Examen

L'épreuve dure trois heures.
Les quatre exercices sont indépendants.
Ni documents, ni appareils électroniques.
La note finale sera sur 50 points.

Exercice 1

Barème indicatif : 15 points (3+3+3+3+3)

Toutes les variables aléatoires sont définies sur un espace de probabilités $(\Omega, \mathscr{F}, \mathbf{P})$.

1. Soit $p \in [0,1]$ un réel. Soient $n, m \geqslant 1$ des entiers. Soient A, B des variables aléatoires indépendantes de lois binomiales, respectivement $\mathcal{B}(n,p)$ et $\mathcal{B}(m,p)$.

Calculer $\mathbf{E}[A \,|\, A+B]$.

2. On note $D = \{(x,y) \in \mathbb{R}^2 : 0 \leqslant y \leqslant 1 - |x|\}$. Soit (X,Y) un couple de variables aléatoires réelles dont la loi admet par rapport à la mesure de Lebesgue sur $\mathbb{R}^2$ la densité

$$f : (x,y) \mapsto c\, x^2 y^2\, \mathbf{1}_D(x,y),$$

où c est un réel positif.

Calculer $\mathbf{E}[X \,|\, Y]$ et $\mathbf{E}[Y \,|\, X]$.

3. Soit T une variable aléatoire de loi exponentielle de paramètre 1. Soit $U = T - \lfloor T \rfloor$ la partie fractionnaire de T.

Calculer $\mathbf{E}[T \,|\, U]$.

4 . Soit (F,G,H) un vecteur aléatoire gaussien centré de matrice de covariance

$$\begin{pmatrix} 3 & 2 & 1 \\ 2 & 3 & 2 \\ 1 & 2 & 3 \end{pmatrix}$$

Calculer $\mathbf{E}[F+G \,|\, G+H]$.

5. Soient K, L, M des variables aléatoires indépendantes et identiquement distribuées, qu'on suppose intégrables.

Calculer $\mathbf{E}\big[K \,\big|\, \mathbf{E}[K+L \,|\, K+L+M]\big]$ et $\mathbf{E}\big[K+L+M \,\big|\, \mathbf{E}[K+L \,|\, K]\big]$.

Solution de l'exercice 1

1. La variable aléatoire $A+B$ est de loi binomiale $\mathcal{B}(n+m,p)$. Soit $g : \mathbb{N} \to \mathbb{R}$ une fonction mesurable bornée. On a

$$\mathbf{E}[Ag(A+B)] = \sum_{a=0}^{n} \sum_{b=0}^{m} \binom{n}{a} \binom{m}{b} p^{a+b} a\, g(a+b).$$

Réorganisons la somme selon la valeur de la somme $s = a + b$. Conservons a comme autre variable. Les contraintes sur a sont que a est un entier compris entre 0 et n, et qui doit être plus petit que s et plus grand que $m - s$. Cette dernière condition équivaut à $s - a \leqslant m$, si bien que

$$\mathbf{E}[Ag(A+B)] = \sum_{s=0}^{n+m} p^s g(s) \sum_{\substack{0 \leqslant a \leqslant s \\ a \leqslant n, s-a \leqslant m}} \binom{n}{a} \binom{m}{s-a} a$$

L'entier s étant fixé, calculons la deuxième somme en utilisant l'identité $a\binom{n}{a} = n\binom{n-1}{a-1}$ puis en faisant le changement de variable $c = a - 1$:

$$\begin{aligned}
\sum_{\substack{0 \leqslant a \leqslant s \\ a \leqslant n, s-a \leqslant m}} \binom{n}{a} \binom{m}{s-a} a &= n \sum_{\substack{1 \leqslant a \leqslant s \\ a \leqslant n, s-a \leqslant m}} \binom{n-1}{a-1} \binom{m}{(s-1)-(a-1)} \\
&= n \sum_{\substack{0 \leqslant c \leqslant s-1 \\ c \leqslant n-1, (s-1)-c \leqslant m}} \binom{n-1}{c} \binom{m}{(s-1)-c}
\end{aligned}$$

Dans la dernière somme, on reconnaît le nombre de manière de choisir $s - 1$ éléments parmi $n + m - 1$ dont $n - 1$ sont bleus et m sont rouges, compté selon le nombre c d'éléments bleus. Cette somme vaut donc

$$\binom{m+n-1}{s-1} = \frac{s}{n+m} \binom{n+m}{s},$$

si bien que l'espérance que nous sommes en train de calculer vaut

$$\mathbf{E}[Ag(A+B)] = \frac{n}{n+m} \sum_{s=0}^{n+m} \binom{n+m}{s} p^s s\, g(s) = \frac{n}{n+m} \mathbf{E}[(A+B)g(A+B)].$$

Nous en déduisons que

$$\mathbf{E}[A \,|\, A+B] = \frac{n}{n+m}(A+B).$$

Autre solution. Le couple (A, B) a même loi que le couple $(X_1+\ldots+X_n, X_{n+1}+\ldots+X_{n+m})$, où $X_1, \ldots, X_{n+m}$ sont des variables aléatoires indépendantes identiquement distribuées de loi de Bernoulli de paramètre p. Ainsi, si nous trouvons une fonction $h : \mathbb{R} \to \mathbb{R}$ telle que

$$\mathbf{E}[X_1 + \ldots + X_n \,|\; X_1 + \ldots + X_{n+m}] = h(X_1 + \ldots + X_{n+m}),$$

alors nous pourrons affirmer que

$$\mathbf{E}[A \,|\, A+B] = h(A+B).$$

Or d'une part,

$$\mathbf{E}[X_1 + \ldots + X_n \,|\; X_1 + \ldots + X_{n+m}] = \sum_{k=1}^{n} \mathbf{E}[X_k \,|\; X_1 + \ldots + X_{n+m}].$$

