# Examen

*Le sujet fait trois pages.*
*L'examen dure trois heures.*
*Les cinq exercices sont indépendants.*
*Ni documents, ni calculatrices, ni téléphones.*
*Le barème est indicatif. La note finale sera sur* 50 *points.*

## Exercice 1

*Barème indicatif : 15 points (3+3+3+3+3).*
*Les cinq questions sont indépendantes.*

**1.** On munit l'ensemble $[0,1]$ de la tribu borélienne $\mathscr{B}$ et de la mesure de Lebesgue $\lambda$. Sur $([0,1],\mathscr{B},\lambda)$, on définit la variable aléatoire $X$ par par $X(t)=t$ pour tout $t\in[0,1]$. On définit la sous-tribu $\mathscr{C}=\sigma\big(\big\{[0,\frac{1}{2}],[\frac{1}{3},1]\big\}\big)$ de $\mathscr{B}$.

Calculer et représenter graphiquement la variable aléatoire $\mathbb{E}[X|\mathscr{C}]$.

**2.** Sur un espace de probabilité $(\Omega,\mathscr{F},\mathbb{P})$, on se donne un vecteur gaussien $(X,Y)$ centré tel que $\mathrm{Var}(X)=1$, $\mathrm{Var}(Y)=2$ et $\mathrm{Cov}(X,Y)=1$.

Calculer $\mathbb{E}[X|X+Y]$.

**3.** Sur un espace de probabilité $(\Omega,\mathscr{F},\mathbb{P})$, on se donne un couple $(X,Y)$ de variables aléatoires réelles dont la loi admet, par rapport à la mesure de Lebesgue sur $\mathbb{R}^2$, la densité $f$ donnée par

$$\forall x,y\in\mathbb{R},\quad f(x,y)=c\,(1+y)\mathbf{1}_{\mathbb{R}_+}(x)\mathbf{1}_{[0,1]}(|x|+|y|).$$

Déterminer la constante $c$ et calculer $\mathbb{E}[Y|X]$.

**4.** Sur un espace de probabilité $(\Omega,\mathscr{F},\mathbb{P})$, on se donne un couple $(X,Y)$ de variables aléatoires réelles tel que pour tout entier $n\geqslant 1$ et tous réels positifs $0\leqslant a\leqslant b$, on ait

$$\mathbb{P}\big(X=n,\,Y\in[a,b]\big)=\int_a^b e^{-2^n t}\,\mathrm{d}t.$$

Calculer $\mathbb{E}[X|Y]$.

**5.** Sur un espace de probabilité $(\Omega,\mathscr{F},\mathbb{P})$, on se donne deux variables aléatoires $X$ et $Y$ indépendantes. On suppose que $X$ suit la loi uniforme sur l'intervalle $[0,\pi]$ et $Y$ la loi gaussienne centrée de variance 1.

Calculer $\mathbb{E}[\cos(X+Y^2)|Y]$.

## Solution de l'exercice 1

1. La tribu $\mathscr{C}$ contient les parties $A=[0,\frac{1}{3}[$ , $B=[\frac{1}{3},\frac{1}{2}]$ et $C=]\frac{1}{2},1]$ qui forment une partition de $[0,1]$. Elle contient donc la tribu engendrée par cette partition. Ceci s'écrit $\mathscr{C}\supseteq\sigma(\{A,B,C\})$. Par ailleurs, puisque $[0,\frac{1}{2}]=A\cup B$ et $[\frac{1}{3},1]=B\cup C$, l'inclusion inverse a aussi lieu, si bien que $\mathscr{C}=\sigma(\{A,B,C\})$.

Dans ce cas, où la sous-tribu par rapport à laquelle on conditionne est engendrée par une partition finie, on a une expression explicite de l'espérance conditionnelle :

$$\mathbb{E}[X|\mathscr{C}]=\frac{\mathbb{E}[X\mathbf{1}_A]}{\mathbb{P}(A)}\mathbf{1}_A+\frac{\mathbb{E}[X\mathbf{1}_B]}{\mathbb{P}(B)}\mathbf{1}_B+\frac{\mathbb{E}[X\mathbf{1}_C]}{\mathbb{P}(C)}\mathbf{1}_C.$$

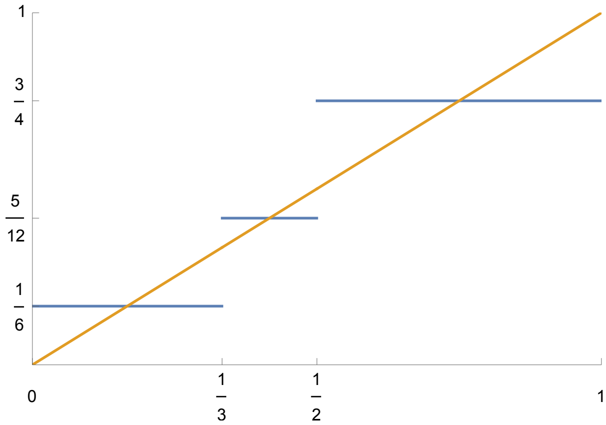
On calcule

$$\mathbb{E}[X\mathbf{1}_A] = \int_0^{\frac{1}{3}} t\,\mathrm{d}t = \frac{1}{18},\quad \mathbb{E}[X\mathbf{1}_B] = \int_{\frac{1}{3}}^{\frac{1}{2}} t\,\mathrm{d}t = \frac{1}{8} - \frac{1}{18} = \frac{5}{72},\quad \mathbb{E}[X\mathbf{1}_C] = \int_{\frac{1}{2}}^{1} t\,\mathrm{d}t = \frac{3}{8}$$

et

$$\mathbb{E}[X|\mathscr{C}] = \frac{1}{6}\mathbf{1}_A + \frac{5}{12}\mathbf{1}_B + \frac{3}{4}\mathbf{1}_C.$$

Voici en orange la variable aléatoire $X$ et en bleu son espérance conditionnelle sachant $\mathscr{C}$.




