Université de Sherbrooke
Faculté des sciences
Département de mathématiques

Processus stochastiques
STT489

Examen final — Solutions
Trimestre d'été 2024

Question 1 **4 + 4 + 4 = 12 points**

Soit X et Y deux variables aléatoires indépendantes telles que $X \sim \text{Exp}(\lambda)$ et $Y \sim \text{Exp}(\mu)$.

a) Montrer que X est sans mémoire. Autrement dit, montrer que pour tous $x, y \in \mathbb{R}_+$,

$$P(X > x + y \mid X > y) = P(X > x).$$

Directement,

$$\begin{aligned} P(X > x + y \mid X > y) &= \frac{P(X > x + y, X > y)}{P(X > y)} \\ &= \frac{P(X > x + y)}{P(X > y)} \\ &= \frac{e^{-\lambda(x+y)}}{e^{-\lambda y}} \\ &= e^{-\lambda x} \\ &= P(X > x) \end{aligned}$$

b) Montrer que $P(X < Y) = \lambda/(\lambda + \mu)$.

En conditionnant sur X, on obtient

$$\begin{aligned} P(X < Y) &= \int_0^\infty P(X < Y \mid X = x) \cdot f_X(x)\, dx \\ &= \int_0^\infty P(Y > x \mid X = x) \cdot f_X(x)\, dx \\ &= \int_0^\infty P(Y > x) \cdot f_X(x)\, dx \quad \text{(car } X \text{ et } Y \text{ sont indépendantes)} \\ &= \int_0^\infty e^{-\mu x} \cdot \lambda \cdot e^{-\lambda x}\, dx \\ &= \lambda \cdot \int_0^\infty e^{-(\lambda+\mu)x}\, dx \\ &= \lambda \cdot \left. \frac{e^{-(\lambda+\mu)x}}{-(\lambda + \mu)} \right|_{x=0}^{x=\infty} \\ &= \frac{\lambda}{\lambda + \mu}. \end{aligned}$$

c) Soit $Z = \min\{X, Y\}$. Montrer que $Z \sim \text{Exp}(\lambda + \mu)$. Indice : Calculer la fonction de répartition ou la fonction de survie de Z.

La fonction de survie de Z est

$$\begin{aligned} S_Z(t) &= P(Z > t) \\ &= P(X > t, Y > t) \\ &= P(X > t) \cdot P(Y > t) \quad \text{(car } X, Y \text{ indépendants)} \\ &= e^{-\lambda t} \cdot e^{-\mu t} \\ &= e^{-(\lambda+\mu)t}, \end{aligned}$$

qui est la fonction de survie d'une exponentielle de paramètre $\lambda + \mu$.

Question 2 **6 + 3 + 3 = 12 points**

a) Soit $\{X_t\}_{t\in\mathbb{R}_+}$ un processus de naissance et de mort, c'est-à-dire une chaîne de Markov à temps continu ayant $E = \mathbb{N}$ pour espace d'états et tel que

- $p_{i,j} = 0$ lorsque $j \notin \{i-1, i+1\}$
- lorsque $X_t = n$, le temps avant la prochaine arrivée suit une loi exponentielle de paramètre λ_n et est indépendante du temps avant le prochain départ, qui, lui, suit une loi exponentielle de paramètre μ_n.

Montrer que les probabilités limites p_j d'un processus de naissance et de mort, lorsqu'elles existent, sont

$$p_0 = \left(1 + \sum_{k=1}^{\infty}\left(\prod_{i=0}^{k-1}\frac{\lambda_i}{\mu_{i+1}}\right)\right)^{-1},$$

$$p_n = p_0 \cdot \prod_{i=0}^{n-1}\frac{\lambda_i}{\mu_{i+1}} \quad \text{pour tout } n \in \mathbb{N}^*.$$

Selon l'égalité des taux, on a

$$\lambda_0 p_0 = \mu_1 p_1,$$

$$\lambda_n p_n + \mu_n p_n = \lambda_{n-1}p_{n-1} + \mu_{n+1}p_{n+1} \text{ pour tout } n \in \mathbb{N}^*$$

