



PRÉPASIMPLE

prepasimple.fr

HEC

Mathématiques

2012

Corrigé de mathématiques approfondies

Mathématiques approfondies — Prépa ECG

Annales de concours · Corrigé

Corrigé — HEC, Mathématiques, 2012

Partie I. Quelques inégalités de concavité

1. a) On étudie le signe de h puis celui de sa dérivée seconde.

Pour $0 < x < 1$, les nombres $\ln x$ et $\ln(1 - x)$ sont strictement négatifs. Les deux termes de $h(x)$ sont donc strictement positifs : $h(x) > 0$.

La fonction h est de classe C^2 sur $]0, 1[$ et :

$$h'(x) = -(\ln x + 1) + \ln(1 - x) + 1 = \ln\left(\frac{1 - x}{x}\right),$$

$$h''(x) = -\frac{1}{1 - x} - \frac{1}{x} = -\frac{1}{x(1 - x)} < 0.$$

Ainsi, h est strictement concave, donc concave, sur $]0, 1[$.

1. b) Les croissances comparées donnent $\lim_{t \rightarrow 0^+} t \ln t = 0$. Au voisinage de 0, le terme $(1 - x) \ln(1 - x)$ tend également vers 0 par continuité. Au voisinage de 1, on procède de même, ou on utilise $h(1 - x) = h(x)$.

Il existe donc un unique prolongement continu à $[0, 1]$, défini par $h(0) = h(1) = 0$.

Pour déterminer la régularité au bord, calculons le taux d'accroissement en 0 :

$$\frac{h(x) - h(0)}{x} = -\ln x - \frac{1 - x}{x} \ln(1 - x).$$

Comme $\ln(1 - x)/x \rightarrow -1$ lorsque $x \rightarrow 0^+$, le second terme tend vers 1, tandis que $-\ln x \rightarrow +\infty$. Le taux d'accroissement tend donc vers $+\infty$. Par symétrie, le taux d'accroissement en 1 tend vers $-\infty$.

Le prolongement n'est pas dérivable avec une dérivée finie aux extrémités. Il n'est donc pas de classe C^1 sur $[0, 1]$.

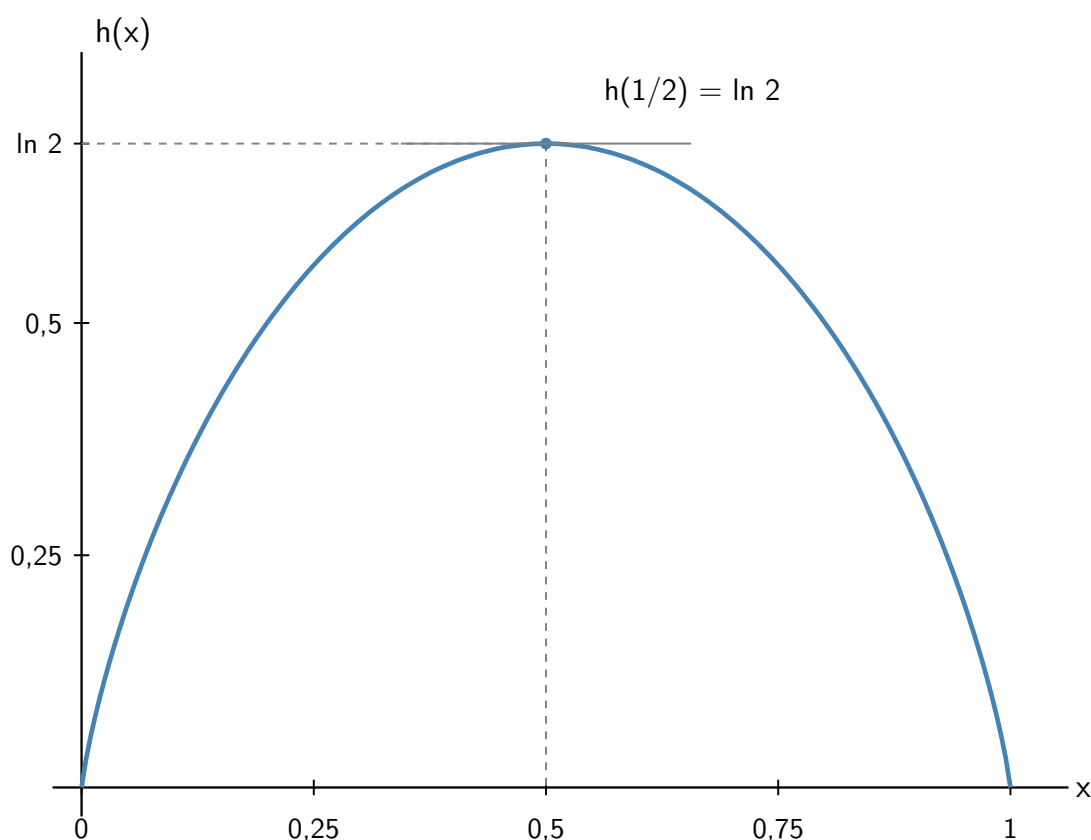
1. c) On utilise les dérivées et la symétrie établies précédemment.

$$h'(x) > 0 \text{ si } 0 < x < 1/2, h'(1/2) = 0, \text{ et } h'(x) < 0 \text{ si } 1/2 < x < 1.$$

La courbe est symétrique par rapport à la droite $x = 1/2$, croît de 0 à $\ln 2$, puis décroît de $\ln 2$ à 0. Son maximum est atteint uniquement en $x = 1/2$ et vaut :

$$h(1/2) = -\frac{1}{2} \ln(1/2) - \frac{1}{2} \ln(1/2) = \ln 2.$$

Elle possède une tangente horizontale au sommet et des demi-tangentes verticales aux extrémités. Le tracé ci-dessous est calculé à partir de l'expression exacte de h ; les unités des deux axes sont identiques.



