Lecture 3: The Poisson process, and the M/M/1 queue – How busy will you keep me, and how big a waiting room?

Concepts checklist

At the end of this lecture, you should be able to:

- Define the Poisson process as a *continuous-time Markov chain* (CTMC); and,
- Define the M/M/1 queue.

Example 2. The Poisson process

Another (important) example of a continuous-time Markov chain (CTMC) is the Poisson process. You have probably seen the Poisson process before, defined as follows.

Definition 4. *Let the sequence $t_1, t_2, \ldots$ be independent exponential random variables with rate λ, and define*

$$\begin{aligned} & T_0 := 0, \\ & T_n := t_1 + \cdots + t_n \textit{ for } n \geq 1, \\ \textit{and} \quad & N(t) := \max\{n : T_n \leq t\} \textit{ for } t \geq 0. \end{aligned}$$

Then, $\{N(t) : t \geq 0\}$ is called the Poisson process.

For example, we can think of the random variables t_n as the times between arrivals of students at a Professor's office (a timely reminder that I have specified office hours!), so

- $T_n = t_1 + \cdots + t_n$ is the arrival time of the n^{th} student, and
- $N(t)$ is the number of arrivals by time t.

In general, the process $\{N(t)\}$ is used to model systems where events of a particular type (often called, points) occur in time, in such a way that the probability of a point occurring in the interval after t is independent of what has happened up to t.

A quantity of interest, to help me plan my day, is the distribution of the number of arrivals I should expect in a day, or a working week. That is, what is the distribution of $N(t)$?

Let us calculate the distribution of $N(t)$. We have $N(t) = n$ if and only if $T_n \leq t < T_{n+1}$; i.e., the n^{th} student arrives before time t but the $(n+1)^{\text{th}}$ after t. Conditioning on the value of $T_n \in [s, s + dt)$ such that $s \leq t$ and noting that $T_{n+1} > t$, we have

$$\begin{aligned} \Pr(N(t) = n) &= \Pr(T_n \leq t < T_{n+1}) \\ &= \int_0^t \frac{\Pr(T_n \leq s)}{ds} \Pr(T_{n+1} > t | T_n \in [s, s + dt)) ds \quad \text{(Law of Total Probability)} \\ &= \int_0^t f_{T_n}(s) \Pr(t_{n+1} > t - s) ds \qquad \text{(independence of } t_{n+1}\text{)}, \end{aligned}$$

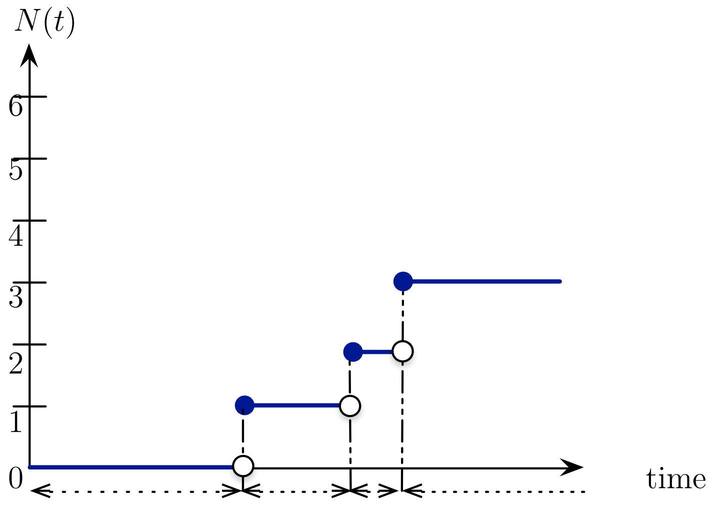

Figure 3: A realisation of the Poisson Process $\{N(t)\}$

where $f_{T_n}(s)$ is the probability density function (pdf) of T_n.

Note that the sum $T_n = t_1 + \cdots + t_n$, where t_i are independent exponential random variables with rate λ, has an Erlang(n, λ) distribution, with pdf

$$f_{T_n}(t) = \lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!} \quad \text{for } t \geq 0.$$

Thus,

$$\begin{aligned}
\Pr(N(t) = n) &= \int_0^t \lambda e^{-\lambda s} \frac{(\lambda s)^{n-1}}{(n-1)!} e^{-\lambda(t-s)} ds \\
&= \frac{\lambda^n}{(n-1)!} e^{-\lambda t} \int_0^t s^{n-1} ds \\
&= \frac{\lambda^n}{(n-1)!} e^{-\lambda t} \frac{t^n}{n} \\
&= e^{-\lambda t} \frac{(\lambda t)^n}{n!}.
\end{aligned}$$

This is the Poisson distribution with parameter λt. Makes sense to call this the Poisson process, right?