La description légèrement inhabituelle de la sous-tribu $\mathscr{C}$ a posé plus de problèmes que je ne m'y attendais. Je n'ai lu une solution complètement juste que dans une petite proportion de copies, moins d'un quart, me semble-t-il — et avec des fractions sous forme irréductible, dans une proportion encore plus petite.

Mais ce qui m'a vraiment le plus étonné a été la variété des réponses fausses que j'ai lues. Je n'ai pas résisté à la tentation de faire un florilège des plus jolies représentations graphiques que j'ai vues, que vous trouverez en annexe (page 11).

Plus sérieusement, quelques vérifications auraient permis d'éliminer rapidement certaines de ces réponses fausses. Tout d'abord, $3$ étant plus grand que $2$, c'est $\frac{1}{3}$ qui est plus petit que $\frac{1}{2}$ et non l'inverse. Ensuite, la variable aléatoire $X$ est comprise entre $0$ et $1$, il en est donc de même de son espérance conditionnelle. (Pourquoi ? Pourquoi est-il vrai que si $X \leqslant a$ p.s., alors $\mathbb{E}[X|\mathscr{G}] \leqslant a$ p.s. ?) Enfin, la moyenne de $X$ étant $\frac{1}{2}$, il en est de même de son espérance conditionnelle. Ainsi, si cette espérance conditionnelle est constante, ce que certains ont pensé, cette constante ne peut être que $\frac{1}{2}$.

2. Posons $Z = X + Y$. Le vecteur aléatoire $(X, Z)$, qui est l'image de $(X, Y)$ par la transformation linéaire $(u, v) \mapsto (u, u + v)$, est encore un vecteur gaussien centré. Cherchons un réel $a$ tel que $\mathrm{Cov}(X + aZ, Z) = 0$. Cette relation équivaut à

$$0 = \mathrm{Cov}(X, Z) + a\mathrm{Var}(Z) = 2 + 5a,$$

c'est-à-dire à $a = -\frac{2}{5}$. Puisque $(X - \frac{2}{5}Z, Z)$ est encore un vecteur gaussien centré, la relation $\mathrm{Cov}(X - \frac{2}{5}Z, Z) = 0$ entraîne que $X - \frac{2}{5}Z$ et $Z$ sont indépendantes. Cette indépendance entraîne

$$\mathbb{E}\big[X - \tfrac{2}{5}Z\big|Z\big] = \mathbb{E}\big[X - \tfrac{2}{5}Z\big] = 0 \ \text{ p.s.}$$

Mais on a par ailleurs

$$\mathbb{E}\big[X - \tfrac{2}{5}Z\big|Z\big] = \mathbb{E}[X|Z] - \tfrac{2}{5}\mathbb{E}[Z|Z] = \mathbb{E}[X|Z] - \tfrac{2}{5}Z \ \text{ p.s.}$$

En comparant ces deux relations, et en se rappelant que $Z = X + Y$, on trouve

$$\mathbb{E}[X|X + Y] = \tfrac{2}{5}(X + Y).$$

J'ai parfois lu l'affirmation que $\mathbb{E}[X|X + Y] = \mathbb{E}[Y|X + Y]$, ce qui conduisait à $\mathbb{E}[X|X + Y] = \frac{1}{2}(X + Y)$. L'erreur était que le couple $(X, X + Y)$ n'a pas la même loi que le couple $(Y, X + Y)$. Le vecteur $(X, Y)$ n'a pas non plus la même loi que le vecteur $(Y, X)$.

3. La constante $c$ est déterminée par le fait que l'intégrale de la densité $f$ sur $\mathbb{R}^2$ par rapport à la mesure de Lebesgue vaut 1. Le domaine sur lequel la fonction $f$ prend des valeurs non nulles est le triangle déterminé par les droites $x = 0$, $y = 1 - x$, et $y = -1 + x$

En utilisant le théorème de Fubini, on a

$$\int_{\mathbb{R}^2} (1+y)\, \mathbf{1}_{\mathbb{R}_+}(x)\mathbf{1}_{[0,1]}(|x|+|y|)\, \mathrm{d}x\mathrm{d}y = \int_0^1 \left( \int_{-1+x}^{1-x} (1+y)\, \mathrm{d}y \right) \mathrm{d}x = \int_0^1 (2-2x)\, \mathrm{d}x = 1.$$

Ainsi, $c = 1$.

Un calcul similaire à celui qui nous a permis de déterminer $c$ montre que pour toute fonction mesurable bornée $f : \mathbb{R} \to \mathbb{R}$, on a

$$\mathbb{E}[f(X)] = \int_0^1 (2-2x) f(x)\, \mathrm{d}x.$$

Cette expression décrit la loi de $X$ et nous pouvons maintenant calculer $\mathbb{E}[Y|X]$ en utilisant la méthode de la fonction muette. Soit donc $g : \mathbb{R} \to \mathbb{R}$ une fonction mesurable bornée. Calculons $\mathbb{E}[Yg(X)]$. Nous trouvons

$$\begin{aligned}
\mathbb{E}[Yg(X)] &= \int_0^1 g(x) \left( \int_{-1+x}^{1-x} y(1+y)\, \mathrm{d}y \right) \mathrm{d}x \\
&= \int_0^1 g(x) \tfrac{2}{3}(1-x)^3\, \mathrm{d}x \\
&= \int_0^1 \tfrac{1}{3}(1-x)^2 g(x)(2-2x)\, \mathrm{d}x \\
&= \mathbb{E}\big[\tfrac{1}{3}(1-X)^2 g(X)\big].
\end{aligned}$$

Nous en déduisons que

$$\mathbb{E}[Y|X] = \tfrac{1}{3}(1-X)^2.$$

Cette question a été une des mieux traitées de cet exercice. Toutes les erreurs de calcul sont possibles et excusables, mais trouver une valeur de $c$ strictement négative doit allumer un signal d'alarme, $c$ étant obtenu comme inverse de l'intégrale d'une fonction positive.

