# Ex 1

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

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

$$\alpha = \frac{b-z}{b-a}$$

$$\Rightarrow \mathbb{E}\left[e^{sZ}\right] \leq \frac{b e^{sa} - \overbrace{\underset{=0}{\mathbb{E}[Z e^{sa}]} + \underset{=0}{\mathbb{E}[Z e^{sb}]}}^{\mathbb{E}[Z]=0.} - a e^{sb}}{b-a}$$

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

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

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

$$= \exp\left(-up + \ln\left(1-p+pe^{u}\right)\right)$$

On pose $\varphi(x) = -xp + \ln\left(1-p+pe^{x}\right)$.

On a $\varphi'(x) = -p + \dfrac{pe^{x}}{1-p+pe^{x}} = -p + \dfrac{p}{(1-p)e^{-x}+p}$

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

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

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

$$\varphi(u) = \varphi(0) + u\varphi'(0) + \frac{u^2}{2}\varphi''(\theta) \qquad \varphi(0) = \varphi'(0) = 0$$

$$\leq \frac{u^2}{8} = \frac{(s(b-a))^2}{8}.$$

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

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

On pose $X_i = Z_i - \mathbb{E}[Z_i]$.

$$\begin{aligned}
\mathbb{P}(\bar{X}_n > t) &= \mathbb{P}(e^{s\bar{X}_n} > e^{st}), \quad \forall s > 0 \\
&\leq e^{-st}\,\mathbb{E}[e^{s\bar{X}_n}] \quad \text{(Markov)} \\
&= e^{-st}\,\mathbb{E}\left[e^{\frac{s}{n}\sum_i X_i}\right] \\
&= e^{-st}\,\mathbb{E}\left[\prod_i e^{\frac{s}{n}X_i}\right] \\
&= e^{-st}\prod_i \mathbb{E}\left[e^{\frac{s}{n}X_i}\right] \quad \left((X_i)_{1\leq i\leq n} \text{ iid}\right) \\
&\leq e^{-st}\prod_i \exp\left(\frac{s^2}{8n^2}\right) \quad \left(\text{lemma avec } a=0,\ b=1\right) \\
&= e^{\frac{s^2}{8n} - st}
\end{aligned}$$

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

$$f'(s) = \frac{s}{4n} - t$$

| $s$ | $0$ | | $4nt$ | | $+\infty$ |
|---|---|---|---|---|---|
| $f$ | | $\searrow$ | | $\nearrow$ | |

On pose $s = 4nt > 0$ :

$$\mathbb{P}(\bar{X}_n > t) \leq \exp\left(-4nt^2 + \frac{16n^2t^2}{8n}\right) = e^{-2nt^2}$$

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

## Ex 2)

On commence par prouver le lemme suivant :

> Soit $V$ et $Z$ deux v.a tq $\mathbb{E}[V|Z] = 0$ et $\exists f, c$ tq $f(Z) \leq V \leq f(Z) + c$.
> $\forall s > 0 \quad \mathbb{E}[e^{sV} | Z] \leq e^{s^2c^2/8}$.

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

$\square$

En utilisant l'inégalité de Markov, on a

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

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

On développe l'espérance :

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

$(V_i)_{1\leq i\leq n-1}$ is a function of $Z_1, \dots, Z_{n-1}$

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

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

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

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

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

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

$$= \exp\left(-\frac{2t^2}{\sum c_i^2}\right), \quad s \triangleq \frac{4t}{\sum c_i^2}.$$

$\square$

On procède de manière analogue avec $(-V_i)_{1\leq i\leq n}$ et $f(Z_n) = -Z - c_n$.

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

On a alors $\sum_{i=1}^{n} V_i = \mathbb{E}[h \mid Z_1] - \mathbb{E}[h]$
$+ \mathbb{E}[h \mid Z_1, Z_2] - \mathbb{E}[h \mid Z_1]$
$\dots$

$$= \underbrace{\mathbb{E}[h \mid Z_1, \dots, Z_n]}_{h \text{ est } \sigma(Z_1, \dots, Z_n)\text{-mesurable}} - \mathbb{E}[h]$$

$$= h(Z_1, \dots, Z_n) - \mathbb{E}[h].$$

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 :

$$\mathbb{E}[V_i \mid F_{i-1}] = \mathbb{E}\left[\mathbb{E}[h \mid F_i] \mid F_{i-1}\right] - \mathbb{E}[h \mid F_{i-1}]$$
$$= \mathbb{E}[h \mid F_{i-1}] - \mathbb{E}[h \mid F_{i-1}] = 0$$