2. Méthode 1 — Étude d'une fonction

Posons $q(u) = u - 1 - \ln u$ sur $]0, +\infty[$. Sa dérivée est :

$$q'(u) = 1 - \frac{1}{u} = \frac{u - 1}{u}.$$

La fonction q décroît sur $]0, 1]$ puis croît sur $[1, +\infty[$. Elle possède un minimum global égal à $q(1) = 0$. Ainsi, $\ln u \leq u - 1$, avec égalité uniquement si $u = 1$.

Méthode 2 — Tangente à une fonction concave

Le logarithme est de classe C^2 sur $]0, +\infty[$ et sa dérivée seconde vaut $-1/u^2 < 0$. Il est donc strictement concave. Sa courbe est située sous sa tangente en 1, dont l'équation est $v = u - 1$.

La stricte concavité assure que la tangente ne rencontre la courbe qu'en son point de contact. On obtient à nouveau $\ln u \leq u - 1$, avec égalité exactement pour $u = 1$.

La première méthode permet de lire directement le cas d'égalité ; la seconde montre l'origine géométrique de l'inégalité.

3. Méthode 1 — Application de la question 2

Les deux quotients y/x et $(1 - y)/(1 - x)$ sont strictement positifs. Comme x et $1 - x$ sont strictement positifs, la question 2 donne :

$$x \ln \left(\frac{y}{x} \right) \leq y - x,$$

$$(1 - x) \ln \left(\frac{1 - y}{1 - x} \right) \leq x - y.$$

En additionnant, $d(x, y) \leq 0$. Les écarts entre chaque membre de droite et chaque membre de gauche sont positifs ou nuls ; leur somme ne peut être nulle que si chacun est nul. Cela impose $y/x = 1$ et $(1 - y)/(1 - x) = 1$, soit $x = y$. Réciproquement, $d(x, x) = 0$.

Méthode 2 — Concavité du logarithme

Les poids x et $1 - x$ sont strictement positifs et de somme 1. L'inégalité de concavité pour une combinaison finie donne :

$$d(x, y) \leq \ln \left(x \frac{y}{x} + (1 - x) \frac{1 - y}{1 - x} \right) = \ln 1 = 0.$$

La stricte concavité du logarithme donne l'égalité si et seulement si les deux arguments

sont égaux :

$$\frac{y}{x} = \frac{1-y}{1-x} \iff y(1-x) = x(1-y) \iff y = x.$$

Dans les deux approches, $d(x, y) = 0$ exactement sur la diagonale de $]0, 1[^2$. La seconde méthode reconnaît directement une moyenne pondérée.

4. a) Une fonction concave dérivable a une dérivée décroissante. Pour $x > y$, le théorème des accroissements finis, applicable à ℓ sur $[y, x]$, fournit $c \in]y, x[$ tel que :

$$\ell(x) - \ell(y) = \ell'(c)(x - y) \leq \ell'(y)(x - y).$$

Pour $x < y$, il fournit $c \in]x, y[$ tel que la même égalité soit satisfaite. Cette fois, $\ell'(c) \geq \ell'(y)$ et $x - y < 0$, donc la multiplication inverse l'inégalité et donne encore le résultat demandé. Pour $x = y$, les deux membres sont nuls.

Ainsi, pour tout $(x, y) \in \mathbb{R}^2$, $\ell(x) - \ell(y) \leq \ell'(y)(x - y)$.

4. b) Fixons $y \in \mathbb{R}$. D'après 4. a), pour tout $x \in \mathbb{R}$:

$$\ell(r(x))f(x) \leq (\ell(y) + \ell'(y)(r(x) - y))f(x),$$

car $f(x) \geq 0$. Tous les intégrandes sont continus. On intègre d'abord sur un segment $[A, B]$:

$$\int_A^B \ell(r(x))f(x) \, dx \leq \ell(y) \int_A^B f(x) \, dx + \ell'(y) \left(\int_A^B r(x)f(x) \, dx - y \int_A^B f(x) \, dx \right).$$

Les trois intégrales impropres intervenant ont une limite finie par hypothèse. On peut donc faire tendre séparément A vers $-\infty$ et B vers $+\infty$. Comme l'intégrale de f vaut 1,

on obtient exactement :

$$\int_{-\infty}^{+\infty} \ell(r(x))f(x) \, dx \leq \ell(y) + \ell'(y) \left(\int_{-\infty}^{+\infty} r(x)f(x) \, dx - y \right).$$

Aucun échange entre une intégrale et une limite de fonctions n'est utilisé : il s'agit de la définition des intégrales impropres convergentes.

4. c) Posons $m = \int_{-\infty}^{+\infty} r(x)f(x) \, dx$. Ce réel existe par hypothèse et appartient au domaine de ℓ , qui est $\mathbb{R}$.

En prenant $y = m$ dans 4. b), le terme multiplié par $\ell'(m)$ est nul. Ainsi :

$$\int_{-\infty}^{+\infty} \ell(r(x))f(x) \, dx \leq \ell(m) = \ell \left(\int_{-\infty}^{+\infty} r(x)f(x) \, dx \right).$$

5. On reprend la majoration par une tangente, cette fois pour des sommes partielles.

Le réel

$$m = \sum_{n=1}^{+\infty} r(x_n)p_n$$

existe par hypothèse. Pour tout entier $K \geq 1$, 4. a) et $p_n \geq 0$ donnent :

$$\sum_{n=1}^K \ell(r(x_n))p_n \leq \ell(m) \sum_{n=1}^K p_n + \ell'(m) \left(\sum_{n=1}^K r(x_n)p_n - m \sum_{n=1}^K p_n \right).$$

Les trois sommes partielles convergent ; on peut faire tendre K vers $+\infty$. Le dernier terme tend vers $\ell'(m)(m - m) = 0$ et la somme des poids tend vers 1.

