Various feedback on the course:

By Ludovic Schwartz: a convex function $f : \mathcal{X} \to \mathbb{R}$ (or even $[0,1]$) does not necessarily achieve its infimum,

hence we safely and rightfully considered p_δ^* as an approximate arginf in the last proof of Lecture 1

Ex: $f(x) = x^2 + \mathbb{1}_{\{x=0\}}$, infimum equal to 0, never achieved

(take $f(x) = x^2 + \mathbb{1}_{\{x=0\}}$ for an exp-concave example)

By Hassan Saber: in the last proof, it is easy to directly optimize over $\varepsilon \in (0,1)$ the obtained bound:

$$B(\varepsilon) = T \ln(1-\varepsilon) + (N-1)\ln \varepsilon \qquad B'(\varepsilon) = \frac{-T}{1-\varepsilon} + \frac{N-1}{\varepsilon}$$

$B''(\varepsilon) < 0$ thus the ε^* s.t. $B'(\varepsilon^*) = 0$ is the global maximum of B

We have $B'(\varepsilon^*) = 0 \iff \dfrac{\varepsilon^*}{N-1} = \dfrac{1}{T} - \dfrac{\varepsilon^*}{T} \iff \varepsilon^* = \dfrac{1/T}{1/T + 1/(N-1)} = \dfrac{N-1}{T+N-1}$

We thus get the sharper final bound

$$\frac{1}{\eta}\left(T \ln\left(\frac{T+N-1}{T}\right) + (N-1)\ln\left(\frac{T+N-1}{N-1}\right)\right) = \frac{T+N-1}{\eta} H\left(\frac{N-1}{T+N-1}\right)$$

$$= \frac{1}{\eta}\left(\underbrace{(N-1)\ln\left(1+\frac{N-1}{T}\right)}_{\leqslant (N-1)^2/T} + (N-1)\ln\left(1+\frac{T}{N-1}\right)\right)$$

where $H(x) = x\ln x + (1-x)\ln(1-x)$ is the binary entropy

* However * both bounds are $\sim \dfrac{N-1}{\eta} \ln T$ as $T \to +\infty$.

Proposition de meilleure solution la question 2 de l'exercice 3
Auteur : Hassan Saber (non relu en détails par Gilles Stoltz)

* Calcul de $\inf_{\varepsilon \in (0,1)} -\ln\left(\varepsilon^{\alpha} e^{-\beta\varepsilon}\right)$ pour $\alpha, \beta > 0$

Avec $f(\varepsilon) = -\ln\left(\varepsilon^{\alpha} e^{-\beta\varepsilon}\right)$ pour $\varepsilon \in (0,1)$

$$= \alpha(-\ln\varepsilon) + \beta\varepsilon$$

· $f(0) = +\infty$, $f(1) = \beta$

Ainsi $\begin{cases} \text{ou bien} \quad \operatorname{argmin}_{\varepsilon \in (0,1)} f(\varepsilon) = 1 \text{ et } \inf_{\varepsilon \in (0,1)} f(\varepsilon) = \beta \\ \text{ou bien} \quad \varepsilon^* = \operatorname{argmin}_{\varepsilon \in (0,1)} f(\varepsilon) \in (0,1) \text{ et } f'(\varepsilon^*) = 0 \end{cases}$

$$f'(\varepsilon^*) = -\frac{\alpha}{\varepsilon^*} + \beta = 0 \iff \varepsilon^* = \frac{\alpha}{\beta} \quad (1)$$

(1) n'est possible que si $\frac{\alpha}{\beta} < 1$ alors $f(\varepsilon^*) = \alpha\left(\ln\left(\frac{\beta}{\alpha}\right) + 1\right)$

$$\leq \alpha\left(\frac{\beta}{\alpha} - 1 + 1\right)$$
$$\leq \beta$$

Ainsi, on a montré que

$$\inf_{\varepsilon \in (0,1)} f(\varepsilon) = \mathbb{1}_{\frac{\beta}{\alpha} > 1}\, \alpha\left(\ln\left(\frac{\beta}{\alpha}\right) + 1\right) + \mathbb{1}_{\frac{\beta}{\alpha} \leq 1}\, \beta.$$

