Compléments manuscrits pour la boîte à outils de probabilité pour le cours

[redacted]

Preuve du lemme de Hoeffding.

* H est bornée, $(s,\omega) \mapsto e^{sH(\omega)}$ est intégrable en ω à s fixé
  dérivable en s à ω fixé
  et sa dérivée He^{sH} est intégrable en ω

  Donc, par théorème de dérivation des intégrales à paramètre,
  $s \mapsto \mathbb{E}[e^{sH}]$ est dérivable, de dérivée $\mathbb{E}[He^{sH}]$

  Le même raisonnement montre en fait que cette fonction est deux fois dérivable, de dérivée seconde $s \mapsto \mathbb{E}[H^2e^{sH}]$.

* Puis, par composition, en notant $\Psi_H : s \in \mathbb{R} \mapsto \ln \mathbb{E}[e^{sH}]$,
  on a que Ψ_H est deux fois dérivable, avec

$$\Psi_H'(s) = \mathbb{E}[He^{sH}] \Big/ \mathbb{E}[e^{sH}]$$

et

$$\Psi_H''(s) = \frac{\mathbb{E}[H^2e^{sH}]\,\mathbb{E}[e^{sH}] - \left(\mathbb{E}[He^{sH}]\right)^2}{\left(\mathbb{E}[e^{sH}]\right)^2} = \mathbb{E}\left[H^2\, e^{sH-\Psi_H(s)}\right] - \left(\mathbb{E}\left[H\, e^{sH-\Psi_H(s)}\right]\right)^2$$

$$= \mathrm{Var}_{\mathbb{Q}}(H)$$

où $\mathbb{Q}$ est la loi absolument continue par rapport à $\mathbb{P}$ de densité

$$\frac{d\mathbb{Q}}{d\mathbb{P}} : \omega \mapsto e^{sH(\omega) - \Psi_H(s)}$$ (≥ 0 et s'intègre à 1)

* Bref : $\qquad \psi_H''(s) \leq \dfrac{(b-a)^2}{4}$

On intègre : $\qquad \psi_H'(s) \leq \dfrac{(b-a)^2}{4} s + \underbrace{C^{ste}}_{= \mathbb{E}[H]}$ car $\psi_H'(0) = \mathbb{E}[H]$

Encore une fois : $\qquad \psi_H(s) \leq \dfrac{(b-a)^2}{8} s^2 + s\,\mathbb{E}[H] + \underbrace{C^{ste}}_{=0}$ car $\psi_H(0) = 0$

Preuve de l'inégalité de Hoeffding, version probabiliste.

* $$\mathbb{E}\left[e^{sS_n}\right] = \prod_{t=1}^{n} \underbrace{\mathbb{E}\left[e^{s(H_t - \mathbb{E}H_t)}\right]}_{\leq e^{s\mathbb{E}H_t} \cdot e^{s^2/8 (b_t - a_t)^2} e^{-s\mathbb{E}H_t}}$$ par indépendance

$$\leq e^{s^2/8 \sum_{t=1}^{n} (b_t - a_t)^2}$$ par lemme de Hoeffding

On réinjecte : $\qquad \mathbb{P}\{S_n \geq \varepsilon\} \leq \exp\Big(\underbrace{\frac{s^2}{8}\sum_{t=1}^{n}(b_t - a_t)^2 - s\varepsilon}_{\text{minimisé (cf. dérivée seconde} \geq 0)}\Big)$

en $\qquad s = 4\varepsilon \Big/ \sum_t (b_t - a_t)^2$

On substitue cette valeur de s et on obtient le résultat.

Corollaire, formulation statistique.

S'obtient en notant que $\qquad \mathbb{P}\left\{\sum_{t=1}^{n} H_t \geq \sum_{t=1}^{n} \mathbb{E}[H_t] + \varepsilon_\delta\right\} \leq \delta$

lorsque ε_δ est tel que $\qquad e^{-2\varepsilon^2 / \sum_{t=1}^{n} (b_t - a_t)^2} = \delta$

soit $\qquad \varepsilon = \sqrt{\dfrac{1}{2}\sum_{t=1}^{n}(b_t - a_t)^2 \ln 1/\delta}$

Exercice 1. On utilise $\mathbb{P}(A \cup B) \leq \mathbb{P}(A) + \mathbb{P}(B)$

Version proba : $\mathbb{P}\left\{ -\sum_{t=1}^{n} H_t \geq -\sum_{t=1}^{n} \mathbb{E}[H_t] + \varepsilon \right\} \leq e^{-2\varepsilon^2 / \sum_t (b_t - a_t)^2}$ également

d'où $\mathbb{P}\left\{ \left| \sum_{t=1}^{n} H_t - \sum_{t=1}^{n} \mathbb{E} H_t \right| \geq \varepsilon \right\} \leq 2\, e^{-2\varepsilon^2 / \sum_{t=1}^{n} (b_t - a_t)^2}$

Version stat : $\left| \sum_{t=1}^{n} H_t - \sum_{t=1}^{n} \mathbb{E} H_t \right| \leq \sqrt{\frac{1}{2} \sum_{t=1}^{n} (b_t - a_t) \ln \frac{2}{\delta}}$

avec probabilité $\geq 1 - \delta$.

