Inégalité de Hoeffding

Ouvrard, *Probabilités 2*, page 132

Exercice :

1. Soit X une variable aléatoire réelle centrée, bornée par 1.
   (a) Soit un réel t quelconque. Montrer que
   $$\forall x \in [-1,1], \quad \exp(tx) \leq \frac{1-x}{2}\exp(-t) + \frac{1+x}{2}\exp(t)$$
   (b) En déduire les inégalités
   $$\mathbb{E}\exp(tX) \leq \mathrm{ch}(t) \quad \text{et} \quad \mathbb{E}\exp(tX) \leq \exp\left(\frac{t^2}{2}\right)$$
2. Soit $(X_n)_{n\geq 1}$ une suite de v.a.r. indépendantes centrées telles que $|X_n| \leq c_n$ avec $c_n > 0$. On note pour tout $n \geq 1$, $S_n = \sum_{j=1}^{n} X_j$.
   (a) Montrer que pour tout t,
   $$\mathbb{E}\exp(tS_n) \leq \exp\left(\frac{t^2}{2}\sum_{j=1}^{n} c_j^2\right)$$
   (b) Montrer alors avec l'inégalité de Markov que pour tout $t > 0$ et $\varepsilon > 0$,
   $$\mathbb{P}(S_n > \varepsilon) \leq \exp\left(-t\varepsilon + \frac{t^2}{2}\sum_{j=1}^{n} c_j^2\right)$$
   (c) En déduire que pour tout $\varepsilon > 0$,
   $$\mathbb{P}(S_n > \varepsilon) \leq \exp\left(-\frac{\varepsilon^2}{2\sum_{j=1}^{n} c_j^2}\right)$$
   (d) Montrer alors pour tout $\varepsilon > 0$ l'inégalité de Hoeffding,
   $$\boxed{\mathbb{P}(|S_n| > \varepsilon) \leq 2\exp\left(-\frac{\varepsilon^2}{2\sum_{j=1}^{n} c_j^2}\right)}$$

1. (a) Soit t un réel quelconque. Pour $|x| \geq 1$, on a
   $$\frac{1-x}{2} \in [0,1] \quad , \quad \frac{1+x}{2} \in [0,1] \quad \text{et} \quad \frac{1-x}{2} + \frac{1+x}{2} = 1$$
   Puisque on a l'égalité $tx = \frac{1-x}{2}(-t) + \frac{1+x}{2}(t)$, la fonction $x \mapsto e^{tx}$ étant convexe, on a pour tout $x \in [-1,1]$,
   $$e^{tx} \leq \frac{1-x}{2}e^{-t} + \frac{1+x}{2}e^{t}$$
   (b) La v.a. X étant bornée par 1, la v.a. $\exp(tX)$ est bornée et admet donc une espérance. On a donc
   $$\exp(tX) \leq \frac{1-X}{2}e^{-t} + \frac{1+X}{2}e^{t}$$
   On a donc
   $$\mathbb{E}\exp(tX) \leq \frac{\mathbb{E}(1-X)}{2}e^{-t} + \frac{\mathbb{E}(1+X)}{2}e^{t} \leq \frac{e^{-t}+e^{t}}{2} = \mathrm{ch}t$$

Or, on a

$$\mathrm{ch}(t) = \sum_{n=0}^{+\infty} \frac{t^{2n}}{(2n)!} \quad \text{et} \quad \exp\left(\frac{t^2}{2}\right) = \sum_{n=0}^{+\infty} \frac{t^{2n}}{n!2^n}$$

et comme $\forall n \in \mathbb{N}, n!2^n \leq (2n)!$, on a $\mathrm{ch}(t) \leq \exp\left(\frac{t^2}{2}\right)$. On a donc

$$\boxed{\mathbb{E}\exp(tX) \leq \exp\left(\frac{t^2}{2}\right)}$$

2. (a) Soit t quelconque. On applique l'inégalité précédente à la v.a.r. $\frac{X_n}{c_n}$, on a alors

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

puis en spécialisant pour $t' = tc_n$, on a

$$\mathbb{E}\exp(tX_n) \leq \exp\left(\frac{t^2}{2}c_n^2\right)$$

Or les v.a.r. $\exp(tX_n)$ sont indépendantes donc

$$\forall t \in \mathbb{R}, \quad \boxed{\mathbb{E}\exp(tS_n) \leq \exp\left(\frac{t^2}{2}\sum_{j=1}^{n} c_j^2\right)}$$

(b) Soient $t > 0$ et $\varepsilon > 0$. La fonction $x \mapsto e^{tx}$ est croissante, on a donc $[S_n > \varepsilon] \subset [\exp(tS_n) > \exp(t\varepsilon)]$ et donc d'après l'inégalité de Markov,

$$\mathbb{P}(S_n > \varepsilon) \leq \mathbb{P}\left(\exp(tS_n) > \exp(t\varepsilon)\right) \leq \frac{\mathbb{E}\exp(tS_n)}{\exp(t\varepsilon)}$$

D'où

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

En appliquant ceci à $-X_n$, on a

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

(c) Soit $\varepsilon > 0$ et soit $a = \sum\limits_{j=1}^{n} c_j^2$. La fonction $t \mapsto a\frac{t^2}{2} - t\varepsilon$ atteint son minimum pour $t = \frac{\varepsilon}{a} > 0$ et ce minimum vaut $-\frac{\varepsilon^2}{2a}$. D'où

$$\mathbb{P}(S_n > \varepsilon) \leq \exp\left(\min_{t>0}(-t\varepsilon + \frac{t^2}{2}\sum_{j=1}^{n} c_j^2)\right) = \exp\left(-\frac{\varepsilon^2}{2\sum\limits_{j=1}^{n} c_j^2}\right)$$

(d) Soit $\varepsilon > 0$. On a $[|S_n| > \varepsilon] = [S_n > \varepsilon] \cup [-S_n > \varepsilon]$ et ainsi

$$\mathbb{P}(|S_n| > \varepsilon) \leq \mathbb{P}(S_n > \varepsilon) + \mathbb{P}(-S_n > \varepsilon)$$

D'où

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