Poisson Processes

EECS 126 (UC Berkeley)

Spring 2022

1 Motivation

At this point in your probability career, you've likely encountered some very interesting ideas relating to discrete time Markov chains. However, there is a slight problem. The real world operates in continuous time, so many random processes that you might want to study as an engineer or computer scientist must use the language of continuous time systems.

In our attempts to build up to a treatment of continuous time Markov processes, we start with Poisson processes, which in some sense is the simplest continuous time Markov chain that you might wish to study. On the other hand, we will soon see that Poisson processes have many amazing properties that make them an indispensable tool for modeling.

2 Construction

Suppose that you're waiting at a bus station and observe that no matter how long you've been waiting for the next bus, it always seems to be around 5 minutes until the next bus comes. Recall the memoryless property of the exponential distribution:

If $X \sim \text{Exponential}(\lambda)$, then $\mathbb{P}(X > t + s \mid X > t) = \mathbb{P}(X > s)$

Returning to our modeling problem before, you can model the memoryless nature by stating that the time it takes for the next bus to come is exponentially distributed with parameter $\lambda := 5\,\text{min}$.

This leads us to our formal definition:

Definition 1. *Fix some* $\lambda > 0$ *and sample* interarrival times $S_1, S_2, S_3, \ldots \sim_{iid}$ *Exp*(λ). *For each* $n \geq 1$ *define* $T_n = \sum_{j=1}^{n} S_j$ *for* $n \geq 1$. *The function* $N(t) = \max\{n \geq 0 : T_n \leq t\}$ *represents the number of arrivals at time* t. *We call the sequence* $\{N(t)\}_{t \geq 0}$ *a* Poisson process *with rate* λ. *The distribution of such processes is denoted* $PP(\lambda)$.

Tying the definition above back to our earlier analogy, the time between the $i-1$st bus and the ith bus is the interarrival time S_i, and the ith bus appears after T_i units of time. The Poisson process $N(t)$ counts the number of buses that have appeared by time t.

Sometimes, instead of starting the process at 0, we want to consider the number of arrivals in an interval starting at some arbitrary time. To do that, we introduce some new notation:

Definition 2. *For a Poisson process with rate λ, we define $N(t_1, t_2) := N(t_2) - N(t_1)$ for $t_2 \geq t_1$ to be the number of arrivals in the interval $[t_1, t_2]$.*

Note that we are "overloading" N here. When we write $N(t)$ as a function with one argument, we mean the number of arrivals in the interval $[0, t]$, while writing $N(t_1, t_2)$ as a function with two arguments describes the number of arrivals in the interval $[t_1, t_2]$. Finally, we sometimes refer to the number of arrivals in an interval $N(t_1, t_2)$ as an *increment*, especially in context of the stationary and independent increments property that we describe in the next section.

3 Equivalent Characterization

The previous section gave a construction of Poisson processes that matches with our mental picture of counting the number of arrivals in time. The memoryless property of the exponential gives Poisson processes many amazing properties. We then finish off this section with an alternative construction of Poisson processes.

Theorem 1. *A Poisson process $\{N(t)\}_{t \geq 0} \sim PP(\lambda)$ satisfies the following properties:*

1. Stationary increments*: For every $t, s > 0$, $N(t, t+s) \stackrel{d}{=} N(s)$; namely, $N(t, t+s)$ has the same distribution as $N(s)$.*
2. Independent increments*: For $0 < t_1 < \cdots < t_k$, the set of random variables $N(t_1), N(t_1, t_2), \ldots, N(t_{k-1}, t_k)$ are jointly independent.*
3. $N(t) \sim \text{Poisson}(\lambda t)$

The next few sections will prove these theorems.

3.1 Proving Stationary and Independent Increments

To show that the stationary independent increment properties hold, we first need to establish the following claim:

Lemma 1. *Let $\{N(t)\}_{t\geq 0} \sim PP(\lambda)$ and fix some time $t > 0$. Define the random variable Z as the amount of time elapsed until the first arrival after the time t. Then the following properties hold:*

1. *$Z \sim Exp(\lambda)$*
2. *Z is independent of all interarrival times before time t*
3. *Z is independent of $\{N(s)\}_{0\leq s\leq t}$*

We relegate the proof to Appendix A.1.1.

Theorem 2. *A Poisson process has stationary increments.*

The proof can be found in Appendix A.1.2.

Theorem 3. *A Poisson process has independent increments.*

The proof can be found in Appendix A.1.3.

