Feuille d'exercices 1

1. Variance des statistiques d'ordre
Soient $X_1, \dots, X_n$ des variables aléatoires réelles i.i.d. On note $X_{(1)} \le \cdots \le X_{(n)}$ les statistiques d'ordre.

(1) À l'aide de l'inégalité d'Efron–Stein, montrer que pour tout $k \le \lfloor \frac{n}{2} \rfloor$,

$$\mathbf{Var}(X_{(k)}) \le k\mathbf{E}\left[\left(X_{(k)} - X_{(k+1)}\right)^2\right] ,$$

et que pour tout $k > \lfloor \frac{n}{2} \rfloor$,

$$\mathbf{Var}(X_{(k)}) \le (n-k+1)\mathbf{E}\left[\left(X_{(k)} - X_{(k-1)}\right)^2\right] .$$

(2) Pour $X_1, \dots, X_n$ i.i.d. de loi uniforme sur $[0,1]$, comparer la vraie variance et la borne donnée par l'inégalité d'Efron–Stein (ordres de grandeur en fonction de n et k). On pourra utiliser le fait que si $(X_1, \dots, X_n) \sim \mathrm{Unif}([0,1])^{\otimes n}$, alors

$$(X_{(1)}, X_{(2)} - X_{(1)}, \dots, X_{(n)} - X_{(n-1)}, 1 - X_{(n)}) \sim \mathrm{Dir}(1, \dots, 1) ,$$

la loi de Dirichlet à $n+1$ coordonnées de paramètre $(1, \dots, 1)$.

Pour la concentration des statistiques d'ordre, voir Boucheron et Thomas, *Concentration Inequalities for order statistics*, Electronic Communications in Probability (2012).

2. Une inégalité isopérimétrique
Pour $A \subset \{0,1\}^n$, on note $\mathbf{I}(A) = \mathbf{I}(\mathbb{1}_A)$, l'influence totale de $\mathbb{1}_A$ sous la loi $\mu = \mathcal{B}(1/2)^{\otimes n}$. Utiliser la sous-additivité de l'entropie pour montrer que

$$\mathbf{I}(A) \ge 2\mu(A) \log_2 \frac{1}{\mu(A)} ,$$

et montrer que l'égalité est atteinte pour les sous-cubes $A = \{x \in \{0,1\}^n,\ x_1 = \cdots = x_k = 1\}$.

3. Constante de Sobolev logarithmique pour la Bernoulli
L'inégalité de Sobolev logarithmique sur le cube $\{0,1\}^n$ énonce que pour toute fonction $f : \{0,1\}^n \to \mathbb{R}$ et pour tout $p \in]0,1[$, si $(X_1, \dots, X_n) \sim \mathcal{B}(p)^{\otimes n}$ et $Z = f(X_1, \dots, X_n)$, alors

$$\mathbf{Ent}(Z^2) \le c(p)p(1-p) \sum_{i=1}^n \mathbf{E}\left[\left(Z - \bar{Z}^{(i)}\right)^2\right] ,$$

où $\bar{Z}^{(i)} = f(X_1, \dots, X_{i-1}, 1 - X_i, X_{i+1}, \dots, X_n)$ et

$$c(p) = \begin{cases} 2 & \text{si } p = \frac{1}{2}, \\ \frac{1}{1-2p} \log\left(\frac{1-p}{p}\right) & \text{sinon.} \end{cases}$$

Montrer que pour tout $p \ne \frac{1}{2}$, cette borne est atteinte par la fonction f donnée par $f(x) = \left(\frac{1-p}{p}\right)^{\sum x_i}$.

4. Inégalité de Sobolev logarithmique et concentration de la mesure
Soit $(X_1, \dots, X_n) \sim \mathcal{B}(1/2)^{\otimes n}$ et $Z = f(X_1, \dots, X_n)$ avec $f : \{0,1\}^n \to \mathbb{R}$. On suppose qu'il existe $v > 0$ tel que pour tout $x \in \{0,1\}^n$,

$$\frac{1}{4} \sum_{i=1}^n \left(f(x) - f(\bar{x}^{(i)})\right)^2 \le v .$$

où $\bar{x}^{(i)} = (x_1, \dots, x_{i-1}, 1 - x_i, x_{i+1}, \dots, x_n)$. L'objectif de cet exercice est de montrer qu'alors, pour tout $t \ge 0$,

$$\mathbf{P}\left(Z - \mathbf{E}Z \ge t\right) \le \exp\left(-\frac{t^2}{2v}\right) .$$

(1) Montrer que pour tous $x, y \in \mathbb{R}$,

$$\left(e^{y/2} - e^{x/2}\right)^2 \le \frac{(y-x)^2}{8}(e^x + e^y)\,.$$

(2) À l'aide de l'inégalité de Sobolev logarithmique sur le cube, montrer que pour tout $\lambda \ge 0$,

$$\mathbf{Ent}\left(e^{\lambda Z}\right) \le \frac{\lambda^2 v}{2}\mathbf{E}\left[e^{\lambda Z}\right]\,.$$

(3) En posant $\psi(\lambda) = \log \mathbf{E} e^{\lambda(Z - \mathbf{E}Z)}$, remarquer que l'inégalité ci-dessus équivaut à $\lambda\psi'(\lambda) - \psi(\lambda) \le \frac{\lambda^2 v}{2}$ et montrer qu'alors $\psi(\lambda) \le \frac{\lambda^2 v}{2}$.

(4) Conclure en utilisant l'inégalité de Markov appliquée à la variable $e^{\lambda Z}$ et en optimisant en $\lambda \ge 0$.