Applied Probability Progress Test 2

28 November 2019

Question 1: (5 points) Let $X = (X_n)_{n\in\{0,1,2,\dots\}}$ denote a discrete-time, homogeneous Markov chain on the state space $E = \{1, 2, 3, 4, 5\}$ with transition matrix given by

$$\mathbf{P} = \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Find all possible stationary distributions.

Solution: This Markov chain has two finite and closed (and hence positive recurrent) communicating classes given by $C_1 = \{1, 3\}, C_2 = \{5\}$. [Hence we already know that the stationary distribution is not unique.] Moreover, it has a third communicating class given by $\{2, 4\}$, which is not closed and hence transient. We recall from the lectures that the elements of the stationary distribution corresponding to transient states are equal to zero. Hence $\pi_2 = \pi_4 = 0$. For the remaining components of the stationary distribution, we solve two systems of equations:

$$(\pi_1, \pi_3) \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} = (\pi_1, \pi_3) \Leftrightarrow \frac{1}{2}\pi_1 + \frac{1}{2}\pi_3 = \pi_1 \Leftrightarrow \pi_1 = \pi_3,$$

and

$$\pi_5 = 1\pi_5 \Leftrightarrow \pi_5 = \pi_5.$$

[3 marks]

Hence all stationary distributions are given by (any of the following representations):

- $\boldsymbol{\pi} = (\pi_1, 0, \pi_1, 0, \pi_5)$ for all $\pi_1, \pi_5 \geq 0$ such that $2\pi_1 + \pi_5 = 1$, or
- $\boldsymbol{\pi} = (\pi_1, 0, \pi_1, 0, 1 - 2\pi_1)$ for all $0 \leq \pi_1 \leq \frac{1}{2}$, or
- $\boldsymbol{\pi} = \left(\frac{1}{2}(1 - \pi_5), 0, \frac{1}{2}(1 - \pi_5), 0, \pi_5\right)$ for all $0 \leq \pi_5 \leq 1$, or
- $\boldsymbol{\pi} = a(\frac{1}{2}, 0, \frac{1}{2}, 0, 0) + b(0, 0, 0, 0, 1)$ for all $a, b \geq 0$ such that $a + b = 1$.

[2 marks]

Question 2: (5 points) Let $(X_n)_{n\in\{0,1,2,\dots\}}$ denote a discrete-time, homogeneous Markov chain on the state space $E = \{1, 2, 3, 4\}$ with transition matrix given by

$$\mathbf{P} = \begin{pmatrix} 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 \end{pmatrix}.$$

Is the Markov chain time-reversible? Please justify your answer carefully.

Solution: We observe that the Markov chain is irreducible since all states communicate with each other. Hence we only have one (closed) communicating class (which has finitely many elements), hence all states are positive recurrent and there exists a unique stationary distribution. Also, the transition matrix is doubly-stochastic, so we know from the problem class (Question 13, Problem Sheet 2) that the unique stationary distribution is given by $\boldsymbol{\pi} = (\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$. We know that the chain is irreducible iff the detailed balance equations hold: For any $i, j \in E$, we have $\pi_i p_{ij} = \pi_j p_{ji}$ which is indeed true here ($\frac{1}{4}\frac{1}{3} = \frac{1}{4}\frac{1}{3}$ and trivially $0 = 0$), so the chain is time-reversible.
Answer that the chain is time-reversible: [2 marks], justification: [3 marks]

Question 3: (a) (3 points) For $n \in \mathbb{N}$, let $X_1, \dots, X_n$ be independent random variables where $X_i \sim \mathrm{Exp}(\lambda_i)$, for $\lambda_i > 0$ and $i = 1, \dots, n$. Let $Y = \min\{X_1, \dots, X_n\}$. Determine the distribution of Y.

(b) (3 points) Three students are working independently on their Applied Probability homework. All three start at 8am on a certain day and each takes an exponential time with mean 6 hours to complete the homework. What is the earliest time, on average, at which all three students will have completed their homework?

Hint: If $X \sim \mathrm{Exp}(\lambda)$ for $\lambda > 0$, then the density of X is given by $f_X(x) = \lambda e^{-\lambda x}$ for $x > 0$ and 0 otherwise. Also, $\mathbb{E}(X) = 1/\lambda$.