3.2 Erlang Distribution and pmf of $N(t)$

While we already know the densities of the interarrival times $\{S_i\}_{i=1}^{\infty}$ as they are just exponentially distributed, we do not have a density for the arrival times T_n. Note that T_n is just a sum of n i.i.d. exponential random variables, which means that we can compute the distribution by convolving the pdfs. The resulting distribution is called the Erlang distribution, denoted Erlang$(n; \lambda)$.

Let T_n be the nth arrival time for a Poisson process with parameter λ. Then,

$$f_{T_n}(t) = \frac{\lambda^n t^{n-1} e^{-\lambda t}}{(n-1)!}.$$

Proof. The joint distribution $f_{T_1,\ldots,T_n}(t_1, \ldots, t)$ can be rewritten as the joint distribution of interarrival times $f_{S_1,S_2,\ldots,S_n}(t_1, t_2 - t_1, \ldots, t - t_{n-1})$.

$$f_{T_1,T_2,\ldots,T_n}(t_1, t_2, \ldots, t) = f_{S_1,S_2,\ldots,S_n}(t_1, t_2 - t_1, \ldots, t - t_{n-1})\mathbb{1}_{\{t_1\leq\ldots\leq t\}}$$

$$
\begin{aligned}
&= f_{S_1}(t_1)\cdot f_{S_2}(t_2-t_1)\dots f_{S_n}(t-t_{n-1})\mathbb{1}_{\{t_1\le\dots\le t\}}\\
&= (\lambda e^{-\lambda t_1})\cdot(\lambda e^{-\lambda(t_2-t_1)})\dots(\lambda e^{-\lambda(t-t_{n-1})})\mathbb{1}_{\{t_1\le\dots\le t\}}\\
&= \lambda^n e^{-\lambda t}\mathbb{1}_{\{t_1\le\dots\le t\}}
\end{aligned}
$$

where the indicator variable above is due to the fact that if $t_i < t_{i-1}$, then the density of S_i is 0; e.g. the ith arrival can't occur before the $i-1$st arrival. Note, amazingly, there is no dependence on the values of $t_1,\dots,t_{n-1}$.

Finally, we can calculate $f_{T_n}(t)$ by integrating over all possible values of $t_1,t_2,\dots,t_{n-1}$ with the constraint that $t_1 < t_2 < \dots < t_{n-1}$.

$$
\begin{aligned}
f_{T_n}(t) &= \int\cdots\int_{t_1\le t_2\le\dots\le t_{n-1}\le t} f_{T_1,T_2,\dots,T_{n-1},T_n}(t_1,t_2,\dots,t_{n-1},t)\,dt_1dt_2\dots dt_{n-1}\\
&= \int\cdots\int_{t_1\le t_2\le\dots\le t_{n-1}\le t} \lambda^n e^{-\lambda t}\,dt_1dt_2\dots dt_{n-1}
\end{aligned}
$$

Now, observe that the integrand is constant with respect to $t_1,\dots,t_{n-1}$. Also note that the volume of the hypercube $[0,t]^{n-1}$ is t^{n-1}. However, the integral is supported only the part of the hypercube where $t_1 \le t_2 \le \dots \le t_{n-1}$. As there are $(n-1)!$ permutations of these items, and each permutation slices out a equal volume of the hypercube, then the volume of the support is $t^{n-1}/(n-1)!$. This means that the value of the integral is $\lambda^n t^{n-1}e^{-\lambda t}/(n-1)!$, which was the form for the density we claimed. □

Theorem 4. *For a Poisson process $\{N(t)\}_{t\ge0} \sim PP(\lambda)$, we have $N(t) \sim$ Poisson(λt) for $t > 0$. Explicitly, the probability of n arrivals in time t is*

$$\mathbb{P}(N(t)=n) = \frac{(\lambda t)^n e^{-\lambda t}}{n!}.$$

Proof. We can consider two ways to calculate $\mathbb{P}(t < T_{n+1} < t+\delta)$, the probability the $(n+1)$th arrival arrives in $[t,t+\delta]$.

One way is to integrate over the density of T_{n+1}, the Erlang$(n+1;\lambda)$ distribution:

$$\mathbb{P}(t < T_{n+1} < t+\delta) = \int_t^{t+\delta} f_{T_{n+1}}(\tau)d\tau \approx f_{T_{n+1}}(t)\delta$$

Another way of computing $\mathbb{P}(t < T_{n+1} < t + \delta)$ is by splitting up the sample space by the number of arrivals in $[t, t + \delta]$:

$$\begin{aligned}
&\mathbb{P}(t < T_{n+1} < t + \delta) \\
&\quad = \mathbb{P}(N(t) = n, N(t, t+\delta) = 1) + \sum_{k=2}^{n+1} \mathbb{P}(N(t) = n + 1 - k, N(t, t+\delta) = k) \\
&\quad = \mathbb{P}(N(t) = n)\mathbb{P}(N(t, t+\delta) = 1) + \sum_{k=2}^{n+1} \mathbb{P}(N(t) = n + 1 - k)\mathbb{P}(N(t, t+\delta) = k)
\end{aligned}$$

The last equality holds due to the independent increments property.

