1.

(a) Using the Thinning Lemma (7.3), $Y_A(t)$ is a Poisson process of rate $\frac{1}{10} \times \frac{1}{3} = \frac{1}{30}$.
Again, using the Thinning Lemma (7.3), $Y_B(t)$ is a Poisson process of rate $\frac{1}{12} \times \frac{1}{2} = \frac{1}{24}$.
Now using the Superposition Lemma (7.2), $G(t) = Y_A(t) + Y_B(t)$ is a Poisson process of rate $\frac{1}{30} + \frac{1}{24} = \frac{3}{40}$.

(b) We want

$$\begin{aligned}
\mathbb{P}(Y_A(90) = 3, Y_B(90) = 2) &= \mathbb{P}(Y_A(90) = 3) \cdot \mathbb{P}(Y_B(90) = 2) \\
&\qquad \text{(since these processes are independent)} \\
&= e^{-3}\frac{3^3}{3!}e^{-15/4}\frac{(15/4)^2}{2!} \\
&\qquad \text{(since } Y_A(90) \sim \mathrm{Po}(90/30) \text{ and } Y_B(90) \sim \mathrm{Po}(90/24)) \\
&= \frac{27 \times 15 \times 15}{6 \times 4 \times 4 \times 2e^{27/4}} \\
&= \frac{2025}{64e^{27/4}} \\
&\approx 0.037
\end{aligned}$$

(c) This is the expectation of the first arrival time of the process $X_A(t)$. The first arrival time of a Poisson process of rate λ is distributed $\mathrm{Exp}(\lambda)$. So the random variable we are interested in has $\mathrm{Exp}(1/10)$ distribution and its expectation is 10 minutes.

This was the answer I was expecting you to give but there is a slight subtlety. Well done if you spotted this. The distribution given is the right distribution under the assumption that Team A's shots at goal form a Poisson process which would be the case if the match had indefinite length. However, we know the match will finish at 90 minutes so there is no chance that I wait for say 100 minutes before seeing a shot at goal. The only circumstances in which this will make a difference to the time I wait is when Team A has no shots at goal in the entire match. This is a fairly unlikely event (because the rate is reasonably high) so 10 is a very good approximation to the answer but the true answer is a little lower. The calculation is a bit fiddly and goes beyond what I wanted you to get out of this question so I won't give it in detail.

(d) By the memoryless property of the exponential distribution, the extra time to wait until Team A has a shot at goal is again distributed Exp(1/10) so its expectation is 10 minutes. As in part (c) this would be an exact result if the match had indefinite length and is still a very good approximation here.

(e) Conditional on there being 4 goals in the match, the number of goals in the first half is distributed $\text{Bin}(4, 1/2)$. We want the probability 3 or 4 of these goals are in the first half. The probability that a $\text{Bin}(4, 1/2)$ random variable takes the value 3 or 4 is

$$\binom{4}{3}\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^1+\binom{4}{4}\left(\frac{1}{2}\right)^4\left(\frac{1}{2}\right)^0=\frac{5}{16}.$$

(f)
- Team A got off to a good start, scoring from their first shot at goal in the first minute.
- At half-time Team A was leading by one goal to nil, with Team B having had no shots at goal in the first half.
- Team B equalised at 72 minutes.
- Team A had two chances to take the lead in the last minute but both shots were saved ...
- ... and the match ended as a one-one draw.

2.

(a)

$$\mathbb{P}(T_1 \leqslant u) = 1 - \mathbb{P}(T_1 > u) = 1 - \mathbb{P}(\text{no arrivals in } [0, u]) = 1 - e^{-\lambda u}.$$

(b) Similarly,

$$\mathbb{P}(T_1 \leqslant u \mid X(t) = n) = 1 - \mathbb{P}(T_1 > u \mid X(t) = n) = 1 - \mathbb{P}(\text{no arrivals in } [0, u] \mid X(t) = n).$$

We know that conditional on $X(t) = n$, the number of arrivals in $[0, u]$ is distributed $\text{Bin}(n, \frac{u}{t})$. So

$$\mathbb{P}(T_1 \leqslant u \mid X(t) = n) = 1 - \binom{n}{0}\frac{u}{t}^0\left(1-\frac{u}{t}\right)^n = 1 - \left(1-\frac{u}{t}\right)^n.$$

(c) The answer to part (a) involves λ. If we changed the rate of the process by increasing λ then this probability would increase.

The answer to part (b) does not involve λ. If we changed the rate of the process by increasing or decreasing λ then this probability would not change. This is because once we condition on $X(t) = n$, the rate of the process stops being relevant.