D'autre part, par symétrie, la loi du couple $(X_k, X_1, \ldots, X_{n+m})$ ne dépend pas de l'entier k compris entre 1 et $n + m$. Ainsi, il existe une fonction f ne dépendant pas de k telle que

$$\mathbf{E}[X_k \,|\; X_1 + \ldots + X_{n+m}] = f(X_1 + \ldots + X_{n+m}).$$

En sommant cette dernière identité pour k variant de 1 à $n+m$, on trouve que

$$X_1 + \ldots + X_{n+m} = \mathbf{E}[X_1 + \ldots + X_{n+m} | X_1 + \ldots + X_{n+m}] = \sum_{k=1}^{n+m} f(X_1 + \ldots + X_{n+m}),$$

si bien que

$$f(X_1 + \ldots + X_{n+m}) = \frac{1}{n+m}(X_1 + \ldots + X_{n+m}).$$

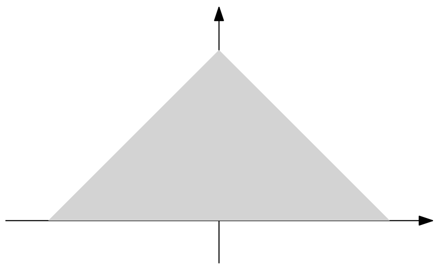
On en déduit donc que

$$\mathbf{E}[X_1 + \ldots + X_n | X_1 + \ldots + X_{n+m}] = \frac{n}{n+m}(X_1 + \ldots + X_{n+m}),$$

si bien que la fonction $h(x) = \frac{n}{n+m}x$ convient.

Cette question n'est pas très facile, et n'a pas été bien traitée en général. La deuxième solution demandait moins de calcul que la première, mais nécessitait d'intoduire des variables aléatoires qui n'étaient pas données par l'énoncé.

2. La première chose à faire est de dessiner le domaine D :



Calculons la loi de X. Elle admet une densité par rapport à la mesure de Lebesgue sur $\mathbb{R}$, qu'on peut calculer en intégrtant par rapport à y la densité de la loi du couple (X, Y). On trouve

$$f_X(x) = \int_{\mathbb{R}^2} f(x,y) \, \mathrm{d}y = c\mathbf{1}_{[-1,1]}(x) \int_0^{1-|x|} x^2 y^2 \, \mathrm{d}y = \frac{c}{3} x^2 (1-|x|)^3 \, \mathbf{1}_{[-1,1]}(x).$$

Soit maintenant $g : \mathbb{R} \to \mathbb{R}$ une fonction mesurable bornée. Calculons $\mathbf{E}[Yg(X)]$. On a

$$\begin{aligned}
\mathbf{E}[Yg(X)] &= \int_D y g(x) f(x,y) \, \mathrm{d}x\mathrm{d}y \\
&= c \int_{-1}^{1} g(x) x^2 \Big(\int_0^{1-|x|} y^3 \, \mathrm{d}y \Big) \, \mathrm{d}x \\
&= \frac{c}{4} \int_{-1}^{1} g(x) x^2 (1-|x|)^4 \, \mathrm{d}x \\
&= c \int_{\mathbb{R}} \frac{3}{4} (1-|x|) g(x) f_X(x) \, \mathrm{d}x \\
&= \frac{3}{4} \mathbf{E}[(1-|X|) g(X)],
\end{aligned}$$

si bien que

$$\mathbf{E}[Y|X] = \tfrac{3}{4}(1-|X|).$$

Considérons une fonction mesurable bornée $g : \mathbb{R} \to \mathbb{R}$ et calculons $\mathbf{E}[Xg(Y)]$. Puisque la fonction f vérifie $f(-x,y) = f(x,y)$, la loi du couple (X, Y) est la même que celle du couple $(-X, Y)$, donc

$$\mathbf{E}[Xg(Y)] = \mathbf{E}[-Xg(Y)] = -\mathbf{E}[Xg(Y)]$$

et cette espérance est nulle, si bien que

$$\mathbf{E}[X|Y] = 0.$$

Cette question a été assez bien traitée en général. Faire le dessin permet de ne pas être surpris du fait qu'une des deux espérances conditionnelles est nulle, et éventuellement même de le prévoir.

3. Calculons la loi de U. Si $g : \mathbb{R} \to \mathbb{R}$ est une fonction mesurable positive, on a

$$\begin{aligned}
\mathbf{E}[g(U)] &= \int_0^\infty g(x - \lfloor x \rfloor) e^{-x} \,\mathrm{d}x \\
&= \sum_{n=0}^\infty \int_n^{n+1} e^{-x} g(x-n) \,\mathrm{d}x \\
&= \sum_{n=0}^\infty \int_0^1 e^{-n-x} g(x) \,\mathrm{d}x \\
&= \int_0^1 \frac{e^{-x}}{1-e^{-1}} g(x) \,\mathrm{d}x.
\end{aligned}$$

Calculons maintenant $\mathbf{E}[Tg(U)]$. Nous avons

$$\begin{aligned}
\mathbf{E}[Tg(U)] &= \int_0^\infty x g(x - \lfloor x \rfloor) e^{-x} \,\mathrm{d}x \\
&= \int_0^1 e^{-x} g(x) \sum_{n=0}^\infty (n+x) e^{-n} \,\mathrm{d}x.
\end{aligned}$$

On utilise le fait que $\sum_{n \geqslant 0} nt^n = \frac{t}{(1-t)^2}$ pour tout t tel que $|t| < 1$ et on trouve

$$\begin{aligned}
\mathbf{E}[Tg(F)] &= \int_0^1 e^{-x} g(x) \left(\frac{x}{1-e^{-1}} + \frac{e^{-1}}{(1-e^{-1})^2} \right) \mathrm{d}x \\
&= \int_0^1 \frac{e^{-x}}{1-e^{-1}} g(x) \big(x + \frac{1}{e-1} \big) \,\mathrm{d}x \\
&= \mathbf{E}\Big[\big(U + \tfrac{1}{e-1} \big) g(U) \Big],
\end{aligned}$$

si bien que

$$\mathbf{E}[T|U] = U + \frac{1}{e-1}.$$

Cette question a été assez peu traitée, alors qu'elle était très proche de questions figurant dans des examens passés, et ne posait pas de difficulté particulière.