The probability of more than one arrivals in a small interval δ and the probability of one arrival in a small interval δ is $\lambda\delta$. Derivations of these claims can be found in Appendix A.1.4.

Then,

$$\begin{aligned}
\mathbb{P}(t < T_{n+1} < t + \delta) &\approx \mathbb{P}(N(t) = n) \cdot \mathbb{P}(N(t, t+\delta) = 1) \\
&\approx \mathbb{P}(N(t) = n)\delta\lambda
\end{aligned}$$

Equating the two expressions, we can solve for $\mathbb{P}(N(t) = n)$, the probability of n arrivals in t time:

$$\mathbb{P}(N(t) = n) = \frac{f_{T_{n+1}(t)}\delta}{\delta\lambda} = \frac{(\lambda t)^n e^{\lambda t}}{n!}$$

For the above proof to be formally correct, we need to carry error terms $o(\delta)$. Little-o notation is used to represent terms that approach 0 faster than linearly as δ approaches 0. In math, this can be expressed as $\lim_{\delta \to 0} \frac{o(\delta)}{\delta} = 0$. For the full proof, refer to Appendix A.1.4. □

3.3 Proving Equivalence

We have shown all the parts of Theorem 1. To recap, we showed that a Poisson process has the stationary and independent increments property, and the number of arrivals in t units of time is distributed according to a Poisson distribution. Amazingly, these properties *uniquely* characterize the Poisson process with parameter λt.

Theorem 5. *Let $S_1, S_2, \ldots$ be some set of almost-surely positive*[1] *interarrival times and define $T_n = \sum_{j=1}^{n} S_j$ and $N(t) = \max\{n \geq 0 : T_n \geq t\}$. If it the process $\{N(t)\}_{t\geq 0}$ has stationary and independent increments, and $\mathbb{P}(N(t) = n) = (\lambda t)^n e^{-\lambda t}/n!$ for all $t \geq 0$. Then it follows that $S_1, S_2, \ldots$ are i.i.d. exponential random variables with parameter λ.*

Proof. Let's compute the complementary cdf of S_n conditioned on $\{S_1 = s_1, \ldots, S_{n-1} = s_{n-1}\}$. Defining $t_{n-1} = \sum_{j=1}^{n-1} s_{n-1}$, we see

$$\begin{aligned}
&\mathbb{P}(S_n > s \mid S_1 = s_1, \ldots, S_{n-1} = s_{n-1}) \\
&\quad = \mathbb{P}(S_n > s \mid \{N(t)\}_{0\leq t\leq t_{n-1}}) \\
&\quad = \mathbb{P}(N(t_{n-1}, t_{n-1} + s) = 0 \mid \{N(t)\}_{0\leq t\leq t_{n-1}}) \\
&\quad = \mathbb{P}(N(t_{n-1}, t_{n-1} + s) = 0) \\
&\quad = \mathbb{P}(N(s) = 0) \\
&\quad = e^{-\lambda s}
\end{aligned}$$

using the fact that specifying $\{S_1, \ldots, S_{n-1}\}$ fully specifies the counting process $N(t)$ until time t_{n-1} and vice versa; we then use the independent [2] and stationary increments property. We conclude that $S_n \sim \text{Exponential}(\lambda)$, independent of $\{S_1, \ldots, S_n\}$. Performing an induction on n in the same manner as in the proof for Theorem 3 lets us conclude $S_1, S_2, \ldots \sim_{\text{iid}} \text{Exponential}(\lambda)$. □

Let's take a moment to reflect upon the power of what we've established. We have just shown two equivalent characterizations of Poisson processes, meaning that we can switch between these two perspectives at will. For example, if we were writing a simulation to model a Poisson process, the first characterization looking at sums of exponential random variables is much more convenient to implement. On the other hand, the second characterization using independent and stationary increments lets us prove theorems about Poisson processes more directly. The utility of having both viewpoints available to us will apparent in the next few sections.

[1] to avoid the having two arrivals at the same time

[2] An eagle-eyed reader might notice that the definition of independent increments requires us to consider finitely many time instants, while the interval $[0, t_{n-1}]$ has uncountably many elements. It turns out the technical definition of independence of a random process is to consider an arbitrary finite collection of time instants, so our logic still holds. These technicalities are beyond the scope of EECS 126, but if you're interested in the details, EECS 226A covers them more in depth.