Exercice 2.

TCL : $\frac{\sqrt{n}}{\sigma} \left(\frac{1}{n} \sum_{t=1}^{n} H_t - \mu \right) \xrightarrow{\mathcal{L}} \mathcal{N}(0,1)$

soit, en notant $z_{1-\delta}$ le quantile d'ordre $1-\delta$ de la loi $\mathcal{N}(0,1)$, ie, $\mathbb{P}\{ \mathcal{N}(0,1) \leq z_{1-\delta} \} = 1-\delta$,

il vient : $\mathbb{P}\left\{ \frac{\sqrt{n}}{\sigma} \left(\frac{1}{n} \sum_{t=1}^{n} H_t - \mu \right) \leq z_{1-\delta} \right\} \longrightarrow 1-\delta$

$= \mathbb{P}\left\{ \frac{1}{n} \sum_{t=1}^{n} H_t - \mu \leq \frac{\sigma}{\sqrt{n}} z_{1-\delta} \right\}$

Note : on peut montrer (exercice calculatoire, cf. indications pour l'exo 8) que

$$z_{1-\delta} \underset{\delta \to 0}{\sim} \sqrt{2 \ln 1/\delta}$$

Hoeffding correspond au remplacement de σ par son majorant $\frac{b-a}{2}$

et au passage d'une borne asymptotique à une borne non asympt.

A LA MAISON, éventuellement en petits groupes, avec RENDU par mail :

- A coder 1
- Exercice 3
- Exercice 4 (pour les très courageux !)

Preuve du lemme de Hoeffding conditionnel.

Par l'absurde, donc : on suppose que $\mathbb{P}(A) > 0$ et on va aboutir à une contradiction, de sorte que $\mathbb{P}(A) = 0$, ce qui est exactement le résultat du lemme.

Hoeffding donne alors : $\ln \mathbb{E}_A\left[e^{sH}\right] \leqslant s\, \mathbb{E}_A[H] + \frac{s^2}{8}(b-a)^2$

Or, par définition de A : $= \ln \mathbb{E}_A\left[\mathbb{E}[e^{sH} | \mathcal{F}]\right] > \ln \mathbb{E}_A\left[e^{s\mathbb{E}[H|\mathcal{F}]}\, e^{s^2/8\,(b-a)^2}\right]$

$$= \frac{s^2}{8}(b-a)^2 + \ln \underbrace{\mathbb{E}_A\left[e^{s\mathbb{E}[H|\mathcal{F}]}\right]}_{\substack{\geqslant e^{s\mathbb{E}_A[\mathbb{E}[H|\mathcal{F}]]} = e^{s\mathbb{E}_A[H]} \\ (\text{Jensen})}}$$

Le minorant serait $>$ au majorant !

Preuve de l'inégalité de Hoeffding-Azuma.

$$S_n = \sum_{t=1}^{n} H_t - \mathbb{E}[H_t | \mathcal{F}_{t-1}]$$

$$\mathbb{P}\{S_n \geqslant \varepsilon\} \leqslant \frac{\mathbb{E}[e^{sS_n}]}{\ldots} \quad e^{-s\varepsilon}\, \mathbb{E}\left[e^{sS_{n-1}}\, \mathbb{E}\left[e^{s(H_n - \mathbb{E}[H_n|\mathcal{F}_{n-1}])} \,|\, \mathcal{F}\right]\right]$$

Chernoff/Markov $\forall s > 0$ $\qquad \leqslant e^{s^2/8\,(b_n-a_n)^2}$

$$\underset{[\text{récurrence}]}{\leqslant} \exp\Big(\underbrace{-s\varepsilon + \frac{s^2}{8} \sum_{t=1}^{n} (b_t - a_t)^2}_{}\Big)$$

et même optimisation en s que précédemment.

Exercice 5 : A LA MAISON (simple jeu d'écriture) / Pas de RENDU attendu.

Théorème 4 : Preuve extraite de la preuve de l'inégalité de Hoeffding.

une optimisation en s donne que le minimum (cf. dérivée 2nde positive) est atteint pour $-\varepsilon + sv$, soit $s = \varepsilon/v$; ce qui donne la borne $e^{-\varepsilon^2/2v}$ attendue.

