Inégalité de Hoeffding

Leçons : 253, 260, 262

On se place dans $(\Omega, \mathscr{F}, \mathbb{P})$ un espace probabilisé.

Théorème 1

Soit $(X_n)_n$ suite de variables aléatoires centrées telles que $|X_n| \leqslant c_n$ presque sûrement. Soit $a_n = \sum_{j=1}^{n} c_j^2$ et $S_n = \sum_{j=1}^{n} X_j$. Alors si $\varepsilon > 0$,

$$\mathbb{P}(|S_n| > \varepsilon) \leqslant 2\exp\left(\frac{-\varepsilon^2}{2a_n}\right)$$

Lemme 2

Soit X variable aléatoire centrée telle que $|X| \leqslant 1$ presque sûrement. Alors $L_X(t) = \mathbb{E}[e^{tX}] \leqslant e^{\frac{t^2}{2}}$.

Démonstration. Si $t \in \mathbb{R}$ et $x \in [-1,1]$ alors $tx = \dfrac{1-x}{2} \times (-t) + \dfrac{1+x}{2} \times t$ donc par convexité de la fonction exp, $e^{tx} \leqslant \dfrac{1-x}{2}e^{-t} + \dfrac{1+x}{2}e^{t}$.

Appliquant ce résultat à e^{tX}, on obtient, comme $|X| \leqslant 1$ presque sûrement, $L_X(t) \leqslant \mathbb{E}\left[\dfrac{1-X}{2}\right]e^{-t} + \mathbb{E}\left[\dfrac{1+X}{2}\right]e^{t} = \mathrm{ch}(t)$ car X est centrée.

Enfin, $\mathrm{ch}(t) = \sum_{n=0}^{+\infty} \dfrac{t^{2n}}{(2n)!} \leqslant \sum_{n=0}^{+\infty} \dfrac{t^{2n}}{2^n n!} = e^{\frac{t^2}{2}}$ car $(2n)! = n! \times (n+1) \times \cdots \times (2n) \geqslant 2^n n!$ □

Démonstration (du théorème)**.** Par indépendance des X_j, on a en remarquant que $\dfrac{X_j}{c_j}$ vérifie les conditions du lemme,

$$\forall t \in \mathbb{R}, L_{S_n}(t) = \prod_{j=1}^{n} L_{X_j}(t) = \prod_{j=1}^{n} L_{\frac{X_j}{c_j}}(tc_j) \leqslant \prod_{i=1}^{n} \exp\left(\frac{t^2 c_j^2}{2}\right) = \exp\left(\frac{t^2 a_n}{2}\right)$$

Soit $\varepsilon > 0$. Selon l'inégalité de Markov,

$$\mathbb{P}(S_n > \varepsilon) = \mathbb{P}(e^{tS_n} > e^{t\varepsilon}) \leqslant \frac{\mathbb{E}\left[e^{tS_n}\right]}{e^{t\varepsilon}} \leqslant \exp\left(\frac{t^2 a_n}{2} - t\varepsilon\right)$$

Or si $\varphi : t \mapsto \dfrac{t^2 a_n}{2} - t\varepsilon$, c'est une fonction polynômiale de degré 2 de coefficient dominant positif et $\varphi'(t) = ta_n - \varepsilon$ donc φ atteint son minimum en $\dfrac{\varepsilon}{a_n}$.

Ainsi, $\mathbb{P}(S_n > \varepsilon) \leqslant \exp\left(\dfrac{\varepsilon^2}{a_n^2}\dfrac{a_n}{2} - \dfrac{\varepsilon^2}{a_n}\right) = \exp\left(\dfrac{-\varepsilon^2}{2a_n}\right)$.

En appliquant ce résultat à $-S_n$, on obtient

$\mathbb{P}(|S_n| > \varepsilon) \leqslant \mathbb{P}(S_n > \varepsilon) + \mathbb{P}(S_n < -\varepsilon) \leqslant 2\exp\left(\dfrac{-\varepsilon^2}{2a_n}\right)$, ce qu'il fallait démontrer. □

Proposition 3

On suppose de plus qu'il existe $\alpha, \beta > 0$ tels que $2\alpha - \beta > 0$ et $a_n \leqslant n^{2\alpha-\beta}$ pour tout $n \in \mathbb{N}$. Alors presque sûrement $\dfrac{S_n}{n^\alpha}$ tend vers 0.

Démonstration. Selon le théorème précédent, $\mathbb{P}(|S_n| \geqslant \varepsilon n^\alpha) \leqslant \exp\left(\dfrac{-\varepsilon^2 n^{2\alpha}}{2a_n}\right) \leqslant \exp\left(\dfrac{-\varepsilon^2 n^\beta}{2}\right)$, ce dernier terme étant le terme général positif d'une série convergente (par exemple parce que il est négligeable devant $\dfrac{1}{n^2}$).

Donc, selon le lemme de Borel-Cantelli, $\dfrac{S_n}{n^\alpha}$ converge presque sûrement. $\square$

Référence : Jean-Yves OUVRARD (2009). *Probabilités (Master Agrégation)*. T. 2. Cassini, p. 210