4 Merging

Now that we have our two viewpoints, we can start using them to prove some properties of Poisson processes. Suppose instead of buses, we are sitting at a bench watching cars and trucks pass by, where both the cars and trucks appear according to independent Poisson processes. How does the total number of vehicles that pass by behave? It turns out that this is also a Poisson process!

Theorem 6 (Poisson Merging). *Let $N \sim PP(\lambda)$ and $M \sim PP(\mu)$ independent of N. Then, $N + M \sim PP(\lambda + \mu)$.*

In order to show that $L := N + M$ is a Poisson process, we need to show that it has independent and stationary increments, and the number of arrivals in any interval has a Poisson distribution. Details of this proof can be found in Appendix A.2.

The merging property makes Poisson processes incredibly useful to for modeling. If two independent processes have the independent and stationary increments properties, we can model them as Poisson processes separately. Poisson merging tell us that the combined process will be a sum of the rate parameters and will also have independent and stationary increments. We get all of this information for free without needing to find the new rate parameter.

5 Splitting

We just saw that given some independent Poisson processes, we can combine them into a single Poisson process. This leads to an interesting follow-up question: is the reverse possible? Can we take a single Poisson process and divide it into a set of independent Poisson processes? Amazingly, the answer is yes.

Consider having packet arrivals come according to a Poisson process. If each arrival is independently routed to a different server, then the number of packets hitting each server are also Poisson processes. This is known as poisson splitting and is illustrated in Figure 1.

Theorem 7 (Poisson Splitting). *Let $N \sim PP(\lambda)$ be a Poisson process and let $B_1, B_2, \ldots \sim_{iid} \text{Bernoulli}(P)$ independent of N. Let $N_0(t) = |\{i : B_i = 0, i \leq N(t)\}|$ and $N_1(t) = |\{i : B_i = 1, i \leq N(t)\}|$ (e.g. each arrival to $N(t)$*

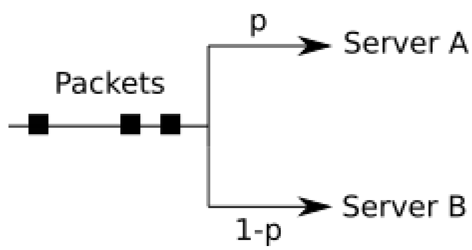

Figure 1: Each arrival is independently routed to server A with probability p and to B with probability $1-p$.

is routed to $N_0(t)$ *or* $N_1(t)$ *according to independent Bernoulli trials). Then* $N_0(t) \sim PP(\lambda p)$*,* $N_1(t) \sim PP(\lambda(1-p))$*, and the processes* N_0 *and* N_1 *are independent of each other.*

The proof can be found in Appendix A.3.

Splitting is the second half of what makes Poisson processes so useful. We can envision a scenario where we are a network engineer and we find that we have requests of different types hitting our load balancer according to Poisson processes, and then our load balancer randomly routes the packets to different servers, which themselves may route the packets with different probabilities. Yet however complex our system is, if we model the incoming traffic as Poisson processes, then we can easily compute how much load we expect each server to experience by using Poisson merging and splitting. This is the basis for the study of *Jackson networks* and other queuing systems.

6 Random Incidence Property

Consider a Poisson Process $\{N(t)\}_{t\geq 0} \sim PP(\lambda)$ that has been running for a long time. We can consider the following questions:

1. What is the expected interarrival time in the Poisson process?
2. Suppose we fix an arbitrary[3] time in the process, t. What is the expected length of the interarrival interval which t falls into?

From a first glance, these seem like the same question, since both are asking for the expected interarrival time. One might conclude that both

[3] e.g. from a distribution independent from the Poisson process

are $1/\lambda$, the expectation of an Exponential(λ) random variable. Perhaps surprisingly, this is not the case for the second question. Let's dig into why.

