

Dev n° 6

Ref: Ouvrad
 "Probabilité 2"
 Site: Mathieu Dutor

Titre: Inégalité de Hoeffding

Le but de ce développement est de montrer le Théorème suivant:

Thm: (Inégalité de Hoeffding)

Soit $(X_n)_{n \in \mathbb{N}^*}$ une suite de variable aléatoires réelles indépendantes centrées telle que, $\forall n \in \mathbb{N}^*, \exists c_n > 0$ t.q. $|X_n| \leq c_n$ presque sûrement.
 $\forall n \in \mathbb{N}^*$, on pose $S_n = \sum_{j=1}^n X_j$.

Alors, on a $\forall \varepsilon > 0$,

$$P(|S_n| > \varepsilon) \leq 2 \exp\left(-\frac{\varepsilon^2}{2 \sum_{j=1}^n c_j^2}\right).$$

Pour montrer ce grand thm nous avons besoin de démontrer un lemme intermédiaire.

Lemme:

Soit X une v.a. réelle centrée et presque sûrement bornée par 1.

Alors,

$$\forall t \in \mathbb{R}, \mathbb{E}[\exp(tX)] \leq \exp(t^2/2).$$

Montrons ce lemme

Démo: (Lemme)

Soit $t \in \mathbb{R}$.

• Quelques remarques:

* Par tout $x \in [-1; 1]$, on a: $\frac{1}{2}(1-x) \in [0; 1]$
 $\frac{1}{2}(1+x) \in [0; 1]$

$$\text{Et } \frac{1}{2}(1-x) + \frac{1}{2}(1+x) = 1 \quad \text{et} \quad \frac{1}{2}(1-x)(-1) + \frac{1}{2}(1+x) = x.$$

* La fonction $x \mapsto \exp(tx)$ est convexe

• Retour à la preuve

Comme $\exp$ est une fonction convexe, on a:

$\forall x \in [-1; 1]$

$$\exp(tx) = \exp\left[t\left(\frac{1}{2}(1-x) + \frac{1}{2}(1+x)\right)\right]$$

$$\leq \frac{1}{2}(1-x) \exp(-t) + \frac{1}{2}(1+x) \exp(t)$$

On a la v.a. X est bornée p.s., donc $\mathbb{E}[\exp tX]$ existe $\forall t \in \mathbb{R}$.

Alors,

$\forall t \in \mathbb{R}$,

$$\exp(tX) \leq \frac{1}{2}(1-X) \exp(-t) + \frac{1}{2}(1+X) \exp(t)$$

Rappel $\mathbb{E}[cst] = cst$

$$\exp(tx) \leq \frac{1}{2} (1-x) \exp(-t) + \frac{1}{2} (1+x) \exp(t)$$

Rappel $\mathbb{E}[\text{cst}] = \text{cst}$
 $\mathbb{E}[\text{v.a. centrée}] = 0$

Donc,

$$\mathbb{E}[\exp(tx)] \leq \frac{1}{2} [\exp(-t) + \exp(t)] \quad \text{comme } x \text{ est centrée}$$

↙ def de ch t avec $\text{ch}: \mathbb{R} \rightarrow [1, +\infty[$

$$= \text{ch}(t).$$

De plus, on sait que $\text{ch}(t) = \sum_{n \geq 0} \frac{t^{2n}}{(2n)!}$ et $\exp(\frac{t^2}{2}) = \sum_{n \geq 0} \frac{t^{2n}}{n! 2^n}$.

Et,

$$\text{ch}(t) = \sum_{n \geq 0} \frac{t^{2n}}{(2n)!} \leq \sum_{n \geq 0} \frac{t^{2n}}{n! 2^n} = \exp\left(\frac{t^2}{2}\right)$$

Donc,

$$\mathbb{E}[\exp(tx)] \leq \exp(t^2/2) \quad \square$$

Montrons maintenant l'inégalité d'Hoeffding.

Démo: (Thm)

Soit $t \in \mathbb{R}$,

- On applique le lemme à X_n/c_n : ↖ borne par 1

$$t' \in \mathbb{R}, \quad \mathbb{E}\left[\exp\left(t' \frac{X_n}{c_n}\right)\right] \leq \exp\left(\frac{t'^2}{2}\right)$$

Prenons $t' = c_n t$,

$$\leadsto \mathbb{E}[\exp(t X_n)] \leq \exp\left(\frac{t^2 c_n^2}{2}\right)$$

Les v.a X_n sont indépendantes donc

$$\begin{aligned} \mathbb{E}[\exp(t S_n)] &= \mathbb{E}\left[\exp\left(t \sum_{j=1}^n X_j\right)\right] \\ &= \prod_{j=1}^n \mathbb{E}[\exp(t X_j)] \\ &\leq \exp\left(\frac{t^2}{2} \sum_{j=1}^n c_j^2\right) \end{aligned}$$

- Maintenant, prenons $t > 0$ et $\varepsilon > 0$.