4. Pour tout $n \geqslant 1$, on a

$$\mathbb{P}(X = n) = \int_0^{+\infty} e^{-2^n t}\, \mathrm{d}t = 2^{-n}.$$

Ainsi, $X$ prend ses valeurs dans $\mathbb{N}^*$, suit une loi géométrique, et est intégrable. De plus, pour tous $0 \leqslant a \leqslant b$, on a, grâce au théorème de convergence monotone, l'égalité

$$\mathbb{P}(Y \in [a,b]) = \sum_{n \geqslant 1} \mathbb{P}(X = n, Y \in [a,b]) = \int_a^b \sum_{n \geqslant 1} e^{-2^n t}\, \mathrm{d}t.$$

La loi de $Y$ admet donc par rapport à la mesure de Lebesgue sur $\mathbb{R}$ la densité $f$ donnée par

$$f(t) = \sum_{n \geqslant 1} e^{-2^n t}.$$

Soit maintenant $g : \mathbb{R} \to \mathbb{R}$ une fonction mesurable bornée. Calculons $\mathbb{E}[Xg(Y)]$. Nous avons

$$\begin{aligned}
\mathbb{E}[Xg(Y)] &= \sum_{n \geqslant 1} n \int_0^{+\infty} g(y) e^{-2^n y}\, \mathrm{d}y = \int_0^{+\infty} g(y) \sum_{n \geqslant 1} n e^{-2^n y}\, \mathrm{d}y \\
&= \int_0^{+\infty} \frac{\sum_{n \geqslant 1} n e^{-2^n y}}{\sum_{n \geqslant 1} e^{-2^n y}}\, g(y) \sum_{n \geqslant 1} e^{-2^n y}\, \mathrm{d}y
\end{aligned}$$

On en déduit que

$$\mathbb{E}[X|Y] = \frac{\sum_{n\geqslant 1} ne^{-2^nY}}{\sum_{n\geqslant 1} e^{-2^nY}} \quad \text{p.s.}$$

Cette question n'a pas eu beaucoup de succès, ce qui peut s'expliquer par la forme peu agréable du résultat.

5. Puisque $Y^2$ est mesurable par rapport à $Y$ et $X$ est indépendante de $Y$, on a, d'après une formule du cours,

$$\mathbb{E}[\cos(X+Y^2)|Y] = h(Y),$$

où $h(t) = \mathbb{E}[\cos(X+t^2)]$. Calculons cette fonction : nous avons

$$h(t) = \frac{1}{\pi}\int_0^\pi \cos(x+t^2)\,\mathrm{d}x = -\tfrac{2}{\pi}\sin(t^2).$$

Ainsi,

$$\mathbb{E}[\cos(X+Y^2)|Y] = -\tfrac{2}{\pi}\sin(Y^2).$$

Notons que l'information concernant la loi de $Y$ était superflue.

Cette question a été souvent assez bien traitée. J'ai souvent lu $\sin(\pi+Y^2)$ qui n'avait pas été simplifié en $-\sin(Y^2)$. Il était par ailleurs possible d'utiliser une formule trigonométrique pour séparer $X$ et $Y$ dans $\cos(X+Y^2)$, et raisonner sans utiliser le résultat mentionné dans la solution ci-dessus.

## Exercice 2

*Barème indicatif : 5 points.*

Existe-t-il sur un espace de probabilité filtré $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geqslant 0}, \mathbb{P})$ une martingale $(M_n)_{n\geqslant 0}$ telle que $\mathbb{E}[M_n^2] < \infty$ pour tout $n \geqslant 0$ et telle que la suite $(M_n)_{n\geqslant 0}$ converge dans $L^1$ mais ne converge pas dans $L^2$ ?

## Solution de l'exercice 2

Donnons-nous un espace de probabilité $(\Omega, \mathscr{F}, \mathbb{P})$ sur lequel est définie une suite de variables aléatoires indépendantes $(X_k)_{k\geqslant 1}$ telles que $X_1 = 0$ et pour tout $k \geqslant 2$, on ait

$$\mathbb{P}(X_k = k) = \mathbb{P}(X_k = -k) = \frac{1}{k^3} \quad \text{et} \quad \mathbb{P}(X_k = 0) = 1 - \frac{2}{k^3}.$$

Posons $M_0 = 0$ et, pour tout $n \geqslant 1$, $M_n = X_1 + \ldots + X_n$. Notons enfin, pour tout $n \geqslant 0$, $\mathscr{F}_n = \sigma(X_0, \ldots, X_n)$.

Pour tout $n \geqslant 0$, la variable aléatoire $M_n$ est bornée (par $1 + \ldots + n$), donc intégrable. Elle est de plus $\mathscr{F}_n$-mesurable. Enfin, puisque les variables $(X_k)_{k\geqslant 1}$ sont centrées et indépendantes, la suite $(M_n)_{n\geqslant 0}$ est une martingale par rapport à la filtration $(\mathscr{F}_n)_{n\geqslant 0}$.

Montrons que la suite $(M_n)_{n\geqslant 0}$ converge dans $L^1$, en utilisant le critère de Cauchy. On a, pour tout $k \geqslant 2$, $\mathbb{E}[|X_k|] = \frac{2}{k^2}$, si bien que pour tous entiers $0 \leqslant n \leqslant m$, on a

$$\mathbb{E}\big[|M_m - M_n|\big] \leqslant \sum_{k=n+1}^{m} \frac{2}{k^2}$$

et puisque la série de terme général $\frac{2}{k^2}$ converge, ceci tend vers 0 lorsque $n$ tend vers l'infini. La suite $(M_n)_{n\geqslant 0}$ est donc de Cauchy dans $L^1$, donc convergente dans $L^1$.