On obtient :

$$\sum_{n=1}^{+\infty} \ell(r(x_n))p_n \leq \ell\left(\sum_{n=1}^{+\infty} r(x_n)p_n\right).$$

Le raisonnement ne nécessite ni réarrangement de séries, ni convergence absolue supplémentaire.

Partie II. Entropie dans le cas continu

6. a) La densité normale centrée réduite est continue, strictement positive et normalisée :

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \exp(-x^2/2),$$

$$\ln \varphi(x) = -\frac{1}{2} \ln(2\pi) - \frac{x^2}{2}.$$

La variable Z possède un moment d'ordre deux, égal à $E(Z^2) = V(Z) + E(Z)^2 = 1$. Par transfert, $\int_{-\infty}^{+\infty} x^2 \varphi(x) dx = 1$. En outre :

$$|\varphi(x) \ln \varphi(x)| \leq \frac{1}{2} |\ln(2\pi)| \varphi(x) + \frac{1}{2} x^2 \varphi(x).$$

Le membre de droite est intégrable ; l'intégrale définissant l'entropie est donc absolument convergente et $\varphi \in \mathcal{F}$. Par linéarité des intégrales convergentes :

$$H(Z) = \frac{1}{2} \ln(2\pi) \int_{-\infty}^{+\infty} \varphi(x) dx + \frac{1}{2} \int_{-\infty}^{+\infty} x^2 \varphi(x) dx = \frac{1}{2} \ln(2\pi) + \frac{1}{2}.$$

Donc $H(Z) = \frac{1}{2} \ln(2\pi e).$

6. b) Comme $a > 0$, pour tout $y \in \mathbb{R}$:

$$F_Y(y) = P\left(X \leq \frac{y-b}{a}\right) = F_X\left(\frac{y-b}{a}\right).$$

Puisque f est continue, F_X est dérivable et $F'_X = f$. La dérivation donne une densité continue et strictement positive de Y :

$$g(y) = \frac{1}{a} f\left(\frac{y-b}{a}\right).$$

Le changement de variable affine $y = ax + b$, effectué sur des segments puis aux bornes infinies, donne $\int_{-\infty}^{+\infty} g(y) \, dy = \int_{-\infty}^{+\infty} f(x) \, dx = 1$.

Sur tout segment $[A, B]$, le même changement de variable donne :

$$\int_A^B g(y) \ln g(y) \, dy = \int_{(A-b)/a}^{(B-b)/a} f(x) \ln f(x) \, dx - \ln a \int_{(A-b)/a}^{(B-b)/a} f(x) \, dx.$$

Les deux intégrales du membre de droite convergent quand les bornes tendent vers les infinis correspondants. L'intégrale de $g \ln g$ converge donc, ce qui prouve $g \in \mathcal{F}$. En changeant le signe :

$$H(Y) = H(X) + \ln a.$$

La convergence impropre supposée dans la définition de $\mathcal{F}$ suffit ; on ne lui substitue pas une hypothèse de convergence absolue non fournie.

6. c) Une variable normale d'espérance μ et d'écart-type $\sigma > 0$ a la loi de $\sigma Z + \mu$.
L'entropie dépend uniquement de la densité, donc de la loi. La question 6. b) s'applique avec $a = \sigma$ et $b = \mu$:

$$H(X) = \frac{1}{2} \ln(2\pi e) + \ln \sigma = \frac{1}{2} \ln(2\pi e \sigma^2).$$

L'entropie est indépendante de la moyenne.

7. a) Pour tout x , $g(x)/f(x) > 0$. La question 2 donne :

$$f(x) \ln \left(\frac{g(x)}{f(x)} \right) \leq g(x) - f(x).$$

On intègre sur un segment, puis on passe aux limites impropres. L'intégrale de gauche converge par l'hypothèse de la question et les intégrales de f et g valent chacune 1.

Ainsi :

$$\int_{-\infty}^{+\infty} f(x) \ln \left(\frac{g(x)}{f(x)} \right) dx \leq 0.$$

Par définition, $D(f, g) \geq 0$.

7. b) Sous l'hypothèse $D(f, g) = 0$, la linéarité des intégrales convergentes donne :

$$\int_{-\infty}^{+\infty} \left(\ln \left(\frac{g(x)}{f(x)} \right) + 1 - \frac{g(x)}{f(x)} \right) f(x) dx = -D(f, g) + 1 - 1 = 0.$$

Posons $v(x) = f(x) \left(\frac{g(x)}{f(x)} - 1 - \ln \left(\frac{g(x)}{f(x)} \right) \right)$. Cette fonction est continue et positive ou nulle par la question 2. Son intégrale vaut 0.

Si $v(x_0) > 0$ en un point, sa continuité fournit un segment de longueur strictement positive autour de x_0 sur lequel $v(x) \geq v(x_0)/2$. Son intégrale sur ce segment serait strictement positive ; par positivité, cela contredirait l'intégrale totale nulle. Ainsi, v est identiquement nulle.

