A8 Probability – Solutions

Question 1

(a): B
(b): S
(c): N
Suggested mark scheme: (a) 3 (b) 2+2+5+5 (c) 8.

Question 2

(a): S
(b)(i)-(iii) S
(b)(iv) N
Suggested mark scheme: (a) 3+4+2 (b) 3+7+2+4.

Question 3

(a) B and S
(b) S and N
Suggested mark scheme: (a) 2+3+4 (b) 4+4+5+3.

Solution (1):
Suggested mark scheme: (a) 3; (b) 2+2+5+5; (c) 8.

(a) $Y_n \to Y$ in distribution if

$$F_n(y) \to F(y)$$

as $n \to \infty$, for all y such that F is continuous at y.

(b) (i) The probability density function of X_i is $f(x) = \frac{d}{dx}(1-x^{-\alpha})$ which is $\alpha x^{-\alpha-1}$. If $\alpha > 1$, then $\mathbb{E}(X_i) = \int_1^\infty x f(x) dx = \int_1^\infty \alpha x^{-\alpha} = [\alpha x^{-\alpha+1}/(-\alpha+1)]_1^\infty = \alpha/(\alpha-1)$.

(ii) Let $x \geq 1$.

$$\begin{aligned}\mathbb{P}(M_n > x) &= \mathbb{P}(X_1 > x, X_2 > x, \ldots, X_n > x) \\ &= \mathbb{P}(X_1 > x)\mathbb{P}(X_2 > x) \ldots \mathbb{P}(X_n > x) \text{ (by independence)} \\ &= (x^{-\alpha})^n \\ &= x^{-\alpha n}.\end{aligned}$$

So M_n has Pareto distribution with parameter $n\alpha$.

(iii) Let $x > 0$.

$$\begin{aligned}\mathbb{P}(nM_n - n > x) &= \mathbb{P}(M_n > 1 + x/n) \\ &= (1 + x/n)^{-\alpha n} \\ &\to e^{-\alpha x} \text{ as } n \to \infty.\end{aligned}$$

So for $x > 0$, $\mathbb{P}(nM_n - n \leq x)$ converges to $1 - e^{-\alpha x}$. Meanwhile for $x \leq 0$, $\mathbb{P}(nM_n - n \leq x) = 0$ for all n.

Hence nM_n converges in distribution to an exponential distribution with parameter α.

(iv) As long as the variance of each X_i is finite, we can apply the CLT. For the variance to be finite, we need EX_i^2 to be finite, for which we need $\alpha > 2$.

Then using the CLT, $(S_n - n\mu)/(\sqrt{n\sigma^2})$ converges to $N(0, 1)$.

Equivalently, $(S_n - n\mu)/(\sqrt{n})$ converges to $N(0, \sigma^2)$.

Hence the statement is true with $c_1 = \mu = \alpha/(\alpha - 1)$ and $c_2 = 1/2$.

(c) Let $u \in \mathbb{R}$ and take any $\epsilon > 0$. From the hint,

$$\mathbb{P}(X_n + Y_n \leq u) \leq \mathbb{P}(X_n \leq u - c + \epsilon) + \mathbb{P}(Y_n \leq c - \epsilon).$$

As $n \to \infty$, we have $\mathbb{P}(X_n \leq u-c+\epsilon) \to \mathbb{P}(X \leq u-c+\epsilon)$ and $\mathbb{P}(Y_n \leq c-\epsilon) \to 0$. Hence

$$\mathbb{P}(X_n + Y_n \leq u) \leq \mathbb{P}(X \leq u - c + \epsilon) + \epsilon \tag{1}$$

for all large enough n.

But also

$$\{X_n + Y_n > u\} \subseteq \{X_n > u - c - \epsilon\} \cup \{Y_n > c + \epsilon\}.$$

So

$$\begin{aligned} \mathbb{P}(X_n + Y_n \leq u) &= 1 - \mathbb{P}(X_n + Y_n > u) \\ &\geq 1 - [\mathbb{P}(X_n > u - c - \epsilon) + \mathbb{P}(Y_n > c + \epsilon)] \\ &= \mathbb{P}(X_n \leq u - c - \epsilon) - \mathbb{P}(Y_n > c + \epsilon). \end{aligned}$$

Now as $n \to \infty$, we have $\mathbb{P}(X_n \leq u - c - \epsilon) \to \mathbb{P}(X \leq u - c - \epsilon)$, and $\mathbb{P}(Y_n > c + \epsilon) \to 0$. Hence

$$\mathbb{P}(X_n + Y_n \leq u) \geq \mathbb{P}(X \leq u - c - \epsilon) - \epsilon \tag{2}$$

for all large enough n.

