Applied Probability and Stochastic Models

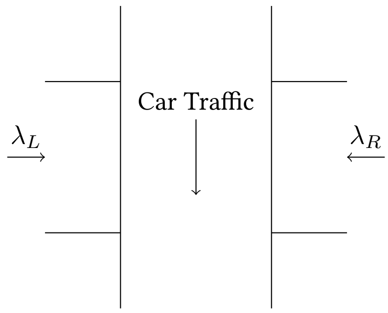
Michael Li

October 23, 2017

Contents

1 Poisson Processes

Now we consider the following problem, modeled after 77 Mass Ave:



We have:

- The number of pedestrians waiting on the left is a poisson process with rate λ_L.
- The number of pedestrians waiting on the right is a poisson process with rate λ_R.
- The car traffic passes through homogeneously.
- There is a traffic light in which on red pedestrians can cross. We assume pedestrians cross instantaneously.

Here is a handy review of the Poisson process:

Review. A *Poisson Process* is a continuous time discrete state markov chain X_t, with an initial state of 0 and rate λ so that it satisfies $X_t \sim Poi(\lambda t)$ for all t.

It can be alternatively defined in the following way: Let us have a continuous time discrete state markov chain X_t that starts at 0. Define the waiting time S_t as the amount of time before X_t increments. Now we let $S_t \sim \exp(\lambda)$. Then X_t is a *Poisson Process.*

It has the following cool properties:

Addition if we add two *independent* Poisson processes with rate λ_1 and λ_2, we get a Poisson process with rate $\lambda_1 + \lambda_2$.

Thinning Now consider a Poisson process X_t with rate λ. Now create a process Y_t that starts at 0, and increments as followed:

When X_t increments, throw a coin that has probability p of being heads up. If it is heads up, Y_t increments. If not, Y_t stays the same.

The *Thinning Property* states that Y_t is a Poisson process with rate λp. Furthermore, it is independent from X_t!

Now we apply this to some situations [Selected from Class]:

1 **Assume light just turned green. What is the expected number of pedestrians waiting after time T?**

Answer. This is just the expected value of X_T, which is $(\lambda_L + \lambda_R)T$. $\square$

2 **Assume the light has just turned green. If we have a button at the crossing, and after the button is pressed the light would turn red after time T_0, what is the probability that there is no one crossing from left to right when the light turns red?**

Answer. Now the total time to the next red is $T + T_0$, where T is the time till the first person presses the button. Thus, we need the person pressing the button to be from the right, and that has probability:

$$\frac{\lambda_R}{\lambda_L + \lambda_R}$$

We can see it in multiple ways, but I prefer the math-theoretical way.

Theorem. Let $X_1, \cdots, X_n$ be exponential variables with rates $\lambda_1, \cdots, \lambda_n$ respectively. Let $\lambda = \lambda_1 + \cdots + \lambda_n$ and denote k such that $X_k = \min X_1, \cdots X_n$. Then $P(k = i) = \frac{\lambda_i}{\lambda}$.

Proof. Consider the joint probability density function $f_{X_k,k}$. We have:

$$f_{X_k,k}(t,i) = \underbrace{\lambda_i e^{-\lambda_i t}}_{\text{PDF that } X_i = t} \times \underbrace{e^{-\lambda_1 t} \cdots e^{-\lambda_n t}}_{\text{Probability that other variables are greater than } t}$$

Where the $\cdots$ range from 1 to n excluding i. Now we take the integral over t to get the result that:

$$P(k = i) = \int f_{X_k,k}(t,i)\mathrm{d}t = \frac{\lambda_i}{\lambda}$$

$\square$

Applying it to the waiting times of the two poisson processes on the left/right, which are exponentially distributed, gives the answer.

Ok , so we have solved the first part. Now after the right person has pressed the button, in fixed time T_0, we need to have no people from the left side. This has probability $e^{-\lambda_L T_0}$. Note that by the definition of a Markov Chain, condition on the time that a person on the right pressed the button, the left side is an *independent* Poisson process with rate λ_L. This is called the *Strong Markov Property.*

Theorem (Strong Markov Property)**.** Let (X_t) be a poisson process with rate λ and let T be a stopping time. Then conditional on T, the process $(X_{T+s} - X_T)$ for $s \geq 0$ is also a poisson process of rate λ and is independent from X_s with $s \leq T$.