On pose $W_i \triangleq \sup_z \mathbb{E}[h \mid Z_1, \dots, Z_i = z] - \mathbb{E}[h \mid F_{i-1}]$
$X_i \triangleq \inf_{z'} \mathbb{E}[h \mid Z_1, \dots, Z_i = z'] - \mathbb{E}[h \mid F_{i-1}]$

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

De plus,

$$W_i - X_i \leq \sup_{z} \sup_{z'} \mathbb{E}\left[h \mid Z_1, \dots, Z_i = z\right] - \mathbb{E}\left[h \mid Z_1, \dots, Z_i = z'\right]$$

$$= \sup_{z} \sup_{z'} \int_{Z_{i+1}, \dots, Z_n} h\left(Z_1, \dots, z, Z_{i+1}, \dots, Z_n\right) - h\left(Z_1, \dots, z', Z_{i+1}, Z_n\right) dP$$

$$\leq \sup_{z} \sup_{z'} \int_{Z_{i+1}, \dots, Z_n} c_i \, dP = c_i$$

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

$\sup \sum \leq \sum \sup$

$-\sup \sum \geq -\sum \sup$

Soit $h(Z_1, \dots, Z_n) = \hat{R}_n(F) = \mathbb{E}\left[\sup_{f \in F} \frac{1}{n} \sum \varepsilon_i f(Z_i) \,\middle|\, D_n\right]$

On a :

$$\sup_{Z_1, \dots, Z_n, Z_i'} \left| h(Z_1, \dots, Z_n) - h(Z_1, \dots, Z_i', \dots, Z_n) \right|$$

$$= \sup_{(Z_i), Z_i'} \left| \mathbb{E}\left[ \underbrace{\sup_{f \in F} \frac{1}{n} \sum \varepsilon_i f(Z_i)}_{(1)} - \sup_{f \in F} \frac{1}{n} \sum_{k=1}^{n} \varepsilon_k f(Z_k') \,\middle|\, D_n \right] \right| \quad Z_k' = Z_k, \ \forall k \neq i$$

Notons $f^*$ la fonction qui maximise le sup dans (1). On a alors $\frac{1}{n} \sum \varepsilon_k f^*(Z_k') \leq \sup_{f \in F} \frac{1}{n} \sum \varepsilon_k f(Z_k')$. Ainsi :

$$\leq \sup_{(Z_i), Z_i'} \left| \mathbb{E}\left[ \frac{1}{n} \sum \varepsilon_k \left( f^*(Z_i) - f^*(Z_k') \right) \,\middle|\, D_n \right] \right|$$

$$= \sup_{Z_i, Z_i'} \left| \mathbb{E}\left[ \frac{1}{n} \varepsilon_i \left( f^*(Z_i) - f^*(Z_i') \right) \,\middle|\, D_n \right] \right|$$

$\leq \frac{1}{n}$ ($f^*$ est $\{0,1\}$-valuée.)

On applique McDiarmid :

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

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

$\Leftrightarrow R_n(F) \leq \hat{R}_n(F) + t$, a.p. $\geq 1 - e^{-2t^2 n}$.

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

$$2tn^2 = \ln(1/\delta) \to t = \left(\frac{\ln(1/\delta)}{2n}\right)^{1/2}.$$

$\square$

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

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

On peut appliquer Mc Diarmid :

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

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

$$\mathbb{E}[\phi] = \mathbb{E}_z\left[\sup_f\left(\mathbb{E}[f(z_1)] - \frac{1}{n}\sum f(z_i)\right)\right]$$
$$= \mathbb{E}_z\left[\sup_f \mathbb{E}_{z'}\left[\frac{1}{n}\sum f(z_i') - \frac{1}{n}\sum f(z_i)\right]\right]$$

$\sup \mathbb{E} \leq \mathbb{E} \sup$

$$\leq \mathbb{E}_{z,z'}\left[\sup_f \frac{1}{n}\sum f(z_i') - f(z_i)\right]$$

si $x, x'$ iid $\Rightarrow \mathcal{L}(f(x) - f(x')) = \mathcal{L}(\varepsilon(f(x) - f(x')))$.

