Université Paris-Dauphine
J. Féjoz

Examen de janvier 2015

Deux heures — Sans document ni appareil électronique.
Chaque question numérotée sera notée sur environ deux points.
Les réponses doivent être concises.

Question indépendante A. — Montrer que $f : x \mapsto \int_0^\pi e^{x \cos t} \cos t \, dt$ est continue sur $\mathbb{R}$.

Question indépendante B. — Soit $(X_1, X_2) : \Omega \to \mathbb{R}^2$ gaussien. Montrer que, si ses composantes sont de covariance nulle, elles sont indépendantes.

Soient $(\Omega, \mathcal{A}, P)$ un espace probabilisé et $T : \Omega \to \Omega$ une bijection mesurable d'inverse mesurable. Un événement $A \in \mathcal{A}$ est dit *presque invariant* si les fonctions indicatrices $\mathbb{1}_A$ et $\mathbb{1}_{T^{-1}(A)}$ sont égales presque sûrement.

1. — Montrer que l'ensemble $\mathcal{I}$ des événements presque invariants est une tribu (appelée la *tribu presque invariante* de T); pour toute A presque invariante, on pourra introduire, si nécessaire, l'ensemble négligeable $N(A)$ en dehors duquel $\mathbb{1}_A = \mathbb{1}_{T^{-1}(A)}$.

2. — Donner un exemple où Ω est un ensemble à deux éléments muni de la probabilité uniforme, T une bijection, et $\mathcal{I} = \{\varnothing, \Omega\}$.

[N.B.: L'application T est bien définie comme une application de Ω dans lui-même. Elle induit une transformation de $\mathcal{A}$ (par $A \mapsto T^{-1}(A)$), une transformation des fonctions f (par $f \mapsto f \circ T$), ainsi qu'une transformation des probabilités (par $P \mapsto T(P) = P \circ T^{-1}$. Ces transformations induites sont encore notées avec la lettre T, mais quand, dans la question précédente, il est demandé de donner un exemple de T vérifiant certaines conditions, il faut bien définir l'application T de Ω dans lui-même. Par ailleurs, si Ω possède deux éléments, on peut bien sûr donner à ces éléments n'importe quel nom, par exemple 0 et 1 (des esprits libres pourront les appeller $\varnothing$ et $\{36\}$, mais le raisonnement n'en sera pas éclairci).]

Soit f une variable aléatoire réelle sur $(\Omega, \mathcal{A})$.

3. — Montrer que, si f est presque invariante par T, c'est-à-dire si $f \circ T = f$ p.s., f est $\mathcal{I}$-mesurable.

4. — Montrer la réciproque; on pourra commencer par la cas où f est une fonction indicatrice.

[N.B.: Les deux questions précédentes n'impliquent pas que l'on suppose, dans la suite, f invariante par T!]

On suppose dorénavant que $f \in L^2(\mathcal{A}) = L^2(\Omega, \mathcal{A}, P)$. On rappelle que $L^2(\mathcal{A})$ possède un produit scalaire hilbertien réel $\langle g|h\rangle = \int g\,h\,dP$, que $L^2(\mathcal{I}) = L^2(\Omega, \mathcal{I}, P)$ s'identifie à un sous-espace fermé de $L^2(\mathcal{A})$, et que l'espérance conditionnelle $E(f|\mathcal{I})$ de f sachant $\mathcal{I}$ est la projection orthogonale de f sur $L^2(\mathcal{I})$; en conséquence, $E(f|\mathcal{I})$ est caractérisée, parmi les éléments de $L^2(\mathcal{I})$, par le fait que $\langle E(f|\mathcal{I})|g\rangle = \langle f|g\rangle$ pour tout $g \in L^2(\mathcal{I})$.

5. — Montrer: $E(E(f|\mathcal{I})) = E(f)$. Cela dépend-il de la sous-tribu $\mathcal{I}$?

Dorénavant, on suppose que la probabilité P est invariante (par T), c'est-à-dire $T(P) = P$, et l'objectif est de montrer que

$$S_N(f) = \frac{f + f\circ T + \cdots + f\circ T^{N-1}}{N} \xrightarrow{L^2}_{N\to\infty} E(f|\mathcal{I}), \tag{1}$$

où $T^0 = id$ et $T^n = T\circ\cdots\circ T$ est la n-ième composée itérée de T.

6. — Montrer que $E(f\circ T) = E(f)$ et $E(f\circ T|\mathcal{I}) = E(f|\mathcal{I})$.