Puisque $f(x) > 0$ pour tout x , le cas d'égalité de la question 2 impose $g(x)/f(x) = 1$ pour tout x . Donc $f = g$ sur $\mathbb{R}$. Réciproquement, $f = g$ donne bien $D(f, g) = 0$.

Partie III. Entropie dans le cas discret

8. a) La fonction h_N est de classe C^2 sur l'ouvert $]0, 1[^N$, car chaque application

$$x_k \mapsto -x_k \ln x_k \text{ y est de classe } C^2.$$

Pour $k \in \{1, \dots, N\}$:

$$\frac{\partial h_N}{\partial x_k}(x) = -\ln x_k - 1.$$

Ainsi :

$$\nabla h_N(x) = \begin{pmatrix} -\ln x_1 - 1 \\ \vdots \\ -\ln x_N - 1 \end{pmatrix}.$$

Chaque terme de la somme ne dépend que d'une coordonnée; toutes les dérivées secondes mixtes sont nulles et :

$$\frac{\partial^2 h_N}{\partial x_k^2}(x) = -\frac{1}{x_k},$$

$$\nabla^2 h_N(x) = \text{diag} \left(-\frac{1}{x_1}, \dots, -\frac{1}{x_N} \right).$$

Les vecteurs de la base canonique sont des vecteurs propres; les valeurs propres sont les nombres $-1/x_k$, tous strictement négatifs. Aucune utilisation d'un déterminant n'est nécessaire.

8. b) La contrainte est linéaire. Un point critique sous cette contrainte est un point où le gradient de h_N appartient à l'espace engendré par $(1, \dots, 1)$, gradient de la somme des coordonnées.

Il existe donc un réel λ tel que, pour tout k , $-\ln x_k - 1 = \lambda$. Toutes les coordonnées sont alors égales à $\exp(-1 - \lambda)$. Leur somme valant 1, elles valent chacune $1/N$.

Le point :

$$x^* = \left(\frac{1}{N}, \dots, \frac{1}{N} \right)$$

appartient bien à $]0, 1[^N$, puisque $N \geq 2$, satisfait la contrainte et a un gradient égal à $(\ln N - 1)(1, \dots, 1)$. C'est donc l'unique point critique sous contrainte. Sa hessienne vaut $-NI_N$ et ses valeurs propres sont toutes égales à $-N < 0$.

8. c) Question adaptée au programme actuel : la formule de Taylor est appliquée à une fonction d'une variable le long du segment reliant x^* à x .

Méthode 1 — Taylor sur un segment

Soit $x \in]0, 1[^N$ tel que

$$\sum_{k=1}^N x_k = 1.$$

Posons $v = x - x^*$ et $g(t) = h_N(x^* + tv)$ pour $t \in [0, 1]$. Le segment est inclus dans $]0, 1[^N$; comme ses coordonnées sont strictement comprises entre 0 et 1, g est de classe C^2 sur un ouvert contenant $[0, 1]$.

Par dérivation le long d'une droite :

$$g'(0) = \nabla^t(h_N(x^*))v = (\ln N - 1) \sum_{k=1}^N \left(x_k - \frac{1}{N}\right) = 0,$$

$$g''(t) = - \sum_{k=1}^N \frac{v_k^2}{1/N + tv_k} \leq 0.$$

La formule de Taylor avec reste intégral à l'ordre 1 donne :

$$h_N(x) - h_N(x^*) = g(1) - g(0) = g'(0) + \int_0^1 (1-t)g''(t) dt \leq 0.$$

Si $x \neq x^*$, au moins un v_k est non nul, donc $g''(t) < 0$ pour tout $t \in [0, 1]$ et l'intégrale est strictement négative. Le maximum global est ainsi unique et vaut :

$$h_N(x^*) = -N \frac{1}{N} \ln(1/N) = \ln N.$$

Méthode 2 — Application de la question 5

Cette méthode exige une précaution de domaine : la fonction $-u \ln u$ n'est pas définie sur tout $\mathbb{R}$, contrairement à la fonction ℓ de la question 5.

Pour le point x fixé, posons $a = \frac{1}{2} \min\{1, Nx_1, \dots, Nx_N\} > 0$ et $b = 1 + \max\{1, Nx_1, \dots, Nx_N\}$. Sur $]0, +\infty[$, notons $\psi(u) = -u \ln u$.

Définissons sur $\mathbb{R}$:

$$\ell(u) = \begin{cases} \psi(a) + \psi'(a)(u - a) & \text{si } u < a, \\ \psi(u) & \text{si } a \leq u \leq b, \\ \psi(b) + \psi'(b)(u - b) & \text{si } u > b. \end{cases}$$

Les valeurs et les dérivées se raccordent en a et b . La dérivée est constante sur chaque intervalle extérieur et décroissante sur $[a, b]$, car $\psi''(u) = -1/u < 0$. Ainsi, ℓ est de classe C^1 et concave sur $\mathbb{R}$.

Appliquons 5 avec $r(u) = u$, les poids $1/N$ pour les indices de 1 à N puis 0, et les valeurs Nx_k pour les premiers indices, puis 1. Toutes les séries sont finies et les hypothèses sont satisfaites. Leur moyenne pondérée vaut

$$\sum_{k=1}^N x_k = 1;$$

les arguments utiles appartiennent à $[a, b]$. On obtient :

$$\frac{1}{N} \sum_{k=1}^N (-Nx_k \ln(Nx_k)) \leq \psi(1) = 0.$$

Le membre de gauche vaut $h_N(x) - \ln N$. On conclut encore $h_N(x) \leq \ln N$.

La stricte concavité de ψ sur $[a, b]$ assure que l'égalité pour cette combinaison à poids strictement positifs n'est possible que si les Nx_k sont tous égaux. Cela donne $x = x^*$, et réciproquement l'égalité y est réalisée.

