

E x 1

1) La fonction \exp est convexe. Ainsi, $\forall z \in [a, b]$,

$$\exp(sz) \leq \frac{b-z}{b-a} \exp(sa) + \frac{z-a}{b-a} \exp(sb)$$

$$x = \frac{b-z}{b-a} \quad \begin{matrix} \text{---} \\ 0 \end{matrix} \quad \begin{matrix} \text{---} \\ 0 \end{matrix} \quad \mathbb{E}[z] = 0.$$

$$\Rightarrow \mathbb{E}[e^{sz}] \leq \frac{be^{sa} - \mathbb{E}[ze^{sa}] + \mathbb{E}[ze^{sb}] - ae^{sb}}{b-a}$$

$$= \frac{e^{sa} (b - e^{s(b-a)}a)}{b-a}$$

$$= e^{-s(b-a)p} (1-p + pe^{s(b-a)}), \quad p \triangleq \frac{a}{b-a}$$

$$= e^{-up} (1-p + pe^u), \quad u \triangleq s(b-a)$$

$$= \exp(-up + \ln(1-p + pe^u))$$

On pose $\Psi(x) = -xp + \ln(1-p + pe^x)$.

$$\text{On a } \Psi'(x) = -p + \frac{pe^x}{1-p+pe^x} = -p + \frac{p}{(1-p)e^{-x} + p}$$

$$\Psi''(x) = \frac{p(1-p)e^{-x}}{\left((1-p)e^{-x} + p\right)^2} = t(1-t) \leq \frac{1}{4}$$

$$t = \frac{(1-p)e^{-x}}{(1-p)e^{-x} + p}$$

On utilise la formule de Taylor-Lagrange : $\exists \theta \in [0, u]$,

$$\begin{aligned} Y(u) &= Y(0) + u Y'(0) + \frac{u^2}{2!} Y''(\theta) \quad Y(0) = k'(0) = 0 \\ &\leq \frac{u^2}{2!} = \frac{s(b-a)^2}{8}. \end{aligned}$$

On a donc bien $E[e^{sz}] \leq \exp\left(\frac{s^2(b-a)^2}{8}\right)$

2) $(Z_i)_{\substack{i \in \mathbb{N} \\ i \leq n}}$ iid, $P(0 \leq Z_i \leq 1) = 1$.

On pose $X_i = Z_i - E[Z_i]$.

$$\begin{aligned} P(\bar{X}_n > t) &= P(e^{s\bar{X}_n} > e^{st}), \quad \forall s > 0 \\ &\leq e^{-st} E[e^{s\bar{X}_n}] \quad (\text{Markov}) \\ &= e^{-st} E\left[e^{\frac{s}{n} \sum_i X_i}\right] \\ &= e^{-st} \prod_i E\left[e^{\frac{s}{n} X_i}\right] \\ &= e^{-st} \prod_i E\left[e^{\frac{s}{n} X_i}\right] \quad ((X_i)_{\substack{i \in \mathbb{N} \\ i \leq n}}) \\ &\leq e^{-st} \prod_i \exp\left(\frac{s^2}{8n^2}\right) \quad (\text{Lemma avec } a=0, b=n) \\ &= e^{\frac{s^2}{8n} - st} \end{aligned}$$

On cherche s qui minimise $f: s \mapsto \frac{s^2}{8n} - st$.

$$f'(s) = \frac{s}{hn} - t$$

s	0	ht	+oo

↓ ↗

On pose $s = hnt > 0$:

$$P(\bar{X}_n > t) \leq \exp\left(-h_n t^2 + \frac{\ln h_n^2 t^2}{R_n}\right) = e^{-2nt^2}$$

On procéde de manière similaire avec $-\bar{X}_n = E[Z_i] - \bar{Z}_n$
et $[a,b] = [-1,0]$.

Ex 2)

On commence par prouver le lemme suivant :

Soit V et Z deux r.v. tq $E[V|Z] = 0$ et

$\exists f, c$ tq $f(Z) \leq V \leq f(Z) + c$.

$\forall s > 0 \quad E[e^{sV}|Z] \leq e^{sc^2/8}$.