Poisson process as a continuous-time Markov chain

Consider a process $\mathcal{X} = (X(t), t \geq 0)$, with the random variable $X(t)$ denoting the state of the system at time t.

The *state space* is $\mathcal{S} = \{0, 1, 2, \ldots\} = \mathbb{Z}_+$.

If $X(t) = n$, then n *arrivals/events* have occurred in the time interval $[0, t]$.

As the events occur at rate λ (i.e., the time between two consecutive events is $\exp(\lambda)$), we say that the *transition rate* $q_{i,i+1}$ from state i to state $i+1$ is λ:

$$\begin{aligned}
q_{i,i+1} &= \lambda \quad \text{for } i \in \mathcal{S}, \\
q_{ij} &= 0 \quad \text{for } j \in \mathcal{S}, j \neq \{i, i+1\}.
\end{aligned}$$

The generator $Q = (q_{ij})_{i,j \in \mathcal{S}}$ is given by

$$Q = \begin{bmatrix} -\lambda & \lambda & 0 & 0 & 0 & \cdots \\ 0 & -\lambda & \lambda & 0 & 0 & \cdots \\ 0 & 0 & -\lambda & \lambda & 0 & \cdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}.$$

Given the states, events, and rates, we can write down a state transition diagram:

$$0 \xrightarrow{\lambda} 1 \xrightarrow{\lambda} 2 \xrightarrow{\lambda} 3 \xrightarrow{\lambda} \cdots$$

The Poisson process is also known as the *simple pure birth process*.

The Poisson process is a CTMC, as

$$\begin{aligned} \Pr(X(t+s) = k \mid X(u) = i, X(s) = j, u < s) &= e^{-\lambda t} \frac{(\lambda t)^{k-j}}{(k-j)!} \\ &= \Pr(X(t+s) = k \mid X(s) = j), \end{aligned}$$

for all $j, k \in \mathcal{S}$ and $t \geq 0$.

Further, note that this also equals $\Pr(X(t) = k \mid X(0) = j)$, and hence the process is *time homogeneous*.

This is helpful, as it tells me something about the distribution of the number of students I can expect (once I've estimated λ); but it doesn't tell me if this will annoy Nigel, Matt, Sanjeeva, and possibly others! That will depend on how many are queued for help in the corridor...

Example 3. A continuous-time single-server queue (the M/M/1 queue)

Consider a simple, single-server queue. The state of the process $\mathcal{X} = (X(t), t \geq 0)$ is the number of customers (students) in the queue at time t, including the person being served (helped).

Assume there are arrivals to the queue according to a Poisson process with rate λ; and, services occur as a Poisson process with rate μ whenever there is at least one customer.

- The state space is $\mathcal{S} = \{0, 1, 2, \ldots\} = \mathbb{Z}_+$, where each state $i \in \mathcal{S}$ means there are i people in the system.

- The transition rates are:

$$\begin{aligned} q_{i,i+1} &= \lambda \quad \text{for } i \in \mathcal{S}, \\ q_{i,i-1} &= \mu \quad \text{for } i \in \mathcal{S} \backslash \{0\}, \\ q_{ij} &= 0 \quad \text{for } j \in \mathcal{S} \backslash \{i, i+1, i-1\}. \end{aligned}$$

Thus, the generator Q is:

$$Q = \begin{bmatrix} -\lambda & \lambda & 0 & 0 & 0 & \cdots \\ \mu & -(\lambda+\mu) & \lambda & 0 & 0 & \cdots \\ 0 & \mu & -(\lambda+\mu) & \lambda & 0 & \cdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots \end{bmatrix}.$$

- The state transition diagram is

$$\begin{tikzcd} 0 \arrow[r, bend left, "\lambda"] & 1 \arrow[l, bend left, "\mu"] \arrow[r, bend left, "\lambda"] & 2 \arrow[l, bend left, "\mu"] \arrow[r, bend left, "\lambda"] & 3 \arrow[l, bend left, "\mu"] \arrow[r, bend left, "\lambda"] & \cdots \arrow[l, bend left, "\mu"] \end{tikzcd}$$

Using the earlier analysis, you can (and should attempt to) show that in states $1, 2, \ldots$, the process waits an exponentially-distributed amount of time with rate $\lambda + \mu$, and at that time sees a decrease in the queue length of one with probability $\mu/(\lambda + \mu)$ and an increase in the queue length by one otherwise.

Is this a good model to inform the possible number of students taking Random Processes III waiting to see me? What physical constraint is missing? How could the model be modified to accomodate that?