La première méthode donne directement la stricte maximalité par la hessienne ; la seconde illustre l'inégalité de moyenne de la partie I, tout en respectant son domaine d'application.

8. d) Il faut aussi traiter les lois possédant des probabilités nulles, absentes du domaine ouvert de h_N .

Soit $J = \{k : p_k > 0\}$ et m son cardinal. Comme les probabilités ont pour somme 1, $1 \leq m \leq N$.

Si $m = 1$, la variable est certaine et $H(X) = 0 < \ln N$. Si $m \geq 2$, les probabilités sur J sont strictement comprises entre 0 et 1 et de somme 1. La question 8. c), avec m à la place de N , donne :

$$H(X) = - \sum_{k \in J} p_k \ln p_k \leq \ln m \leq \ln N.$$

Pour avoir $H(X) = \ln N$, il faut $m = N$ et l'égalité dans 8. c), donc $p_1 = \dots = p_N = 1/N$. Réciproquement, ces probabilités donnent une entropie égale à $\ln N$.

La loi de plus grande entropie est donc uniquement la loi uniforme sur $\{1, \dots, N\}$.

9. a) La convergence de la série

$$\sum_{n \geq 1} p_n$$

implique $p_n \rightarrow 0$. Les termes sont strictement positifs et inférieurs à 1, car tous les autres termes sont également strictement positifs.

Pour $t \rightarrow 0^+$, en posant $u = 1/\sqrt{t} \rightarrow +\infty$, on a :

$$\sqrt{t} |\ln t| = \frac{2 \ln u}{u} \rightarrow 0$$

par croissance comparée du logarithme et d'une puissance. Par conséquent, $\sqrt{p_n} |\ln p_n| \rightarrow 0$. Il existe un entier n_0 , que l'on peut prendre supérieur ou égal à 2, tel que, pour tout $n \geq n_0$, $\sqrt{p_n} |\ln p_n| \leq 1$.

9. b) Si $n \geq n_0$ et $p_n \leq 1/n^3$, on écrit :

$$p_n |\ln p_n| = \sqrt{p_n} (\sqrt{p_n} |\ln p_n|) \leq \sqrt{p_n} \leq \frac{1}{n^{3/2}}.$$

C'est la majoration demandée.

9. c) Pour $n \geq n_0$, distinguons deux cas.

Si $p_n \leq 1/n^3$, la question 9. b) donne $p_n |\ln p_n| \leq 1/n^{3/2}$.

Si $p_n > 1/n^3$, puisque $p_n < 1$ et que le logarithme est croissant :

$$0 < |\ln p_n| = -\ln p_n < 3 \ln n,$$

$$\text{et donc } p_n |\ln p_n| \leq 3p_n \ln n.$$

Dans tous les cas :

$$p_n |\ln p_n| \leq \max \left\{ \frac{1}{n^{3/2}}, 3p_n \ln n \right\}.$$

9. d) Les deux quantités dans le maximum sont positives ou nulles. Le maximum est donc inférieur ou égal à leur somme. De plus, $\ln n \leq n$ pour $n \geq 1$. Ainsi, pour $n \geq n_0$:

$$0 \leq p_n |\ln p_n| \leq \frac{1}{n^{3/2}} + 3np_n.$$

La série

$$\sum 1/n^{3/2}$$

converge comme série de Riemann d'exposant strictement supérieur à 1, et

$$\sum np_n$$

converge par hypothèse. Le critère de comparaison des séries à termes positifs donne la convergence de

$$\sum p_n |\ln p_n|.$$

Une variable aléatoire à valeurs dans $\mathbb{N}^*$ et d'espérance finie possède donc une entropie finie. Pour les lois appartenant à $\mathcal{S}$, c'est exactement le résultat établi. Si certains p_n sont nuls, on utilise la convention $p_n |\ln p_n| = 0$: les majorations précédentes s'appliquent aux indices où $p_n > 0$, et restent vraies aux autres indices puisque le terme est alors nul.

10. a) La loi est géométrique sur $\mathbb{N}^*$, de paramètre $\theta \in]0, 1[$. En effet, les probabilités sont $\theta(1 - \theta)^{n-1}$ et leur somme vaut 1 par la série géométrique. On a :

$$E(X) = \frac{1}{\theta}.$$

L'existence de l'entropie découle de 9. d), car $E(X) < +\infty$. Puisque $0 < p_n < 1$, $|\ln p_n| = -\ln p_n$ et :

$$\ln p_n = \ln \theta + (n - 1) \ln(1 - \theta).$$

Les sommes

$$\sum p_n$$

et

$$\sum (n - 1)p_n$$

sont convergentes, la seconde valant $E(X) - 1 = (1 - \theta)/\theta$. Par linéarité :

$$H(X) = -\ln \theta - \frac{1-\theta}{\theta} \ln(1-\theta).$$

Avec la fonction de la question 1, cela s'écrit aussi $H(X) = h(\theta)/\theta$.

10. b) Code adapté en Python. Le paramètre fourni doit vérifier $0 < \theta < 1$.

Méthode 1 — Attente du premier succès

Python

```
import numpy as np

def X(theta):
    n = 1
    while np.random.random() >= theta:
        n = n + 1
    return n
```

Chaque tirage uniforme représente un essai indépendant : il y a succès si le tirage est strictement inférieur à θ . La variable n compte le rang de cet essai. La fonction renvoie n si les $n-1$ premiers essais échouent et le suivant réussit, ce qui a pour probabilité $(1-\theta)^{n-1}\theta$.

