + \frac{\tau^2\lambda^2}{2}\Big)\,;$$

le maximum de la fonction de λ sous l'exponentielle est atteint en $\lambda = a/\tau^2$; en reportant sa valeur, on trouve

$$P(X \geq a) \leq \exp\Big(-\frac{a^2}{2\tau^2}\Big).$$

V.2. Cas convexe de classe C^1.

On reprend la démarche utilisée au **V.1.a**, cette fois en utilisant l'inégalité (29) à la place de (30), inégalité dans laquelle on majore $I_q(f)$ par $\rho^2_{\max}$, ce qui est licite car q est une densité et $\|\nabla f(x)\|^2 \leq \rho^2_{\max}$ sur $[0,1]^n$. Ensuite, tout se déroule comme dans les trois questions précédentes et l'on aboutit donc à l'inégalité de concentration (38).

FIN