En additionnant les équations 1 à n, on obtient

$$\lambda_{n-1}p_{n-1} = \mu_n p_n.$$

Ainsi,

$$p_n = \frac{\lambda_{n-1}}{\mu_n}p_{n-1} = \frac{\lambda_{n-1}\lambda_{n-2}}{\mu_n\mu_{n-1}}p_{n-2} = \cdots = \left(\prod_{i=0}^{n-1}\frac{\lambda_i}{\mu_{i+1}}\right)p_0$$

Or,

$$\begin{aligned}
\sum_{k=0}^{\infty}p_k = 1 &\Longrightarrow p_0 + \sum_{k=1}^{\infty}p_k = 1 \\
&\Longrightarrow p_0 + \sum_{k=1}^{\infty}\left(\left(\prod_{i=0}^{k-1}\frac{\lambda_i}{\mu_{i+1}}\right)p_0\right) = 1 \\
&\Longrightarrow p_0\left(1 + \sum_{k=1}^{\infty}\left(\prod_{i=0}^{k-1}\frac{\lambda_i}{\mu_{i+1}}\right)\right) = 1 \\
&\Longrightarrow p_0 = \left(1 + \sum_{k=1}^{\infty}\left(\prod_{i=0}^{k-1}\frac{\lambda_i}{\mu_{i+1}}\right)\right)^{-1}
\end{aligned}$$

b) Soit $\{X_t\}_{t\in\mathbb{R}_+}$ un processus de naissance et de mort où $\lambda_n = \lambda$ et $\mu_n = n\mu$, avec $\lambda, \mu \in \mathbb{R}_+^*$.

i) Démontrer que les probabilités limites p_i existent.

Les probabilités limites existent si

$$\sum_{k=1}^{\infty}\left(\prod_{i=0}^{k-1}\frac{\lambda_i}{\mu_{i+1}}\right)$$

converge. Directement,

$$\begin{aligned}
\sum_{k=1}^{\infty}\left(\prod_{i=0}^{k-1}\frac{\lambda_i}{\mu_{i+1}}\right) &= \sum_{k=1}^{\infty}\left(\prod_{i=0}^{k-1}\frac{\lambda}{(i+1)\mu}\right)\\
&= \sum_{k=1}^{\infty}\frac{\lambda^k}{k!\,\mu^k}\\
&= \sum_{k=1}^{\infty}\frac{(\lambda/\mu)^k}{k!}\\
&= e^{\lambda/\mu} - 1.
\end{aligned}$$

ii) Déterminer les probabilités limites pour chacun des états si $\lambda = \mu = 1$.

On a

$$\begin{aligned}
p_0 &= \left(1 + \sum_{k=1}^{\infty}\left(\prod_{i=0}^{k-1}\frac{\lambda_i}{\mu_{i+1}}\right)\right)^{-1}\\
&= \left(1 + \sum_{k=1}^{\infty}\left(\prod_{i=0}^{k-1}\frac{1}{i+1}\right)\right)^{-1}\\
&= \left(1 + \sum_{k=1}^{\infty}\left(\frac{1}{k!}\right)\right)^{-1}\\
&= \big(1 + (e-1)\big)^{-1}\\
&= \frac{1}{e}
\end{aligned}$$

et

$$\begin{aligned}
p_n &= p_0 \cdot \prod_{i=0}^{n-1}\frac{\lambda_i}{\mu_{i+1}}\\
&= \frac{1}{e}\cdot\prod_{i=0}^{n-1}\frac{1}{i+1}\\
&= \frac{1}{e\cdot n!}
\end{aligned}$$

Question 3 **3 + 3 + 3 + 3 = 12 points**

Aujourd'hui, entre 6h et 21h, des visiteurs se présenteront au parc national du Mont-Mégantic selon un processus de Poisson $\{N_t\}_{t\in\mathbb{R}_+}$ d'intensité $\lambda = 16$ visiteurs par heure, où t représente le temps écoulé depuis 6h du matin (autrement dit, $t = 0$ représente 6h du matin).

a) Déterminer la probabilité qu'exactement 30 visiteurs se présentent au parc entre 13h et 15h30.

