Inégalité de Hoeffding

Théorème : Soit $(X_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires réelles indépendantes centrées telle que, pour tout $n\in\mathbb{N}^*$, il existe $c_n>0$ vérifiant $|X_n|\leqslant c_n$ presque sûrement. Pour tout $n\in\mathbb{N}^*$, on pose $S_n=\sum_{j=1}^{n} X_j$. On a, pour tout $\varepsilon>0$, $\mathbb{P}(|S_n|>\varepsilon)\leqslant 2\exp\left(-\frac{\varepsilon^2}{2\sum_{j=1}^{n} c_j^2}\right)$.

Démonstration :

• Lemme : Soit X une variable aléatoire réelle centrée et presque sûrement bornée par 1. On a, pour tout $t\in\mathbb{R}$, $\mathbb{E}[\exp(tX)]\leqslant \exp(t^2/2)$.

-> Preuve du lemme : Soit $t\in\mathbb{R}$. Pour tout $x\in[-1,1]$, 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 $x\longmapsto \exp(tx)$ étant convexe, on a, pour tout $x\in[-1,1]$,

$$\exp(tx)=\exp\left[t\left(\tfrac{1}{2}(1-x)\times(-1)+\tfrac{1}{2}(1+x)\right)\right]\leqslant \tfrac{1}{2}(1-x)\exp(-t)+\tfrac{1}{2}(1+x)\exp(t).$$

La variable aléatoire X étant bornée p.s, $\mathbb{E}[\exp(tX)]$ existe pour tout $t\in\mathbb{R}$.

Pour tout $t\in\mathbb{R}$, on a alors $\exp(tX)\leqslant \frac{1}{2}(1-X)\exp(-t)+\frac{1}{2}(1+X)\exp(t)$, donc, comme X est centrée, on a $\mathbb{E}[\exp(tX)]\leqslant \frac{1}{2}(\exp(-t)+\exp(t))=\operatorname{ch} t$. Or $\operatorname{ch} t=\sum_{n\geqslant 0}\frac{t^{2n}}{(2n)!}$,

et $\exp(t^2/2)=\sum_{n\geqslant 0}\frac{t^{2n}}{n!\,2^n}$, donc $\operatorname{ch} t\leqslant \exp(t^2/2)$, car $n!\,2^n\leqslant (2n)!$ pour tout $n\geqslant 0$.

On a donc $\mathbb{E}[\exp(tX)]\leqslant \exp(t^2/2)$, ce qui donne le lemme.

Soit $t\in\mathbb{R}$. On applique le lemme à la variable aléatoire $\frac{X_n}{c_n}$, ce qui donne, pour tout $t'\in\mathbb{R}$, $\mathbb{E}\left[\exp\left(t'\frac{X_n}{c_n}\right)\right]\leqslant \exp\left(\frac{t'^2}{2}\right)$, donc, en prenant $t'=tc_n$, on obtient $\mathbb{E}[\exp(tX_n)]\leqslant \exp\left(\frac{t^2}{2}c_n^2\right)$.

Les variables aléatoires $\exp(tX_n)$ étant indépendantes, on a alors, pour tout $n\in\mathbb{N}^*$,

$$\mathbb{E}[\exp(tS_n)]=\prod_{j=1}^{n}\mathbb{E}[\exp(tX_j)]\leqslant \exp\left(\frac{t^2}{2}\sum_{j=1}^{n}c_j^2\right).$$

Soient à présent $t > 0$ et $\varepsilon > 0$. On a $\mathbb{P}(S_n > \varepsilon) \leqslant \mathbb{P}(\exp(tS_n) > \exp(t\varepsilon))$ par croissance de $x \mapsto \exp(tx)$

$$\leqslant \frac{\mathbb{E}[\exp(tS_n)]}{\exp(t\varepsilon)} \quad \text{par Markov}$$

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

On pose $a = \sum_{j=1}^{n} c_j^2$. Soit $f : \begin{array}{ccc} \mathbb{R}_+ & \longrightarrow & \mathbb{R} \\ t & \longmapsto & \frac{t^2}{2}a - t\varepsilon \end{array}$. Pour tout $t \geqslant 0$, 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 $\mathbb{P}(S_n > \varepsilon) \leqslant \exp\left(-\frac{\varepsilon^2}{2a}\right)$

$$\leqslant \exp\left(-\frac{\varepsilon^2}{2\sum_{j=1}^{n} c_j^2}\right),$$ ce qui achève la preuve.