4. Calculons la matrice de covariance du vecteur gaussien $(F+G, G+H)$. On a

$$\mathrm{Var}(F+G) = \mathrm{Var}(G+H) = 3+3+4 = 10$$

et

$$\mathrm{Cov}(F+G, G+H) = 2+1+3+2 = 8.$$

On a donc
$$\mathrm{Cov}\Big((F+G)-\tfrac{4}{5}(G+H), G+H\Big)=0.$$
On en déduit que $(F+G)-\frac{4}{5}(G+H)$ est indépendant de $G+H$, donc
$$\mathbf{E}[(F+G)-\tfrac{4}{5}(G+H)|G+H]=\mathbf{E}[(F+G)-\tfrac{4}{5}(G+H)]=0.$$
Or cette espérance conditionnelle vaut
$$\mathbf{E}[F+G|G+H]-\frac{4}{5}(G+H),$$
si bien que
$$\mathbf{E}[F+G|G+H]=\frac{4}{5}(G+H).$$

Aux erreurs de calcul près, cette question a été la mieux traitée de l'exercice. J'aimerais qu'écrire $\frac{8}{10}$ provoque chez la plupart des étudiantes et étudiants, par réflexe, un besoin de simplifier la fraction en $\frac{4}{5}$, mais force est de constater que ce n'est pas le cas. On peut en rire, et aussi réfléchir au fait que laisser $\frac{8}{10}$ ne donne pas un bon signal de maturité mathématique, et peut influencer négativement la suite de la correction.

5. Commençons par calculer la deuxième espérance conditionnelle. On a
$$\mathbf{E}[K+L|K]=\mathbf{E}[K|K]+\mathbf{E}[L|K]=K+\mathbf{E}[L].$$
La tribu engendrée par la variable aléatoire $K+\mathbf{E}[L]$ est la même que la tribu engendrée par K, donc
$$\mathbf{E}[K+L+M|\mathbf{E}[K+L|K]]=\mathbf{E}[K+L+M|K]=K+\mathbf{E}[L]+\mathbf{E}[M]=K+2\mathbf{E}[K].$$

Pour la première, on commence par calculer $\mathbf{E}[K+L|K+L+M]$. On utilise le même argument que dans la question 1. Les couples $(K,K+L+M)$, $(L,K+L+M)$ et $(M,K+L+M)$ ont même loi, donc
$$\mathbf{E}[K|K+L+M], \mathbf{E}[L|K+L+M] \text{ et } \mathbf{E}[M|K+L+M]$$
sont égales, de somme égale à $K+L+M$. Ainsi, ces trois variables aléatoires valent $\frac{1}{3}(K+L+M)$, donc
$$\mathbf{E}[K+L|K+L+M]=\frac{2}{3}(K+L+M).$$
Cette variable aléatoire engendre la même tribu que $K+L+M$, donc
$$\mathbf{E}[K|\mathbf{E}[K+L|K+L+M]]=\mathbf{E}[K|K+L+M]=\frac{1}{3}(K+L+M).$$

Cette question a été peu traitée, mais plus que je ne m'y attendais, ce qui a constitué une bonne surprise.
Globalement, cet exercice, qui était attendu (parce qu'il était annoncé, et très similaire aux exercices de la même nature figurant dans les sujets récents de partiel et d'examen), n'a pas été très bien traité. Une erreur de calcul arrive à tout le monde, et on peut ne pas trouver, par exemple, comment mener le calcul de la première question. Mais sur les 15 points proposés ici, en obtenir 10 ou 12 ne devrait pas constituer l'exception, et permettrait, avec les 10 points de l'exercice "classique" sur les chaînes de Markov (ici exercice 4), d'obtenir au moins 20 points, et de ne pas être loin de valider l'UE.

Exercice 2

Barème indicatif : 20 points (3+3+3+3+2+3+3)

Dans cet exercice, toutes les variables aléatoires sont définies sur un espace de probabilités filtré $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geqslant 0}, \mathbf{P})$.

1. Soit Z une variable aléatoire positive. Montrer d'une part que pour tout $\omega \in \Omega$ on a

$$\int_0^\infty \mathbf{1}_{\{Z\geqslant a\}}(\omega)\ \mathrm{d}a = Z(\omega) \quad \text{et} \quad \int_0^\infty a\mathbf{1}_{\{Z\geqslant a\}}(\omega)\ \mathrm{d}a = \tfrac{1}{2}Z^2(\omega),$$

et d'autre part que

$$\int_0^\infty a\mathbf{P}(Z \geqslant a)\ \mathrm{d}a = \tfrac{1}{2}\mathbf{E}[Z^2].$$