7. — Interpréter (1) lorsque f est la fonction indicatrice d'un événement $A \in \mathcal{A}$ tel que $\sigma(\{A\})$ et $\mathcal{I}$ sont indépendantes, en termes de fréquence limite de passage dans A pour les suites $(T^n(x))_{n\geqslant 0}$, $x \in \Omega$.

8. — Montrer que, pour tout N, $S_N(E(f|\mathcal{I})) = E(f|\mathcal{I})$.

Notons $f_0 = f - E(f|\mathcal{I})$. Il s'agit donc de montrer que la suite $(S_N(f_0))$ converge vers 0 dans L^2. On pose

$$F_N = \frac{1}{N^2}\sum_{0\leqslant n,m\leqslant N-1} f_0\circ T^{n-m} \quad (N \geqslant 1).$$

9. — Montrer que $\|S_N(f_0)\|_{L^2}^2 = \langle F_N|f_0\rangle$ et que $(\|F_N\|_{L^2})_N$ est bornée.

10. — Conclure dans le cas où (F_N) tend vers 0 dans L^2.

Supposons par l'absurde que (F_N) ne tende pas vers 0. On admettra (c'est une conséquence du théorème de Banach-Alaoglu) qu'il existe un suite extraite (F_{N_j}) et une fonction $F_\infty \in L^2(\mathcal{A})$ non nulle telles que, pour toute $g \in L^2(\mathcal{A})$, $\langle F_{N_j}|g\rangle \longrightarrow_{j\to\infty} \langle F_\infty|g\rangle$.

11. — Montrer, par un argument de série télescopique, qu'il existe $C > 0$ tel que, quand N tend vers l'infini,

$$\|F_N\circ T - F_N\|_{L^2} \leqslant \frac{C}{N}.$$

On admettra que donc $F_\infty \circ T = F_\infty$ et $E(F_\infty|\mathcal{I}) = F_\infty$.

12. — ★ Montrer que $E(F_N|\mathcal{I}) = 0$ pour tout N, en déduire que $F_\infty = 0$ et conclure.

13. — Trouver un exemple dans lequel les variables $f \circ T^n$ $(n \geqslant 0)$ ne sont pas indépendantes; on pourra utiliser la loi des grands nombres.

Solution. —

Question indépendante A. — Soit (x_n) une suite réelle convergeant vers $x \in \mathbb{R}$. Notons

$$f_n(t) = e^{x_n \cos t} \cos t.$$

Pour tout t, comme la fonction $x \mapsto e^{x \cos t} \cos t$ est continue, $(f_n(t))$ converge vers $e^{x \cos t} \cos t$. De plus, (x_n) est bornée, disons $|x_n| \leqslant C$, donc

$$|f_n| \leqslant e^C \in \mathcal{L}^1([0, \pi]).$$

Par convergence dominée,

$$f(x_n) = \int_0^\pi e^{x_n \cos t} \cos t \, dt \to \int_0^\pi e^{x \cos t} \cos t \, dt = B(x).$$

Donc f est continue sur $\mathbb{R}$.

Question indépendante B. — La matrice de covariance K de $X = (X_1, X_2)$ est diagonale:

$$K = \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix}.$$

Donc la fonction caractéristique de X vérifie

$$\Phi_X(\xi) = \prod_j \Phi_{X_j}(\xi_j).$$

Donc les X_j sont indépendantes.

1. — Soit A presque invariante. On a

$$\mathbb{1}_A = \mathbb{1}_{T^{-1}(A)} \quad \text{p.s.},$$

donc

$$1 - \mathbb{1}_A = 1 - \mathbb{1}_{T^{-1}(A)} \quad \text{p.s.},$$

soit

$$\mathbb{1}_{A^c} = \mathbb{1}_{T^{-1}(A)^c} = \mathbb{1}_{T^{-1}(A^c)} \quad \text{p.s.},$$

donc A^c est presque invariante.

Soient $A_1, A_2, ...$ une famille dénombrable de parties presque invariantes. Pour tout i, il existe une partie $N_i = N(A_i)$ de probabilité nulle telle que

$$A_i \cap N_i^c = T^{-1}(A_i) \cap N_i^c.$$

La partie $N = \cup N_i$ est aussi de probabilité nulle, et

$$(\cup_i A_i) \cap N^c = \cup_i (A_i \cap N^c) = \cup_i (T^{-1}(A_i) \cap N^c) = T^{-1}(\cup_i A_i) \cap N^c,$$

donc

$$\mathbb{1}_{\cup_i A_i} = \mathbb{1}_{T^{-1}(\cup_i A_i)} \quad \text{p.s.},$$

donc $\cup_i A_i$ est presque invariante. Donc $\mathcal{I}$ est une tribu.