La probabilité d'effectuer plus de K essais vaut $(1-\theta)^K$, qui tend vers 0. L'algorithme se termine donc presque sûrement. La valeur renvoyée appartient à $\mathbb{N}^*$ et suit bien la loi demandée.

Méthode 2 — Générateur de la loi usuelle

Python

```
import numpy as np

def X(theta):
    return np.random.geometric(theta)
```

Dans NumPy, `geometric(theta)` renvoie le rang du premier succès, avec support $\{1, 2, \dots\}$ et probabilité $\theta(1 - \theta)^{n-1}$ au rang n . Il ne s'agit donc pas du nombre d'échecs précédant le premier succès. Cette convention coïncide avec celle de l'énoncé.

Les deux codes ont été exécutés pour $\theta = 0,2, 0,5$ et $0,8$. Les valeurs renvoyées étaient des entiers strictement positifs et les fréquences observées cohérentes avec les probabilités théoriques. Ces observations contrôlent les programmes ; la justification de leur loi repose sur les arguments précédents.

La première méthode explique le mécanisme probabiliste ; la seconde convient à une simulation plus directe.

10. c) L'espérance de Y est finie, égale à $1/\theta$; la question 9. d) assure donc que $H(Y)$ existe. L'entropie de X existe par 10. a).

La série

$$\sum q_n \ln p_n$$

converge absolument car :

$$|\ln p_n| \leq |\ln \theta| + (n-1)|\ln(1-\theta)|,$$

et

$$\sum q_n = 1,$$

$$\sum (n-1)q_n = E(Y) - 1 < +\infty.$$

On peut donc utiliser la linéarité et l'égalité des espérances :

$$\sum_{n=1}^{+\infty} q_n \ln p_n = \ln \theta + \left(\frac{1}{\theta} - 1 \right) \ln(1 - \theta) = -H(X).$$

D'autre part,

$$\sum q_n \ln q_n = -H(Y),$$

cette série étant absolument convergente. Par différence :

$$\sum_{n=1}^{+\infty} q_n \ln \left(\frac{p_n}{q_n} \right) = -H(X) + H(Y).$$

Donc

$$H(Y) - H(X) = \sum_{n=1}^{+\infty} q_n \ln(p_n/q_n).$$

Le raisonnement montre également que la convergence de cette dernière série est déjà assurée par les autres hypothèses.

10. d) Pour chaque n , $q_n > 0$ et $p_n/q_n > 0$. D'après la question 2 :

$$q_n \ln \left(\frac{p_n}{q_n} \right) \leq p_n - q_n.$$

En sommant de 1 à K , puis en passant à la limite dans les séries convergentes, il vient :

$$H(Y) - H(X) \leq \sum_{n=1}^{+\infty} p_n - \sum_{n=1}^{+\infty} q_n = 0.$$

Pour préciser l'égalité, posons :

$$a_n = p_n - q_n - q_n \ln \left(\frac{p_n}{q_n} \right) \geq 0.$$

Les sommes partielles de

$$\sum a_n$$

convergent vers $H(X) - H(Y)$. Si cette limite est nulle, tous les a_n sont nuls, car ils sont positifs ou nuls. Le cas d'égalité de la question 2 impose alors $p_n = q_n$ pour tout n .

Réciproquement, l'identité des lois donne l'égalité des entropies.

Ainsi, parmi les lois de $\mathcal{S}$ ayant l'espérance $1/\theta$, la loi géométrique de paramètre θ est l'unique loi d'entropie maximale. Sa valeur maximale est celle calculée en 10. a).

Partie IV. Entropie et taux de rendement asymptotique

- 11.** a) Question adaptée au programme actuel : on décrit les événements associés aux transformations, sans faire appel à une théorie abstraite de la mesurabilité.

L'exponentielle prend des valeurs strictement positives. Pour $t \leq 0$, les événements $[Z_n \leq t]$ et $[Z \leq t]$ sont donc vides.

Pour $t > 0$, l'exponentielle est strictement croissante et de fonction réciproque le logarithme. On a donc :

$$[Z_n \leq t] = [X_n \leq \ln t],$$

$$[Z \leq t] = [X \leq \ln t].$$

Ces expressions permettent notamment d'écrire leurs fonctions de répartition :

$F_{Z_n}(t) = 0$ si $t \leq 0$, et $F_{Z_n}(t) = F_{X_n}(\ln t)$ si $t > 0$, avec la même formule pour Z .

11. b) Pour $m \in \mathbb{N}^*$, les événements $A_m = [|X| \geq m]$ forment une suite décroissante.

Leur intersection est vide puisque X prend, pour chaque issue, une valeur réelle finie.

La continuité monotone des probabilités donne :

$$\lim_{m \rightarrow +\infty} P(A_m) = P\left(\bigcap_{m \geq 1} A_m\right) = 0.$$

Pour tout $\alpha > 0$, il existe donc un entier $s \geq 1$ tel que $P(|X| \geq s) < \alpha$.

11. c) Pour deux événements A et B , la formule du crible donne

$P(A \cup B) = P(A) + P(B) - P(A \cap B) \leq P(A) + P(B)$. En l'appliquant d'abord à $K_1 \cup K_2$ et K_3 , puis à K_1 et K_2 , on obtient :

$$P(K_1 \cup K_2 \cup K_3) \leq P(K_1) + P(K_2) + P(K_3).$$

Définissons :

$$K_1 = [|X| \geq s], \quad K_2 = [|X_n - X| \geq 1], \quad K_3 = [|X_n - X| \geq \varepsilon e^{-1-s}].$$

Sur le complémentaire de leur union, on a $|X| < s$ et $|X_n - X| < 1$, donc $|X_n| < s + 1$.