(d) In part (b) we worked out the conditional cdf. To find the conditional pdf from the conditional cdf we differentiate:

$$\begin{aligned} f_{T_1|X(t)=n}(u) &= F'_{T_1|X(t)=n}(u) \\ &= \frac{d}{du}\left(1 - \left(1 - \frac{u}{t}\right)^n\right) \\ &= \frac{n}{t}\left(1 - \frac{u}{t}\right)^{n-1} && \text{(by the chain rule)} \end{aligned}$$

3. In each part of this question we are told that $X(t) = n$ for some t and n. We consider random variables U_i for $1 \leqslant i \leqslant n$ where the U_i are independent and each is distributed uniformly on $[0, t]$. We know from lectures (Corollary 7.9) that if r is a symmetric function of the T_i (as the functions in this question are) then

$$\mathbb{E}(r(T_1, \ldots, T_n) \mid X(t) = n) = \mathbb{E}(r(U_1, \ldots, U_n)).$$

(a) By Corollary 7.9:

$$\mathbb{E}(T_1 + T_2 + T_3 \mid X(10) = 3) = \mathbb{E}(U_1 + U_2 + U_3)$$

where $U_i \sim U[0, 10]$. So $\mathbb{E}(U_i) = 5$ and

$$\mathbb{E}(T_1 + T_2 + T_3 \mid X(10) = 3) = \mathbb{E}(U_1) + \mathbb{E}(U_2) + \mathbb{E}(U_3) = 15.$$

(b) By Corollary 7.9:

$$\mathbb{E}(T_1^2 T_2^2 T_3^2 T_4^2 \mid X(1) = 4) = \mathbb{E}(U_1^2 U_2^2 U_3^2 U_4^2)$$

where $U_i \sim U[0, 1]$ and the U_i are independent. So

$$\begin{aligned} \mathbb{E}(U_1^2 U_2^2 U_3^2 U_4^2) &= \mathbb{E}(U_1^2)\mathbb{E}(U_2^2)\mathbb{E}(U_3^2)\mathbb{E}(U_4^2) && \text{(by independence)} \\ &= \left(\frac{1}{3}\right)^4 && (\text{as } \mathbb{E}(U_i^2) = \textstyle\int_0^1 u^2\,du = \frac{1}{3}) \\ &= \frac{1}{81} \end{aligned}$$

4.

(a) Processing occurs every T minutes and costs £k each time. This gives a cost of £$\frac{k}{T}$ per minute.

Following the hint, let's suppose that there are n requests waiting at time T. The ith request to arrive has been waiting for time $T - T_i$ and so incurred a cost $c(T-T_i)$. Hence, the total waiting cost incurred is a random variable with expectation

$$\mathbb{E}(\text{total waiting cost} \mid X(T) = n) = \mathbb{E}\left(\sum_{i=1}^{n} c(T - T_i) \mid X(T) = n\right)$$

We can use the same trick as Question 3 (replacing T_i with $U_i \sim U[0,T]$) to get that

$$\mathbb{E}\left(\sum_{i=1}^{n} c(T - T_i) \mid X(T) = n\right) = \mathbb{E}\left(\sum_{i=1}^{n} c(T - U_i)\right) = \frac{ncT}{2}.$$

Now, conditioning on the number of requests waiting at time T, the expected total waiting cost is

$$\begin{aligned}
\mathbb{E}(\text{total waiting cost}) &= \sum_{n \geqslant 0} \mathbb{E}\left(\sum_{i=1}^{n} c(T - T_i) \mid X(T) = n\right) \mathbb{P}(X(T) = n) \\
&= \frac{cT}{2} \sum_{n \geqslant 0} n e^{-\lambda T} \frac{(\lambda T)^n}{n!} \\
&= \frac{cT^2\lambda}{2}
\end{aligned}$$

(where the last identity comes from the fact that the expectation of a $\mathrm{Po}(\lambda T)$ random variable is λT).

So the expectation of the total waiting cost per minute is $\frac{cT\lambda}{2}$. The total expected cost per minute (processing and waiting) is

$$\frac{k}{T} + \frac{c\lambda T}{2}$$

as required.

(b) Sketching the graph of this function against T we see that it has a single minimum. Differentiating we get that the minimum is the solution to

$$\frac{c\lambda}{2} - \frac{k}{T^2} = 0$$

So we should take $T = \sqrt{\frac{2k}{c\lambda}}$.

5.

(a) Under the first assumption we have that, by the lack of memory of the Poisson process, buses pass as a Poisson process (even if we arrive at a random time). The time we waits is T_1 (or S_1) for the process starting from when we arrives and so it is an Exp(6) random variable. The expectation is therefore 1/6 hours, that is 10 minutes.

(b) Under the second assumption assumption ii) the fact that we arrive at a random time means that the time we wait is a random variable uniformly distributed on $[0, 10]$. Such a random variable has expectation 5.

This result appears paradoxical since in both cases the average interval between buses is 10 minutes. Intuitively the average time we must wait should be half this. This argument is correct for assumption (b). However in the Poisson process case we are more likely to arrive at the bus stop during a long interval between buses (because in total the longer intervals take up more time than the shorter ones) and so the average time we wait is longer.

Please let me know if you have any comments or corrections

Robert Johnson
r.johnson@qmul.ac.uk