Directement,

$$\begin{aligned} P\big(N_{9,5} - N_7 = 30\big) &= P(\text{Poisson}(2{,}5 \cdot 16) = 30) \\ &= \frac{e^{-40} \cdot (40)^{30}}{30!} \approx 0{,}0185. \end{aligned}$$

b) En supposant qu'il y a eu au moins 21 visiteurs dans la journée, déterminer la probabilité que le temps écoulé entre l'arrivée du 20e visiteur et du 21e visiteur soit de plus de 10 minutes.

Soit T_i le temps entre l'arrivée des visiteurs $i-1$ et i. On sait que $T_i \sim \text{Exp}(16)$. Ainsi,

$$\begin{aligned} P(T_{21} > 1/6) &= P(\text{Exp}(16) > 1/6) \\ &= e^{-16/6} \approx 0{,}0695. \end{aligned}$$

c) On suppose que la probabilité qu'un visiteur soit âgé de plus de 60 ans est de $p = 10$ %. Quelle est la probabilité qu'entre 10h et 11h, il y ait à la fois exactement 2 visiteurs âgés de plus de 60 ans et exactement 10 visiteurs âgés de 60 ans et moins ?

Soit $N1_t$ le nombre de visiteurs de plus de 60 ans arrivées au temps t et soit $N2_t$ le nombre de visiteurs de 60 ans et moins arrivées au temps t. Ainsi, $\{N1_t\}_{t\in\mathbb{R}_+} \sim PP(\lambda p = 1{,}6)$ et $\{N2_t\}_{t\in\mathbb{R}_+} \sim PP(\lambda(1-p) = 14{,}4)$ et ces processus sont indépendants. Ainsi,

$$\begin{aligned} P\Big(\underbrace{N1_5 - N1_4}_{\sim\text{Poisson}(1,6)} = 2\Big) &= \frac{e^{-1,6} \cdot 1{,}6^2}{2!} \\ &= \frac{16}{25} \cdot e^{-1,6} \approx 0{,}2584. \end{aligned}$$

et

$$P\Big(\underbrace{N2_5 - N2_4}_{\sim\text{Poisson}(14,4)} = 10\Big) = \frac{e^{-14,4} \cdot 14{,}4^{10}}{10!} \approx 0{,}0589.$$

d'où

$$\begin{aligned} P(N1_5 - N1_4 = 1, N2_5 - N2_4 = 10) &= P(N1_5 - N1_4 = 1) \cdot P(N2_5 - N2_4 = 10) \\ &= \frac{16}{25} \cdot e^{-1,6} \cdot \frac{e^{-14,4} \cdot 14{,}4^{10}}{10!} \approx 0{,}0152. \end{aligned}$$

d) Soit $s, t \in [0, 15]$ avec $s < t$ et soit $k, n \in \mathbb{N}$ avec $k \leq n$. Calculer $P(N_s = k \mid N_t = n)$.

Directement,

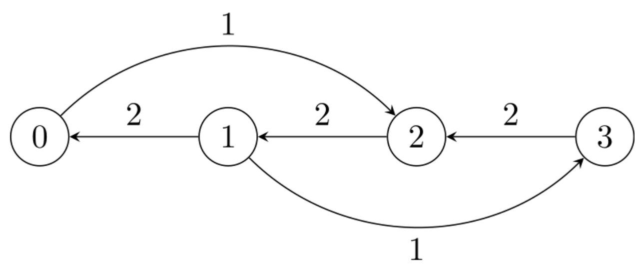
$$\begin{aligned}
P(N_s = k \mid N_t = n) &= \frac{P(N_s = k, N_t = n)}{P(N_t = n)} \\
&= \frac{P(N_t = n \mid N_s = k) \cdot P(N_s = k)}{P(N_t = n)} \\
&= \frac{P(N_t - N_s = n - k \mid N_s = k) \cdot P(N_s = k)}{P(N_t = n)} \\
&= \frac{P(N_{t-s} = n - k) \cdot P(N_s = k)}{P(N_t = n)} \\
&= \frac{\frac{e^{-16(t-s)} \cdot \left(16(t-s)\right)^{n-k}}{(n-k)!} \cdot \frac{e^{-16s} \cdot (16s)^k}{k!}}{\frac{e^{-16t} \cdot (16t)^n}{n!}} \\
&= \frac{n!}{k! \cdot (n-k)!} \cdot \left(\frac{(t-s)^{n-k} s^k}{t^{n-k} t^k}\right) \\
&= \binom{n}{k} \cdot \left(\frac{s}{t}\right)^k \cdot \left(\frac{t-s}{t}\right)^{n-k}.
\end{aligned}$$