Les deux arguments X_n et X appartiennent à $[-s - 1, s + 1]$.

La dérivée de l'exponentielle y est majorée par e^{s+1} . L'inégalité des accroissements finis donne donc, issue par issue :

$$|Z_n - Z| = |e^{X_n} - e^X| \leq e^{s+1}|X_n - X| < e^{s+1}\varepsilon e^{-1-s} = \varepsilon.$$

Par contraposée, $[|Z_n - Z| \geq \varepsilon] \subset K_1 \cup K_2 \cup K_3$. La croissance des probabilités et la

majoration précédente donnent :

$$P(|Z_n - Z| \geq \varepsilon) \leq P(|X| \geq s) + P(|X_n - X| \geq 1) + P(|X_n - X| \geq \varepsilon e^{-1-s}).$$

11. d) Fixons $\varepsilon > 0$ et une précision $\eta > 0$. Par 11. b), choisissons $s \geq 1$ tel que

$$P(|X| \geq s) < \eta/3.$$

Le nombre $\delta = \varepsilon e^{-1-s}$ est strictement positif et fixé. La convergence de X_n vers X en probabilité entraîne l'existence d'un rang n_1 à partir duquel :

$$P(|X_n - X| \geq 1) < \eta/3 \quad \text{et} \quad P(|X_n - X| \geq \delta) < \eta/3.$$

L'inégalité de 11. c) donne alors, pour tout $n \geq n_1$, $P(|Z_n - Z| \geq \varepsilon) < \eta$.

Ainsi, pour tout $\varepsilon > 0$, $P(|Z_n - Z| \geq \varepsilon) \rightarrow 0$. Par définition :

$$\exp(X_n) \xrightarrow{P} \exp(X).$$

Ce raisonnement établit ici la propriété de composition par l'exponentielle à partir d'une majoration explicite.

12. a) Toutes les mises sont dépensées. Lorsque le cheval gagnant est $G_n = k$, seule la mise sur ce cheval rapporte de l'argent. Le montant reçu est $R_{n-1} f_k c_k$, avec $R_0 = r_0$. Ainsi :

$$M_n = f_{G_n} c_{G_n}.$$

Posons $m(k) = f_k c_k$. La fonction m est la même à chaque course et ses valeurs sont strictement positives. On a donc $M_n = m(G_n) > 0$.

Les variables G_n sont mutuellement indépendantes. Le lemme des coalitions assure que leurs transformations séparées par m restent mutuellement indépendantes. Comme les G_n ont la même loi et que la fonction de transformation est identique, les M_n ont également la même loi.

Pour toute valeur réelle u , cette loi s'écrit précisément :

$$P(M_n = u) = \sum_{\{k \in \{1, \dots, N\} : f_k c_k = u\}} p_k.$$

Cette écriture tient compte des éventuelles égalités entre plusieurs produits $f_k c_k$. On ne suppose donc pas à tort que ces produits sont distincts.

12. b) Posons $L_i = \ln M_i$ et $\mu_L = E(L_1)$. Les L_i sont indépendantes et de même loi par transformation des M_i . Leur espérance et leur variance sont finies par hypothèse ; en fait, leurs valeurs appartiennent à un ensemble fini de nombres réels.

Comme

$$R_n/r_0 = \prod_{i=1}^n M_i > 0,$$

on peut prendre le logarithme et écrire :

$$T_n = \exp \left(\frac{1}{n} \sum_{i=1}^n L_i \right) - 1.$$

Par indépendance et linéarité de l'espérance :

$$E \left(\frac{1}{n} \sum_{i=1}^n L_i \right) = \mu_L,$$

$$V \left(\frac{1}{n} \sum_{i=1}^n L_i \right) = \frac{V(L_1)}{n}.$$

L'inégalité de Bienaymé-Tchebychev donne, pour tout $\delta > 0$:

$$P \left(\left| \frac{1}{n} \sum_{i=1}^n L_i - \mu_L \right| \geq \delta \right) \leq \frac{V(L_1)}{n\delta^2} \rightarrow 0.$$

Cela prouve la convergence en probabilité de la moyenne des L_i vers la constante μ_L , y compris si $V(L_1) = 0$. La question 11, ou le théorème de composition en probabilité par la fonction continue $t \mapsto e^t - 1$, donne :

$$T_n \xrightarrow{P} \tau, \quad \tau = \exp(E(\ln M_1)) - 1.$$

La constante τ est le taux de rendement asymptotique demandé.

13. a) Par 12. a), $\ln M_1 = \ln(f_{G_1} c_{G_1})$. Cette variable est à valeurs dans un ensemble fini, donc son espérance existe. Par transfert pour une variable finie :

$$E(\ln M_1) = \sum_{k=1}^N p_k \ln(f_k c_k).$$

La formule de 12. b) devient :

$$\tau = \exp \left(\sum_{k=1}^N p_k \ln(f_k c_k) \right) - 1.$$

13. b) Méthode 1 — Inégalité logarithmique

L'exponentielle étant strictement croissante, maximiser τ équivaut à maximiser :

$$L(f) = \sum_{k=1}^N p_k \ln(f_k c_k),$$

sur les proportions $f_k > 0$ de somme 1. Les probabilités p_k sont elles-mêmes strictement comprises entre 0 et 1 et de somme 1 : la stratégie $f_k = p_k$ est donc admissible.

