Inégalité de Hoeffding et applications

Geoffrey Deperle

Leçons associées :
- 261 : Loi d'une variable aléatoire : caractérisations, exemples, applications.
- 262 : Convergences d'une suite de variables aléatoires. Théorèmes limite. Exemples et applications.
- 264 : Variables aléatoires discrètes. Exemples et applications.
- 266 : Illustration de la notion d'indépendance en probabilités.

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

Lemme 1. *Soit X une variable réelle, centrée, bornée par 1 presque sûrement alors*

$$\forall t \in \mathbb{R}^{+*}, \mathbb{E}[\exp(tX)] \leq \exp\left(\frac{t^2}{2}\right)$$

Preuve : Pour $x \in [-1,1]$ et $t \in \mathbb{R}$, on a $\begin{cases} \frac{1-x}{2} + \frac{1+x}{2} = 1 \\ \frac{1-x}{2} \in [0,1] \end{cases}$ donc $\frac{1-x}{2}(-t) + \frac{1+x}{2}t = tx$ est une combinaison convexe de $-t$ et t.
Par convexité de exp,

$$\forall t \in \mathbb{R}, \exp(tx) \leq \frac{1-x}{2}\exp(-t) + \frac{1+x}{2}\exp(t)$$

Pour $u \in \mathbb{R}$, uX est bornée p.s. donc $\exp(uX)$ est bornée p.s. donc intégrable. Donc pour $t \in \mathbb{R}$,

$$\begin{aligned} L_X(t) &:= \mathbb{E}[\exp(tX)] \\ &\leq \frac{1}{2}\mathbb{E}[(1-X)\exp(-t)] + \frac{1}{2}\mathbb{E}[(1+X)\exp(t)] \\ &\leq \frac{1}{2}(\exp(-t) + \exp(t)) \qquad \text{car } \mathbb{E}[X] = 0 \\ &\leq \cosh(t) \end{aligned}$$

Or, comme $\forall k \in \mathbb{N}^*, 2^k k! = 2 \times 4 \times 6 \times \cdots \times 2(k-1) \times 2k \leq (2k)!$.
Donc pour tout $t \in \mathbb{R}$,

$$\cosh(t) = \sum_{k=0}^{+\infty} \frac{t^{2k}}{(2k)!} \leq \sum_{k=0}^{+\infty} \frac{t^{2k}}{2^k k!} = \exp\left(\frac{t^2}{2}\right)$$

d'où le résultat. $\square$

Théorème. *Soit* $(X_n)_{n\in\mathbb{N}}$ *une suite de variable aléatoires réelles indépendantes, centrées, bornées p.s. :* $\forall n \in \mathbb{N}, \exists c_n > 0 / |X_n| \leq c_n$.
Soit $S_n = \sum_{k=1}^n X_j$, *alors*

$$\forall n \in \mathbb{N}, \varepsilon > 0, \mathbb{P}(|S_n| \geq \varepsilon) \leq 2 \exp\left(\frac{\varepsilon^2}{2\sum_{k=1}^n c_k^2}\right)$$

Preuve : Soit $n \in \mathbb{N}$, $t \in \mathbb{R}^{+*}$,

Étape 1 : Donnons une majoration de $\mathbb{E}[\exp(tS_n)]$

La variable aléatoire $\frac{X_n}{c_n}$ est bornée p.s. par 1 et centrée, donc d'après le lemme

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

Avec $t' = tc_n$, on a donc $\mathbb{E}[\exp(tX_n)] \leq \exp\left(\frac{t^2}{2}c_n^2\right)$. Ainsi,

$$\begin{aligned}
\mathbb{E}[\exp(tS_n)] &= \prod_{k=1}^n \mathbb{E}[\exp(tX_k)] \quad \text{car les } \exp(tX_n) \text{ sont indépendantes} \\
&\leq \prod_{k=1}^n \exp\left(\frac{t^2}{2}c_k^2\right) \\
&\leq \exp\left(\frac{t^2}{2}\sum_{k=1}^n c_k^2\right)
\end{aligned}$$

Étape 2 : Établissons l'inégalité avec l'inégalité de Markov

Soit $\varepsilon > 0$, la fonction $x \mapsto \exp(tx)$ est strictement croissante donc $[S_n > \varepsilon] = [\exp(tS_n) > \exp(t\varepsilon)]$ d'où

$$\begin{aligned}
\mathbb{P}(S_n > \varepsilon) &\leq \mathbb{P}(\exp(tS_n) > \exp(t\varepsilon)) \\
&\leq \frac{\mathbb{E}[\exp(tS_n)]}{\exp(t\varepsilon)} \\
&\leq \exp\left(\frac{t^2}{2}\sum_{k=1}^n c_k^2 - t\varepsilon\right)
\end{aligned}$$

Étape 3 : Optimisons l'inégalité

En posant $a = \sum_{k=1}^n c_k^2$, la fonction $t \mapsto \frac{a}{2}t^2 - \varepsilon t$ atteint son minimum en $\frac{-\varepsilon^2}{2a}$ pour $t = \frac{\varepsilon}{a} > 0$. Par croissance de l'exponentielle, on a donc

$$\mathbb{P}(S_n > \varepsilon) \leq \exp\left(\frac{\varepsilon^2}{2\sum_{k=1}^n c_k^2}\right)$$

Conclusion

On a $[|S_n| > \varepsilon] = [S_n > \varepsilon] \sqcup [-S_n > \varepsilon]$, on a

$$\forall n \in \mathbb{N}, \varepsilon > 0, \mathbb{P}(|S_n| \geq \varepsilon) \leq 2 \exp\left(\frac{\varepsilon^2}{2\sum_{k=1}^n c_k^2}\right)$$

$\square$

Intéressons nous à trois applications.