On se donne une martingale $(X_n)_{n\geqslant 0}$. Pour tout entier $n \geqslant 0$, on pose

$$Y_n = |X_n| \ \text{ et } \ M_n = \max(Y_0, \dots, Y_n).$$

2. Soit a un réel positif. On note $T = \inf\{n \geqslant 0 : Y_n \geqslant a\}$, avec la convention habituelle selon laquelle $\inf \varnothing = \infty$. Montrer que pour tout entier $n \geqslant 0$, on a

$$a\mathbf{P}(M_n \geqslant a) \leqslant \sum_{k=0}^n \mathbf{E}\big[Y_k \mathbf{1}_{\{T=k\}}\big].$$

3. En déduire que pour tout réel positif a et tout entier $n \geqslant 0$, on a

$$a\mathbf{P}(M_n \geqslant a) \leqslant \mathbf{E}\big[Y_n \mathbf{1}_{\{M_n\geqslant a\}}\big].$$

4. Montrer que pour tout entier $n \geqslant 0$, on a

$$\mathbf{E}\big[M_n^2\big] \leqslant 4\mathbf{E}\big[X_n^2\big].$$

On suppose désormais que la martingale $(X_n)_{n\geqslant 0}$ est bornée dans L^2, c'est-à-dire qu'il existe un réel C tel que pour tout entier $n \geqslant 0$, on ait $\mathbf{E}\big[X_n^2\big] \leqslant C$.

5. Montrer que la martingale $(X_n)_{n\geqslant 0}$ est bornée dans L^1.

La suite $(X_n)_{n\geqslant 0}$ converge donc presque sûrement (on ne demande pas de le démontrer) vers une limite que l'on notera X_∞.

6. Montrer que la variable aléatoire

$$M = \sup\big\{|X_n|, n \geqslant 0\big\}$$

est de carré intégrable. Que peut-on en déduire sur la convergence de $(X_n)_{n\geqslant 0}$ vers X_∞ ?

7. On vient de démontrer que si la martingale $(X_n)_{n\geqslant 0}$ est bornée dans L^2, alors la variable aléatoire $M = \sup\big\{|X_n|, n \geqslant 0\big\}$ appartient à L^2. Cette assertion reste-t-elle vraie si l'on remplace L^2 par L^1 ?

Autrement dit : est-il vrai que si la martingale $(X_n)_{n\geqslant 0}$ est bornée dans L^1, alors la variable aléatoire $M = \sup\big\{|X_n|, n \geqslant 0\big\}$ appartient à L^1 ?

Si vous pensez que oui, donnez une démonstration, et si vous pensez que non, donnez un contre-exemple et démontrez que c'en est un.

Solution de l'exercice 2

1. Soit $\omega \in \Omega$. La première intégrale vaut

$$\int_0^{Z(\omega)} \mathrm{d}a = Z(\omega)$$

et la seconde

$$\int_0^{Z(\omega)} a\ \mathrm{d}a = \tfrac{1}{2}Z(\omega)^2.$$

L'espérance de $\frac{1}{2}Z^2$ vaut donc l'espérance de la deuxième intégrale, et le théorème de Fubini permet d'intervertir l'intégrale et l'espérance, pour trouver

$$\tfrac{1}{2}\mathbf{E}[Z^2] = \int_0^\infty a\mathbf{E}[\mathbf{1}_{\{Z \geqslant a\}}]\ \mathrm{d}a = \int_0^\infty a\mathbf{P}(Z \geqslant a)\ \mathrm{d}a.$$

J'ai lu quelques fois pour la première égalité le raisonnement faux suivant : on sait (parce que c'est une formule qu'on connaît) que $\mathbf{E}[Z] = \int_0^\infty \mathbf{P}(Z \geqslant a)\ \mathrm{d}a$, et on en "déduit" (en enlevant l'espérance ?) que $Z = \int_0^\infty \mathbf{1}_{\{Z \geqslant a\}}\ \mathrm{d}a$. Or deux variables aléatoires peuvent avoir la même espérance sans être égales.

2. On a

$$a\mathbf{P}(M_n \geqslant a) = a\mathbf{P}(T \leqslant n) = \sum_{k=0}^n \mathbf{E}[a\mathbf{1}_{\{T=k\}}].$$

Et puisque sur l'événement $\{T = k\}$ on a $Y_k \geqslant a$, on a

$$a\mathbf{P}(M_n \geqslant a) \leqslant \sum_{k=0}^n \mathbf{E}[Y_k\mathbf{1}_{\{T=k\}}].$$

Cette question a été plutôt bien traitée en général, quoique souvent avec des arguments un peu confus et très longs.

3. La suite $(Y_n)_{n\geqslant 0}$, image par une fonction convexe de la martingale $(X_n)_{n\geqslant 0}$, est une sous-martingale. Puisque de plus l'événement $\{T = k\}$ appartient à $\mathscr{F}_k$, on a pour tous $k \leqslant n$

$$\mathbf{E}[Y_k\mathbf{1}_{\{T=k\}}] \leqslant \mathbf{E}[\mathbf{E}[Y_n|\mathscr{F}_k]\mathbf{1}_{\{T=k\}}] = \mathbf{E}[Y_n\mathbf{1}_{\{T=k\}}].$$

Ainsi, on a

$$a\mathbf{P}(M_n \geqslant a) \leqslant \sum_{k=0}^n \mathbf{E}[Y_n\mathbf{1}_{\{T=k\}}] = \mathbf{E}[Y_n\mathbf{1}_{\{T\leqslant n\}}] = \mathbf{E}[Y_n\mathbf{1}_{\{M_n\geqslant a\}}].$$

Cette question, par contre, n'a preque jamais été traitée correctement. Seule une petite minorité de copies ont mentionné que la suite $(Y_n)_{n\geqslant 0}$ est une sous-martingale, ce qui est le point clé.

