Examen du 28 mars 2022

Salle Nœther, 15h15 – 16h45
Documents autorisés mais Internet interdit
Justifier les réponses

Exercice 1 (Deux inégalités de Poincaré)**.**
Soit U une variable aléatoire de loi uniforme sur $[0,1]$.

1. Montrer que U a même loi que $\Phi(X)$ avec $X \sim \mathcal{N}(0,1)$ et $\Phi(x) := \mathbb{P}(X \leq x)$, et en déduire que
$$\mathrm{Var}(f(U)) \leq \frac{1}{2\pi}\mathbb{E}(f'(U)^2)$$
pour toute fonction $f : \mathbb{R} \to \mathbb{R}$ de classe $\mathcal{C}^1$ bornée et à dérivée bornée.
2. Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction de classe $\mathcal{C}^1$, périodique de période 1. Considérons sa série de Fourier
$$S_n(f)(x) := \mathbb{E}(f(U)) + \sum_{k=1}^{n}(a_k \sin(2\pi kx) + b_k \cos(2\pi kx)).$$
Montrer que si $\lim_{n\to\infty} S_n(f) = f$ et $\lim_{n\to\infty} S_n(f)' = f'$ dans $L^2([0,1])$ alors
$$\mathrm{Var}(f(U)) \leq \frac{1}{4\pi^2}\mathbb{E}(f'(U)^2),$$
et préciser les cas d'égalité.
3. Montrer que pour toute fonction $f : \mathbb{R} \to \mathbb{R}$ de classe $\mathcal{C}^1$,
$$\mathrm{Var}(f(U)) \leq \frac{1}{\pi^2}\mathbb{E}(f'(U)^2),$$
et que la constante $\frac{1}{\pi^2}$ est optimale. Comparer avec l'inégalité de la première question.

Élements de solution - Exercice 1.

1. Comme $\Phi : \mathbb{R} \to]0,1[$ est continue et strictement croissante, elle est inversible. Pour tout $u \in [0,1]$, $\mathbb{P}(\Phi(X) \leq u) = \mathbb{P}(X \leq \Phi^{-1}(u)) = \Phi(\Phi^{-1}(u)) = u$, donc $\Phi(X)$ suit la loi uniforme sur $[0,1]$. L'inégalité de Poincaré optimale pour la loi gaussienne $\mathcal{N}(0,1)$ et la fonction bornée et à dérivée bornée $f \circ \Phi$ s'écrit
$$\mathrm{Var}(f(\Phi(X))) \leq \mathbb{E}(f'(\Phi(X))^2\Phi'(X)^2).$$
Maintenant comme Φ' prend ses valeurs dans $[0, \frac{1}{\sqrt{2\pi}}]$ et comme $\Phi(X)$ a même loi que U, on obtient
$$\mathrm{Var}(f(U)) \leq \frac{1}{2\pi}\mathbb{E}(f'(U)^2).$$
2. L'inégalité à démontrer est invariante par translation, et on peut supposer que $\mathbb{E}(f(U)) = 0$. On a alors
$$\mathrm{Var}(f(U)) = \mathbb{E}(f(U)^2) = \frac{1}{2}\sum_{k=1}^{\infty}(a_k^2 + b_k^2)$$
car $\int_0^1 \sin(2\pi kx)\cos(2\pi k'x)\mathrm{d}x = 0$ et $\int_0^1 \sin(2\pi kx)^2\mathrm{d}x = \int_0^1 \cos(2\pi kx)^2\mathrm{d}x = \frac{1}{2}$ si $k, k' \geq 1$ tandis que $\int_0^1 \sin(2\pi kx)\sin(2\pi k'x)\mathrm{d}x = \int_0^1 \cos(2\pi kx)\cos(2\pi k'x)\mathrm{d}x = 0$ si $k, k' \geq 1$ avec $k \neq k'$. D'autre part
$$\mathbb{E}(f'(U)^2) = \frac{1}{2}\sum_{k=1}^{\infty}(2\pi k)^2(a_k^2 + b_k^2) \geq \frac{(2\pi)^2}{2}\sum_{k=1}^{\infty}(a_k^2 + b_k^2).$$
L'égalité est atteinte ssi $a_k = b_k = 0$ pour $k \geq 2$: f est combinaison linéaire de 1, $\sin(2\pi\cdot)$, et $\cos(2\pi\cdot)$.
3. Considérons la restriction de f à $[0,1]$, symétrisons là en une fonction paire sur $[-1,1]$, puis périodisons là (période 2). Cette dernière fonction g, malgré ses points irréguliers, vérifie une inégalité de Poincaré pour la loi uniforme sur $[-1,1]$, de constante $\frac{(1-(-1))^2}{4\pi^2} = \frac{1}{\pi^2}$, déduite de la question précédente par régularisation, translation, et dilatation. D'autre part on peut vérifier que $\mathrm{Var}(g(V)) = \mathrm{Var}(f(U))$ tandis que $\mathbb{E}(g'(V)^2) = \mathbb{E}(f'(U)^2)$ où V suit la loi uniforme sur $[-1,1]$. Enfin l'égalité est atteinte pour $f(x) = \cos(2\pi x)$. Elle est strictement meilleure que l'inégalité de la première question.