<u>Exercice 7.</u> Chernoff / Cor. 5 (ici, $v=1$) : $\mathbb{P}\{N \geq x\} \leq e^{-x^2/2}$

Calcul direct :

$$\mathbb{P}\{N \geq x\} = \int_x^{+\infty} \Phi(t)\,dt$$ où $\Phi(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}$ est la densité de la $\mathcal{N}(0,1)$

$$= \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} e^{-\frac{1}{2}(t+x)^2}\,dt$$

$(t+x)^2 \geq t^2 + x^2$; $(t+x)^2 \geq x^2 + 2tx$

$$\leq e^{-x^2/2} \underbrace{\frac{1}{\sqrt{2\pi}} \int_0^{+\infty} e^{-t^2/2}\,dt}_{= \mathbb{P}\{N \geq 0\} = 1/2} = \frac{1}{2} e^{-x^2/2}$$

$$\leq e^{-x^2/2} \frac{1}{\sqrt{2\pi}} \int_0^{+\infty} e^{-tx}\,dt = \frac{e^{-x^2/2}}{x\sqrt{2\pi}}$$

vs. Chernoff : $e^{-x^2/2}$ simplement.

<u>Proposition 3 / Optimisation de la fin de la preuve.</u>

On a donc obtenu : $$\mathbb{E}[M] \leq \frac{1}{\lambda} \ln\left(N e^{\lambda^2 v/2}\right) = \frac{1}{\lambda} \ln N + \lambda v/2$$

Minimum atteint pour λ égalisant les deux termes (cf à aire fixé, le rectangle de périmètre minimal est le carré) : $\lambda^2 = 2\ln N / v$ $\quad \lambda = \sqrt{(2\ln N)/v}$

et la borne est $2 \times \frac{\lambda v}{2} = \lambda v = \sqrt{2v \ln N}$.

<u>A LA MAISON</u>, { exo8 éventuellement en petits groupes, A coder 2 chacun pour soi, } avec <u>RENDU</u> par mail :

- A coder 2
- Exercice 8 (difficile ! voire très difficile / nécessite d'avoir suivi le cours de statistiques)
  ↳ Vu la Prop. 3, il suffit de montrer que

$$\liminf \mathbb{E}\left[\max_{j \leq K} N_j\right] / \sqrt{2\ln K} \geq 1$$

1) On note $(\ )^+$ la partie positive et l'on a

$$\mathbb{E}\left[\max_{j\leq K} N_j\right] \geq \mathbb{E}\left[\left(\max_{j\leq K} N_j\right)^+\right] - 1$$

2) On note Ψ_K la fonction de répartition de $\left(\max_{j\leq K} N_j\right)^+$ et Φ celle de la loi $\mathcal{N}(0,1)$; alors $\Psi_K(0) = 2^{-K}$ et $\forall t \geq 0$, $\Psi_K(t) = (\phi(t))$

D'où l'inverse généralisée :

$$\Psi_K^{(-1)}(u) = \inf\{x \in \mathbb{R} : \Psi_K(x) \geq u\}$$

$$= \begin{cases} 0 & \text{si } 0 < u \leq 2^{-K} \\ \Phi^{-1}(u^{1/K}) & \text{si } u \geq 2^{-K} \end{cases}$$

3) Ainsi,

$$\mathbb{E}\left[\left(\max_{j\leq K} N_j\right)^+\right] = \int_0^1 \Psi_K^{(-1)}(u)\,du \geq \int_\delta^1 \Phi^{-1}(u^{1/K})\,du$$

($\delta > 0$ fixé & pour K assez grand)

$$\geq (1-\delta)\,\Phi^{-1}(\delta^{1/K})$$

on redémontrera que si $U \sim \mathrm{Unif}[0,1]$ alors $\left(\max_{j\leq K} N_j\right)^+$ a même loi que $\Psi_K^{(-1)}(U)$

4) Il suffit donc de montrer que $\liminf_{K\to+\infty} \dfrac{(1-\delta)\,\Phi^{-1}(\delta^{1/K})}{\sqrt{2\ln K}} \geq 1-\delta$ (pour tout δ fixé)

ce qui procède de calculs :

Par deux IPP : $1-\Phi(x) \sim e^{-x^2/2}/x\sqrt{2\pi}$ quand $x \to +\infty$

Ainsi, $\ln(1-v) = \ln(1-\Phi(\Phi^{-1}(v))) \sim -\Phi^{-1}(v)^2/2$

càd, $\Phi^{-1}(v) \sim \sqrt{2\ln(1/1-v)}$ quand $v \to 1$

De sorte que : $(1-\delta)\,\Phi^{-1}(\delta^{1/K}) \sim (1-\delta)\sqrt{2\ln K}$, ce qui conclut.