On a $\mathbb{P}(S_n > \varepsilon) \leq \mathbb{P}[\exp(t S_n) > \exp(t \varepsilon)]$ par croissance de $x \mapsto \exp(tx)$

$$\leq \frac{\mathbb{E}[\exp(t S_n)]}{\exp(t \varepsilon)} \quad \text{Markov}$$

$$\leq \exp\left(\frac{t^2}{2} \sum_{j=1}^n c_j^2 - t \varepsilon\right) \quad \leftarrow \text{on a pris } t \text{ quelconque donc optimisons cette formule}$$

On pose $a = \sum_{j=1}^n c_j^2$ et $f: \mathbb{R}^+ \rightarrow \mathbb{R}, t \mapsto \frac{t^2}{2} a - t \varepsilon$. ↖ polynôme 2nd degré avec $a > 0$ donc $\Downarrow$

On a $\forall t \in \mathbb{R}^+, f'(t) = at - \varepsilon \Rightarrow f' = 0 \Leftrightarrow t = \frac{\varepsilon}{a} \leadsto f$ minimum en $t = \varepsilon/a$.

$$f\left(\frac{\varepsilon}{a}\right) = \frac{\varepsilon^2}{2a} - \frac{\varepsilon^2}{a} = -\frac{\varepsilon^2}{2a}$$

Donc $\mathbb{P}(S_n > \varepsilon) \leq \exp\left(-\frac{\varepsilon^2}{2a}\right)$.

$$\begin{aligned} \text{Donc } IP(S_n > \varepsilon) &\leq \exp\left(-\frac{\varepsilon^2}{2a}\right) \\ &\leq \exp\left(-\frac{\varepsilon^2}{2 \sum_{j=1}^n c_j^2}\right) \quad \text{def } a \end{aligned}$$

• Avec le même raisonnement on trouve : $IP(S_n < -\varepsilon) \leq \exp\left(-\frac{\varepsilon^2}{2a}\right)$.

• Conclusion: $IP(|S_n| > \varepsilon) = IP(S_n > \varepsilon) + IP(S_n < -\varepsilon) \leq 2 \exp\left(-\frac{\varepsilon^2}{2 \sum_{j=1}^n c_j^2}\right)$

Donc on a l'inégalité souhaitée $\square$

Annexe:

• Markov :

$$X \geq 0 \text{ v.a.r.} \quad \forall a > 0, \quad IP(X > a) \leq \frac{E[X]}{a}$$

$$\begin{aligned} \text{Démonstration: } X &\geq X \mathbb{1}_{X > a} \geq a \mathbb{1}_{X > a} \Rightarrow E[X] \geq E[X \mathbb{1}_{X > a}] \\ &\geq E[a \mathbb{1}_{X > a}] \\ &\Rightarrow IP(X > a) \leq \frac{E[X]}{a} \quad \square \end{aligned}$$

• Lemme dit que la transformée de Laplace d'une v.a. centrée et p.s. bornée par 1 est $\leq$ transformée de Laplace de $\mathcal{P}(0,1)$

• Fonction convexe $\Leftrightarrow f((1-t)x + ty) \leq (1-t)f(x) + tf(y)$.

Définition — Une fonction f d'un intervalle réel I vers $\mathbb{R}$ est dite convexe lorsque, pour tous x_1 et x_2 de I et tout t dans $[0; 1]$ on a :

$$f(tx_1 + (1-t)x_2) \leq tf(x_1) + (1-t)f(x_2).$$

• $|X_n| \leq c_n$ p.s. $\Leftrightarrow IP(|X_n| \leq c_n) = 1$
 $\Rightarrow IP\left(\bigcap_{j=1}^n |X_j| \leq c_j\right) = 1$

Démonstration:

$$\begin{aligned} IP\left(\bigcap_{j=1}^n \{|X_j| \leq c_j\}\right) &= 1 - IP\left(\bigcup_{j=1}^n \{|X_j| > c_j\}\right) \\ &\leq \sum_{j=1}^n \underbrace{IP(|X_j| > c_j)}_{=0} \end{aligned}$$

$$\text{Donc } IP\left(\bigcap_{j=1}^n \{|X_j| \leq c_j\}\right) \geq 1 - 0 = 1 \quad \square$$

Inégalité de Hoeffding

Thm: (≤ Hoeffding)

Soit $(X_n)_{n \in \mathbb{N}^*}$ suite de v.a.r indé. centrées t.q.