Exercice 2 (Une inégalité de concentration)**.**
Soit Y une variable aléatoire réelle intégrable définie sur $(\Omega, \mathcal{F}, \mathbb{P})$. Soit $(\mathcal{F}_k)_{0\leq k\leq n}$ une filtration, c'est-à-dire une famille de sous-tribus de $\mathcal{F}$ telles que $\{\varnothing, \Omega\} = \mathcal{F}_0 \subset \cdots \subset \mathcal{F}_n = \mathcal{F}$. Pour tout $1 \leq k \leq n$ on pose

$$d_k := \mathbb{E}(Y \mid \mathcal{F}_k) - \mathbb{E}(Y \mid \mathcal{F}_{k-1}) \quad \text{et} \quad \mathrm{osc}(d_k) := \sup d_k - \inf d_k.$$

Le but de l'exercice est d'établir l'inégalité de concentration suivante : pour tout $r \geq 0$,

$$\mathbb{P}(|Y - \mathbb{E}Y| \geq r) \leq 2\exp\left(-\frac{2r^2}{\mathrm{osc}(d_1)^2 + \cdots + \mathrm{osc}(d_n)^2}\right).$$

1. Soit U une variable aléatoire à valeur dans l'intervalle $[a, b]$ avec $-\infty < a < b < +\infty$, telle que $\mathbb{E}U = 0$. Posons $p := \frac{-a}{b-a}$ et $f(u) := -pu + \log(1 - p + p\mathrm{e}^u)$. Montrer que $p \in [0, 1]$ et que pour tout $t \geq 0$,
$$\mathbb{E}(\mathrm{e}^{tU}) \leq \mathrm{e}^{f(t(b-a))}.$$
Indication : établir que pour tout $t \geq 0$ et tout $u \in [a, b]$ on a $\mathrm{e}^{tu} \leq \frac{u-a}{b-a}\mathrm{e}^{tb} + \frac{b-u}{b-a}\mathrm{e}^{ta}$.
2. Montrer que $f'' \leq \frac{1}{4}$ et $f(0) = f'(0) = 0$ et en déduire que $\mathbb{E}(\mathrm{e}^{tU}) \leq \mathrm{e}^{\frac{t^2}{8}(b-a)^2}$.
3. En déduire que pour tout $1 \leq k \leq n$,
$$\mathbb{E}(\mathrm{e}^{td_k} \mid \mathcal{F}_{k-1}) \leq \mathrm{e}^{\frac{t^2}{8}\mathrm{osc}(d_k)^2}.$$
4. En déduire que
$$\mathbb{E}(\mathrm{e}^{t(Y-\mathbb{E}Y)}) \leq \mathrm{e}^{\frac{t^2}{8}(\mathrm{osc}(d_1)^2+\cdots+\mathrm{osc}(d_n)^2)}.$$
5. En déduire l'inégalité de concentration annoncée.
6. En déduire que si $X_1, \ldots, X_n$ sont des v.a. indépendantes et centrées dans $[-1, 1]$ alors pour tout $r \geq 0$,
$$\mathbb{P}\left(\left|\frac{X_1 + \cdots + X_n}{\sqrt{n}}\right| \geq r\right) \leq 2\mathrm{e}^{-\frac{r^2}{2}}.$$

Élements de solution - Exercice 2. Inégalité de concentration de Azuma-Hoeffding et martingale de Doob.