2. — Notons $\Omega = \{x_1, ..., x_n\}$. Si T est une permutation circulaire, par exemple

$$\begin{pmatrix} 1 & 2 & \cdots & n-1 & n \\ 2 & 3 & \cdots & n & 1 \end{pmatrix}$$

ses seules parties invariantes sont $\varnothing$ et Ω. Comme P est uniforme, les parties presque invariantes de T sont en fait invariantes. Alors $\mathcal{I} = \{\varnothing, \Omega\}$.

3. — Supposons que $f \circ T = f$ p.s. et soit $B \in \mathcal{B}(\mathbb{R})$. Il s'agit de montrer que $f^{-1}(B)$ est presque invariante. Soit N une partie de probabilité nulle en dehors de laquelle $f \circ T = f$. En restriction à N^c,

$$\mathbb{1}_{f^{-1}(B)} = \mathbb{1}_{(f \circ T)^{-1}(B)} = \mathbb{1}_{T^{-1}(f^{-1}(B))},$$

donc $f^{-1}(B)$ est presque invariante. Donc f est $\mathcal{I}$-mesurable.

4. — Réciproquement, supposons que f soit $\mathcal{I}$-invariante. Dans le cas où f est une fonction étagée positive, il existe des parties $A_1, ..., A_n \in \mathcal{I}$ et des réels $a_1, ..., a_n > 0$ tels que

$$f = \sum_i a_i \, \mathbb{1}_{A_i}.$$

Alors

$$f \circ T = \sum_i a_i \, \mathbb{1}_{T^{-1}(A_i)} = \sum_i a_i \, \mathbb{1}_{A_i} = f \quad \text{p.s..}$$

Si f est positive, il existe une suite croissante (f_n) de fonctions étagés positives $\mathcal{I}$-mesurables, qui converge vers f. D'après le cas précédent, pour tout i,

$$f_i \circ T = f_i \quad \text{p.s..}$$

et l'égalité est réalisée en dehors d'une partie négligeable N_i. En dehors de la partie elle-aussi négligeable $N = \cup_i N_i$, $f_i = f_i \circ T$ pour tout i, donc, en passant à la limite quand i tend vers l'infini, $f = f \circ T$ en dehors de N. Donc $f = f \circ T$ p.s.

Si enfin f est de signe quelconque, ses parties positives et négatives sont presque invariantes d'après le cas précédent. Donc f aussi, puisque l'union de deux parties négligeables est négligeable.

Rappel. — $L^2(\mathcal{I})$ est un espace vectoriel au même titre que $L^2(\mathcal{A})$. Comme $\mathcal{I}$ est une sous-tribu de $\mathcal{A}$, une fonction $\mathcal{I}$-mesurable est automatiquement $\mathcal{A}$-mesurable. Toute classe de fonctions $\mathcal{I}$-mesurables (pour la relation d'égalité presque sûre) est donc incluse dans une unique classe de fonctions $\mathcal{A}$-mesurables. Donc il existe une application $L^2(\mathcal{I}) \to L^2(\mathcal{A})$, et cette application est linéaire, injective et isométrique. De plus, si une suite (f_n) de $L^2(\mathcal{I})$ converge vers f dans $L^2(\mathcal{A})$, comme la norme $L^2(\mathcal{A})$ d'un élément de $L^2(\mathcal{I})$ est aussi sa norme $L^2(\mathcal{I})$, (f_n) converge dans $L^2(\mathcal{I})$, et sa limite doit être la même, f. Donc $f \in L^2(\mathcal{I})$. Donc $L^2(\mathcal{I})$ est un sous-espace fermé. De plus, comme $E(f|\mathcal{I})$ est la projection orthogonale de f

sur $L^2(\mathcal{I})$, $E(f|\mathcal{I})$ est caractérisée, parmi les éléments de $L^2(\mathcal{I})$, par le fait que, pour tout $g \in L^2(\mathcal{I})$, $\langle f - E(f|\mathcal{I})|g\rangle = 0$, c'est-à-dire

$$\langle E(f|\mathcal{I})|g\rangle = \langle f|g\rangle. \tag{2}$$

5. — En prenant $g = 1 \in L^2(\mathcal{I})$ dans (2), on obtient $E(E(f|\mathcal{I})) = E(f)$. Cette identité reste vérifiée quand on remplace $\mathcal{I}$ par n'importe quelle sous-tribu $\mathcal{B}$ de $(\Omega, \mathcal{A})$, parce que la démonstration n'utilise en rien le fait que $\mathcal{I}$ soit la tribu presque invariante d'une transformation T.