Même preuve que pour le lemme de Hoeffding avec
 $E[V e^{f(Z)}|Z] = 0$, $E[V e^{s(f(Z)+c)}] = 0$, et $[a,b] = [f(Z), f(Z)+c]$

□

En utilisant l'inégalité de Markov, on a

$$P\left(\sum_{i=1}^n V_i > t\right) = P\left(e^{s \sum V_i} > e^{st}\right), \quad \forall s > 0$$

$$\leq e^{-st} \mathbb{E}[e^{s \sum V_i}] \quad (\text{Markov})$$

On développe l'espérance :

$$\mathbb{E}[e^{s \sum V_i}] = \mathbb{E}\left[\mathbb{E}\left[e^{s V_n} e^{s \sum_{i=1}^{n-1} V_i} \mid Z_1, \dots, Z_{n-1}\right]\right]$$

$$(V_i)_{1 \leq i \leq n-1} \text{ is a function of } Z_1, \dots, Z_{n-1} = \mathbb{E}\left[e^{s \sum_{i=1}^{n-1} V_i} \mathbb{E}[e^{s V_n} \mid Z_1, \dots, Z_{n-1}]\right]$$

En utilisant le lemme précédent avec $f(Z_n) = Z_n$ et $c = c_n$, on obtient

$$\leq \mathbb{E}\left[e^{s \sum_{i=1}^{n-1} V_i} e^{s^2 c_n^2 / 8}\right]$$

$$\leq e^{s^2 \sum c_i^2 / 8}.$$

Par récurrence,

$$\text{Ainsi, } P\left(\sum V_i > t\right) \leq e^{-st} e^{s^2 \sum c_i^2 / 8}$$

$$= e^{-st + s^2 \sum c_i^2 / 8}$$

$$= \exp\left(-\frac{2t^2}{\sum c_i^2}\right), \quad s \triangleq \frac{ht}{\sum c_i^2}.$$

□

On procède de manière analogue avec $(-V_i)_{1 \leq i \leq n}$ et

$$f(Z_n) = -Z_n - c_n.$$

$$2) \text{ On pose } V_i = \mathbb{E}[h(Z_1, \dots, Z_n) | Z_1, \dots, Z_{i-1}] - \mathbb{E}[h(Z_1, \dots, Z_n) | Z_1, \dots, Z_{i-1}]$$

$$\begin{aligned} \text{On a alors } \sum_{i=1}^n V_i &= \mathbb{E}[h | Z_1] - \mathbb{E}[h] \\ &\quad + \mathbb{E}[h | Z_1, Z_2] - \mathbb{E}[h | Z_1] \\ &\quad \dots \\ &= \underbrace{\mathbb{E}[h | Z_1, \dots, Z_n]}_{h \text{ est } \sigma(Z_1, \dots, Z_n)\text{-mesurable}} - \mathbb{E}[h] \\ &= h(Z_1, \dots, Z_n) - \mathbb{E}[h]. \end{aligned}$$

On pose $F_i \triangleq \sigma(Z_1, \dots, Z_i)$. Comme $F_{i-1} \subseteq F_i$, V_i est F_i -mesurable. De plus :

$$\begin{aligned} \mathbb{E}[V_i | F_{i-1}] &= \mathbb{E}[\mathbb{E}[h | F_i] | F_{i-1}] \\ &\quad - \mathbb{E}[h | F_{i-1}] \\ &= \mathbb{E}[h | F_{i-1}] - \mathbb{E}[h | F_{i-1}] \end{aligned}$$

$$\begin{aligned} \text{On pose } W_i &\triangleq \sup_{\mathbb{Z}} \mathbb{E}[h | Z_1, \dots, Z_i = z] - \mathbb{E}[h | F_{i-1}] \\ X_i &\triangleq \inf_{\mathbb{Z}'} \mathbb{E}[h | Z_1, \dots, Z_i = z'] - \mathbb{E}[h | F_{i-1}] \end{aligned}$$

On obtient donc $X_i \leq V_i \leq W_i$.

De plus,

$$W_i - X_i \leq \sup_z \sup_{z'} E[h(z_1, \dots, z_i = z)] - E[h(z_1, \dots, z_i = z')]$$
$$\left(= \sup_z \sup_{z'} \int_{Z_{i+1}, \dots, Z_n} h(z_1, \dots, z_i, z_{i+1}, \dots, z_n) - h(z_1, \dots, z'_i, z_{i+1}, \dots, z_n) dP \right)$$
$$\leq \sup_z \sup_{z'} \int_{Z_{i+1}, \dots, Z_n} c_i dP = c_i$$

On en conclut que $X_i \leq V_i \leq X_i + c_i$. Comme X_i est une fonction des Z_i , appliquer Azuma donne le résultat désiré.