Montrons maintenant que la suite $(M_n)_{n\geqslant 0}$ ne converge pas dans $L^2$. Pour tout $k \geqslant 2$, on a $\mathbb{E}[X_k^2] = \frac{2}{k}$, donc

$$\mathbb{E}[M_n^2] = \mathbb{E}[M_0^2] + \sum_{k=1}^{n} \mathbb{E}[(M_k - M_{k-1})^2] = \sum_{k=1}^{n} \mathbb{E}[X_k^2] = \sum_{k=2}^{n} \frac{2}{k}$$

qui tend vers l'infini lorsque $n$ tend vers l'infini. La suite $(M_n)_{n\geqslant 0}$ n'est donc pas bornée dans $L^2$, et elle ne converge en particulier pas dans $L^2$.

La réponse est donc : oui, cela existe, nous venons d'en construire un exemple.

Voici un autre exemple. Considérons l'espace de probabilité filtré $([0,1[, \mathscr{B}, (\mathscr{D}_n)_{n\geqslant 0}, \lambda)$, où pour tout $n \geqslant 0$, on a posé

$$\mathscr{D}_n = \sigma\big(\{[k2^{-n}, (k+1)2^{-n}[ : 0 \leqslant k \leqslant 2^n - 1\}\big).$$

Soit $X$ une variable aléatoire réelle sur cet espace de probabilité qui est telle que $\mathbb{E}[|X|] < \infty$ mais $\mathbb{E}[X^2] = +\infty$. On peut par exemple prendre

$$X(t) = \frac{1}{\sqrt{1-t}}, \ t \in [0,1[.$$

Notons, pour tout $n \geqslant 0$, $M_n = \mathbb{E}[X|\mathscr{D}_n]$. La variable aléatoire $M_n$ ne prend qu'un nombre fini de valeurs, elle est donc en particulier de carré intégrable. De plus, la martingale $(M_n)_{n\geqslant 0}$ est fermée par $X$, donc elle converge dans $L^1$ (et presque sûrement) vers l'espérance conditionnelle de $X$ vers $\mathscr{D}_\infty = \mathscr{B}$, c'est-à-dire vers $X$.

Si cette martingale convergeait dans $L^2$, elle ne pourrait que converger vers $X$, mais $X$ n'appartient pas à $L^2$. La martingale $(M_n)_{n\geqslant 0}$ ne converge donc pas dans $L^2$.

Cet exercice a été très rarement bien traité, dans une ou deux copies seulement. Beaucoup ont essayé de montrer que c'était impossible, parfois en confondant l'hypothèse, qui dit que chaque variable aléatoire $M_n$ est dans $L^2$, avec une hypothèse qui aurait dit que la suite $(M_n)_{n\geqslant 0}$ est bornée dans $L^2$.

## Exercice 3

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

On travaille sur un espace de probabilité dont la mesure est notée $\mathbb{P}$. Soient $p$ et $q$ des réels strictement positifs tels que $p + q = 1$. On se donne une suite $(Z_n)_{n\geqslant 1}$ de variables aléatoires indépendantes et identiquement distribuées telles que pour tout $n \geqslant 1$, on ait $\mathbb{P}(Z_n = 1) = p$ et $\mathbb{P}(Z_n = -1) = q$.

On définit par récurrence deux suites $(W_n)_{n\geqslant 0}$ et $(X_n)_{n\geqslant 0}$ de variables aléatoires en posant

$$\begin{aligned} W_0 = 0 \ \text{ et pour tout } \ n \geqslant 1, \ W_n &= W_{n-1} + Z_n, \\ Y_0 = 0 \ \text{ et pour tout } \ n \geqslant 1, \ \ Y_n &= Y_{n-1} + Z_n + \mathbf{1}_{\{Y_{n-1}=0\}}(1 - Z_n). \end{aligned}$$

**1.** Montrer que $(W_n)_{n\geqslant 0}$ est une chaîne de Markov sur $\mathbb{Z}$.
**2.** Montrer que $(Y_n)_{n\geqslant 0}$ est une chaîne de Markov sur $\mathbb{N}$.
**3.** Montrer que $\mathbb{P}$-presque sûrement, pour tout $n \geqslant 0$, on a $W_n \leqslant Y_n$.

Sur $\mathbb{N}$, on définit de le noyau de transition $P$ en posant, pour tout $i \geqslant 1$, $P(i, i+1) = p$ et $P(i, i-1) = q$, et $P(0,1) = 1$. On se donne une chaîne de Markov $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geqslant 0}, (\mathbb{P}_x)_{x\in\mathbb{N}}, (X_n)_{n\geqslant 0})$ sur $\mathbb{N}$ de noyau de transition $P$.

**4.** Montrer que cette chaîne de Markov est irréductible, puis déterminer une mesure $\mu$ sur $\mathbb{N}$ invariante pour cette chaîne et telle que $\mu(1) = 1$.

**5.** On suppose $p \neq \frac{1}{2}$. Déterminer, en discutant suivant la valeur de $p$, si la chaîne de Markov sur $\mathbb{N}$ de noyau de transition $P$ est récurrente ou transiente.

## Solution de l'exercice 3

1. Notons $Q$ le noyau de transition sur $\mathbb{Z}$ défini par $Q(i, i+1) = p$ et $Q(i, i-1) = q$. Remarquons que pour tous $i, j \in \mathbb{Z}$ et tout $n \geqslant 1$, on a $Q(i,j) = \mathbb{P}(Z_n = j - i)$.

Donnons-nous un entier $n \geqslant 0$ et des éléments $x_0, \ldots, x_n \in \mathbb{Z}$. Nous avons

$$\begin{aligned} \mathbb{P}(W_0 = x_0, \ldots, W_n = x_n) &= \delta_{x_0,0}\mathbb{P}(Z_1 = x_1 - x_0, \ldots, Z_n = x_n - x_{n-1}) \\ &= \delta_{x_0,0}Q(x_0, x_1)\ldots Q(x_{n-1}, x_n), \end{aligned}$$

ce qui montre que $(W_n)_{n\geqslant 0}$ est une chaîne de Markov sur $\mathbb{Z}$ de noyau de transition $Q$.