Question 4 **2 + 8 + 2 = 12 points**

Un petit magasin d'informatique dispose d'un espace pour exposer jusqu'à trois ordinateurs à vendre. Les clients viennent à certains moments selon un processus de Poisson avec un taux de 2 par semaine pour acheter un ordinateur et en achèteront un si au moins un est disponible (si aucun n'est disponible, le client quitte le magasin sans effectuer d'achat). Lorsque le magasin n'a plus qu'un seul ordinateur, il passe une commande pour deux ordinateurs supplémentaires. La commande prend un temps exponentiellement distribué avec une moyenne de 1 semaine pour arriver. Bien entendu, pendant que le magasin attend la livraison, les ventes peuvent ramener le stock à 1 ordinateur, puis à 0.

a) Modéliser cette situation à l'aide d'une chaîne de Markov à temps continu. Votre modélisation doit comprendre la définition des variables, l'identification des états puis des taux de transition de l'état i vers l'état j (ce qui peut être présenté à l'aide à l'aide d'un diagramme).

Soit X_t le nombre d'ordinateurs à vendre au temps t, où $E = \{0, 1, 2, 3\}$. Alors $\{X_t\}_{t \in \mathbb{R}_+}$ est une chaîne de Markov à temps continu ayant pour diagramme de transition



b) Déterminer les probabilités limites pour chacun des états de ce processus.

Pour chaque état, le taux de départ et le taux d'arrivée est

État	**Taux de sortie**	**Taux d'arrivée**
0	p_0	$2p_1$
1	$3p_1$	$2p_2$
2	$2p_2$	$p_0 + 2p_3$
3	$2p_3$	p_1

Selon le principe d'égalité des taux, on obtient

$$p_0 = 2p_1, \quad p_2 = \frac{3}{2}p_1, \quad p_3 = \frac{1}{2}p_1.$$

Par ailleurs,

$$\begin{aligned} p_0 + p_1 + p_2 + p_3 = 1 &\Longrightarrow p_1\left(2 + 1 + \frac{3}{2} + \frac{1}{2}\right) = 1 \\ &\Longrightarrow p_1 = \frac{1}{5}, \end{aligned}$$

d'où

$$p_0 = \frac{2}{5}, \quad p_1 = \frac{1}{5}, \quad p_2 = \frac{3}{10}, \quad p_3 = \frac{1}{10}.$$

c) À quel taux le magasin effectue-t-il des ventes ?

La probabilité qu'un client arrivant au magasin achète un ordinateur est $1 - p_0 = 0{,}6$. Puisque l'arrivée des clients suit un taux de 2 par semaine, alors le taux des ventes des ordinateurs est $2 \cdot 0{,}6 = 1{,}2$ ventes par semaine.