Ex 3

$$\text{Soit } h(z_1, \dots, z_n) = \hat{R}_n(F) = \mathbb{E} \left[\sup_{f \in F} \frac{1}{n} \sum_i \varepsilon_i f(z_i) \mid D_n \right]$$

On a :

$$\sup_{z_1, \dots, z_n, z'_i} |h(z_1, \dots, z_n) - h(z_1, \dots, z'_i, \dots, z_n)|$$

$$= \sup_{(z_i), z'_i} \left| \mathbb{E} \left[\sup_{f \in F} \frac{1}{n} \sum_i \varepsilon_i f(z_i) - \sup_{f \in F} \frac{1}{n} \sum_{k=1}^n \varepsilon_k f(z'_k) \mid D_n \right] \right|$$

$$\leq \sup_{(z_i), z'_i} \left| \mathbb{E} \left[\sum_{k=1}^n \frac{1}{n} \sup_f \varepsilon_k f(z_k) - \sup_f \varepsilon_k f(z'_k) \right] \right| \quad \text{L} \sup \Sigma \leq \Sigma \sup.$$

$$= \sup_{z_i, z'_i} \left| \mathbb{E} \left[\sup_f \frac{1}{n} \varepsilon_i (f(z_i) - f(z'_i)) \right] \right|$$

$$\left(\sup_{z_i, z'_i} \mathbb{E} \left[\frac{1}{n} \left| \sup_f \varepsilon_i (f(z_i) - f(z'_i)) \right| \right] \right)$$

$$\left(\frac{1}{n} \text{ (f est } [0, 1]\text{-valide)} \right) \triangleq c_i$$

) Senser, l.) conv.

On applique McDiarmid :

$$\mathbb{P}(\hat{R}_n(F) - R_n(F) < -t) \leq \exp\left(-\frac{2t^2}{\sum c_i^2}\right)$$

$$\Leftrightarrow \hat{R}_n(F) - R_n(F) > t \quad , \quad \text{a.p.} \geq 1 - \exp\left(-\frac{2t^2}{\sum \frac{1}{n^2}}\right) = 1 - \exp(-2t^2/n)$$

$$\Leftrightarrow R_n(F) \leq \bar{R}_n(F) + t, \quad a.p. \geq -e^{-2t^2 n}.$$

On pose $\delta \triangleq e^{-2t^2 n}$ et on cherche t en fonction de δ :

$$2t_n^2 = \ln(\frac{1}{\delta}) \Rightarrow t = \left(\frac{\ln(\frac{1}{\delta})}{2n} \right)^{\frac{1}{2}}.$$

□

2) On pose $\phi(z_1, \dots, z_n) = \sup_{f \in F} \left(\mathbb{E}[f(z_1)] - \frac{1}{n} \sum f(z_i) \right)$

$$|\phi(z_1, \dots, z_n) - \phi(z_1, \dots, z'_1, \dots)| \leq \left| \sup_f \frac{1}{n} \left(-f(z_i) + f(z'_i) \right) \right| \leq \frac{1}{n}.$$

On peut appliquer McDiarmid:

$$\phi(z_1, \dots, z_n) \leq \mathbb{E}[\phi(z_1, \dots, z_n)] + \sqrt{\frac{\ln(\frac{1}{\delta})}{2n}}, \quad a.p. \geq -\frac{\delta}{2}.$$

Majorons $\mathbb{E}[\phi]$:

$$\begin{aligned} \mathbb{E}[\phi] &= \mathbb{E}\left[\sup_f \left(\mathbb{E}[f(z_1)] - \frac{1}{n} \sum f(z_i) \right) \right] \\ &= \mathbb{E}\left[\sup_f \mathbb{E}_{z'} \left[\frac{1}{n} \sum f(z'_i) - \frac{1}{n} \sum f(z_i) \right] \right] \\ &\leq \mathbb{E}\left[\sup_f \frac{1}{n} \sum f(z'_i) - f(z_i) \right] \quad \sup \mathbb{E} \leq \mathbb{E} \sup \text{ (à)} \end{aligned}$$