Pour montrer qu'un processus est une chaîne de Markov selon la définition du cours, il est nécessaire de définir d'abord son noyau. Aussi, bien que ce ne soit pas explicitement dit, la question appelait la définition d'un noyau de transition, dont on montrait ensuite que c'est celui de la chaîne.

2. Notons $R$ le noyau de transition sur $\mathbb{N}$ défini en posant $R(0,1) = 1$ et, pour tout $i \geqslant 1$, $R(i,i+1) = p$ et $R(i,i-1) = q$.

Donnons-nous un entier $n \geqslant 0$ et un élément $x \in \mathbb{N}$. Calculons

$$\mathbb{P}(Y_{n+1} = x | Y_0, \dots, Y_n).$$

La variable aléatoire $Y_{n+1}$ est une fonction de $Y_n$ et de $Z_{n+1}$. La variable aléatoire $Y_n$ est une fonction de $Z_1, \dots, Z_n$, et elle est donc indépendante de $Z_{n+1}$. La probabilité conditionnelle que nous calculons s'exprime donc comme suit :

$$\mathbb{P}(Y_{n+1} = x | Y_0, \dots, Y_n) = \mathbb{E}[\mathbf{1}_{\{x\}}(Y_n + Z_{n+1} + \mathbf{1}_{\{Y_n=0\}}(1 - Z_{n+1})) | Y_0, \dots, Y_n] = h(Y_n),$$

où pour tout $y \in \mathbb{N}$, nous avons posé

$$h(y) = \mathbb{E}[\mathbf{1}_{\{x\}}(y + Z_{n+1} + \mathbf{1}_{\{y=0\}}(1 - Z_{n+1}))] = \mathbb{P}(y + Z_{n+1} + \mathbf{1}_{\{y=0\}}(1 - Z_{n+1}) = x).$$

Si $y = 0$, cette probabilité vaut 1 si $x = 1$ et 0 sinon. Si $y \geqslant 1$, cette probabilité vaut $p$ si $x = y+1$ et $q$ si $x = y - 1$. Dans tous les cas, elle vaut $R(y, x)$.

Finalement, nous avons montré que

$$\mathbb{P}(Y_{n+1} = x | Y_0, \dots, Y_n) = R(Y_n, x),$$

si bien que $(Y_n)_{n\geqslant 0}$ est une chaîne de Markov sur $\mathbb{N}$ de noyau de transition $R$.

3. On a $W_0 \leqslant Y_0$ et, pour tout $n \geqslant 1$, en utilisant le fait que $Z_n \leqslant 1$,

$$Y_n - W_n = (Y_{n-1} - W_{n-1}) + \mathbf{1}_{\{Y_{n-1}=0\}}(1 - Z_n) \geqslant Y_{n-1} - W_{n-1},$$

si bien que par récurrence sur $n$, on a presque sûrement pour tout $n \geqslant 0$ l'inégalité souhaitée.

4. Pour tout entier $i \geqslant 1$, on a $P^i(i, 0) > 0$ et $P^i(0, i) > 0$, parce que $i \to i - 1 \to \dots \to 0$ et $0 \to 1 \to \dots \to i$. La chaîne est donc irréductible.

Cherchons une mesure $\mu$ réversible pour cette chaîne telle que $\mu(1) = 1$. Une telle mesure doit satisfaire, outre $\mu(1) = 1$, les équations

$$\mu(0) = q\mu(1) \text{ et, pour tout } i \geqslant 1,\ p\mu(i) = q\mu(i+1).$$

L'unique telle mesure est donnée par

$$\mu(0) = q \text{ et, pour tout } i \geqslant 1,\ \mu(i) = \left(\tfrac{p}{q}\right)^{i-1}.$$

5. Si $p < \frac{1}{2}$, alors $\frac{p}{q} < 1$ et $\mu(\mathbb{N}) < \infty$. Nous avons donc une chaîne de Markov irréductible qui admet une mesure invariante finie (donc une mesure de probabilité invariante). Un théorème du cours assure qu'une telle chaîne de Markov est récurrente.

Si $p > \frac{1}{2}$, alors $p > q$ et $\mathbb{E}[Z_n] = p - q > 0$. La loi forte des grands nombres assure alors que

$$\frac{1}{n} W_n = \frac{1}{n}(Z_1 + \dots + Z_n) \xrightarrow[n\to\infty]{\mathbb{P}-\text{p.s.}} p - q.$$

Il s'ensuit que la suite $(W_n)_{n\geqslant 0}$ converge $\mathbb{P}$-presque sûrement vers $+\infty$. Puisqu'on a $Y_n \geqslant W_n$ pour tout $n$ presque sûrement, ceci entraîne également la convergence presque sûre de la suite $(Y_n)_{n\geqslant 0}$ vers $+\infty$. Or la suite $(Y_n)_{n\geqslant 0}$ a sous $\mathbb{P}$ la loi de la chaîne de Markov sur $\mathbb{N}$ de noyau de transition $R = P$ issue de 0, c'est-à-dire la loi de $(X_n)_{n\geqslant 0}$ sous $\mathbb{P}_0$.

Ainsi, la suite $(X_n)_{n\geqslant 0}$ converge $\mathbb{P}_0$-presque sûrement vers $+\infty$. En particulier, $\mathbb{P}_0$-presque sûrement, la suite $(X_n)_{n\geqslant 0}$ ne prend qu'un nombre fini de fois la valeur 0. L'état 0 est donc transient pour notre chaîne de Markov, et puisqu'elle est irréductible, il en est de même de tous les autres états.