Question 5 **6 points**

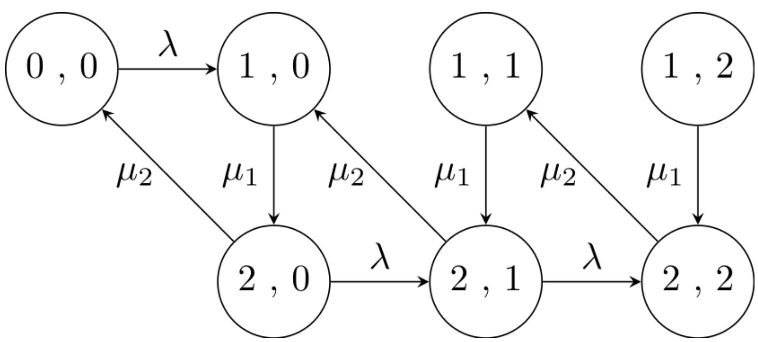
Un petit atelier de réparation de vélo ne comprenant qu'un mécanicien dispose de deux postes de travail : le poste 1 pour le diagnostic et le poste 2 pour la réparation. Lorsque le mécanicien prend en charge un client, il va d'abord effectuer un diagnostic, puis il va immédiatement procéder à la réparation. Le temps de réalisation d'un diagnostic suit une loi exponentielle de paramètre μ_1, puis le temps de réparation suit une loi exponentielle de paramètre μ_2. Les arrivées des clients potentiels suivent un processus de Poisson de paramètre λ. Il faut toutefois noter qu'un client potentiel va entrer dans la boutique seulement si le poste 1 est libre et qu'il y a moins de 4 clients dans le magasin. Autrement dit, un client potentiel ne va pas entrer dans l'atelier si le mécanicien est en train d'effectuer un diagnostic, il ne va pas entrer non plus s'il y a déjà 4 clients dans le magasin.

Modéliser cette situation à l'aide d'une chaîne de Markov à temps continu. Votre modélisation doit comprendre la définition des variables, l'identification des états puis des taux de transition de l'état i vers l'état j (ce qui peut être présenté à l'aide à l'aide d'un diagramme).

Les différents états possibles sont :

- $(0,0)$: L'atelier est vide.
- $(1,0)$: L'atelier a un client au poste 1 et aucun en attente.
- $(2,0)$: L'atelier a un client au poste 2 et aucun en attente.
- $(1,1)$: L'atelier a un client au poste 1 et un client en attente.
- $(2,1)$: L'atelier a un client au poste 2 et un client en attente.
- $(1,2)$: L'atelier a un client au poste 1 et deux clients en attente. (état superflu, n'est pas atteint)
- $(2,2)$: L'atelier a un client au poste 2 et deux clients en attente.

Soit X_t l'état de cet atelier au temps t. Le diagramme de transition de la chaîne de Markov $\{X_t\}_{t\in\mathbb{R}_+}$ est



Question 6 **8 points**

Soit $\{X_t\}_{t\in\mathbb{R}_+}$ une chaine de Markov homogène à temps continu ayant E pour espace d'états et soit $p_{i,j}(t)$ ses probabilités de transition. Montrer que, pour tout $s, t > 0$,

$$p_{i,j}(s+t) = \sum_{k\in E} p_{i,k}(s)p_{k,j}(t)$$

Directement,

$$\begin{aligned}
&p_{i,j}(s+t) \\
&= P(X_{s+t} = j \mid X_0 = i) \\
&= \sum_{k\in E} P(X_{s+t} = j, X_s = k \mid X_0 = i) \\
&= \sum_{k\in E} P(X_{s+t} = j \mid X_s = k, X_0 = i) \cdot P(X_s = k \mid X_0 = i) \\
&= \sum_{k\in E} P(X_{s+t} = j \mid X_s = k) \cdot P(X_s = k \mid X_0 = i) \quad \text{(propriété markovienne)} \\
&= \sum_{k\in E} P(X_t = j \mid X_0 = k) \cdot p_{i,k}(s) \quad \text{(chaine homogène)} \\
&= \sum_{k\in E} p_{k,j}(t) \cdot p_{i,k}(s) \\
&= \sum_{k\in E} p_{i,k}(s) \cdot p_{k,j}(t)
\end{aligned}$$

Question 7 **2 + 2 + 4 + 6 = 14 points**

Des joueurs de l'équipe de ultimate frisbee Amatheus ont remarqué que la probabilité de remporter une partie dépend de leur série de victoires ou de défaites en cours. En effet, la probabilité de remporter la prochaine partie est de :

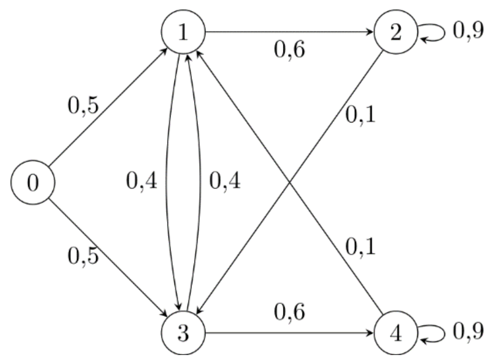
— 50 % si elle n'a aucune série en cours (première partie),
— 60 % si elle a une série d'une victoire consécutive,
— 90 % si elle a une série de deux victoires consécutives ou plus,
— 40 % si elle a une série d'une défaite consécutive,
— 10 % si elle a une série de deux défaites consécutives ou plus.