$\forall n \in \mathbb{N}^*, \exists c_n > 0$ t.q. $|X_n| \leq c_n$ p.s.

On pose $\forall n \in \mathbb{N}^*, S_n = \sum_{j=1}^n X_j$.

Alors, on a $\forall \varepsilon > 0$,
$$\mathbb{P}(|S_n| > \varepsilon) \leq 2 \exp\left(-\frac{\varepsilon^2}{2 \sum_{j=1}^n c_j^2}\right)$$

Lemme:

Soit X v.a.r centrée et p.s. bornée par 1.

Alors, $\forall t \in \mathbb{R}, \mathbb{E}[e^{tX}] \leq e^{t^2/2}$

Démon: (Lemme)

Soit $t \in \mathbb{R}$.

R.q.: $\forall x \in [-1, 1], \frac{1}{2}(1-x) \in [0, 1]$ et $\frac{1}{2}(1+x) \in [0, 1]$

- $x \mapsto e^{tx}$ convexe.

On a donc;

$\forall x \in [-1, 1],$

$$\exp(tx) = \exp\left[t\left(-\frac{1}{2}(1-x) + \frac{1}{2}(1+x)\right)\right]$$

$$\leq \frac{1}{2}(1-x) \exp(-t) + \frac{1}{2}(1+x) \exp(t)$$

Comme la v.a. X est bornée p.s. on a donc que

$\mathbb{E}[e^{tX}]$ existe $\forall t \in \mathbb{R}$ et:

$$e^{tX} \leq \frac{1}{2}(1-X) e^{-t} + \frac{1}{2}(1+X) e^t$$

$$\Rightarrow \mathbb{E}[e^{tX}] \leq \frac{1}{2} [e^{-t} + e^t] = \cosh(t)$$

$$= \sum_{n \geq 0} \frac{t^{2n}}{(2n)!} \leq \sum_{n \geq 0} \frac{t^{2n}}{n! 2^n} = e^{t^2/2} \quad \square$$

Démon: (thm)

Soit $t \in \mathbb{R}$.

* On applique le lemme à X_n/c_n :

$$t' \in \mathbb{R}, \mathbb{E}[\exp(t' X_n/c_n)] \leq \exp(t'^2/2)$$

Prendons $t' = c_n t$

$$\leadsto \mathbb{E}[\exp(t X_n)] \leq \exp\left(\frac{t^2 c_n^2}{2}\right)$$

Les v.a. X_n sont indé. donc:

$$\mathbb{E}[\exp(t S_n)] = \prod_{j=1}^n \mathbb{E}[\exp(t X_j)] \leq \exp\left(\frac{t^2}{2} \sum_{j=1}^n c_j^2\right)$$

* Prenons $t > 0$ et $\varepsilon > 0$:

$$\text{On a: } \mathbb{P}(S_n > \varepsilon) \leq \mathbb{P}[\exp(t S_n) > \exp(t \varepsilon)] \quad (\exp \nearrow)$$

$$\leq \frac{\mathbb{E}[\exp(t S_n)]}{\exp(t \varepsilon)} \quad (\text{Markov})$$

$$\leq \exp\left(\frac{t^2}{2} \underbrace{\sum_{j=1}^n c_j^2}_{:= a} - t \varepsilon\right)$$

Posons $f: t \in \mathbb{R}^+ \mapsto \frac{t^2}{2} a - t \varepsilon$.

On a $\forall t \in \mathbb{R}^+, f'(t) = at - \varepsilon \Rightarrow f' = 0 \Leftrightarrow t = \varepsilon/a$

$\leadsto$ min de f en $t = \varepsilon/a$ et $f(\varepsilon/a) = -\varepsilon^2/2a$.

$$\text{Donc, } \mathbb{P}(S_n > \varepsilon) \leq \exp(-\varepsilon^2/2a)$$

$$\text{De même: } \mathbb{P}(S_n < -\varepsilon) \leq \exp(-\varepsilon^2/2a)$$

* Conclusion:

$$\mathbb{P}(|S_n| > \varepsilon) \leq 2 \exp\left(-\frac{\varepsilon^2}{2 \sum_{j=1}^n c_j^2}\right)$$

□