6. — Comme P est T-invariante,

$$\begin{aligned} E(f \circ T) &= \int f \circ T\, dP = \int f\, d(T(P)) &&\text{(formule de transfert)} \\ &= \int f\, dP = E(f) &&\text{(invariance de } P\text{).} \end{aligned}$$

Par ailleurs, en remplaçant f par $f \circ T - f$ dans (2), on obtient

$$\langle E(f \circ T - f|\mathcal{I})|g\rangle = \langle f \circ T - f|g\rangle.$$

Or, le membre de droite vaut

$$\langle f \circ T - f|g\rangle = \langle f \circ T|g\rangle - \langle f|g\rangle = 0$$

(puisque $g \in L^2(\mathcal{I})$, g est invariante presque partout, donc l'égalité précédente résulte du changement de variable $x = T^{-1}(y)$, comme pour montrer la première identité de cette question). Donc

$$\langle E(f \circ T - f|\mathcal{I})|g\rangle = 0 \quad (\forall g \in L^2(\mathcal{I})),$$

donc, puisque cette dernière égalité caractérise la fonction nulle parmi les fonctions de $L^2(\mathcal{I})$,

$$E(f \circ T|\mathcal{I}) - E(f|\mathcal{I}) = 0,$$

soit

$$E(f \circ T|\mathcal{I}) = E(f|\mathcal{I}).$$

7. — Si $f = \mathbb{1}_A$, $A \in \mathcal{A}$, $S_N(f)(x)$ est la probabilité que $T^n(x)$ soit dans A, quand n varie entre 0 et $N-1$. Autrement dit, c'est la fréquence de passage dans A de la suite des composées itérées de x par T.

Si $\mathcal{I}$ est indépendante de

$$\sigma(\{A\}) = \{\varnothing, A, A^c, \Omega\} = \sigma(f),$$

on a $E(f|\mathcal{I}) = E(f) = E(\mathbb{1}_A) = P(A)$.

Sous ces hypothèses, le théorème ergodique affirme donc que la fréquence limite de passage dans A de l'orbite de x converge vers la probabilité de A dans L^2.

8. — Par définition, $E(f|\mathcal{I})$ possède un représentant $\mathcal{I}$-mesurable. Donc, d'après la question 4, $E(f|\mathcal{I})$ est preque invariante: $E(f|\mathcal{I}) \circ T = E(f|\mathcal{I})$ p.s. Donc, par récurrence, $E(f|\mathcal{I}) \circ T^n = E(f|\mathcal{I})$ p.s. pour tout $n \in \mathbb{N}$. Donc, pour tout $N \in \mathbb{N}$,

$$S_N(E(f|\mathcal{I})) = \frac{1}{N} \sum_{0 \leqslant n \leqslant N-1} E(f|\mathcal{I}) \circ T^n = \frac{1}{N} \sum_{0 \leqslant n \leqslant N-1} E(f|\mathcal{I}) = E(f|\mathcal{I}).$$

9. — On s'intéresse dorénavant à $f_0 = f - E(f|\mathcal{I})$, qui peut s'interpréter comme la "partie aléatoire" de f. On a

$$\begin{aligned}
\|S_N(f_0)\|_{L^2}^2 &= \frac{1}{N^2} \int \Bigg(\sum_{0 \leqslant n \leqslant N-1} f_0 \circ T^n \Bigg)^2 dP \\
&= \frac{1}{N^2} \int \sum_{0 \leqslant n,m \leqslant N-1} (f_0 \circ T^n \, f_0 \circ T^m)\, dP \\
&= \frac{1}{N^2} \int \sum_{0 \leqslant n,m \leqslant N-1} \big(f_0 \circ T^{n-1} \, f_0 \circ T^{m-1}\big)\, dP \\
&\quad \text{(formule de transfert, avec } T(P) = P) \\
&= \frac{1}{N^2} \int \sum_{0 \leqslant n,m \leqslant N-1} \big(f_0 \circ T^{n-m}\big)\, f_0 \, dP \qquad \text{(récurrence)} \\
&= \langle F_N | f_0 \rangle.
\end{aligned}$$

De plus, d'après l'inégalité triangulaire,

$$\|F_N\|_{L^2} \leqslant \frac{1}{N^2} \sum_{0 \leqslant n,m \leqslant N-1} \|f_0 \circ T^{n-m}\|_{L^2}.$$

