Annexe : l'inégalité de Hoeffding

On se place ici dans (Ω,P) un espace probabilisé fini.

1 Lemmes

Proposition 1. Inégalité de Markov
Soit Z une variable aléatoire réelle presque sûrement positive. On a

$$\forall a > 0 \;\; P(Z \geq a) \leq \frac{\mathbb{E}[Z]}{a}$$

Proposition 2. .
Soit I un ensemble, $\Phi : I \to \mathbb{R}_+$ croissante, Y une variable aléatoire réelle telle que $P(Y \in I) = 1$.
Alors, en appliquant l'inégalité de Markov à la variable aléatoire $\Phi(Y)$:

$$\forall b \in I,\, \Phi(b) > 0 \implies P(Y \geq b) \leq \frac{\mathbb{E}[\Phi(Y)]}{\Phi(b)}$$

Proposition 3. Lemme d'Hoeffding
Soit Y une variable aléatoire réelle centrée. On suppose que $P(a \leq Y \leq b) = 1$ avec $a < b$ deux réels. Alors

$$\forall s > 0 \;\; \mathbb{E}[\mathrm{e}^{sY}] \leq \exp\Big(\frac{s^2(b-a)^2}{8}\Big)$$

Démonstration. Remarquons dans un premier temps que les cas où $a \geq 0$ ou $b \leq 0$ sont immédiats.
En effet, si $a \geq 0$, alors Y est positive et d'espérance nulle : elle est donc nulle et la propriété est immédiate. De même pour $b \leq 0$.

On suppose donc que $a < 0$ et $b > 0$.

soit $s > 0$ et $x \in [a,b]$. on peut écrire

$$\mathrm{e}^{sx} = \exp\Big(\frac{b-x}{b-a}sa + \frac{x-a}{b-a}sb\Big)$$

et donc par convexité de la fonction qui à t associe e^{st} :

$$\mathrm{e}^{sx} \leq \frac{b-x}{b-a}\mathrm{e}^{sa} + \frac{x-a}{b-a}\mathrm{e}^{sb}$$

en appliquant cette inégalité avec Y et en passant à l'espérance, il vient, comme Y est centrée et par linéarité de l'espérance :

$$\mathbb{E}[\mathrm{e}^{sY}] \leq \frac{b}{b-a}\mathrm{e}^{sa} + \frac{-a}{b-a}\mathrm{e}^{sb}$$

on pose $p = \frac{-a}{b-a}$, d'où :

$$\mathbb{E}[\mathrm{e}^{sY}] \leq (1-p)\mathrm{e}^{sa} + p\mathrm{e}^{sb}$$

alors, en posant l'application :

$$\Psi : \begin{cases} \mathbb{R}_+ & \to \quad \mathbb{R} \\ u & \mapsto \quad -pu + \ln(1 - p + p\mathrm{e}^u) \end{cases}$$

qui est de classe C^2, on remarque que

$$(1-p)\mathrm{e}^{sa} + p\mathrm{e}^{sb} = \mathrm{e}^{\Psi(s(b-a))}$$

Etudions alors Ψ. soit $u \in \mathbb{R}_+$. on a

$$\Psi'(u) = -p + \frac{p\mathrm{e}^u}{1-p+p\mathrm{e}^u}$$

et

$$\Psi''(u) = \frac{p\mathrm{e}^u - p^2\mathrm{e}^u}{(1-p+p\mathrm{e}^u)^2} = \frac{p\mathrm{e}^u}{1-p+p\mathrm{e}^u}\Big(1 - \frac{p\mathrm{e}^u}{1-p+p\mathrm{e}^u}\Big) = X(1-X) \leq \frac{1}{4}$$

de plus, les expressions précédentes nous permettent de dire que $\Psi(0) = \Psi'(0) = 0$.
Alors, en vertu de l'inégalité de Taylor-Lagrange avec Ψ qui est de classe C^2 entre les points u et 0, il vient :

$$\big|\Psi(u) - \Psi(0) - u\Psi'(0)\big| \leq \frac{u^2}{2}\frac{1}{4}$$

d'où, comme Ψ est positive car croissante et nulle en 0 :

$$\Psi(u) \leq \frac{u^2}{8}$$

et ceci pour tout $u \geq 0$.
ainsi, la majoration précédemment établie nous permet de conclure :

$$\mathbb{E}[\mathrm{e}^{sY}] \leq \exp\Big(\frac{s^2(b-a)^2}{8}\Big)$$

2 L'inégalité

Proposition 4. Une inégalité de Hoeffding
Soit $t \in \mathbb{N}^*$ et $(X_1,...,X_t)$ des variables aléatoires réelles indépendantes à valeurs presque sûrement dans [0, 1].
On pose $S_t = \frac{1}{t}\sum_{k=1}^{t} X_k$. Alors

$$\forall u > 0 \ \ P\big(\mathbb{E}[s_t] \geq S_t + u\big) \leq \mathrm{e}^{-2tu^2}$$

Démonstration. pour tout k entre 1 et t, on pose $Y_k = \mathbb{E}[X_k] - X_k$.
Les Y_k sont centrées et à valeurs dans $\big[-\mathbb{E}[X_k], 1 - \mathbb{E}[X_k]\big]$ avec $1 - \mathbb{E}[x_k] - \big(-\mathbb{E}[X_k]\big) = 1$.
On a

$$\mathbb{E}[S_t] - S_t = \frac{1}{t}(Y_1 + ... + Y_t)$$

soit $u > 0$ et $s > 0$. on calcule

$$\begin{aligned}
P\big(\mathbb{E}[S_t] - S_t \geq u\big) &\leq \mathbb{E}\big[\exp\big(s(\mathbb{E}[S_t] - S_t)\big)\big]\mathrm{e}^{-su} \quad (\mathit{Prop.}\ 2)\\
&= \mathbb{E}\big[\exp\big(\frac{s}{t}(Y_1 + \ ... \ + Y_t)\big)\big]\mathrm{e}^{-su}\\
&= \mathrm{e}^{-su}\prod_{k=1}^{t}\mathbb{E}[\exp\big(\frac{s}{t}Y_k\big)] \quad (\text{indépendance des } Y_k)\\
&\leq \mathrm{e}^{-su}\prod_{k=1}^{t}\exp\big(\frac{s^2}{8t^2}\big) \quad (\text{Prop. 3 avec } b-a=1)\\
&= \exp(-su + \frac{s^2}{8t})
\end{aligned}$$

on choisit alors $s = 4ut$ et on obtient l'inégalité souhaitée.