4. Intégrons l'inégalité précédente par rapport à a de 0 à l'infini. En vertu du résultat de la première question, nous trouvons

$$\tfrac{1}{2}\mathbf{E}[M_n^2] \leqslant \int_0^\infty \mathbf{E}[Y_n\mathbf{1}_{\{M_n\geqslant a\}}]\ \mathrm{d}a.$$

Le théorème de Fubini nous permet d'échanger l'intégrale et l'espérance, pour trouver dans le membre de droite

$$\mathbf{E}\left[Y_n \int_0^\infty \mathbf{1}_{\{M_n\geqslant a\}}\ \mathrm{d}a\right] = \mathbf{E}[Y_nM_n].$$

L'inégalité de Cauchy–Schwarz nous donne alors

$$\tfrac{1}{2}\mathbf{E}[M_n^2] \leqslant \mathbf{E}[Y_n^2]^{\frac{1}{2}}\mathbf{E}[M_n^2]^{\frac{1}{2}}.$$

En élevant au carré et en simplifiant, on obtient l'inégalité souhaitée.

Beaucoup de copies sont arrivées à l'inégalité $\cdots \leqslant \mathbf{E}[Y_n M_n]$ et, à partir de là, ont essayé sans trop savoir comment de faire comme si elles parvenaient à en déduire le résultat. Soyez prudent(e)s lorsque vous démontrez une égalité ou une inégalité qui est donnée : si vous ne savez pas comment vous trouvez le résultat, c'est probablement que vous n'avez pas vu l'argument à utiliser, et cela se verra. Il vaut mieux dire : je n'arrive pas à aller plus loin (ou : "non abouti" — sans t, soit dit en passant) que de recopier le résultat souhaité comme s'il s'en déduisait par magie.

5. Pour tout $n \geqslant 0$, on a, par l'inégalité de Jensen ou par l'inégalité de Cauchy–Schwarz appliquée à $|X_n|$ et 1,

$$\mathbf{E}[|X_n|] \leqslant \mathbf{E}[X_n^2]^{\frac{1}{2}} \leqslant \sqrt{C}.$$

La martingale $(X_n)_{n\geqslant 0}$ est donc bornée dans L^1.

Cette question a été souvent bien traitée. Attention : il n'est pas vrai, pour une variable positive Z, qu'on ait l'inégalité $\mathbf{E}[Z] \leqslant \mathbf{E}[Z^2]$, considérez par exemple le cas où Z est constante égale à $\frac{1}{2}$.

6. La suite $(M_n)_{n\geqslant 0}$ est croissante, et converge presque sûrement vers M. Du résultat de la question 4, nous déduisons que pour tout $n \geqslant 0$,

$$\mathbf{E}[M_n^2] \leqslant 4C,$$

et du théorème de convergence monotone appliqué à la suite $(M_n^2)_{n\geqslant 0}$, nous déduisons que

$$\mathbf{E}[M^2] = \lim_{n\to\infty} \mathbf{E}[M_n^2] \leqslant 4C < \infty.$$

La variable aléatoire M est donc de carré intégrable.

Or la convergence presque sûre de $(X_n)_{n\geqslant 0}$ vers X est dominée par M. Cette convergence a donc lieu dans L^2, ce qu'un résultat du cours assurait, mais au terme d'une démonstration différente.

Cette question n'a pas été bien traitée et a souvent révélé des confusions entre "suite de variables aléatoires de carré intégrable" et "suite de variables aléatoires bornée dans L^2".
Par ailleurs, lorsque l'énoncé demande à la question 6 de déduire quelque chose de ce qui précède, en l'occurrence de tout ce qui précède, le but n'est pas de vous faire appliquer un théorème du cours. C'est bien de savoir qu'une martingale bornée dans L^2 converge dans L^2, mais le but ici est de le redémontrer, autrement que dans le cours.

7. Considérons la marche aléatoire simple symétrique sur $\mathbb{Z}$ issue de 1 et arrêtée lorsqu'elle touche 0. Nous savons que cette martingale $(X_n)_{n\geqslant 0}$ converge presque sûrement vers 0, mais pas dans L^1. Par ailleurs, pour tout $n \geqslant 0$, nous avons $\mathbf{E}[|X_n|] = \mathbf{E}[X_n] = 1$, donc cette martingale est bornée dans L^1.

Si la variable aléatoire M était intégrable, alors la martingale $(X_n)_{n\geqslant 0}$ serait dominée dans L^1, et sa convergence presque sûre vers 0 aurait aussi lieu dans L^1, ce qui n'est pas le cas. Donc dans ce cas, la variable aléatoire M n'est pas intégrable.

La réponse est donc négative.

Dans plusieurs réponses s'est là aussi manifestée la confusion entre "suite de v.a. intégrables" et "suite de v.a. bornée dans L^1".

Exercice 3

Barème indicatif : 20 points (2+3+3+3+3+3+3)

Soit E un espace d'états, dont on suppose qu'il a au moins deux éléments. Soit P un noyau de transition irréductible sur E.

Pour tout $x \in E$, on définit $\widetilde{P}(x,x) = 0$, et pour tout $y \in E$ tel que $y \neq x$,

$$\widetilde{P}(x,y) = \frac{P(x,y)}{1 - P(x,x)}.$$

1. Montrer que $\widetilde{P}$ est un noyau de transition sur E.

On considère la chaîne de Markov canonique $(E^{\mathbb{N}}, \mathscr{C}, (\mathscr{C}_n)_{n\geqslant 0}, (\mathbf{P}_x)_{x\in E}, X = (X_n)_{n\geqslant 0})$ associée au noyau P. On définit les variables aléatoires $T_0 = 0$ et, pour tout $k \geqslant 0$,

$$T_{k+1} = \inf \big\{ n > T_k : X_n \neq X_{n-1} \big\}.$$

2. Montrer que pour tout $x \in E$, la variable aléatoire T_1 est finie $\mathbf{P}_x$-presque sûrement.

3. Montrer que pour tout $k \geqslant 0$, la variable aléatoire T_k est un temps d'arrêt. On pourra commencer par montrer que si S est un temps d'arrêt, alors $T = \inf\{n > S : X_n \neq X_{n-1}\}$ est un temps d'arrêt.