Soit X_n : « Série de victoires ou de défaites en cours après n parties », où

$$X_n = \begin{cases} 0 & \text{si l'équipe n'a aucune série après } n \text{ parties} \\ 1 & \text{si l'équipe a une série d'une victoire consécutive après } n \text{ parties} \\ 2 & \text{si l'équipe a une série de deux victoires consécutives ou plus après } n \text{ parties} \\ 3 & \text{si l'équipe a une série d'une défaite consécutive après } n \text{ parties} \\ 4 & \text{si l'équipe a une série de deux défaites consécutives ou plus après } n \text{ parties} \end{cases}$$

Notons que $X_0 = 0$. La chaîne de Markov $\{X_n\}_{n\in\mathbb{N}}$ a pour espace d'états $E = \{0, 1, 2, 3, 4\}$. Voici sa matrice de transition et son diagramme de transition.

$$P = \begin{bmatrix} 0 & 0{,}5 & 0 & 0{,}5 & 0 \\ 0 & 0 & 0{,}6 & 0{,}4 & 0 \\ 0 & 0 & 0{,}9 & 0{,}1 & 0 \\ 0 & 0{,}4 & 0 & 0 & 0{,}6 \\ 0 & 0{,}1 & 0 & 0 & 0{,}9 \end{bmatrix}$$



a) Quelle est la période de l'état 1 ?

Puisque $p_{1,1}^{(2)} = 0{,}16 > 0$ et $p_{1,1}^{(3)} = 0{,}048 > 0$, alors la période de l'état 1 est $d(1) = \text{pgcd}\{2, 3, \ldots\} = 1$.

b) Quelle est la probabilité que l'équipe Amatheus remporte ses quatre premières parties ?

La probabilité que l'équipe Amatheus remporte ses cinq premières parties est

$$0{,}5 \cdot 0{,}6 \cdot 0{,}9^2 = 0{,}243.$$

c) Après la partie 20, l'équipe Amatheus a une série d'une défaite consécutive. Quelle est la probabilité qu'après la partie 24, elle ait une série d'au moins deux victoires consécutives ?

On cherche $P(X_{24} = 2 \mid X_{20} = 3)$. Voici les chemins menant de l'état 3 à l'état 2 en 4 transitions, ainsi que les probabilités associées :

— $3 \to 1 \to 2 \to 2 \to 2$: $0{,}4 \cdot 0{,}6 \cdot 0{,}9 \cdot 0{,}9 = 0{,}1944$
— $3 \to 1 \to 3 \to 1 \to 2$: $0{,}4 \cdot 0{,}4 \cdot 0{,}4 \cdot 0{,}6 = 0{,}0384$
— $3 \to 4 \to 1 \to 2 \to 2$: $0{,}6 \cdot 0{,}1 \cdot 0{,}6 \cdot 0{,}9 = 0{,}0324$
— $3 \to 4 \to 4 \to 1 \to 2$: $0{,}6 \cdot 0{,}9 \cdot 0{,}1 \cdot 0{,}6 = 0{,}0324$

Ainsi,

$$P(X_{24} = 2 \mid X_{20} = 3) = 0{,}1944 + 0{,}0384 + 0{,}0324 + 0{,}0324 = 0{,}2976.$$

d) Une saison comporte un très grand nombre de parties. Quelle proportion du temps l'équipe Amatheus aura-t-elle une série d'au moins 2 victoires consécutives ? Autrement dit, calculer la probabilité limite $\pi_4 = \lim_{n\to\infty} p_{i,j}^{(n)}$. Pour simplifier les calculs, noter par symétrie que $\pi_1 = \pi_3$ et $\pi_2 = \pi_4$.