Cette question (et en fait tout cet exercice) avait été traité en cours. Très peu de gens ont su dire que lorsque $p < \frac{1}{2}$, la mesure invariante déterminée à la question 4 est de masse totale finie, ce qui entraîne que la chaîne est récurrente.

## Exercice 4

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

Sur un espace d'états $E$, on se donne un noyau de transition markovien $P$. On se donne une chaîne de Markov $(\Omega, \mathscr{F}, (\mathscr{F}_n)_{n\geqslant 0}, (\mathbb{P}_x)_{x\in E}, (X_n)_{n\geqslant 0})$ sur $E$ de noyau de transition $P$. On suppose cette chaîne irréductible.

Soit $B$ une partie de $E$ qui n'est ni vide ni égale à $E$. On pose $H = \inf\{n \geqslant 0 : X_n \in B\}$ et, pour tout $x \in E$, $h(x) = \mathbb{E}_x[H]$.

**1.** Montrer que la fonction $h$ satisfait la relation

$$\forall x \in B, h(x) = 0 \ \text{ et } \ \forall x \in E \setminus B,\ h(x) = 1 + \sum_{y\in E} P(x,y)h(y).$$

**2.** Calculer la fonction $h$ dans le cas de la marche aléatoire simple symétrique sur $\mathbb{Z}$, c'est-à-dire de la chaîne de Markov sur $\mathbb{Z}$ de noyau de transition $P$ défini en posant, pour tout $i \in \mathbb{Z}$, $P(i, i+1) = P(i, i-1) = \frac{1}{2}$, et où on a posé $B = \mathbb{Z}_- = \{0, -1, -2, \ldots\}$.

**3.** On se donne deux éléments $x$ et $y$ distincts de $E \setminus B$. On pose $T = \inf\{n \geqslant 0 : X_n = y\}$ et $S = \inf\{n \geqslant T : X_n \in B\}$. Montrer que $\mathbb{E}_x[S] = \mathbb{E}_x[T] + h(y)$, puis que $h(x) \leqslant h(y) + \mathbb{E}_x[T]$.

On suppose désormais que $E$ est un ensemble fini et on note $A = E \setminus B$. Pour tous $x, y \in E$, on pose

$$m(x,y) = \mathbb{P}_x(X_H = y).$$

**4.** Calculer $m(x,y)$ lorsque $y$ appartient à $A$, puis lorsque $x$ appartient à $B$. Montrer que pour tous $a \in A$ et $b \in B$ on a

$$m(a,b) = \sum_{z\in E} P(a,z)m(z,b).$$

**5.** On fixe $a \in A$ et $b \in B$. Montrer que le processus $\big(m(X_{\min(H,n)}, b)\big)_{n\geqslant 0}$ est sous $\mathbb{P}_a$ une martingale par rapport à la filtration $(\mathscr{F}_n)_{n\geqslant 0}$.

## Solution de l'exercice 4

1. Soit $x \in E$. Si $x \in B$, alors $H = 0$ $\mathbb{P}_x$-presque sûrement, donc $h(x) = 0$. Supoosons maintenant que $x \notin B$. Alors on a $H \geqslant 1$. Plus précisément, en définissant la variable aléatoire $\widehat{H} = \inf\{n \geqslant 0 : \widehat{X}_n \in B\}$ sur l'espace canonique $E^{\mathbb{N}}$, on a sur $\Omega$ l'égalité $\mathbb{P}_x$-presque sûre de variables aléatoires

$$H = 1 + \widehat{H}(\theta(X)).$$

En appliquant la propriété de Markov (faible) au temps 1 à cette égalité, on trouve

$$\mathbb{E}_x[H] = 1 + \mathbb{E}_x[\mathbb{E}_x[\widehat{H}(\theta(X))|\mathscr{F}_1]] = 1 + \mathbb{E}_x[\mathbb{E}_{X_1}[\widehat{H}(X)]] = 1 + \mathbb{E}_x[\mathbb{E}_{X_1}[H]].$$

Or

$$\mathbb{E}_{X_1}[H] = \sum_{y\in E} \mathbf{1}_{\{X_1=y\}}\mathbb{E}_y[H] = \sum_{y\in E} \mathbf{1}_{\{X_1=y\}} h(y).$$

Nous trouvons donc, grâce au théorème de convergence monotone,

$$h(x) = \mathbb{E}_x[H] = 1 + \sum_{y\in E} \mathbb{P}_x(X_1 = y)h(y) = 1 + \sum_{y\in E} P(x,y)h(y),$$

ce qui est l'égalité cherchée.

2. La fonction $h$ est nulle sur $\mathbb{Z}_-$ et satisfait, pour tout $i \geqslant 1$, la relation

$$h(i) = 1 + \tfrac{1}{2}(h(i-1) + h(i+1)).$$

La fonction $h$ prend, par définition, ses valeurs dans $[0, +\infty]$. Supposons que $h(1)$ soit différent de $+\infty$. Alors il découle de la relation ci-dessus que $h$ prend des valeurs finies sur $\mathbb{Z}$ tout entier. De plus, la relation satisfaite par $h$ peut se réécrire

$$\forall i \geqslant 1,\ h(i+1) - h(i) = h(i) - h(i-1) - 2.$$

On va montrer que cette relation impose que $h$ prend des valeurs strictement négatives, ce qui est impossible. Une manière de le faire est de résoudre explicitement cette relation. En effet, nous avons, pour tout $i \geqslant 1$,

$$h(i+1) - h(i) = h(1) - h(0) - 2i = h(1) - 2i,$$

si bien que pour tout $i \geqslant 1$,

$$\begin{aligned} h(i+1) &= h(0) + (h(1) - h(0)) + \ldots + (h(i+1) - h(i)) \\ &= 0 + h(1) + (h(1) - 2) + \ldots + (h(1) - 2i) \\ &= (i+1)h(1) - i(i+1), \end{aligned}$$

ou encore, pour tout $i \geqslant 0$,

$$h(i) = -i^2 + i(h(1) + 1).$$

Cette expression est négative pour $i$ suffisamment grand, ce qui est absurde. Ainsi, $h(1) = +\infty$.