On cherche un encadrement de $\inf_{\eta>0} Q(\eta)$

où $Q(\eta) = \frac{\eta}{8}(M-m)^2 T + \frac{1}{\eta} \inf_{\varepsilon\in(0,1)} \{ -\ln \varepsilon^{N-1} e^{-\eta T(M-m)\varepsilon} \}$

On pose $\alpha = N-1$, $\beta = \eta T(M-m)$ et $\gamma = \frac{T(M-m)}{N-1}$

• D'après ce qui précède, si $\eta\gamma \leq 1$, $Q(\eta) = \frac{\eta}{8}(M-m)^2 T + T(M-m)$

d'où $\inf_{\eta\gamma\leq 1} Q(\eta) = T(M-m)$

• si $\eta\gamma > 1$, $Q(\eta) = (\eta\gamma)\frac{(M-m)(N-1)}{8} + \frac{1}{(\eta\gamma)} T(M-m)\left(\ln(\eta\gamma)+1\right)$

d'où $\inf_{\eta\gamma>1} Q(\eta) = \inf_{\eta>1} \eta \frac{(M-m)(N-1)}{8} + \frac{T(M-m)}{\eta}\left(\ln(\eta)+1\right)$

$= \frac{(M-m)(N-1)}{4} \inf_{\eta>1} \frac{1}{2}\eta + \frac{a^2}{\eta}\left(\ln(\eta)+1\right)$

avec $a^2 = \frac{4T}{N-1}$

On pose alors $g(\eta) = \frac{1}{2}\eta + \frac{a^2}{\eta}\left(\ln(\eta)+1\right)$

On cherche alors $\inf_{\eta>1} g(\eta)$.

 Lecture 1 — January 24, 2018 Lecture 1 — January 24, 2018

g est $\mathcal{C}^1$ et $g'(\eta) = \frac{1}{2} - \frac{a^2}{\eta^2}\big(\ln(\eta)+1\big) + \frac{a^2}{\eta^2}$

$$= \frac{\frac{1}{2}\eta^2 - a^2\ln(\eta)}{\eta^2}$$

On pose $h(\eta) = \frac{1}{2}\eta^2 - a^2\ln(\eta)$ pour $\eta \geqslant 1$.

h est $\mathcal{C}^1$ et $h'(\eta) = \eta - \frac{a^2}{\eta}$

$$h''(\eta) = 1 + \frac{a^2}{\eta^2} \geqslant 0 \qquad h'(a) = 0$$

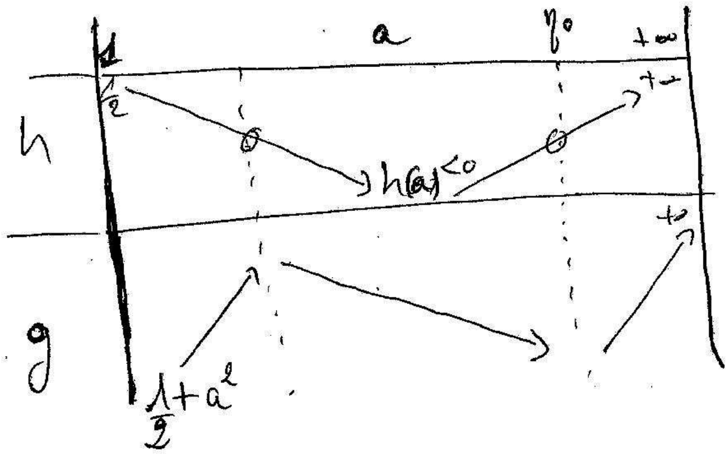
D'où

	1		a		$+\infty$
h''		$+$		$+$	
h'		$-$	0	$+$	
h	$1/2$	$\searrow$	$h(a)$	$\nearrow$	$+\infty$

