

Inégalité de Hoeffding:Théorème:

Soit (X_n) une suite de v.a. réelles indépendantes centrées tq $\forall n \in \mathbb{N}^*$,
 $\exists C_n > 0$ vérifiant $|X_n| \leq C_n$ p.s. $\forall n \in \mathbb{N}^*$, on pose $S_n = \sum_{j=1}^n X_j$.
 On a $\forall \varepsilon > 0$, $P(|S_n| > \varepsilon) \leq 2 \cdot \exp\left(-\frac{\varepsilon^2}{2 \sum_{j=1}^n C_j^2}\right)$

Lemme: Soit X une v.a. réelle centrée et p.s. bornée par 1.
 Alors $\forall t \in \mathbb{R}$, $E[e^{tX}] \leq \exp(t^2/2)$.

Preuve lemme:

Soit $t \in \mathbb{R}$. $\forall x \in [-1, 1]$, on a $\frac{1}{2}(1-x) \in [0, 1]$
 $\frac{1}{2}(1+x) \in [0, 1]$

et $\frac{1}{2}(1-x) + \frac{1}{2}(1+x) = 1$. La fonction $x \mapsto \exp(tx)$ étant convexe,
 on a $\forall x \in [-1, 1]$:

$$\exp(tx) = \exp\left(t\left(\frac{1}{2}(1-x)x(-1) + \frac{1}{2}(1+x)\right)\right) \leq \frac{1}{2}(1-x)e^{-t} + \frac{1}{2}(1+x)e^t.$$

La v.a. X étant bornée p.s., $E[e^{tX}]$ existe $\forall t \in \mathbb{R}$.

$\forall t \in \mathbb{R}$, on a alors $\exp(tX) \leq \frac{1}{2}(1-X)\exp(-t) + \frac{1}{2}(1+X)\exp(t)$,

et comme X est centrée, on a :

$$E[\exp(tX)] \leq \frac{1}{2}(e^{-t} + e^t) = \text{ch}(t).$$

On $\text{ch}(t) = \sum_{n \geq 0} \frac{t^{2n}}{(2n)!}$ et $\exp(t^2/2) = \sum_{n \geq 0} \frac{t^{2n}}{n! 2^n}$ donc $\text{ch}(t) \leq e^{t^2/2}$

car $n! 2^n \leq (2n)!$ $\forall n \geq 0$. On a donc $E[e^{tX}] \leq \exp(t^2/2)$ ■

Preuve théorème:

Soit $t \in \mathbb{R}$. On applique le lemme à la variable aléatoire $\frac{X_n}{C_n}$, ce qui donne
 $\forall t' \in \mathbb{R}$: $E[\exp(t' \frac{X_n}{C_n})] \leq \exp(\frac{t'^2}{2})$ donc en prenant $t' = t \cdot C_n$,
 on obtient $E[\exp(t \cdot X_n)] \leq \exp(\frac{t^2}{2} \cdot C_n^2)$. Les variables aléatoires
 $\exp(t \cdot X_n)$ étant indépendantes, on a alors $\forall n \in \mathbb{N}^*$:

$$E[e^{tS_n}] = \prod_{j=1}^n E[e^{tX_j}] \leq \exp\left(\frac{t^2}{2} \cdot \sum_{j=1}^n C_j^2\right).$$

Soient à présent $t > 0$ et $\varepsilon > 0$. On a $IP(S_n > \varepsilon) \leq IP(e^{tS_n} > e^{t\varepsilon}) \xrightarrow{\pi} e^{t\varepsilon} \neq$

$$\leq \frac{E[e^{tS_n}]}{e^{t\varepsilon}} \quad \text{par Markov}$$

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

Posons $a = \sum_{j=1}^n c_j^2$. Soit $f: \mathbb{R}^+ \rightarrow \mathbb{R}$

$$t \mapsto \frac{t^2}{2} \cdot a - t\varepsilon$$

$\forall t \geq 0$, on a $f'(t) = at - \varepsilon$

donc f' ne s'annule qu'en $t = \frac{\varepsilon}{a}$. La fonction f atteint alors son minimum en $t = \frac{\varepsilon}{a}$ et vaut $\frac{\varepsilon^2}{2a} - \frac{\varepsilon^2}{a} = -\frac{\varepsilon^2}{2a}$ en ce point.