La relation satisfaite par $h$ écrite au début de cette solution entraîne que $h(i) \geqslant \frac{1}{2}h(i-1)$, si bien que par récurrence, $h(i) = +\infty$ pour tout $i \geqslant 1$.

3. Les variables aléatoires $T$ et $S$ sont des temps d'arrêt. Définissons sur l'espace canonique $E^{\mathbb{N}}$ les temps d'arrêt

$$\widehat{T} = \inf\{n \geqslant 0 : \widehat{X}_n = y\} \quad \text{et} \quad \widehat{S} = \inf\{n \geqslant \widehat{T} : \widehat{X}_n \in B\},$$

de sorte que sur $\Omega$, on ait les relations $T = \widehat{T}(X)$ et $S = \widehat{S}(X)$.

Le temps $S$ est le premier temps d'atteinte de $B$ après le premier temps d'atteinte de $y$. On a donc la relation

$$S = T + \widehat{H}(\theta_T(X))\mathbf{1}_{\{T<\infty\}}.$$

En appliquant la propriété de Markov forte au temps $T$, on trouve

$$\begin{aligned} \mathbb{E}_x[S] &= \mathbb{E}_x[T] + \mathbb{E}_x\big[\mathbb{E}_x\big[\widehat{H}(\theta_T(X))\mathbf{1}_{\{T<\infty\}}\big|\mathscr{F}_T\big]\big] \\ &= \mathbb{E}_x[T] + \mathbb{E}_x\big[\mathbb{E}_y[H]\mathbf{1}_{\{T<\infty\}}\big] \\ &= \mathbb{E}_x[T] + \mathbb{P}_x(T<\infty)h(y). \end{aligned}$$

Si $T$ est fini $\mathbb{P}_x$-presque sûrement, c'est l'égalité cherchée. Si par contre $\mathbb{P}_x(T=\infty) > 0$, alors on a $\mathbb{E}_x[T] = \infty$, et puisque $S \geqslant T$, on a aussi $\mathbb{E}_x[S] = \infty$. Puisque $h(y) \geqslant 0$, l'égalité cherchée fait sens et reste vraie dans ce cas.

Pour démontrer la dernière inégalité, on observe que $S$, qui est le premier temps d'atteinte de $B$ *après le premier temps de passage en* $y$, est supérieur ou égal au premier temps d'atteinte de $B$, qui est $H$. Ainsi, $\mathbb{P}_x$-presque sûrement, $H \geqslant S$. En prenant l'espérance, on trouve

$$h(x) \leqslant \mathbb{E}_x[S] = \mathbb{E}_x[T] + h(y).$$

4. Le fait que $E$ soit fini entraîne, puisque la chaîne est irréductible, qu'elle est récurrente. Ainsi, pour tout $x \in E$ et tout $b \in B$, le temps d'atteinte de $b$ partant de $x$ est fini presque sûrement. Le temps $H$, qui est inférieur ou égal à ce temps d'atteinte, est donc lui aussi fini $\mathbb{P}_x$-presque sûrement. La définition de $m$ a donc un sens, puisque $X_H$ est bien définie sous $\mathbb{P}_x$.

Par définition de $H$, on a $\mathbb{P}_x(X_H \in B) = 1$, donc si $y \in A$, on a $m(x,y) = 0$.

Si $x \in B$, on a $\mathbb{P}_x(H = 0) = 1$, donc $\mathbb{P}_x(X_H = x) = 1$, donc $m(x,x) = 1$ et $m(x,y) = 0$ pour tout $y \neq x$.

Donnons-nous maintenant $a \in A$ et $b \in B$. En particulier, $a \notin B$, et $\mathbb{P}_a(H \geqslant 1) = 1$. On a donc $\mathbb{P}_a$-presque sûrement l'égalité

$$\mathbf{1}_{\{X_H=b\}} = \mathbf{1}_{\{\widehat{X}_{\widehat{H}}=b\}}(\theta(X)).$$

La propriété de Markov faible au temps 1 nous donne donc

$$\mathbb{P}_a(X_H = y) = \mathbb{E}_x\big[\mathbb{E}_{X_1}[X_H = b]\big],$$

ce qui s'écrit
$$m(a,b) = \sum_{z\in E} P(a,z)m(z,b).$$

5. Notons $f : E \to \mathbb{R}$ la fonction définie par $f(x) = m(x,b)$. Pour tout $n \geqslant 0$, notons $M_n = f(X_{n\wedge H})$.

La fonction $f$ est positive et majorée par 1, donc chaque variable aléatoire $M_n$ est bornée par 1, donc intégrable.

De plus, pour tout $n \geqslant 0$, la variable aléatoire $X_{n\wedge T}$ est mesurable par rapport à $\mathscr{F}_n$, car un processus adapté reste adapté lorsqu'on l'arrête à un temps d'arrêt.

Donnons-nous $n \geqslant 0$ et calculons $\mathbb{E}[M_{n+1}|\mathscr{F}_n]$. En utilisant le fait que $H$ est un temps d'arrêt, le fait que le processus $X$ arrêté au temps $H$ est adapté, et la relation satisfaite par $m$, on trouve