1. Comme $\mathbb{E}U = 0$ on a forcément $a \leq 0$ et $b \geq 0$ ce qui donne successivement $p \geq 0$ puis $p \leq 1$. Ensuite, pour tout $t \geq 0$ et tout $u \in [a, b]$, la convexité de $x \mapsto \mathrm{e}^{tx}$ donne $\mathrm{e}^{tu} \leq \frac{u-a}{b-a}\mathrm{e}^{tb} + \frac{b-u}{b-a}\mathrm{e}^{ta}$, d'où
$$\mathbb{E}(\mathrm{e}^{tU}) \leq \mathbb{E}\left(\frac{U-a}{b-a}\right)\mathrm{e}^{tb} + \mathbb{E}\left(\frac{b-U}{b-a}\right)\mathrm{e}^{ta} = p\mathrm{e}^{tb} + \mathrm{e}^{ta}(1-p) = \mathrm{e}^{\log(p\mathrm{e}^{tb}+\mathrm{e}^{ta}(1-p))} = \mathrm{e}^{at+\log(p\mathrm{e}^{t(b-a)}+(1-p))}.$$
2. On a $f(0) = 0$, $f'(u) = -p + \frac{p\mathrm{e}^u}{1-p+p\mathrm{e}^u}$ donc $f'(0) = 0$, et $f''(u) = (1-p)p\frac{\mathrm{e}^u}{(1-p+p\mathrm{e}^u)^2} \leq p(1-p) \leq \frac{1}{4}$ pour $u \in [a, b]$, et il en découle par une formule de Taylor que $f(u) \leq f(0) + f'(0)u + \frac{u^2}{8} = \frac{u^2}{8}$. Note : $0 \in [a, b]$.
3. Pour tout $1 \leq k \leq n$, la variable aléatoire $d_k = \mathbb{E}(Y \mid \mathcal{F}_k) - \mathbb{E}(Y \mid \mathcal{F}_{k-1})$ vérifie $\mathbb{E}(d_k \mid \mathcal{F}_{k-1}) = 0$. Si $\mathrm{osc}(d_k) = \infty$ alors l'inégalité visée est triviale, tandis que si $\mathrm{osc}(d_k) < \infty$ alors d_k prend ses valeurs dans un intervalle fini de longueur $\mathrm{osc}(d_k)$ et l'inégalité visée découle de la question précédente avec $U = d_k$.
4. On utilise une somme télescopique et l'emboîtement des espérances conditionnelles :
$$\mathbb{E}(\mathrm{e}^{t(Y-\mathbb{E}Y)}) = \mathbb{E}(\mathrm{e}^{t(d_n+\cdots+d_1)}) = \mathbb{E}(\mathrm{e}^{t(d_{n-1}+\cdots+d_1)}\mathbb{E}(\mathrm{e}^{td_n} \mid \mathcal{F}_{n-1})) \leq \cdots \leq \mathrm{e}^{\frac{t^2}{8}(\mathrm{osc}(d_1)^2+\cdots+\mathrm{osc}(d_n)^2)}.$$
5. Par l'inégalité de Markov, on a, pour tout $r \geq 0$, en notant $c := \mathrm{osc}(d_1)^2 + \cdots + \mathrm{osc}(d_n)^2$,
$$\mathbb{P}(Y - \mathbb{E}Y \geq r) \overset{t\geq 0}{=} \mathbb{P}(\mathrm{e}^{t(Y-\mathbb{E}Y)} \geq \mathrm{e}^{tr}) \leq \mathrm{e}^{-tr+c\frac{t^2 r^2}{8}} \leq \mathrm{e}^{\inf_{t>0}(-tr+c\frac{t^2}{8})} \overset{t=\frac{4r}{c}}{=} \mathrm{e}^{-2\frac{r^2}{c}}.$$
Utilisée pour $-Y$ cela donne $\mathbb{P}(Y - \mathbb{E}Y \leq -r) \leq \mathrm{e}^{-2\frac{r^2}{c}}$, d'où $\mathbb{P}(|Y - \mathbb{E}Y| \geq r) \leq 2\mathrm{e}^{-2\frac{r^2}{c}}$.
6. On pose $\mathcal{F}_k = \sigma(X_1, \ldots, X_k)$ pour $1 \leq k \leq n$, et $Y = \frac{X_1+\cdots+X_n}{\sqrt{n}}$. On a alors $d_k = \frac{X_k}{\sqrt{n}}$ et $\mathrm{osc}(d_k) \leq \frac{1}{\sqrt{n}}$.

– oOo –