4. Soit $x \in E$. En admettant la relation

$$\forall k \geqslant 0, \quad T_{k+1} = (T_k + T_1 \circ \theta_{T_k})\, \mathbf{1}_{\{T_k<\infty\}} + \infty\, \mathbf{1}_{\{T_k=\infty\}},$$

montrer que pour tout $k \geqslant 0$, la variable aléatoire T_k est finie $\mathbf{P}_x$-presque sûrement.

5. Soit $x \in E$. Montrer que sous $\mathbf{P}_x$, la suite $(X_{T_k})_{k\geqslant 0}$ est une chaîne de Markov sur E issue de x et de noyau de transition $\widetilde{P}$.

6. Montrer que si $\widetilde{P}$ est transient, alors P est transient.

7. Soit $d \geqslant 3$ un entier. On suppose que $E = \mathbb{Z}^d$ et que le noyau de transition P est celui de la marche aléatoire simple symétrique, défini en posant, pour tous $x, y \in \mathbb{Z}^d$,

$$P(x,y) = \frac{1}{2d} \ \text{ si } \ \|x-y\| = 1, \quad \text{et} \ \ P(x,y) = 0 \ \text{ sinon},$$

où la norme est la norme euclidienne usuelle. On admettra que P est irréductible.

On définit l'application $\pi : \mathbb{Z}^d \to \mathbb{Z}^3$ en posant, pour tout $x = (x^1, \ldots, x^d) \in \mathbb{Z}^d$,

$$\pi(x^1, \ldots, x^d) = (x^1, x^2, x^3).$$

On admettra que pour tout $x \in E$, la suite $(\pi(X_n))_{n\geqslant 0}$ est sous $\mathbf{P}_x$ une chaîne de Markov sur $\mathbb{Z}^3$, dont le noyau de transition, noté Q, ne dépend pas de x.

Calculer le noyau Q puis le noyau $\widetilde{Q}$. Que peut-on en conclure ?

Solution de l'exercice 3

1. Notons d'abord que pour tout $x \in E$, on ne peut avoir $P(x,x) = 1$, sans quoi x ne mènerait à aucun autre élément que lui-même, et serait donc le seul élément de sa classe de communication, ce qui contredirait les hypothèses. La définition de $\widetilde{P}$ a donc un sens. De plus, puisque P est un noyau, la fonction $\widetilde{P}$ est à valeurs positives, et pour tout $x \in E$,

$$\sum_{y \in E} \widetilde{P}(x,y) = \frac{1}{1-P(x,x)} \sum_{y \neq x} P(x,y) = \frac{1}{1-P(x,x)}(1-P(x,x)) = 1.$$

La fonction $\widetilde{P}$ est donc un noyau de transition sur E.

Peu de copies ont justifié clairement le fait que $P(x,x) < 1$. Dans beaucoup de copies, j'ai un raisonnement qui aboutissait à la conclusion que $P(x,x) = 0$, ce qui n'est pas le cas.

2. Soit $n \geqslant 0$ un entier. On a

$$\mathbf{P}_x(T_1 > n) = \mathbf{P}_x(X_0 = \ldots = X_n) = \mathbf{P}_x(X_0 = x, \ldots, X_n = x) = P(x,x)^n.$$

Ainsi, T_1 suit sous $\mathbf{P}_x$ une loi géométrique de paramètre $P(x,x)$ (la convention étant que cette loi est la masse de Dirac en 1 si $P(x,x) = 0$). En particulier, T_1 est fini presque sûrement, ce qu'on retrouve en écrivant

$$\mathbf{P}_x(T_1 = \infty) = \lim_{n \to \infty} \mathbf{P}_x(T_1 > n) = \lim_{n \to \infty} P(x,x)^n = 0,$$

puisque comme nous l'avons déjà observé, $P(x,x) < 1$.

J'ai lu beaucoup de réponses vagues et assez verbeuses à cette question, qui soulignaient quand même, dans les bons cas, l'importance du fait que la chaîne soit irréductible et que E ait au moins deux éléments.

3. Montrons que si S est un temps d'arrêt, alors

$$T = \inf\{n > S : X_n \neq X_{n-1}\}$$

est un temps d'arrêt. Puisque $T_0 = 0$ est un temps d'arrêt, le résultat s'ensuivra immédiatement par récurrence sur k.

Soit $n \geqslant 0$ un entier. L'écriture

$$\{T > n\} = \{S \geqslant n\} \cup \bigcup_{k=0}^{n-1} \Big(\{S = k\} \cap \{X_k = X_{k+1} = \ldots = X_n\} \Big)$$

montre que l'événement $\{T > n\}$, et donc son complémentaire $\{T \leqslant n\}$, appartiennent à $\mathscr{C}_n$. La variable aléatoire T est donc bien un temps d'arrêt.

J'ai souvent lu des tentatives d'écrire l'événement $\{T > n\}$ ou $\{T = n\}$ sous une forme qui permette de déterminer à quelle tribu il appartient, et c'est une bonne chose. Souvent toutefois, cette écriture faisait intervenir S dans les bornes d'une union ou d'une intersection, ce qui n'a pas de sens.

4. Montrons par récurrence sur k que T_k est fini $\mathbf{P}_x$-presque sûrement. Pour $k = 0$ c'est vrai par définition et pour T_1, nous l'avons établi à la question 2.