$$\begin{aligned}
\mathbb{E}[M_{n+1}|\mathscr{F}_n] &= \mathbb{E}[f(X_{(n+1)\wedge H})|\mathscr{F}_n] \\
&= \mathbb{E}[f(X_{(n+1)\wedge H})\mathbf{1}_{\{H\geqslant n+1\}}|\mathscr{F}_n] + \mathbb{E}[f(X_{(n+1)\wedge H})\mathbf{1}_{\{H\leqslant n\}}|\mathscr{F}_n] \\
&= \mathbb{E}[f(X_{n+1})|\mathscr{F}_n]\mathbf{1}_{\{H\geqslant n+1\}} + \mathbb{E}[f(X_{n\wedge H})|\mathscr{F}_n]\mathbf{1}_{\{H\leqslant n\}} \\
&= \mathbb{E}[f(X_{n+1})|\mathscr{F}_n]\mathbf{1}_{\{H\geqslant n+1\}} + f(X_{n\wedge H})\mathbf{1}_{\{H\leqslant n\}} \\
&= \sum_{z\in E} P(X_n,z)f(z)\mathbf{1}_{\{H\geqslant n+1\}} + f(X_{n\wedge H})\mathbf{1}_{\{H\leqslant n\}} \\
&= f(X_n)\mathbf{1}_{\{H\geqslant n+1\}} + f(X_{n\wedge H})\mathbf{1}_{\{H\leqslant n\}} \\
&= f(X_{n\wedge H}) \\
&= M_n.
\end{aligned}$$

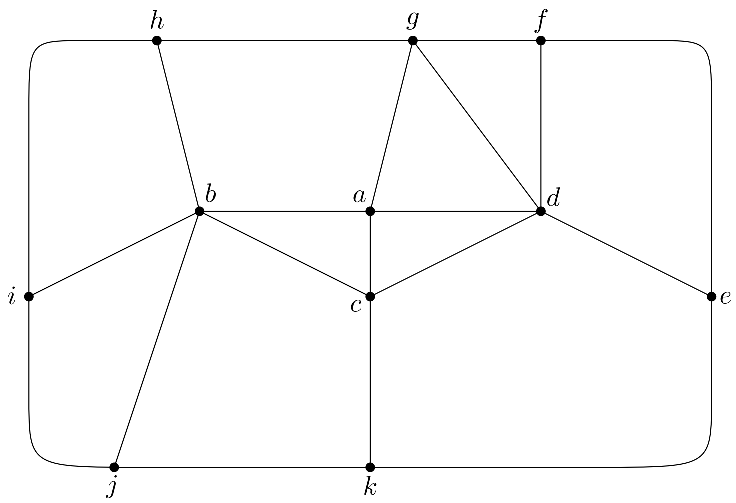
Le processus $(M_n)_{n\geqslant 0}$ est donc bien une martingale.

Cet exercice a été plutôt bien traité dans une proportion assez élevée des copies, d'une manière qui montrait une bonne compréhension de la propriété de Markov. La cinquième question, un peu plus délicate, a été moins traitée que les autres.

## Exercice 5

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

On considère la marche aléatoire sur le graphe ci-dessous, inspiré d'un plan de métro.




**1.** Déterminer toutes les probabilités invariantes pour cette marche aléatoire.
**2.** Partant du sommet $e$, combien de temps met la marche, en moyenne, à revenir en $e$ ?
**3.** Entre deux visites en $g$, combien de fois la marche visite-t-elle, en moyenne, le sommet $k$ ?
**4.** Notons $(X_n)_{n\geqslant 0}$ la marche aléatoire. La quantité

$$\frac{1}{n}\sum_{i=0}^{n-1} \mathbf{1}_{\{a,b,c,d\}}(X_i)$$

admet-elle sous $\mathbb{P}_f$ une limite presque sûre lorsque $n$ tend vers l'infini, et si oui, laquelle ?

**5.** Soit $u : \{a, b, c, d, e, f, g, h, i, j, k\} \to \mathbb{R}$ une fonction. La quantité $\mathbb{E}_a[u(X_n)]$ admet-elle une limite lorsque $n$ tend vers l'infini, et si oui, laquelle ?

——— FIN DU SUJET ———

**Solution de l'exercice 5**

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, au moins 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 est donnée par

$$\mu = (\mu_a, \mu_b, \mu_c, \mu_d, \mu_e, \mu_f, \mu_g, \mu_h, \mu_i, \mu_j, \mu_k) = (4, 5, 4, 5, 3, 3, 4, 3, 3, 3, 3).$$

La masse totale de $\mu$ est 40, si bien qu'elle se normalise en l'unique probabilité invariante

$$\pi = (\tfrac{1}{10}, \tfrac{1}{8}, \tfrac{1}{10}, \tfrac{1}{8}, \tfrac{3}{40}, \tfrac{3}{40}, \tfrac{1}{10}, \tfrac{3}{40}, \tfrac{3}{40}, \tfrac{3}{40}, \tfrac{3}{40}).$$

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

$$\mathbb{E}_e[T_e] = \frac{1}{\pi(e)} = \frac{40}{3}.$$

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

$$\nu(k) = \frac{\pi(k)}{\pi(g)} = \frac{3}{4}.$$

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

$$\pi(\{a, b, c, d\}) = \frac{9}{20}.$$

5. Il est possible pour la marche aléatoire issue de $a$ de revenir en $a$ en un temps 2 (en faisant par exemple $a, b, a$) et en un temps 3 (en faisant par exemple $a, b, c, a$). Puisque 2 et 3 sont premiers entre eux, la chaîne est apériodique, et le théorème de convergence vers l'équilibre assure que l'espérance considérée converge, lorsque $n$ tend vers l'infini, vers

$$\int u \, \mathrm{d}\pi = \frac{1}{40}\big(4u(a)+5u(b)+4u(c)+5u(d)+3u(e)+3u(f)+4u(g)+3u(h)+3u(i)+3u(j)+3u(k)\big).$$

À part les quelques erreurs habituelles de comptage pour déterminer la bonne normalisation de la mesure dans la première question, cet exercice attendu a été très bien traité dans une grande partie des copies.

Pour les années prochaines, voici une petite remarque qui peut être utile : la masse totale de la mesure qui à chaque sommet associe son nombre de voisins est égale à deux fois le nombre d'arêtes du graphe (Pourquoi ?). Ici, le graphe a 20 arêtes, et la masse totale de la mesure était 40.

**Annexe.** Quelques réponses à la première question de l'exercice 1.