Say that t falls between the ith and $i+1$th arrivals, which happen at time T_i and T_{i+1}, respectively. If there has been no arrivals before t, then set $T_i = 0$. We can write

$$T_{i+1} - T_i = (t - T_i) + (T_{i+1} - t)$$

We know that the second term, $T_{i+1} - t$, is exponentially distributed with parameter λ, by Lemma 1, so $\mathbb{E}[T_{i+1} - t] = 1/\lambda$. We then want to find the distribution of the first term, $t - T_i$. To do that, we find the the complementary cdf $\mathbb{P}(t - T_i > \tau)$ for $0 \leq \tau \leq t$ as

$$\begin{aligned}
\mathbb{P}(t - T_i > \tau) &= \mathbb{P}\left(N(t-\tau, t) = 0\right) \\
&= \mathbb{P}\left(N(\tau) = 0\right) \\
&= \frac{(\lambda\tau)^0 e^{-\lambda\tau}}{0!} \\
&= e^{-\lambda\tau}
\end{aligned}$$

using the stationary increments property. However, when $\tau > t$, then the complementary cdf is 0 because T_i cannot be negative; thus we have complementary cdf

$$\mathbb{P}(t - T_i > \tau) = \begin{cases} e^{-\lambda\tau} & 0 \leq \tau \leq t \\ 0 & \tau > t \end{cases}$$

We see that $t - T_i$ is approximately Exponential(λ), and we can compute the exact expectation using the tail-sum formula as

$$\begin{aligned}
\mathbb{E}[t - T_i] &= \int_0^\infty \mathbb{P}(t - T_i > \tau)\, d\tau \\
&= \int_0^t e^{-\lambda\tau}\, d\tau \\
&= \frac{1 - e^{-\lambda t}}{\lambda}
\end{aligned}$$

so

$$\mathbb{E}[T_{i+1} - T_i] = \mathbb{E}[t - T_i] + \mathbb{E}[T_{i+1} - t] = \frac{1 - e^{-\lambda t}}{\lambda} + \frac{1}{\lambda} \xrightarrow{t \to \infty} \frac{2}{\lambda}$$

In the limit as $t \to \infty$, corresponding to a Poisson process that has been running for an infinitely long time before the observation at time t, the expected length of the interarrival time tends to $2/\lambda$. This is twice the expectation of the interarrival time! How can we explain this?

One way of rationalizing this is that in a Poisson process, there will be longer and shorter interarrival intervals. Since we are choosing our t randomly, it's more likely to land in a longer interval than a short one regardless of the distribution of choosing t. Thus, each interval is chosen with a weighted probability. Longer intervals have a better chance to be chosen, while shorter intervals have a lower chance. If you take the weighted average, we should expect an interval picked under this biased scheme to have a greater expectation than normal! This is known as the Random Incidence Property, abbreviated as RIP.

A Appendix

A.1 Independent and Stationary Increments

A.1.1 Memoryless Property of Poisson Processes (Lemma 1)

Proof. To find the distribution of Z, we can condition on the number of arrivals at time t and the previous arrival time.

$$\begin{aligned}
\mathbb{P}(Z > z \mid N(t) = i, T_i = \tau) &= \mathbb{P}(S_{i+1} > z + (t - \tau) \mid S_{i+1} > t - \tau, T_i = \tau) \\
&= \mathbb{P}(S_{i+1} > z + (t - \tau) \mid S_{i+1} > t - \tau) \\
&= \mathbb{P}(S_{i+1} > z) \\
&= e^{-\lambda z}
\end{aligned}$$

Above, we use the fact that S_{i+1} is independent of $T_i = \sum_{j=1}^{i} S_i$, interarrival times being independent, to remove the conditioning on T_i, then leverage the memoryless property of the Exponential distribution. Because the conditional ccdf of Z does not depend on i or τ, we conclude that this is the unconditional ccdf of Z. Shown explicitly,

$$\begin{aligned}
\mathbb{P}(Z > z) &= \sum_{i=1}^{\infty} \int_0^{\infty} \mathbb{P}(Z > z \mid N(t) = i, T_i = \tau)\mathbb{P}(N(t) = i, T_i = \tau) d\tau \\
&= e^{-\lambda z} \sum_{i=1}^{\infty} \int_0^{\infty} \mathbb{P}(N(t) = i, T_i = \tau) d\tau = e^{-\lambda z}
\end{aligned}$$

As this is the complementary cdf of an Exponential distribution, we conclude $Z \sim \text{Exponential}(\lambda)$. This proves property (1) above.

The previous derivation showed that the conditional distribution of Z given $N(t) = i$ and $T_i = \tau$ did not depend on i and τ, implying that Z is independent of these quantities. In fact, we can repeat the previous derivation where we additionally condition on $T_1 = \tau_1, \ldots, T_{i-1} = \tau_{i-1}$; however, the same steps shows that the distribution of Z still does not depend on these quantities, so we conclude that Z is independent of the set of arrival times.