Solution:

(a) (3 points) Consider the survival function of Y: Let $y > 0$, then

$$\begin{aligned}\mathbb{P}(Y > y) &= \mathbb{P}(\min\{X_1, \dots, X_n\} > y) = \mathbb{P}(X_1 > y, \dots, X_n > y)\\ &\overset{\text{by independence of the } X_i}{=} \prod_{i=1}^{n} \mathbb{P}(X_i > y) = \prod_{i=1}^{n} e^{-\lambda_i y} = e^{-\sum_{i=1}^{n} \lambda_i y},\end{aligned}$$

which is the survival function of a random variable with $\mathrm{Exp}(\sum_{i=1}^{n} \lambda_i)$ distribution.

(b) (3 points) We know that, if $X \sim \mathrm{Exp}(\lambda)$ for $\lambda > 0$, then $f_X(x) = \lambda e^{-\lambda x}, x > 0$ and $\mathbb{E}(X) = 1/\lambda$. Here we have that $\mathbb{E}(X) = 1/\lambda = 6$, so $\lambda = 1/6$.

Method 1: Define random variables as follows: X_i:=time it takes for the ith student to complete their homework, then X_i are independent and $\mathrm{Exp}(1/6)$ distributed for $i = 1, 2, 3$. We define $T := \max\{X_1, X_2, X_3\}$. We need to find $\mathbb{E}(T)$.

First we derive the c.d.f. of T: Let $t > 0$, then

$$\begin{aligned}F_T(t) &= \mathbb{P}(T \le t) = \mathbb{P}(\max\{X_1, X_2, X_3\} \le t) = \mathbb{P}(X_1 \le t, X_2 \le t, X_3 \le t)\\ &\overset{\text{by independence of the } X_i}{=} \mathbb{P}(X_1 \le t)\mathbb{P}(X_2 \le t)\mathbb{P}(X_3 \le t) = (1 - e^{-\lambda t})^3.\end{aligned}$$

Hence, the corresponding density is given by

$$f_T(t) = \frac{d}{dt} F_T(t) = 3\lambda e^{-\lambda t}(1 - e^{-\lambda t})^2,$$

for $t > 0$ and 0 otherwise. Then

$$\begin{aligned}\mathbb{E}(T) &= \int_0^\infty t f_T(t) dt = \int_0^\infty 3\lambda t e^{-\lambda t} dt - 3\int_0^\infty 2\lambda t e^{-2\lambda t} dt + \int_0^\infty 3\lambda t e^{-3\lambda t} dt\\ &= \frac{3}{\lambda} - \frac{3}{2\lambda} + \frac{1}{3\lambda} = \frac{11}{6\lambda} = 11[\text{hours}].\end{aligned}$$

Method 2: Alternatively, you can argue as follows: Define $T_1 = \min(X_1, X_2, X_3)$ as the time it takes for the first student to complete the homework, T_2 is defined as the additional time it takes the second student to complete the homework, and T_3 is defined as the additional time it takes the last student to complete the homework after the first two students have already finished. Then $T = T_1 + T_2 + T_3$. By (a) we know that $T_1 \sim \text{Exp}(3\lambda)$. When the first student finishes, then by the lack of memory property of the exponential distribution, the two remaining students start from fresh, hence $T_2 \sim \text{Exp}(2\lambda)$ and also $T_3 \sim \text{Exp}(\lambda)$. [Note that T_1, T_2, T_3 are independent (by the lack of memory property).] Then, using the linearity of the expectation, we have

$$\mathbb{E}(T) = \sum_{i=1}^{3} \mathbb{E}(T_i) = \frac{1}{3\lambda} + \frac{1}{2\lambda} + \frac{1}{\lambda} = \frac{11}{6\lambda} = 11[\text{hours}].$$

So the earliest time at which all students will have completed their homework, on average, will be 7pm.

Question 4: (4 points) Define a compound Poisson process.

Solution: Let $\{N_t\}_{t\geq 0}$ be a Poisson process of rate $\lambda > 0$. **[1 mark]**

In addition, let $Y_1, Y_2, \ldots$ be a sequence of independent and identically distributed random variables, **[1 mark]**
that are independent of $\{N_t\}$. **[1 mark]**
Then $(X_t)_{t\geq 0}$ with

$$X_t = \sum_{i=1}^{N_t} Y_i$$

is a compound Poisson process. **[1 mark]**