Application ((261, 264))**.** *Soit* $(X_i)_{i\in[\![1,n]\!]}$ *des variables aléatoires iid de loi* $X_1 \sim \mathcal{B}(p)$ *avec* $p \in]0,1[$ *et* $\alpha \in]0,1[$, *alors un intervalle de confiance par excès de niveau* $1-\alpha$ *de paramètre* p *est*

$$I_{1-\alpha} = \left[\frac{1}{n}S_n - \sqrt{\frac{2}{n}\ln\left(\frac{2}{\alpha}\right)}, \frac{1}{n}S_n + \sqrt{\frac{2}{n}\ln\left(\frac{2}{\alpha}\right)}\right]$$

Preuve : On applique l'inégalité de Hoeffding à $(X_n - p)$ qui est borné par 1

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

Soit $\varepsilon' = \sqrt{\frac{2}{n}\ln\left(\frac{2}{\alpha}\right)}$ tel que $\alpha = 2\exp(-\frac{n\varepsilon'^2}{2})$ d'où

$$\mathbb{P}\left(\left|\frac{1}{n}S_n - p\right| < \varepsilon'\right) = 1 - \mathbb{P}\left(\left|\frac{1}{n}S_n - p\right| > \varepsilon\right) \geq 1-\alpha$$

d'où le résultat. $\square$

Application ((261, 266))**.** *Soit* $I = [0,1]^d$ *muni de la mesure de Lebesgue, soit* f *une fonction intégrable et bornée sur* I.
Soit (Y_n) *une suite de variable aléatoire iid de loi uniforme sur* I. *On a*

$$\forall \varepsilon, \mathbb{P}\left(\left|\frac{1}{n}\sum_{k=1}^{n} f(Y_k) - \int_I f\right| > \varepsilon\right) \leq 2\exp\left(-\frac{n\varepsilon^2}{8||f||_\infty^2}\right)$$

Preuve : Posons pour $n \in \mathbb{N}$, $X_n = f(Y_n) - \displaystyle\int_I f$.
On a $\mathbb{E}[X_n] = 0$ d'après la formule de transfert et $|X_n| \leq 2||f||_\infty$ donc d'après l'inégalité de Hoeffding avec $c_n = 2||f||_\infty$.
Pour tout $\varepsilon > 0$,

$$\begin{aligned}\mathbb{P}\left(\left|\frac{1}{n}\sum_{k=1}^{n} f(Y_k) - \int_I f\right| > \varepsilon\right) &= \mathbb{P}\left(\left|\sum_{k=1}^{n} X_k\right| \geq n\varepsilon\right)\\ &\leq 2\exp\left(\frac{n^2\varepsilon^2}{2\sum_{k=1}^{n} 4||f||_\infty^2}\right)\\ &\leq 2\exp\left(\frac{n\varepsilon^2}{8||f||_\infty^2}\right)\end{aligned}$$

$\square$

Application ((262))**.** *Soit* $\alpha > 0$, *s'il existe* $\beta > 0$ *tel que pour tout* $n \in \mathbb{N}^*$, $\displaystyle\sum_{k=1}^{n} c_k^2 \leq n^{2\alpha-\beta}$, *alors*

$$\frac{S_n}{n^\alpha} \xrightarrow[\text{p.s.}]{n\to+\infty} 0$$

Preuve : Soit $\varepsilon > 0$, par inégalité de Hoeffding, on a

$$\mathbb{P}(|S_n| > n^\alpha \varepsilon) \leq 2\exp\left(-\frac{n^{2\alpha}\varepsilon^2}{2\sum\limits_{k=1}^{n} c_k^2}\right)$$

Or comme $\sum\limits_{k=1}^{n} c_k^2 \leq n^{2\alpha-\beta}$, on a

$$\mathbb{P}(|S_n| > n^\alpha \varepsilon) \leq 2\exp(-\varepsilon n^\beta)$$

Or, comme $\sum \exp(-\varepsilon n^\beta)$ converge, la série de terme général $\mathbb{P}(|S_n| > n^\alpha \varepsilon)$ converge. Donc d'après le lemme de Borel-Cantelli, on a $\mathbb{P}(\limsup_n [|S_n| > n^\alpha \varepsilon]) = 0$.

Or, comme $\mathbb{Q}^{+*}$ est dénombrable, on a $\mathbb{P}\left(\bigcup\limits_{s\in\mathbb{Q}^{+*}} \limsup\limits_n [|S_n| > n^\alpha \varepsilon]\right) = 0$.

Donc en passant au complémentaire, on a

$$\mathbb{P}\left(\bigcap_{s\in\mathbb{Q}^{+*}} \liminf_n [|S_n| > n^\alpha \varepsilon]\right) = 1$$

Donc la suite $\left(\frac{S_n}{n}\right)$ converge presque sûrement vers 0. $\square$

Références

[1] Jean-Yves Ouvrard. *Probabilités 2*. Cassini, 2009.