Now, note that the set of interarrival times before t and $\{N(s)\}_{s\leq t}$ are both fully determined by specifying the arrival times, so we conclude that Z is independent of these as well. Thus properties (2) and (3) above hold as well. □

A.1.2 Stationary Increments Proof (Theorem 2)

Proof. Note that Lemma 1 implies that the distribution of the number of arrivals in interval $[t, t+s]$ is precisely the distribution of the number of independent Exponential(λ) random variables it takes before their sum exceeds s. This matches the definition of $N(s)$ for a Poisson process with rate λ, so we conclude $N(t, t+s) \stackrel{d}{=} N(s)$. □

A.1.3 Independent Increments Proof (Theorem 3)

Proof. Recall that our goal is to show that for $0 < t_1 < \cdots < t_k$, the set of random variables $N(t_1), N(t_1, t_2), \ldots, N(t_{k-1}, t_k)$ are jointly independent. Let $t = t_{k-1}$ for $1 \leq i \leq k$ and define $Z, Z_2, Z_3, \ldots$ as in the proof of the stationary increments property above. Observe that the lemma shows that Z is independent of $N(t_1), N(t_1, t_2), \ldots, N(t_{k-2}, t)$. In the previous proof, we have seen that the subsequent interarrival times $Z_2, Z_3, \ldots$ are independent of $N(t_1), N(t_1, t_2), \ldots, N(t_{k-2}, t)$ as well. It follows that $N(t, t_k)$ is independent of $N(t_1), N(t_1, t_2), \ldots, N(t_{k-2}, t)$, as $N(t, t_k)$ is a function of $Z, Z_2, Z_3, \ldots$. Thus the last increment $N(t_{k-1}, t_k)$ is independent of the previous ones.

Now, we can inductively claim that the entire set is independent. To see this, we can consider the joint pmf:

$$\begin{aligned}
&\mathbb{P}(N(t_1) = n_1, \ldots, N(t_{k-1}, t_k) = n_k) \\
&\qquad = \mathbb{P}(N(t_1) = n_1, \ldots, N(t_{k-2}, t_{k-1}) = n_{k-1}) \cdot \mathbb{P}(N(t_{k-1}, t_k) = n_k) \\
&\qquad = \mathbb{P}(N(t_1) = n_1) \cdot \ldots \cdot \mathbb{P}(N(t_{k-1}, t_k) = n_k)
\end{aligned}$$

so the independent increments property holds. ∎

A.1.4 pmf of $N(t)$ (Theorem 4)

Proof. We can consider two ways to calculate $\mathbb{P}(t < T_{n+1} < t+\delta)$, the probability the $(n+1)$th arrival arrives in $[t, t+\delta]$.

One way is to integrate over the density of T_{n+1}, the Erlang$(n+1; \lambda)$ distribution:

$$\mathbb{P}(t < T_{n+1} < t+\delta) = \int_t^{t+\delta} f_{T_{n+1}}(\tau) d\tau = f_{T_{n+1}}(t)(\delta + o(\delta))$$

Another way of computing $\mathbb{P}(t < T_{n+1} < t+\delta)$ is by splitting up the sample space by the number of arrivals in $[t, t+\delta]$:

$$\begin{aligned}
&\mathbb{P}(t < T_{n+1} < t+\delta) \\
&\quad = \mathbb{P}(N(t) = n, N(t, t+\delta) = 1) + \sum_{k=2}^{n+1} \mathbb{P}(N(t) = n+1-k, N(t, t+\delta) = k) \\
&\quad = \mathbb{P}(N(t) = n)\mathbb{P}(N(t, t+\delta) = 1) + \sum_{k=2}^{n+1} \mathbb{P}(N(t) = n+1-k)\mathbb{P}(N(t, t+\delta) = k)
\end{aligned}$$

The last equality holds due to the independent increments property.

Next, we calculate the individual terms that make up the expression. Observe, using the Erlang density derived above, that

$$\begin{aligned}
\mathbb{P}(N(t, t+\delta) \geq 1) &= \mathbb{P}(N(\delta) \geq 1) \\
&= \mathbb{P}(T_1 \leq \delta) \\
&= \int_0^{\delta} \lambda e^{-\lambda t}\, dt \\
&= 1 - e^{-\lambda\delta} \\
&= 1 - (1 - \lambda\delta + o(\delta)) \\
&= \lambda\delta + o(\delta) \\
\mathbb{P}(N(t, t+\delta) \geq 2) &= \mathbb{P}(N(\delta) \geq n) \\
&= \mathbb{P}(T_2 \leq \delta) \\
&= \int_0^{\delta} \lambda^2 t e^{-\lambda t}\, dt
\end{aligned}$$

$$
\begin{aligned}
&= 1 - e^{-\lambda\delta}(1 - \lambda\delta) \\
&= 1 - (1 - \lambda\delta + o(\delta))(1 + \lambda\delta) \\
&= o(\delta)
\end{aligned}
$$