Supposons maintenant établi que T_k est fini $\mathbf{P}_x$-presque sûrement pour un entier $k \geqslant 1$. Alors d'une part,

$$\{T_{k+1} < \infty\} = \{T_k < \infty\} \cap \{T_{k+1} - T_k < \infty\}.$$

D'autre part, on a l'égalité $\mathbf{P}_x$-presque sûre

$$T_{k+1} - T_k = T_1 \circ \theta_{T_k}.$$

On a donc

$$\mathbf{P}_x(T_{k+1} < \infty) = \mathbf{P}_x(T_{k+1} - T_k < \infty) = \mathbf{P}_x(T_1 \circ \theta_{T_k} < \infty) = \mathbf{E}_x[\mathbf{1}_{\mathbb{R}_+}(T_1 \circ \theta_{T_k})].$$

La propriété de Markov forte au temps T_k nous permet de calculer cette espérance comme

$$\mathbf{P}_x(T_{k+1} < \infty) = \mathbf{E}_x\Big[\mathbf{E}_{X_{T_k}}[\mathbf{1}_{\mathbb{R}_+}(T_1)]\Big].$$

Or

$$\begin{aligned}\mathbf{E}_{X_{T_k}}[\mathbf{1}_{\mathbb{R}_+}(T_1)] &= \sum_{y\in E} \mathbf{1}_{\{X_{T_k}=y\}} \mathbf{E}_y[\mathbf{1}_{\mathbb{R}_+}(T_1)] \\ &= \sum_{y\in E} \mathbf{1}_{\{X_{T_k}=y\}} \mathbf{P}_y(T_1 < \infty) \\ &= \sum_{y\in E} \mathbf{1}_{\{X_{T_k}=y\}} \\ &= 1,\end{aligned}$$

donc

$$\mathbf{P}_x(T_{k+1} < \infty) = 1.$$

Cette question a parfois été bien traitée. Outre le fait qu'il faut utiliser la propriété de Markov forte, le point clé est que dans la récurrence, on a besoin de savoir que T_1 est fini $\mathbf{P}_y$-presque sûrement pour tout y.
J'ai vu plusieurs fois des confusions entre "T_1 est fini presque sûrement" et "T_1 est d'espérance finie". Enfin, notons que dans les copies qui avaient abouti à la conclusion que $P(x, x) = 0$ pour tout x, je n'ai jamais vu tirée la conclusion qui devrait en découler, à savoir que $T_k = k$ pour tout k.

5. Donnons-nous un état $x \in E$, un entier $n \geqslant 0$ et $x_0, \dots, x_n \in E$, et calculons

$$\mathbf{P}_x(X_{T_0} = x_0, X_{T_1} = x_1, \dots, X_{T_n} = x_n).$$

Par définition des temps d'arrêt T_k, s'il existe $i \in \{0, \dots, n-1\}$ tel que $x_i = x_{i+1}$, alors cette probabilité est nulle. Sinon, elle vaut

$$\begin{aligned}&\sum_{k_1,\dots,k_n\geqslant 1} \mathbf{P}_x(X_{T_0} = x_0, T_1 = k_1, X_{T_1} = x_1, \dots, T_n = k_1 + \dots + k_n, X_{T_n} = x_n) \\ &\quad= \sum_{k_1,\dots,k_n\geqslant 1} \mathbf{P}_x(X_0 = \dots = X_{k_1-1} = x_0, X_{k_1} = \dots = X_{k_1+k_2-1} = x_1, \dots, \\ &\qquad\qquad X_{k_1+\dots+k_{n-1}} = \dots = X_{k_1+\dots+k_n-1} = x_{n-1}, X_{k_1+\dots+k_n} = x_n) \\ &\quad= \sum_{k_1,\dots,k_n\geqslant 1} \delta_{x,x_0} P(x_0, x_0)^{k_1-1} P(x_0, x_1) P(x_1, x_1)^{k_2-1} P(x_1, x_2) \dots P(x_{n-1}, x_{n-1})^{k_n-1} P(x_{n-1}, x_n)\end{aligned}$$

qui vaut exactement

$$\delta_{x,x_0} \widetilde{P}(x_0, x_1) \dots \widetilde{P}(x_{n-1}, x_n).$$

6. Supposons $\widetilde{P}$ transient. Soit $x \in E$ un état. Alors, par définition du fait que x est transient, et puisque la suite $(X_{T_k})_{k\geqslant 0}$ est sous $\mathbf{P}_x$ une chaîne de Markov issue de x de noyau de transition $\widetilde{P}$,

$$\mathbf{P}_x(\forall k \geqslant 1, X_{T_k} \neq x) > 0.$$

Or sur l'événement $\{\forall k \geqslant 1, X_{T_k} \neq x\}$, la chaîne $(X_n)_{n\geqslant 0}$ ne passe que T_1 fois en x. Ainsi,

$$\mathbf{P}_x(N_x < \infty) > 0,$$

ce qui implique que x est transient pour la chaîne $(X_n)_{n\geqslant 0}$de noyau P. Puisque le noyau P est irréductible, il est transient.

7. Pour déterminer le noyau Q, il suffit de calculer, pour tous $u, v \in \mathbb{Z}^3$,

$$\mathbf{P}_u(\pi(X_1) = v).$$

Si u et v ne sont pas égaux ou voisins, cette probabilité est nulle. Si u et v sont voisins, alors

$$\mathbf{P}_u(\pi(X_1) = v) = \frac{1}{2d}.$$

Enfin, si $u = v$, cette probabilité vaut

$$\mathbf{P}_u(\pi(X_1) = u) = 1 - \frac{6}{2d},$$

puisque u admet 6 voisins dans $\mathbb{Z}^3$.