• Si $h(a) = \frac{1}{2}a^2\big(1 - \ln(a^2)\big) \geqslant 0 \iff a^2 \leq \exp(1)$

alors $h \geqslant 0$ et $g' \geqslant 0$ et $\inf_{\eta \geqslant 1} g(\eta) = g(1) = \frac{1}{2} + a^2$

• Si $a^2 > \exp(1)$, $h(a) < 0$ et



(✱)

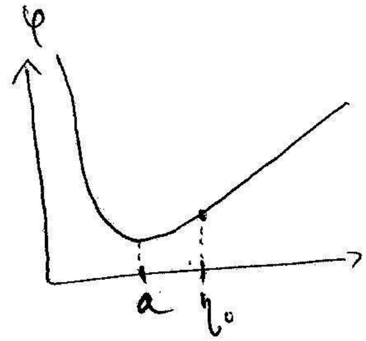
On cherche alors à estimer $g(\eta_0)$.

$g'(\eta_0) = 0$ d'où $\frac{1}{2}\eta_0^2 = a^2 \ln(\eta_0)$

$$\text{et} \quad g(\eta_0) = \frac{1}{2}\eta_0 + \frac{a^2 \ln(\eta_0)}{\eta_0} + \frac{a^2}{\eta_0}$$

$$= \eta_0 + \frac{a^2}{\eta_0}$$

Or avec $\varphi : \eta \to \eta + \frac{a^2}{\eta}$, on a :



En regardant (✱), on voit que $a < \eta_0$. Ainsi si on trouve $\eta_1 > \eta_0$ on aura $g(\eta_0) = \varphi(\eta_0) \leq \varphi(\eta_1)$.

 Lecture 1 — January 24, 2018 Lecture 1 — January 24, 2018

Borne inférieure $g(\eta_0)$ ...

$$\frac{1}{2}\eta_0^2 = a^2 \ln(\eta_0) \Rightarrow \eta_0 = a\sqrt{2\ln(\eta_0)}$$

d'où $$\frac{1}{2}\eta_0^2 = a^2 \ln(\eta_0) = a^2 \ln(a) + \frac{a^2}{2}\ln(2) + \frac{1}{2}a^2 \ln\ln(\eta_0)$$

Or $\eta_0 > a$ et $a > \exp(1)$ d'où $\ln\ln(\eta_0) > 0$

et $$\frac{1}{2}\eta_0^2 \geqslant a^2 \left(\ln(a) + \frac{\ln(2)}{2}\right)$$

i.e. $$\eta_0 \geqslant \eta_1 := a\sqrt{2\ln(a) + \ln(2)}$$

Or $a \leq \eta_1 \leq \eta_0$ donc $\varphi(a) \leq \varphi(\eta_1) \leq \varphi(\eta_0) = g(\eta_0)$

d'où $$g(\eta_0) \geqslant a\sqrt{2\ln(a) + \ln(2)} + \frac{a}{\sqrt{2\ln(a)+\ln(2)}}$$

Borne supérieure sur $g(\eta_0)$. $\varepsilon > 0$

$$\begin{aligned}
h(a^{1+\varepsilon}) &= \frac{1}{2}a^{2+2\varepsilon} - a^2 \ln(a^{1+\varepsilon}) \\
&= \frac{1}{2}a^2\left[(a^2)^{\varepsilon} - (1+\varepsilon)\ln(a^2)\right] \\
&= \frac{1}{2}a^2\left[e^{\varepsilon \ln(a^2)} - (1+\varepsilon)\ln(a^2)\right] \\
&\geqslant \frac{1}{2}a^2\left[1 + \varepsilon\ln(a^2) + \frac{\varepsilon^2(\ln(a^2))^2}{2} - (1+\varepsilon)\ln(a^2)\right]
\end{aligned}$$