On cherche à résoudre le système

$$\pi_j = \sum_{i=1}^{4} \pi_i p_{i,j} \ \text{ pour } j \in \{1,2,3,4\} \quad \text{et} \quad \sum_{i=1}^{4} \pi_i = 1$$

Les équations sont

① $\pi_1 = 0{,}4\pi_3 + 0{,}1\pi_4$
② $\pi_2 = 0{,}6\pi_1 + 0{,}9\pi_2$
③ $\pi_3 = 0{,}4\pi_1 + 0{,}1\pi_2$
④ $\pi_4 = 0{,}6\pi_3 + 0{,}9\pi_4$
⑤ $\pi_1 + \pi_2 + \pi_3 + \pi_4 = 1$

Par symétrie, $\pi_1 = \pi_3$ et $\pi_2 = \pi_4$. Des équations ② et ⑤, on a

$$\pi_2 = 6\pi_1 \quad \text{et} \quad \pi_1 + \pi_2 = 0{,}5,$$

ce qui donne $\pi_1 = \pi_3 = \frac{1}{14}$ et $\pi_2 = \pi_4 = \frac{6}{14}$. Ainsi, la proportion du temps que l'équipe Amatheus aura une série d'au moins 2 victoires consécutives est $3/7 \approx 0{,}4286$.

Question 8 **6 + 6 = 12 points**

Un manuscrit est envoyé à une entreprise de révision ayant pour employés Albert, Béatrice et Carole.

- Si le manuscrit est révisé par Albert, alors le nombre d'erreurs non détectées suit variable aléatoire de Poisson avec une moyenne de 2.
- Si le manuscrit est révisé par Béatrice, alors le nombre d'erreurs non détectées suit variable aléatoire uniforme discrète variant entre 1 et 5 (en incluant les bornes).
- Si le manuscrit est révisé par Carole, alors le nombre d'erreurs non détectées suit variable aléatoire binomiale de paramètre $n = 6$ et $p = 1/2$.

Soit X le nombre d'erreurs non détectées dans le manuscrit révisé et soit Y la variable identifiant a qui a été attiré le travail de révision ($Y = 1$ pour Albert, $Y = 2$ pour Béatrice et $Y = 3$ pour Carole). Supposons que Albert a 50 % d'effectuer le travail de révision, alors que cette probabilité est de 25 % pour Béatrice et Carole.

a) Démontrer que $E(X) = 2{,}5$ en effectuant un calcul d'espérance par conditionnement sur Y.

Selon l'énoncé,

$$P(Y = 1) = 1/2\,, \quad P(Y = 2) = 1/4\,, \quad P(Y = 3) = 1/4.$$

Par ailleurs,

$$X \mid Y = 1 \sim \text{Poisson}(2) \implies E(X \mid Y = 1) = 2,\ Var(X \mid Y = 1) = 2$$
$$X \mid Y = 2 \sim \text{Unif}(1, 5) \implies E(X \mid Y = 2) = 3,\ Var(X \mid Y = 2) = 24/12 \approx 2$$
$$X \mid Y = 3 \sim \text{Bin}(6, 1/2) \implies E(X \mid Y = 3) = 3,\ \ Var(X \mid Y = 3) = 3/2 = 1{,}5$$

Directement,

$$\begin{aligned}
E(X) &= E\big(E(X \mid Y)\big) \\
&= \sum_{y=1}^{3} E(X \mid Y = y) \cdot P(Y = y) \\
&= E(X \mid Y = 1) \cdot P(Y = 1) + E(X \mid Y = 2) \cdot P(Y = 2) + E(X \mid Y = 3) \cdot P(Y = 3) \\
&= 2 \cdot \frac{1}{2} + 3 \cdot \frac{1}{4} + 3 \cdot \frac{1}{4} \\
&= \frac{5}{2} = 2{,}5.
\end{aligned}$$