Finalement, $Q(u, v)$ est nul si $\|u - v\| > 1$, vaut $\frac{1}{2d}$ si $\|u - v\| = 1$, et vaut $1 - \frac{3}{d}$ si $u = v$.

Le noyau $\widetilde{Q}$ est donc donné par

$$\widetilde{Q}(u, v) = \frac{1}{6}$$

si $\|u - v\| = 1$, et 0 sinon. Autrement dit, $\widetilde{Q}$ est le noyau de la marche aléatoire symétrique simple sur $\mathbb{Z}^3$.

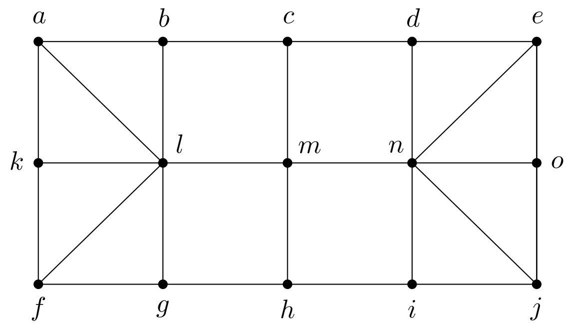
Cette marche est transiente, donc, d'après ce qui précède, la chaîne de Markov $(\pi(X_n))_{n\geqslant 0}$ est transiente. La chaîne $(X_n)_{n\geqslant 0}$ est donc elle-même transiente, sinon elle passerait un nombre infini de fois en chaque état de $\mathbb{Z}^d$, donc $(\pi(X_n))_{n\geqslant 0}$ passerait une infinité de fois en chaque état de $\mathbb{Z}^3$, ce qui serait une contradiction.

Ces trois dernières questions ont été très peu traitées.

Exercice 4

Barème indicatif : 10 points (2+2+2+2+2).

On étudie la marche aléatoire sur le graphe représenté ci-dessous.



1. Montrer que cette marche aléatoire admet une unique mesure de probabilité invariante, qu'on notera π, et calculer $\pi(a)$, $\pi(l)$ et $\pi(m)$.

2. On fait partir la marche aléatoire de l'état n. Après combien de temps en moyenne la marche aléatoire revient-elle en n pour la première fois ?

3. On fait partir la marche aléatoire de l'état k. Combien de fois en moyenne la marche visite-t-elle le sommet l avant de revenir pour la première fois à son point de départ ?

4. Notons $(X_n)_{n\geqslant 0}$ la marche aléatoire issue de c. Posons $M = \{l, m, n\}$. La quantité

$$\frac{1}{t}\sum_{i=0}^{t-1} \mathbf{1}_M(X_i)$$

admet-elle sous $\mathbf{P}_c$ une limite presque sûre lorsque t tend vers l'infini, et si oui laquelle ?

5. Soit u une fonction à valeurs réelles sur l'ensemble des sommets de notre graphe. La quantité

$$\mathbf{E}_a[u(X_t)]$$

admet-elle une limite lorsque t tend vers l'infini, et si oui, laquelle ?

Solution de l'exercice 4

1. Le graphe est connexe, donc la marche aléatoire sur ce graphe est une chaîne de Markov irréductible. Tous les états sont donc de même nature (récurrents ou transients), et comme l'espace d'états est fini, l'un d'entre eux est récurrent, donc ils sont tous récurrents. La chaîne admet donc une mesure invariante unique à multiplication près par une constante strictement positive, donc une unique mesure de probabilité invariante.

Pour déterminer cette probabilité invariante, on utilise le fait que la mesure qui à chaque sommet associe son nombre de voisins est réversible, donc invariante, pour la marche aléatoire sur le graphe. Cette mesure associe respectivement les masses 3, 6 et 4 aux sommets a, l et m.

La masse totale de cette mesure est 52 (c'est deux fois le nombre d'arêtes). Ainsi, les masses des sommets a, l, m pour l'unique probabilité invariante π sont

$$\pi(a) = \frac{3}{52},\ \pi(l) = \frac{3}{26},\ \pi(m) = \frac{1}{13}.$$

2. Le temps moyen de retour est donné par l'inverse de la masse attribuée par la probabilité invariante :

$$\mathbf{E}_n[T_n] = \frac{1}{\pi(n)} = \frac{26}{3}.$$

3. L'unique mesure invariante ν qui associe à k la masse 1 associe à chaque autre sommet une masse égale au nombre moyen de visites en ce sommet entre deux visites en k. On a $\nu = \pi/\pi(k)$, donc le nombre moyen de visites en l entre deux visites en k vaut

$$\nu(o) = \frac{\pi(l)}{\pi(k)} = 2.$$

4. Puisque la chaîne est irréductible et récurrente, le théorème ergodique assure que la quantité considérée converge $\mathbf{P}_a$-presque sûrement vers

$$\pi(M) = \frac{4}{13}.$$

5. Partant de a, on peut y revenir en 2 pas ou en 3 pas. L'état a, et donc tout la chaîne, est donc apériodique. Le théorème de convergence vers l'équilibre s'applique donc, et on a la convergence

$$\lim_{n\to\infty} \mathbf{E}_a[f(X_n)] = \int_{\{a,\dots,o\}} f \,\mathrm{d}\pi.$$

Les valeurs numériques n'étaient pas particulièrement agréables, mais il s'agit de la seule "difficulté" de cet exercice, puisqu'il est entièrement annoncé, dans le moindre détail. Il est dommage de perdre des points sur cet exercice parce qu'on l'a traité trop rapidement. Par ailleurs, il est là aussi bienvenu de réduire les fractions.

Fin du sujet