Exercice 9.

$(\omega, \varepsilon) \mapsto \mathbb{1}_{\{T(\omega) \geq \varepsilon\}}$ est mesurable (par composition : $(t, \varepsilon) \mapsto \mathbb{1}_{\{t \geq \varepsilon\}}$ est mesurable)

et $\mathbb{P}$, d (lebesgue) sont deux mesures positives σ-finies.

Donc par théorème de Fubini-Tonelli,

$$\int_{\Omega \times [0,+\infty[} \mathbb{1}_{\{T(\omega) \geq \varepsilon\}} \, d\mathbb{P}(\omega) \, d\varepsilon = \int T(\omega) \, d\mathbb{P}(\omega) = \mathbb{E}[T]$$

on intègre d'abord en ε puis en ω

$$= \int_0^{+\infty} \mathbb{P}\{T \geq \varepsilon\} \, d\varepsilon$$

on intègre d'abord en ω puis en ε

Exercice 10.

$$\mathbb{E}[\max_{j \leq K} N_j] \leq \mathbb{E}[(\max_{j \leq K} N_j)^+] = \int_0^{+\infty} \mathbb{P}\{\max_{j \leq K} N_j \geq \varepsilon\} \, d\varepsilon$$

$$\leq K \int_0^{+\infty} e^{-\varepsilon^2/2v} \, d\varepsilon$$

$$= K \sqrt{2\pi v} \times \frac{1}{2}$$

$$= K \sqrt{\pi v/2}$$

on reconnaît une densité gaussienne, à un facteur $\sqrt{2\pi v}$ près.

TRÈS GROSSIER !

ε_0 tq $K e^{-\varepsilon_0^2/2v} = 1$, $\varepsilon_0^2/2v = \ln K$, $\varepsilon_0 = \sqrt{2v \ln K}$

et on majore selon $\mathbb{E}[(\max_{j \leq K} N_j)^+] \leq \varepsilon_0 + \int_{\varepsilon_0}^{+\infty} K e^{-\varepsilon^2/2v}$

$$= \varepsilon_0 + \underbrace{K e^{-\varepsilon_0^2/2v}}_{=1} \int_0^{+\infty} \underbrace{e^{-x^2/2v}}_{\leq 1} \cdot e^{-x\varepsilon_0/v} \, dx$$

$$\leq \underset{\sqrt{2v \ln K}}{\varepsilon_0} + \frac{v}{\varepsilon_0} \qquad \sqrt{v/2\ln K} = o(1)$$

L'endroit où l'on perd est en particulier le fait qu'on passe aux parties positives :
$$\mathbb{E}\Big[\max_{j \leq K} N_j\Big] \leq \mathbb{E}\Big[\big(\max_{j\leq K} N_j\big)^+\Big]$$
Alors que dans la preuve de la Proposition 3, les parties négatives comptaient ! Mais ce qu'on perd est un $o(1)$, ce qui est naturel : quand $K \to +\infty$, $\max_{j \leq K} N_j$ n'est quasi jamais négatif !!

Exercice 11.

$$N_j = \frac{1}{n}\sum_{t=1}^{n} X_{t,j} \quad \text{où} \quad (X_{t,j})_{t,j} \text{ iid de support } [a,b]$$

Les N_j sont alors (cf. Théorème 4) sous-gaussiennes, de paramètre $v = (b-a)^2/4n$. (avec $b \geq 0$)

De sorte que (1) donne : $\mathbb{P}\Big\{\max_{j \geq K} N_j \geq \varepsilon\Big\} \leq \underbrace{K e^{-\varepsilon^2/2v}}_{= f(\varepsilon) = \delta}$

et $f^{-1}(\delta) = \sqrt{2v \ln \frac{K}{\delta}}$

Puis
$$\mathbb{E}\Big[\max_{j\leq K} N_j\Big] \leq \inf_{\delta \in (0,1)} \left\{ (1-\delta)\sqrt{2v \ln\frac{K}{\delta}} + \delta b \right\}$$
$$\underset{\text{(on substitue } v\text{)}}{\leq} \inf_{\delta\in(0,1)} \left\{ \frac{b-a}{\sqrt{2n}} \sqrt{\ln K/\delta} + \delta b \right\}$$

Un bon choix de δ est donc de l'ordre de $1/\sqrt{n}$ (pour rendre les deux termes approximativement du même ordre de grandeur) $\to$ borne finale en $\sqrt{v \ln n}$ ... ie, détérioration des ordres de grandeur en n...