Pour une stratégie admissible quelconque, les cotes se simplifient dans la différence :

$$L(f) - L(p) = \sum_{k=1}^N p_k \ln \left(\frac{f_k}{p_k} \right).$$

D'après la question 2 :

$$p_k \ln \left(\frac{f_k}{p_k} \right) \leq f_k - p_k.$$

En sommant,

$$L(f) - L(p) \leq \sum f_k - \sum p_k = 0.$$

Les poids étant strictement positifs, l'égalité impose, terme par terme, $f_k/p_k = 1$. Le maximum est donc unique.

La stratégie optimale consiste à miser sur chaque cheval une proportion égale à sa probabilité de victoire :

$$f_k^* = p_k \quad \text{pour tout } k \in \{1, \dots, N\}.$$

La valeur optimale du taux est :

$$\tau_{\text{opt}} = \exp \left(\sum_{k=1}^N p_k \ln(p_k c_k) \right) - 1.$$

Avec l'entropie de la loi des gagnants, cette expression devient :

$$\tau_{\text{opt}} = \exp \left(\sum_{k=1}^N p_k \ln c_k - H(G_1) \right) - 1.$$

Méthode 2 — Gradient, hessienne et contrainte linéaire

La fonction L est de classe C^2 sur l'ouvert convexe $]0, +\infty[^N$. Ses dérivées sont :

$$\frac{\partial L}{\partial f_k}(f) = \frac{p_k}{f_k}, \quad \nabla^2 L(f) = \text{diag} \left(-\frac{p_1}{f_1^2}, \dots, -\frac{p_N}{f_N^2} \right).$$

Les valeurs propres de cette matrice diagonale sont strictement négatives. Pour un point critique sous la contrainte

$$\sum f_k = 1,$$

le gradient doit être colinéaire à $(1, \dots, 1)$: il existe λ tel que $p_k/f_k = \lambda$ pour tout k .

Ces rapports étant positifs, $\lambda > 0$. En sommant les relations $p_k = \lambda f_k$, on obtient $1 = \lambda$, puis $f_k = p_k$. Ce point est admissible, son gradient vaut $(1, \dots, 1)$ et sa hessienne est $\text{diag}(-1/p_1, \dots, -1/p_N)$, à valeurs propres strictement négatives.

Pour vérifier la maximalité globale sous contrainte, fixons $f \neq p$ admissible et posons $w = f - p$, $q(t) = L(p + tw)$ sur $[0, 1]$. Le segment est inclus dans l'ouvert positif et q est de classe C^2 sur un voisinage de ce segment. On a :

$$q'(0) = \sum_{k=1}^N w_k = 0,$$

$$q''(t) = - \sum_{k=1}^N \frac{p_k w_k^2}{(p_k + t w_k)^2} < 0.$$

La stricte inégalité provient du fait qu'au moins un w_k est non nul. Taylor avec reste intégral donne $q(1) - q(0) = \int_0^1 (1-t)q''(t) dt < 0$. Donc $L(f) < L(p)$ et le maximum global est bien unique.

On retrouve la stratégie $f^* = p$ et la valeur τ_{opt} calculée ci-dessus. L'inégalité logarithmique fournit une comparaison immédiate de toutes les stratégies ; le calcul différentiel explique la même optimisation dans le cadre des contraintes linéaires.

13. c) Posons $q_k = 1/c_k$. Les cotes étant strictement supérieures à 1, chaque q_k appartient à $]0, 1[$. L'hypothèse de cette question donne

$$\sum_{k=1}^N q_k = 1.$$

Pour chaque k , la question 2 appliquée à $q_k/p_k > 0$ donne :

$$p_k \ln \left(\frac{q_k}{p_k} \right) \leq q_k - p_k.$$

En sommant et en utilisant les deux normalisations :

$$\sum_{k=1}^N p_k \ln \left(\frac{q_k}{p_k} \right) \leq 0.$$

Puisque $\ln(p_k c_k) = \ln(p_k/q_k)$, il vient :

$$\sum_{k=1}^N p_k \ln(p_k c_k) \geq 0.$$

Le cas d'égalité de la question 2 montre, comme en 13. b), que l'égalité a lieu si et seulement si $p_k = q_k = 1/c_k$ pour tout k .

Par croissance de l'exponentielle, $\tau_{\text{opt}} \geq 0$, et :

$$\tau_{\text{opt}} = 0 \iff p_k = \frac{1}{c_k} \text{ pour tout } k.$$

C'est donc exactement lorsque les probabilités de victoire coïncident avec les inverses des cotes que le parieur ne peut obtenir de taux asymptotique strictement positif. Dans

ce cas, la meilleure stratégie a un taux nul et toutes les autres stratégies admissibles ont un taux strictement négatif, par l'unicité du maximum établie en 13. b). Si au moins une de ces égalités est fausse, la stratégie $f_k = p_k$ donne un taux asymptotique strictement positif.

Relevé des adaptations

- 8. c) La mention de Taylor-Lagrange est remplacée par la formule de Taylor à l'ordre 1 pour $g(t) = h_N(x^* + t(x - x^*))$. La réduction explicite à une variable préserve la méthode proposée et reste dans le référentiel. L'autre possibilité, l'utilisation de la question 5, est conservée.
- 10. b) Code adapté en Python : la demande de simulation dans un ancien langage est remplacée par une fonction `def X(theta) :`. Le support géométrique $\mathbb{N}^*$ et le comportement probabiliste sont préservés.
- 11. a) Question adaptée au programme actuel : la vérification abstraite du statut de variable aléatoire des applications exponentielles est remplacée par la description des événements $[Z_n \leq t]$ et $[Z \leq t]$. Elle conserve l'étude de ces transformations sans mobiliser de théorie de la mesurabilité.