Or, la formule de transfert et l'invariance de P montrent (de même que dans la question 5) que, si $k \geqslant 1$,

$$\|f_0 \circ T^k\|_{L^2}^2 = \int (f_0 \circ T^k)^2 = \int (f_0 \circ T^{k-1})^2 = \cdots = \int f_0^2 = \|f_0\|_{L^2}^2.$$

Si $k \leqslant -1$, la propriété précédente appliquée avec T^{-1} au lieu de T montre l'identité analogue. Donc $\|f \circ T^k\|_{L^2} = \|f_0\|_{L^2}$ pour tout $k \in \mathbb{Z}$. Donc

$$\|F_N\|_{L^2} \leqslant \|f_0\|_{L^2},$$

et en particulier la suite (F_N) est bornée dans L^2.

10. — Si (F_N) tend vers 0, d'après l'inégalité de Cauchy-Schwarz,

$$\|S_N(f_0)\|_{L^2}^2 = (F_N | f_0)_{L^2} \leqslant \|F_N\|_{L^2} \, \|f_0\|_{L^2} \to 0,$$

de quoi la conclusion voulue découle.

11. — On a

$$\begin{aligned}F_N \circ T - F_N &= \frac{1}{N^2} \sum_{0 \leqslant n,m \leqslant N-1} \left(f_0 \circ T^{n-m+1} - f_0 \circ T^{n-m}\right) \\ &= \frac{1}{N^2} \sum_{0 \leqslant m \leqslant N-1} \left(\sum_{1 \leqslant n \leqslant N} f_0 \circ T^{n-m} - \sum_{0 \leqslant n \leqslant N-1} f_0 \circ T^{n-m} \right) \\ &= \frac{1}{N^2} \sum_{0 \leqslant m \leqslant N-1} \left(f_0 \circ T^{N-m} - f_0 \circ T^{-m}\right),\end{aligned}$$

donc

$$\|F_N \circ T - F_N\|_{L^2} \leqslant \frac{2\|f_0\|_{L^2}}{N} \to 0.$$

On admet ensuite pouvoir passer à la limite quand N tend vers l'infini, et en déduire que $F_\infty \circ T = F_\infty$. Donc, F_∞ est $\mathcal{I}$-mesurable (question 3), $F_\infty \in L^2(\mathcal{I})$ et $E(F_\infty|\mathcal{I}) = F_\infty$.

12. — Par ailleurs, on a $E(f_0|\mathcal{I}) = 0$ donc, en utilisant la linéarité et la question 6, on voit que $E(F_N|\mathcal{I}) = 0$.

Donc, pour toute $g \in L^2(\mathcal{A})$,

$$0 \equiv E(E(F_{N_j}|\mathcal{I})\, g) = E(E(F_{N_j}|\mathcal{I})\, E(g|\mathcal{I})) = E(F_{N_j}\, E(g|\mathcal{I}))$$

et, d'après ce qui précède, cette dernière quantité (nulle) tend, quand j tend vers l'infini, vers

$$E(F_\infty\, E(g|\mathcal{I})) = E(E(F_\infty|\mathcal{I})\, g).$$

Donc (d'après le théorème de représentation de Riesz) $E(F_\infty|\mathcal{I}) = 0$. D'après la question précédente, ceci montre que $F_\infty = 0$, ce qui est contraire à l'hypothèse. Donc le cas de la question 10 est le seul possible, et le résultat voulu est démontré.

13. — Considérons l'ensemble Ω muni de la probabilité uniforme, ainsi qu'une permutation qui ne soit pas circulaire (contrairement à la permutation de la question 2), par exemple,

$$T = \begin{pmatrix} 123 \\ 213 \end{pmatrix} : \{1,2,3\} \circlearrowleft$$

Alors la tribu (presque) invariante est $\mathcal{I} = \{\varnothing, \{1,2\}, \{3\}, \Omega\}$. Pour toute variable aléatoire $f : \Omega \to \mathbb{R}$,

$$E(f|\mathcal{I}) = \frac{f(1)+f(2)}{2} \mathbb{1}_{\{1,2\}} + f(3)\mathbb{1}_{\{3\}} \neq E(f)$$

si $f(3) \neq \frac{f(1)+f(2)}{2}$. Les variables aléatoires $f \circ T^n$ ont même loi. Si elles étaient indépendantes, la loi (faible) des grands nombres dirait que $S_N(f) \to E(f)$ dans L^2, ce qui n'est pas le cas. Donc les $f \circ T^n$ ne sont pas indépendantes.