But ϵ is arbitrary, and as $\epsilon \to 0$, both $\mathbb{P}(X \leq u - c + \epsilon)$ and $\mathbb{P}(X \leq u - c - \epsilon)$ converge to $\mathbb{P}(X \leq u - c)$ (since the distribution function of X is continuous).

Hence by (1) and (2), as $n \to \infty$,

$$\begin{aligned} \mathbb{P}(X_n + Y_n \leq u) &\to \mathbb{P}(X \leq u - c) \\ &= \mathbb{P}(X + c \leq u). \end{aligned}$$

So indeed $X_n + Y_n \to X + c$ in distribution.

Solution (2):
Suggested mark scheme: (a) 3+4+2 (b) 3+7+2+4

(a) (i) For the second matrix, for example, for π to be stationary we need $\pi_1 = \frac{1}{3}\pi_2$, $\pi_2 = \pi_1 + \pi_3$, and $\pi_3 = \frac{2}{3}\pi_2$. (One of these is redundant.) In addition we need $\pi_1 + \pi_2 + \pi_3$. Solving gives

$$(\pi_1, \pi_2, \pi_3) = \left(\frac{1}{6}, \frac{1}{2}, \frac{1}{3}\right).$$

(ii) In the first case, the matrix is irreducible and aperiodic (note that $p_{22} > 0$ so the period must be 1). Hence by the theorem on convergence to equilibrium, $p_{11}^{(n)} \to \pi_1 = 1/6$.

In the second case, the matrix is not aperiodic; the period is 2. Then $p_{11}^{(n)}$ does not converge; in fact $p_{11}^{(m)}$ is 0 whenever m is odd, and $p_{11}^{(m)}$ is 1/3 whenever $m \geq 2$ is even.

(iii) The expected return time m_1 is simply the reciprocal of π_1. Hence in both cases $m_1 = 1/\pi_1 = 6$.

(b) (i) For transitions from $i > 0$, $p_{ij} = 1/i$ if $0 \leq j < i$, and 0 otherwise.

For transitions from 0, $p_{00} = 1/2$ and $p_{0i} = 1/2N$ for $i > 0$.

If $i > 0$, then to move from i to i in two steps, we must go via 0. Then $p_{ii}^{(2)} = p_{i0}p_{0i} = (1/i)(1/2N) = 1/(2Ni)$.

(ii) For $1 \leq i \leq N-1$,

$$\pi_i = \frac{1}{2N}\pi_0 + \sum_{j>i} \frac{1}{j}\pi_j$$

while

$$\pi_{i+1} = \frac{1}{2N}\pi_0 + \sum_{j>i+1} \frac{1}{j}\pi_j.$$

Hence

$$\begin{aligned}\pi_i &= \pi_{i+1} + \frac{1}{i+1}\pi_{i+1}\\ &= \frac{i+2}{i+1}\pi_{i+1}.\end{aligned}$$

So we obtain

$$\pi_{N-1} = \frac{N+1}{N}\pi_N,$$
$$\pi_{N-2} = \frac{N}{N-1}\frac{N+1}{N}\pi_N = \frac{N+1}{N-1}\pi_N,$$
$$\ldots \pi_1 = \frac{N+1}{2}\pi_N.$$

Finally by considering state N, we have $\pi_N = \frac{1}{2N}\pi_0$, so

$$\pi_0 = 2N\pi_N.$$

So overall

$$(\pi_0, \pi_1, \pi_2, \dots, \pi_{N-1}, \pi_N) = \pi_N \left(2N, \frac{N+1}{2}, \frac{N+1}{3}, \dots, \frac{N+1}{N}, \frac{N+1}{N+1}\right).$$

By normalising, we get

$$\pi_N = \frac{1}{(N+1)H_{N+1} + N - 1}.$$

(iii) Then $m_0 = 1/\pi_0 = \dfrac{(N+1)H_{N+1} + N - 1}{2N}$.

