Feuille d'exercices 4

1. Nombre chromatique d'un graphe aléatoire

Soit $G \sim \mathcal{G}(n,p)$ un graphe aléatoire d'Erdös–Renyi : l'ensemble de sommets correspond à $\{1,\dots,n\}$ et, indépendamment pour chaque paire $i<j$ de sommets distincts, on place une arête entre i et j avec probabilité p. Soit $\chi(G)$ le nombre chromatique de G, i.e. le nombre minimal de couleurs permettant de colorier les sommets de telle sorte que deux sommets voisins n'aient jamais la même couleur. À l'aide de l'inégalité des différences bornées, montrer que pour tout $t \geq 0$,

$$\mathbf{P}\left(|\chi(G) - \mathbf{E}\chi(G)| \geq t\right) \leq \exp\left\{-\frac{t^2}{2n}\right\} .$$

Remarque : on peut montrer que pour $p \in]0,1[$ fixé,

$$\mathbf{E}\chi(G) \underset{n\to\infty}{\sim} \frac{n}{2\log_b(n)} ,$$

où $b = \frac{1}{1-p}$ (voir Bollobás, *The chromatic number of random graphs*, 1988).

2. Entropie de Vapnik–Chervonenkis

Soit $\mathcal{X}$ un ensemble et $\mathcal{A}$ une collection de sous-ensembles A de $\mathcal{X}$. Pour $x = (x_1,\dots,x_n) \in \mathcal{X}^n$, on définit la trace de $\mathcal{A}$ sur x comme

$$\mathsf{tr}(x) = \{A \cap \{x_1,\dots,x_n\},\, A \in \mathcal{A}\} .$$

L'entropie de Vapnik–Chervonenkis, ou VC-entropie, est alors donnée par

$$h(x) = \log_2 |\mathsf{tr}(x)| .$$

(1) À l'aide de l'inégalité de Han, montrer que la fonction h est auto-bornée.

(2) Si $X = (X_1,\dots,X_n)$ sont des variables indépendantes à valeurs dans $\mathcal{X}$, que peut-on dire de la concentration de $Z = h(X)$ autour de son espérance ?

(3) Montrer que

$$\log \mathbf{E}|\mathsf{tr}(X)| \leq \mathbf{E}\left[\log_2 |\mathsf{tr}(X)|\right] \leq \log_2 \mathbf{E}|\mathsf{tr}(X)| .$$

3. Fonctions faiblement (a,b)-auto-bornées

Une fonction $f : \mathcal{X}^n \to \mathbb{R}_+$ est dite faiblement (a,b)-auto-bornée, avec $a,b \geq 0$, s'il existe des fonctions $f_i : \mathcal{X}^{n-1} \to \mathbb{R}_+$ telles que pour tout $x \in \mathcal{X}^n$,

$$\sum_{i=1}^n \left(f(x) - f_i(x^{(i)})\right)^2 \leq af(x) + b .$$

Montrer que si $f : \mathcal{X}^n \to \mathbb{R}_+$ est faiblement (a,b)-auto-bornée, et vérifie de plus, pour tout $x \in \mathcal{X}^n$ et $i \in [\![1,n]\!]$, $f_i(x^{(i)}) \leq f(x)$, et si $Z = f(X_1,\dots,X_n)$ avec $X_1,\dots,X_n$ indépendantes à valeurs dans $\mathcal{X}$, alors pour tout $\lambda \in [0, 2/a[$,

$$\log \mathbf{E}e^{\lambda(Z - \mathbf{E}Z)} \leq \frac{(a\mathbf{E}Z + b)\lambda^2}{2(1 - a\lambda/2)} .$$

(On a bien sûr le droit d'aller fouiller dans *Concentration Inequalities* de Boucheron, Lugosi, et Massart.)