$\hookrightarrow s: x \mapsto d \Rightarrow \mathcal{L}(f(x) - f(x'))$
 $= \mathcal{L}(f(x) - f(x')).$

$$\begin{aligned}
&= \mathbb{E}_{z, z', \varepsilon} \left[\sup_{f} \frac{1}{n} \sum \varepsilon_i (f(z'_i) - f(z_i)) \right] \\
&\leq \mathbb{E}_{\varepsilon, z} \left[\sup_f \frac{1}{n} \sum \varepsilon_i f(z'_i) \right] + \mathbb{E}_{\varepsilon, z} \left[\sup_f \frac{1}{n} \sum \varepsilon_i f(z_i) \right] \\
&= \mathbb{E}_{\varepsilon, z} \left[\sup_f \frac{1}{n} \sum \varepsilon_i f(z'_i) \right] + \mathbb{E}_{\varepsilon, z} \left[\sup_f \frac{1}{n} \sum \varepsilon_i f(z_i) \right] \\
&= 2 R_n(F).
\end{aligned}$$

Ainsi, on a $\phi(z_1, \dots, z_n) \leq 2R_n(F) + \sqrt{\frac{\ln(2/\delta)}{2n}}$, avec probabilité d'au moins $1 - \frac{\delta}{2}$.

De plus, avec proba $1 - \frac{\delta}{2}$, on a que $R_n(F) \leq \bar{R}_n(F) + \left(\frac{\ln(2/\delta)}{2n}\right)^{1/2}$.

En combinant ces 2 inégalités, on obtient

$$\phi(z_1, \dots, z_n) \leq 2\bar{R}_n(F) + 3\sqrt{\frac{\ln(2/\delta)}{2n}}.$$

Exh

$$\gamma(C, n) = \max_{\substack{X \\ |X|=n}} |C \cap X| = |\{c \in C \mid c \in X\}|$$

Lemme de Sauer :

$$\forall C, VC(C) = V \vee +\infty, \forall n \leq V \Rightarrow \gamma(C, n) \leq \sum_{i=0}^V \binom{n}{i}$$

Preuve :

On procède par récurrence sur $n+V$. Si $n=0$, alors

$$\gamma(C, n) = |\{c \in C \mid c \cap \emptyset\}| = |\{\emptyset\}| = 1 = \sum_{i=0}^V \binom{0}{i}$$

$$\text{Si } V=1, \text{ alors } \gamma(C, 0) = 1 = \sum_{i=0}^1 \binom{0}{i}.$$

Supposons que la propriété soit vérifiée pour tout $n'+V \leq n+V$.

Posons C un ensemble d'ensembles de \mathbb{R}^d tq $VC(C) = V$.

Soit S un ensemble de n points. Posons C_1, C_2 tq :

$$C = C_1 \triangleq \{c \in C \mid s \notin c\} \cup \{c \cup \{s\} \mid c \in C\}$$

$$C_2 \triangleq \{c \cup \{s\} \in C \mid c \in C\}.$$

En posant $\gamma(C, S) = |C \cap S|$, on a que

$$\gamma(C, S) \leq \gamma(C_1, S) + \gamma(C_2, S).$$

De plus, $\Upsilon(C, S \setminus \{s\}) = \Upsilon(C_n, S)$

En effet, $\forall c \in C$, on veut que tout unique ensemble $c \cap S \setminus \{s\}$ soit associé à un unique ensemble $c' \cap S$, $c' \in C_n$.

Si $c \in C_n$, alors $c = c'$ et on a $S \setminus \{s\} \rightarrow c' \cap S \setminus \{s\} \cup c' \cap \{s\}$

Si $c \in C_n$, alors $c = c'' \cup \{s\}$. Or, $c'' \in C_n$ ou $c'' \setminus \{s\} \in C_n^{\text{ou }} \emptyset$

En effet, si $s \notin c''$, $c'' \in C_n$. Si $s \in c''$:

- $c'' \setminus \{s\} \in C$, alors $c'' \setminus \{s\} \cup \{s\} \in C_n$.

- $c'' \setminus \{s\} \in C$, alors, puisque $s \notin c'' \setminus \{s\}$, $c'' \setminus \{s\} \in C_n$.