Thus, the summation over k in our expression for $P(t < T_{n+1} < t + \delta)$ above is $o(\delta)$ because each of the terms is upper bounded by $\mathbb{P}(T_2 \leq \delta) = o(\delta)$. On the other hand, we have

$$\mathbb{P}(N(t, t + \delta) = 1) = \mathbb{P}(N(t, t + \delta) \geq 1) - \mathbb{P}(N(t, t + \delta) \geq 2) = \lambda\delta + o(\delta)$$

Thus it follows

$$\mathbb{P}(t < T_{n+1} < t + \delta) = \mathbb{P}(N(t) = n)(\lambda\delta + o(\delta)) + o(\delta)$$

Equating these two methods of computing the probability gives

$$
\begin{aligned}
f_{T_{n+1}}(t)(\delta + o(\delta)) &= \mathbb{P}(N(t) = n)(\lambda\delta + o(\delta)) + o(\delta) \\
\mathbb{P}(N(t) = n) &= \frac{1}{\lambda} f_{T_{n+1}}(t) + o(\delta)
\end{aligned}
$$

Then, taking $\delta \to 0$ and plugging in our expression for $f_{T_{n+1}}(t)$ gives $\mathbb{P}(N(t) = n) = \lambda^n t^n e^{-\lambda t}/n!$ as desired. □

A.2 Proof of Merging (Theorem 6)

Before we begin the proof, we introduce a fact that you may recall from earlier in the course: If $X \sim \text{Poisson}(\lambda)$ and $Y \sim \text{Poisson}(\mu)$ are independent, then $X + Y \sim \text{Poisson}(\lambda + \mu)$.

Proof. Exercise for the reader. □

Proof of Poisson Merging. To establish independent increments, consider some series of time instants $0 = t_0 < t_1 < \ldots t_k$ and consider the joint pmf of arrivals between these snapshots:

$$
\begin{aligned}
&\mathbb{P}(L(t_1) = n_1, \ldots, L(t_{k-1}, t_k) = n_k) \\
&= \sum_{i_1,\ldots,i_k} \mathbb{P}\left(N(t_1) = i_1, M(t_1) = n_1 - i_1, \ldots, N(t_{k_1}, t_k) = i_k, M(t_{k-1}, t_k) = n_k - i_k\right) \\
&= \sum_{i_1,\ldots,i_k} \mathbb{P}\left(\bigcap_{j=1}^{k} N(t_{j-1}, t_j) = i_j\right) \mathbb{P}\left(\bigcap_{j=1}^{k} M(t_{j-1}, t_j) = n_j - i_j\right)
\end{aligned}
$$

$$
\begin{aligned}
&= \sum_{i_1,\ldots,i_k} \prod_{j=1}^{k} \left(\mathbb{P}(N(t_{j-1},t_j) = i_j)\, \mathbb{P}(M(t_{j-1},t_j) = n_j - i_j)\right) \\
&= \prod_{j=1}^{k} \sum_{i_j} \left(\mathbb{P}(N(t_{j-1},t_j) = i_j)\, \mathbb{P}(M(t_{j-1},t_j) = n_j - i_j)\right) \\
&= \prod_{j=1}^{k} \mathbb{P}(L(t_{j-1},t_j) = n_j)
\end{aligned}
$$

The above derivation is a mess of indices, but the key idea is that we use Law of Total Probability to consider the values of L in terms of the possible values of N and M, and then we used the independence of the two processes and the independent increments property of N and M to factor the joint pmf. The interchange of the sum and product operators above is just using the distributive property.

To show that the increments are stationary, we perform a similar calculation as above for $s \geq 0$:

$$
\begin{aligned}
\mathbb{P}(L(t,t+s) = n) &= \sum_{i=1}^{n} \mathbb{P}(N(t,t+s) = i)\, \mathbb{P}(M(t,t+s) = n - i) \\
&= \sum_{i=1}^{n} \mathbb{P}(N(s) = i)\, \mathbb{P}(M(s) = n - i) \\
&= \mathbb{P}(L(s) = n)
\end{aligned}
$$

where we again use Law of Total Probability and the stationary increments property of N and M.