b) Démontrer que $Var(X) \approx 2{,}125$ en utilisant au besoin la formule

$$Var(X) = E(Var(X \mid Y)) + Var\big(E(X \mid Y)\big).$$

On a

$$\begin{aligned}
E\big(Var(X \mid Y)\big) &= \sum_{y=1}^{3} Var(X \mid Y = y) \cdot P(Y = y) \\
&= Var(X \mid Y = 1)P(Y = 1) + Var(X \mid Y = 2)P(Y = 2) + Var(X \mid Y = 3)P(Y = 3) \\
&= 2 \cdot \frac{1}{2} + 2 \cdot \frac{1}{4} + \frac{3}{2} \cdot \frac{1}{4} \\
&= \frac{15}{8} = 1{,}875
\end{aligned}$$

et

$$\begin{aligned}
Var\big(E(X \mid Y)\big) &= \sum_{y=1}^{3} \left(E(X \mid Y = y) - \underbrace{E\big(E(X \mid Y)\big)}_{=E(X)} \right)^2 \cdot P(Y = y) \\
&= \left(2 - \frac{5}{2}\right)^2 \cdot \frac{1}{2} + \left(3 - \frac{5}{2}\right)^2 \cdot \frac{1}{4} + \left(3 - \frac{5}{2}\right)^2 \cdot \frac{1}{4} \\
&= \frac{1}{4} \cdot \frac{1}{2} + \frac{1}{4} \cdot \frac{1}{4} + \frac{1}{4} \cdot \frac{1}{4} \\
&= \frac{1}{4} = 0{,}25.
\end{aligned}$$

Ainsi,

$$\begin{aligned}
Var(X) &= E\big(Var(X \mid Y)\big) + Var\big(E(X \mid Y)\big) \\
&= \frac{15}{8} + \frac{1}{4} \\
&= \frac{17}{8} \approx 2{,}125.
\end{aligned}$$

Question 9 **6 + 6 = 12 points**

Une population est formée d'individus tous identiques qui donnent naissance à des individus de même type. Chaque individu produit dans sa vie j enfants avec probabilité p_j, où $p_0 = 0{,}4$, $p_1 = 0{,}2$, $p_2 = 0{,}2$ et $p_3 = 0{,}2$. Soit X_n le nombre d'individus de la génération $n \in \mathbb{N}$.

a) Montrer que $E(X_n) = (1{,}2)^n \cdot E(X_0)$. Note : une démonstration complète est requise (et non la simple énonciation d'un résultat vu en classe).

Le nombre moyen d'enfants par individu est

$$0 \cdot 0{,}5 + 1 \cdot 0{,}2 + 2 \cdot 0{,}2 + 3 \cdot 0{,}2 = 1{,}2.$$

Soit Z_i le nombre d'enfants de l'individu i de la génération X_n. Ainsi,

$$\begin{aligned} E(X_n) &= E\left(\sum_{i=1}^{X_{n-1}} Z_i\right) \\ &= E(Z_i) \cdot E(X_{n-1}) \quad \text{(identité de Wald)} \\ &= 1{,}2 \cdot E(X_{n-1}). \end{aligned}$$

Par récurrence, on trouve $E(X_n) = (1{,}2)^n \cdot E(X_0)$.

b) Sachant que la population comprend trois individus au temps 0, calculer la probabilité qu'elle s'éteigne.

La probabilité π que la lignée d'un individu s'éteigne est la plus petite racine positive de l'équation

$$\begin{aligned} &\pi = 0{,}4 + 0{,}2\pi + 0{,}2\pi^2 + 0{,}2\pi^3 \Longrightarrow 0{,}2\pi^3 + 0{,}2\pi^2 - 0{,}8\pi + 0{,}4 = 0 \\ &\Longrightarrow \pi^3 + \pi^2 - 4\pi + 2 = 0 \\ &\Longrightarrow (\pi - 1)(\pi^2 + 2\pi - 2) = 0 \\ &\Longrightarrow \pi \in \left\{1, \frac{-2+\sqrt{12}}{2}, \frac{-2-\sqrt{12}}{2}\right\} \end{aligned}$$

Ainsi,

$$\pi = \frac{-2+\sqrt{12}}{2} = \sqrt{3} - 1 \approx 0{,}7321.$$

Par conséquent, la probabilité qu'une population de trois individus s'éteignent est

$$\pi^3 = \left(\sqrt{3} - 1\right)^3 = 6\sqrt{3} - 10 \approx 0{,}3923.$$