

Inégalité de Hoeffding

Thomas CHEN

L'inégalité de Hoeffding est une inégalité très classique. Il y a deux points clefs à retenir, outre les techniques classiques en probabilités.

1. Déjà, on établit l'inégalité

$$\forall t \in \mathbb{R}, \cosh(t) \leq e^{t^2/2}.$$

Cette inégalité se montre via les séries entières.

2. Ensuite, on utilise le fait, peu remarqué par pas mal d'étudiants, suivant : s'il existe une quantité M indépendante de λ telle que

$$\forall \lambda \in I, M \leq f(\lambda),$$

alors il est extrêmement utile de considérer λ_0 qui minimise f ! (par passage de l'inégalité à l'infimum). Quand on établit une inégalité, on choisit souvent celle qui est la plus précise.

Exercice 1. 1. Soit $x \in [-1, 1], t \in \mathbb{R}$. Montrer que

$$e^{tx} \leq \frac{1-x}{2} e^{-t} + \frac{1+x}{2} e^t.$$

2. Soit X une variable aléatoire discrète ayant une espérance. On la suppose centrée bornée par 1. Montrer que e^{tX} admet une espérance et que

$$\mathbb{E} \left[e^{tX} \right] \leq e^{t^2/2}.$$

3. Soit X_1, \dots, X_n des variables aléatoires réelles discrètes centrées indépendantes. On suppose que pour tout $i \in \llbracket 1, n \rrbracket$, $|X_i|$ est bornée par un réel a_i . En notant $S_n = \sum_{i=1}^n X_i$, montrer que

$$\forall t \in \mathbb{R}, \mathbb{E} \left[e^{tS_n} \right] \leq \exp \left(\frac{t^2}{2} \sum_{i=1}^n a_i^2 \right).$$

4. En déduire

$$\forall \varepsilon > 0, \mathbb{P}(|S_n| \geq \varepsilon) \leq 2 \exp \left(-\frac{\varepsilon^2}{2 \sum_{j=1}^n c_j^2} \right).$$

Remarque. Si on ne suppose plus les variables centrées, le résultat tient à condition d'estimer $\mathbb{P}(|S_n - \mathbb{E}[S_n]| \geq \varepsilon)$.

Corrigé :

1. exp est convexe donc par l'inégalité de convexité, on a

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

puisque $\frac{1-x}{2} + \frac{1+x}{2} = 1$ et que ces deux termes que l'on somme sont dans $[0, 1]$, x étant lui dans $[-1, 1]$.

2. Soit $t \in \mathbb{R}$. On a $tX \leq |t||X| \leq |t|$ donc par croissance de \exp , on a $0 \leq \exp(tX) \leq \exp(|t|)$. Ainsi, $\exp(tX)$ est positive bornée donc elle admet une espérance. Par ailleurs, X est bornée par 1 donc par la question précédente, on a

$$\exp(tX) \leq \frac{1-X}{2} \exp(-t) + \frac{1+X}{2} \exp(t).$$

Par croissance puis linéarité de l'espérance, puisque X est centrée (donc $\mathbb{E}[X] = 0$), on a

$$\mathbb{E}[\exp(tX)] \leq \frac{1-\mathbb{E}[X]}{2} \exp(-t) + \frac{1+\mathbb{E}[X]}{2} \exp(t) = \frac{1}{2}(\exp(-t) + \exp(t)) = \cosh(t).$$

Or, on a l'inégalité classique

$$\forall t \in \mathbb{R}, \cosh(t) \leq \exp(t^2/2).$$

Pour cela, on va utiliser les séries entières. On a

$$\forall t \in \mathbb{R}, \cosh(t) = \sum_{n=0}^{+\infty} \frac{1}{(2n)!} t^{2n}, \quad \exp(t^2/2) = \sum_{n=0}^{+\infty} \frac{1}{n! 2^n} t^{2n}.$$

Or,

$$\forall n \in \mathbb{N}, 2^n \leq \prod_{k=n+1}^{2n} k$$

donc

$$\frac{1}{(2n)!} \leq \frac{1}{n! 2^n}$$

donc par positivité de t^{2n} pour tout $t \in \mathbb{R}$ et $n \in \mathbb{N}$, on a

$$\cosh(t) \leq \exp(t^2/2).$$

Finalement, on a bien $\mathbb{E}[\exp(tX)] \leq \exp(t^2/2)$.

3. Assurons-nous de l'existence des objets. Soit $t \in \mathbb{R}$. S_n est bornée comme somme finie de variables aléatoires bornées donc $\exp(tS_n)$ l'est aussi et admet donc une espérance. Par indépendance des $(X_i)_i$, on a alors l'indépendance des $(\exp(tX_i))_i$ ce qui donne

$$\mathbb{E}[\exp(tS_n)] = \mathbb{E} \left[\prod_{i=1}^n \exp(tX_i) \right] = \prod_{i=1}^n \mathbb{E}[\exp(tX_i)].$$

On veut utiliser la question précédente. Pour obtenir X_i bornée par 1, il suffit de diviser par a_i donc on a

$$\forall t \in \mathbb{R}, \forall 1 \leq i \leq n, \mathbb{E}[\exp(tX_i/a_i)] \leq \exp(t^2/2).$$

Ainsi,

$$\forall t \in \mathbb{R}, \forall 1 \leq i \leq n, 0 \leq \mathbb{E}[\exp(tX_i)] \leq \exp(a_i^2 t^2/2).$$

Ainsi, par produit, on a

$$\mathbb{E}[\exp(tS_n)] \leq \prod_{i=1}^n \exp(a_i^2 t^2/2) = \exp\left(\frac{t^2}{2} \sum_{i=1}^n a_i^2\right).$$

4. Fixons maintenant $t > 0$ et $\varepsilon > 0$. Par positivité de $\exp(tS_n)$, par Markov, on a

$$\mathbb{P}(S_n \geq \varepsilon) \leq \mathbb{P}(\exp(tS_n) \geq \exp(t\varepsilon)) \leq \frac{\mathbb{E}[\exp(tS_n)]}{\exp(t\varepsilon)} \leq \exp\left(\frac{t^2}{2} \sum_{j=1}^n c_j^2 - t\varepsilon\right).$$

Cela étant vrai pour tout t , il faut trouver un t « optimal ». Posons $a = \sum_{j=1}^n c_j^2$ et $f : t \geq 0 \mapsto \frac{t^2}{2} a - t\varepsilon \in \mathbb{R}$.

f est polynomiale : elle atteint un unique minimum en ε/a et f vaut $-\varepsilon^2/2a$. Ainsi,

$$\mathbb{P}(S_n \geq \varepsilon) \leq \exp(-\varepsilon^2/2a) = \exp\left(\frac{-\varepsilon^2}{2 \sum_{j=1}^n c_j^2}\right).$$