(iv) Consider the first time that the flea enters the set $\{0, 1, \dots, i\}$. Because the transition probability from $j > i$ is uniform on $\{0, 1, \dots, j-1\}$, this jump is equally likely to be to any of $\{0, 1, \dots, i\}$. But we reach i before 0 only if this jump is to i. So the probability is $1/(i+1)$.

Solution (3):
Suggested mark scheme: (a) 2+3+4 (b) 4+4+5+3.

(a) (i) $N_t \sim \text{Poisson}(\lambda t)$. $N_t - N_s \sim \text{Poisson}(\lambda(t-s))$.

(ii) T_1 and $T_2 - T_1$ both have exponential distribution with parameter λ, and they are independent.

(iii) The events $N_t \geq 2$ and $T_2 \leq t$ are the same. Since $N_t \sim \text{Poisson}(\lambda t)$, the probability is $1 - e^{-\lambda t} - \lambda t e^{-\lambda t}$.

To find the density function of T_2, we take the derivative to obtain

$$\begin{aligned} f_{T_2}(t) &= \lambda e^{-\lambda t} - \lambda e^{-\lambda t} + \lambda t e^{-\lambda t} \\ &= \lambda t e^{-\lambda t} \end{aligned}$$

for $t \geq 0$. This is the Gamma distribution with parameters 2 and λ.

(b) (i) We have

$$\begin{aligned} X_{n+1} &= N_{n+1} - (n+1) \\ &= (N_{n+1} - N_n) + (N_n - n) - 1 \\ &= X_n + (N_{n+1} - N_n) - 1. \end{aligned}$$

The quantity $N_{n+1} - N_n$ is independent of the process $N_t, t \leq n$, using the independent increments property of the Poisson process. But $X_i, i \leq n$ is determined by $N_t, t \leq n$.

So $N_{n+1} - N_n$ is independent of $X_n, X_{n-1}, \dots$.

So we have written X_{n+1} as a function of X_n and a random variable independent of $X_n, X_{n-1}, X_{n-2}, \dots$. This implies that X is a Markov chain.

(ii)

$$\begin{aligned} p_{i,i-1} &= \mathbb{P}(X_{n+1} = i-1 | X_n = i) \\ &= \mathbb{P}(N_{n+1} - (n+1) = i-1 | N_n - n = i) \\ &= \mathbb{P}(N_{n+1} - N_n = 0) \end{aligned}$$

(using independence as above)

$$= e^{-\lambda}$$

since $N_{n+1} - N_n \sim \text{Poisson}(\lambda)$.

In a similar way $p_{i,j} = \mathbb{P}(N_{n+1} - N_n = i - j + 1)$. Then

$$p_{i,j} = \begin{cases} \frac{e^{-\lambda}\lambda^{j-i+1}}{(j-i+1)!} & j \geq i-1 \\ 0 & j < i-1 \end{cases}.$$

(iii) Let $\lambda = 1$. Since the chain X starts at 0,

$$\begin{aligned} p_{00}^{(n)} &= \mathbb{P}(X_n = 0) \\ &= \mathbb{P}(N_n - n = 0) \\ &= \mathbb{P}(N_n = n) \\ &= \frac{e^{-n}n^n}{n!} \end{aligned}$$

since $N_n \sim \text{Poisson}(n)$

$$\sim \frac{1}{\sqrt{2\pi}n^{1/2}}$$

as $n \to \infty$, using Stirling's approximation.

Hence $\sum p_{00}^{(n)} = \infty$. But a state i is recurrent iff $\sum p_{ii}^{(n)} = \infty$. Hence 0 is a recurrent state. Since the chain is irreducible, then every state is recurrent.

(iv) Suppose π is a stationary distribution. We have $p_{00}^{(n)} \to 0$. But the chain is irreducible and aperiodic, so we should have $p_{00}^{(n)} \to \pi_0$. This would give $\pi_0 = 0$ and indeed similarly $\pi_i = 0$ for all i. But then π is not a distribution. Hence no such π exists, and so the chain cannot be positive recurrent. It is null recurrent.