 Lecture 1 — January 24, 2018 Lecture 1 — January 24, 2018

d'où
$$h(a^{1+\varepsilon}) \geq \frac{1}{2} a^2 \left[1 + \ln(a^2)\left(\frac{\varepsilon^2 \ln(a^2)}{2} - 1\right)\right]$$
$$\geq \frac{1}{2} a^2 \ln(a^2) \left(\varepsilon^2 \frac{\ln(a^2)}{2} - 1\right)$$

Ainsi si $\dfrac{\varepsilon^2 \ln(a^2)}{2} = 1$, $\varepsilon = \dfrac{1}{\sqrt{\ln(a)}}$ et $h(a^{1+\varepsilon}) \geq 0$

Comme $a < a^{1+\varepsilon}$ et $h(a^{1+\varepsilon}) \geq 0$, on a $\eta_0 \leq \eta_2 = a^{1+\frac{1}{\sqrt{\ln a}}}$

Or $\dfrac{1}{2}\eta_0^2 = a^2 \ln(a) + \dfrac{a^2}{2}\ln(2) + \dfrac{1}{2} a^2 \ln\ln(\eta_0)$

d'où
$$\frac{1}{2}\eta_0^2 \leq a^2 \ln(a) + \frac{a^2}{2}\ln(2) + \frac{1}{2} a^2 \underbrace{\ln\ln(\eta_2)}$$
$$= \ln\left(\left(1 + \frac{1}{\sqrt{\ln(a)}}\right)\ln(a)\right)$$
$$= \ln\left(\ln(a) + \sqrt{\ln(a)}\right)$$
$$\leq \ln\left(2\ln(a)\right)$$

Ainsi
$$\eta_0 \leq \eta_3 := a\sqrt{2\ln(a) + \ln(2\ln(a)) + \ln(2)}$$

et
$$g(\eta_0) = \psi(\eta_0) \leq \psi(\eta_3) = a\sqrt{2\ln(a) + \ln(2\ln(a)) + \ln(2)}$$
$$+ \frac{a}{\sqrt{\cdots}}$$

 Lecture 1 — January 24, 2018 Lecture 1 — January 24, 2018

Alors

$$\min\left(\frac{1}{2}+a^2,\ a\,\Psi\left(\sqrt{\ln(2a^2)}\right)\right) \le \inf_{\eta>1} g(\eta) \le a\,\Psi\left(\sqrt{\ln(2a^2)+\ln(\ln(a^2))}\right)$$

où $\Psi : x \longmapsto x+\frac{1}{x}$.

En remplaçant a par $\sqrt{\frac{4T}{N-1}}$, il vient :

$$\min\left(\frac{(M-m)(N-1)}{8}+(M-m)T\,;\ \frac{(M-m)}{2}\sqrt{T(N-1)}\;\Psi\left(\sqrt{\ln\left(\frac{8T}{N-1}\right)}\right)\right) \le \inf_{\eta\gamma>1} Q(\eta)$$

et $$\inf_{\eta\gamma>1} Q(\eta) \le \frac{(M-m)}{2}\sqrt{T(N-1)}\;\Psi\left(\sqrt{\ln\left(\frac{8T}{N-1}\right)+\ln\ln\left(\frac{4T}{N-1}\right)}\right)$$

En conclusion, on a donc :

$$\begin{cases} \min\left(T(M-m)\,;\ (M-m)\sqrt{T(N-1)}\;\Psi\left(\sqrt{\ln\left(\frac{8T}{N-1}\right)}\right)\right) \le \inf_{\eta>0} Q(\eta) \\ \inf_{\eta>0} Q(\eta) \le \frac{(M-m)}{2}\sqrt{T(N-1)}\;\Psi\left(\sqrt{\ln\left(\frac{8T}{N-1}\right)+\ln\ln\left(\frac{4T}{N-1}\right)}\right) \end{cases}$$

On a alors $$\inf_{\eta>0} Q(\eta) \underset{T\to\infty}{\sim} \frac{(M-m)}{2}\sqrt{T(N-1)}\sqrt{\ln(T)}$$