Par suite, on a $IP(S_n > \varepsilon) \leq \exp\left(-\frac{\varepsilon^2}{2a}\right)$

$$\leq \exp\left(-\frac{\varepsilon^2}{2 \cdot \sum_{j=1}^n c_j^2}\right)$$

Corollaire:

supposons maintenant la suite $(X_i)_{i \in \mathbb{N}}$ infinie, avec les mêmes hypothèses.

[s'il existe $\alpha, \beta > 0$ tq $\sum_{i=1}^n c_i^2 \leq n^{2\alpha-\beta}$, alors $\frac{S_n}{n^\alpha} \xrightarrow{PS} 0$]

Preuve: Par l'inégalité de Hoeffding, on a

$$\sum_{n \in \mathbb{N}} IP(|S_n| > n^\alpha \cdot \varepsilon) \leq 2 \cdot \sum_{n \in \mathbb{N}} e^{-\frac{n^{2\alpha} \cdot \varepsilon^2}{2n^{2\alpha-\beta}}} = 2 \cdot \sum_{n \in \mathbb{N}} e^{-\frac{\varepsilon^2}{2} \cdot n^\beta} < +\infty.$$

La convergence de la série a lieu car $e^{-\frac{\varepsilon^2}{2} \cdot n^\beta} = o(n^{-1})$.

On peut donc appliquer le lemme de Borel Cantelli:

$IP(\limsup_{n \rightarrow +\infty} |S_n| > n^\alpha \varepsilon) = 0$, puis par

$IP\left(\bigcup_{h \in \mathbb{N}} \limsup_{n \rightarrow +\infty} |S_n| > \frac{n^\alpha}{h}\right) = 0$ donc $IP\left(\bigcap_{h \in \mathbb{N}} \liminf_{n \rightarrow +\infty} |S_n| \leq \frac{n^\alpha}{h}\right) = 1$

Ou encore $IP\left(\lim_{n \rightarrow +\infty} \frac{|S_n|}{n^\alpha} = 0\right) = 1$.

Ainsi $\frac{S_n}{n^\alpha} \xrightarrow{PS} 0$.

Δ rappel: $\limsup_{n \rightarrow +\infty} A_n$: "A_n réalise pour une infinité de n"

$\liminf_{n \rightarrow +\infty} A_n$: "A_n réalise pour tout n à partir d'un certain rang"

Application: (intégral de Monte Carlo)

Soit $\mathbb{I} = [0, 1]^d$ muni de la mesure de Lebesgue.

Soit f intégrable et bornée, sur $\mathbb{I}$.

Soit (Y_n) suite iid de v.a. uniforme sur $\mathbb{I}$.

$$\text{Posons } e_n = \frac{1}{n} \cdot \sum_{k=1}^n f(Y_k) - \int_{\mathbb{I}} f$$

$$\text{Alors } \forall \varepsilon > 0, \mathbb{P}(|e_n| > \varepsilon) \leq 2 \cdot \exp\left(-\frac{n \varepsilon^2}{8 \cdot \|f\|_{\infty}^2}\right)$$

Preuve:

$$\text{Posons } X_n = f(Y_n) - \int_{\mathbb{I}} f.$$

X_n est centrée et $|X_n| \leq 2 \cdot \|f\|_{\infty}$

On applique Hoeffding à $\sum_{k=1}^n X_k = n \cdot e_n$:

$$\forall \varepsilon > 0, \mathbb{P}(|e_n| > \varepsilon) = \mathbb{P}(n \cdot |e_n| > n \varepsilon)$$

$$\leq 2 \cdot \exp\left(\frac{-n^2 \varepsilon^2}{2n \cdot 4 \|f\|_{\infty}^2}\right)$$