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

$\sup_{x, y \in E} (x+y) \leq \sup x + \sup y$

$$\leq \mathbb{E}_{\varepsilon, z'} \left[ \sup_f \frac{1}{n} \sum \varepsilon_i f(z'_i) \right] + \mathbb{E}_{\varepsilon, z} \left[ \sup_f \frac{1}{n} \sum -\varepsilon_i f(z_i) \right]$$

$$= \mathbb{E}_{\varepsilon, z'} \left[ \sup_f \frac{1}{n} \sum \varepsilon_i f(z'_i) \right] + \mathbb{E}_{\varepsilon, z} \left[ \sup_f \frac{1}{n} \sum \varepsilon_i f(z_i) \right]$$

$$= 2 \mathcal{R}_n(F).$$

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

De plus, avec proba $1 - \frac{\delta}{2}$, on a que

$$\mathcal{R}_n(F) \leq \hat{\mathcal{R}}_n(F) + \left( \frac{\ln(2/\delta)}{2n} \right)^{1/2}.$$

En combinant ces 2 inégalités, on obtient

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

Ex 4

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

Lemme de Sauer :

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

Preuve :

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

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

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

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

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

Soit $S$ un ensemble de $n$ points. Posons $s \in S$ et $\mathcal{C}_1, \mathcal{C}_2$ tq :

$$\mathcal{C} = \mathcal{C}_1 \triangleq \{c \in \mathcal{C} \mid s \notin c\} \cup \{c \cup \{s\} \in \mathcal{C} \mid c \notin \mathcal{C}\}$$
$$\cup\, \mathcal{C}_2 \triangleq \{c \cup \{s\} \in \mathcal{C} \mid c \in \mathcal{C}\}.$$

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

$$\gamma(\mathcal{C}, S) \leq \gamma(\mathcal{C}_1, S) + \gamma(\mathcal{C}_2, S).$$

De plus, $\gamma(C, S\setminus\{s\}) = \gamma(C_1, S)$

En effet, $\forall_c \in C$, on veut que tout unique ensemble $c \cap S\setminus\{s\}$ soit associé à un unique ensemble $c' \cap S$, $c' \in C_1$.

Si $c \in C_1$, alors $c = c'$ et $c \cap S\setminus\{s\} \to c' \cap S\setminus\{s\} \cup c' \cap \{s\}$

Si $c \in C_2$, alors $c = c'' \cup \{s\}$. Or, $c'' \in C_1$ ou $c''\setminus\{s\} \in C_1$. $\overset{\{s\} \text{ ou } \emptyset}{\parallel}$

En effet, si $s \notin c''$, $c'' \in C_1$. Si $s \in c''$ :

- $c''\setminus\{s\} \notin C$, alors $c''\setminus\{s\} \cup \{s\} \in C_1$.
- $c''\setminus\{s\} \in C$, alors, puisque $s \notin c''\setminus\{s\}$, $c''\setminus\{s\} \in C_1$.

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 $c'' \in C_1 \to (c'' \cap S\setminus\{s\}) \cup \{s\} = c'' \cap S\setminus\{s\} \cup \underbrace{c'' \cap \{s\}}_{\{s\} \text{ ou } \emptyset}$

Si $c''\setminus\{s\} \in C_1 \to c''\setminus\{s\} \cap S\setminus\{s\} \cup c''\setminus\{s\} \cap \{s\}$
$= c'' \cap S\setminus\{s\} \cup \emptyset$

Notons $C'$ l'ensemble des $c$ qui constituent $C_2$ :

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

De la même manière que précédemment, $\gamma(C_2, S) = \gamma(C', S\setminus\{s\})$

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

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

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 tel 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'étiquettes arbitraire associé à $\{x_0, \dots, x_d\}$.

Posons $\omega = (y_1, \dots, y_d)^T$. Ainsi, l'hyperplan d'équation $\omega x + \frac{y_0}{2} = 0$ pulvérise $\{x_0, \dots, x_d\}$ car, $\forall i \in \{1, \dots, d\}$ :

$$\text{sgn}\left(\omega x_i + \frac{y_0}{2}\right) = \text{sgn}\left(y_i + \frac{y_0}{2}\right) = \text{sgn}(y_i)$$

$$\text{et } \text{sgn}\left(\omega x_0 + \frac{y_0}{2}\right) = \text{sgn}\left(\frac{y_0}{2}\right) = \text{sgn}(y_0).$$

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 partionner en 2 ensembles $X_1, X_2$ tq
>
> $$X_1 \cap X_2 = \emptyset, \quad X_1 \cup X_2 = X \quad \text{et} \quad \text{conv}(X_1) \cap \text{conv}(X_2) \neq \emptyset.$$

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

de manière à ce que leur enveloppe convexe s'intersectent.
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 pulvérisé.

On en conclut $d+1 \leq V(A) < d+2 \iff V(A) = d+1$.

$\forall X$, il existe une séparation (celle de Radon) qui n'est pas achevable par un hyperplan

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