Finally, to show $L(t) \sim \text{Poisson}((\lambda+\mu)t)$, recall that $N(t) \sim \text{Poisson}(\lambda t)$ and $M(t) \sim \text{Poisson}(\mu t)$ the sum of independent Poisson random variables is Poisson with the sum of the parameters. $\square$

A.3 Proof of Splitting (Theorem 7)

As with for Poisson merging, we recall another fact from earlier in the course: Let $X \sim \text{Poisson}(\lambda)$ and $B_1, B_2, \ldots \sim_{iid} \text{Bernoulli}(p)$ independent of X. Let $X_0 = |\{i : B_i = 0, i \leq X\}|$ and $X_1 = |\{i : B_i = 1, i \leq X\}|$. Then $X_0 \sim \text{Poisson}(\lambda p)$, $X_1 \sim \text{Poisson}(\lambda(1-p))$, and X_0 and X_1 are independent of each other.

Proof. Exercise for the reader. □

Proof of Poisson Splitting. As with for merging, to show that N_0 and N_1 are Poisson processes, we need to show that the increments are independent and stationary and that the number of arrivals in a given interval a Poisson random variable. We will show that these properties hold for N_0; a similar set of steps could be used to show that N_1 is a Poisson process.

To show that the increments are independent, consider some series of time instants $0 = t_0 < t_1 < \ldots < t_k$ and write the joint pmf

$$\begin{aligned}
&\mathbb{P}\left(N_0(t_1) \le n_1, \ldots, N_0(t_{k-1}, t_k) \le n_k\right) \\
&= \sum_{m_1 \ge n_1, \ldots, m_k \ge n_k} \left[\mathbb{P}\left(\bigcap_{j=1}^{k} N(t_{j-1}, t_j) \le n_j\right) \prod_{j=1}^{k} \left(\binom{m_j}{n_j} p^{n_j} (1-p)^{m_j - n_j}\right)\right] \\
&= \sum_{m_1 \ge n_1, \ldots, m_k \ge n_k} \left[\prod_{j=1}^{k} \left(\mathbb{P}\left(N(t_{j-1}, t_j) \le n_j\right) \binom{m_j}{n_j} p^{n_j} (1-p)^{m_j - n_j}\right)\right] \\
&= \prod_{j=1}^{k} \left[\sum_{m_j \ge n_j} \left(\mathbb{P}\left(N(t_{j-1}, t_j) \le n_j\right) \binom{m_j}{n_j} p^{n_j} (1-p)^{m_j - n_j}\right)\right] \\
&= \prod_{j=1}^{k} \mathbb{P}(N(t_{j-1}, t_j) = n_j)
\end{aligned}$$

where, just as we did for the proof of Poisson merging, used Law of Total Probability to write our expression in terms of the original process $N(t)$ and then use the independent increments property of $N(t)$ to break up the joint probabilities. The tricky step above was noting that because each arrival is routed to N_0 according to an independent Bernoulli trial, then the distribution of $N_0(t_1, t_2)$ conditioned on $N(t_1, t_2) = n$ is a binomial random variable, independent of the of the other intervals.

Stationarity follows by a similar calculation:

$$\begin{aligned}
\mathbb{P}(N_0(t, t+s) = n) &= \sum_{m \ge n} \mathbb{P}(N(t, t+s) = m) \binom{m}{n} p^n (1-p)^{m-n} \\
&= \sum_{m \ge n} \mathbb{P}(N(s) = m) \binom{m}{n} p^n (1-p)^{m-n} \\
&= \mathbb{P}(N_0(s) = n)
\end{aligned}$$

Finally, note that $N(t) \sim \text{Poisson}(\lambda t)$ and observe that arrivals are independently being routed to $N_0(t)$ according to Bernoulli trials. According to the splitting property of Poisson random variables, it follows that $N_0(t) \sim \text{Poisson}(\lambda p t)$. Thus we have shown that $N_0(t)$ is a Poisson process with parameter λp; a similar argument shows that $N_1(t)$ is a Poisson process with parameter $\lambda(1-p)$.

However, we are not done with the proof – we have not yet established that the two processes are independent of each other. Again, we consider some series of time steps $0 = t_0 < t_1 < \ldots < t_k$ and we want to show that the sets $\{N_0(t_{i-1}, t_i) \,:\, 1 \leq i \leq k\}$ and $\{N_1(t_{j-1}, t_j) \,:\, 1 \leq j \leq k\}$ are independent of each other. Note that for $i \neq j$, then independence holds because the arrivals in $N_0(t_{i-1}, t_i)$ and $N_1(t_{j-1}, t_j)$ are independently chosen from $N(t_{i-1}, t_i)$ and $N(t_{j-1}, t_j)$, which are themselves independent because of the independent increments of $N(t)$. Thus it suffices to check that independence holds for $i = j$; however, this holds because we can apply the splitting property of Poisson random variables to $N(t_{i-1}, t_i)$. $\square$