Ainsi, $C = (c'' \cup \{s\}) \cap S \setminus \{s\} = c'' \cap S \setminus \{s\} \cup \{s\} \cap S \setminus \{s\} = c'' \cap S \setminus \{s\}$

Si $c'' \in C_n \rightarrow (c'' \cap S \setminus \{s\}) \cup \{s\} = c'' \cap S \setminus \{s\} \cup c'' \cap \{s\}$
 $\{s\} \cup \emptyset$

Si $c'' \setminus \{s\} \in C_n \rightarrow c'' \setminus \{s\} \cap S \setminus \{s\} \cup c'' \setminus \{s\} \cap \{s\}$

$= c'' \cap S \setminus \{s\} \cup \emptyset$

Notons C' l'ensemble des c qui constituent C_n :

$$C' = \{C \in C \mid s \notin C, C \cup \{s\} \in C\}$$

De la même manière que précédemment, $\Upsilon(C_n, S) =$

$$\Upsilon(C', S \setminus \{s\})$$

Ainsi, si C' pulvérise un échantillon, $C(C_2)$ le pulvérise également, même si on lui rapporte $\{s\}$. On a alors :

$$VC(C') + 1 \leq VC(C) \Leftrightarrow VC(C') \leq V - 1.$$

Par hypothèse de récurrence, on a donc, $\forall S$ tq $|S|=n$:

$$\gamma(C, S) \leq \gamma(C_1, S) + \gamma(C_2, S) = \gamma(C, S \setminus \{s\}) + \gamma(C', S \setminus \{s\})$$

$$\leq \sum_{i=0}^V \binom{n-1}{i} + \sum_{i=0}^{V-1} \binom{n-1}{i}$$

$$= \sum_{i=0}^V \binom{n}{i}$$

Ex 5

On commence par chercher une borne inf à $V(A)$.

On choisit $d+1$ points tels que x_0 soit l'origine et $\{x_1, \dots, x_d\} = \{e_1, \dots, e_d\}$. On pose aussi $\{y_0, \dots, y_d\}$ un ensemble d'échantillons arbitraire associé à $\{x_0, \dots, x_d\}$.

Posons $w = (y_1, \dots, y_d)^T$. Ainsi, l'hyperplan d'égalation $wx + \frac{y_0}{2} = 0$ sépare le $\{x_0, \dots, x_d\}$ car, $\forall i \in \{1, \dots, d\}$:

$$\operatorname{sgn}(wx_i + \frac{y_0}{2}) = \operatorname{sgn}(y_i + \frac{y_0}{2}) = \operatorname{sgn}(y_i)$$

et $\operatorname{sgn}(wx_0 + \frac{y_0}{2}) = \operatorname{sgn}(\frac{y_0}{2}) = \operatorname{sgn}(y_0)$.

Donc $d+1 \leq V(A)$.

On cherche maintenant une borne supérieure en utilisant le théorème de Radon.

Tout ensemble X de $d+2$ points de \mathbb{R}^d peut être partitionné en 2 ensembles X_1, X_2 tels que

$$X_1 \cap X_2 = \emptyset, \quad X_1 \cup X_2 = X \text{ et } \operatorname{conv}(X_1) \cap \operatorname{conv}(X_2) \neq \emptyset.$$

Posons X un ensemble arbitraire de $d+2$ points. Par le théorème de Radon, on peut partitionner X en 2 ensembles

de manière à ce que leur enveloppe convexe s'intersecte). Or, si deux ensembles X_1 et X_2 peuvent être séparés par un hyperplan, alors leur enveloppe convexe peut aussi être séparée par ce même hyperplan^(*). Ainsi, la décomposition fournie par le théorème de Radon ne pourra pas être séparée par un hyperplan, donc X ne sera pas pulvrisé.

$$\text{On en conclut } d+1 \leq V(A) \leq d+2 \Leftrightarrow V(A) = d+1.$$

Si X , il existe une séparation (celle de Radon) qui n'est pas achievable par un hyperplan

() Si X_1 et X_2 sont séparés par un k -plan, alors, on peut utiliser cet k -plan pour construire l'enveloppe convexe. Ainsi, les deux enveloppes sont séparées par le même k -plan.*