Note. Don't remember what's a stopping time? It is just an event such that for all t the event $T \leq t$ only depends on X_s with $s \leq t$. Here the event of a person on the right is a stopping time as the event happening at time t only depends on the present state of the poisson processes on the left and right (Namely the left should have 0 people and the right should have 1).

So we can simply multiply these probabilities together to get that:

$$\frac{\lambda_R}{\lambda_L + \lambda_R} e^{-\lambda_L T_0}$$

□

3 **What is the expected time that a randomly arriving pedestrian must wait until crossing under the rule that pedestrians cross when there are N_0 people waiting?**

Answer. Here we have to clarify one thing : A randomly arriving pedestrian here means that the person is randomly slotted into a time of a day. This means that he is equally likely to be the 1st, 2nd, and the N_0th person to arrive.

For each case that he arrives as the ith, $i = 1, \cdots, N_0$, he needs to wait, on average for $N_0 - i$ people, so in total of $\frac{N_0 - i}{\lambda}$ time. Thus the expected time is just the average:

$$\bar{W} = \frac{1}{N_0}\left(0 + \frac{1}{\lambda} + \cdots + \frac{N_0 - 1}{\lambda}\right) = \frac{N_0 - 1}{2\lambda}$$

□

1.1 Detour: Renewal Theory and Size-bias

4 **Now let's assume the traffic light operates under the rule that when the first person arrives at the crosswalk, it would turn green after T_0. What is the expected wait time of the an observer arriving at the crosswalk before the light turns green?**

Answer. Now this question is harder than it looks, due to the fact that it is *size-biased.* To solve this problem, we need to introduce size-bias formally:

Definition. Consider a stochastic process in discrete state X_t such that the length of waiting times are distributed with pdf $f_W(w)$. Now fix $t = T$, and we randomly pick a $w \in [0, T]$ and look at the length of the waiting time interval that X_w is in, Y. Then Y has the *size-biased distribution of $f_W(w)$*, which is defined as:

$$f_Y(w) = \frac{w f_W(w)}{E(W)}$$

Now we break this down in an example:

Example (Definition Explained)**.** Consider the arrival of buses (a stochastic process), and thus $f_W(w)$ is the distribution of time between successive buses. Then if we randomly arrive at the bus station, the distribution of time between last bus and the next bus is Y, and that is *size-biased.*

This is intuitive as we arrive randomly, we have a higher chance of arriving at a time where time between successive buses are higher, and thus *biases* towards larger sizes.

Example (Poisson Process Mean)**.** Consider the bus process above, and assume it is a Poisson process with rate λ. Then the waiting time is distributed exponentially with mean $\frac{1}{\lambda}$.

Now you arrive at the bus station. By the memoryless property of exponential variable, the expected time to the next bus is still $\frac{1}{\lambda}$ (as given that you have arrived T minutes after the last bus left does not change the waiting time distribution for any T, by the memoryless property).

But if you turn the Poisson process reverse in time (Poisson processes are time reversible), the properties of the Poisson process does not change and the expected time to the last bus is also $\frac{1}{\lambda}$.

So the total expected time between two buses is $\frac{1}{\lambda} + \frac{1}{\lambda} = \frac{2}{\lambda}$, and not $\frac{1}{\lambda}$ (which is the mean length of the waiting time)!

This illustrates the size-biasness. The fact that you are arriving randomly means you are more likely to fall in an interval where the waiting time is longer, and the larger mean proves that.

Now we are going to quantify this in Mathematics:

Theorem. Given a waiting time distribution W and its size-biased variable (waiting time on random arrival) Y, then we have the following results:

- The expected length of waiting time interval on random arrival is:

$$E(Y) = E(W)(1 + \eta^2) \qquad \eta = \frac{\sigma_Y}{E[Y]}$$

- The distribution of time you have to wait till next arrival, X, given that you fell into an waiting time interval of length W is uniformly distributed. That is:

$$X|(Y = W) \sim U[0, W] \quad \Rightarrow \quad E[X|Y = W] = \frac{W}{2}$$

- For a random arrival, the expected time you have to wait till next arrival is:

$$E[X] = \frac{E[Y]}{2} = \frac{E(W)(1 + \eta^2)}{2}$$

Proof. - We have:

$$\begin{aligned}
E(Y) &= \int_{-\infty}^{\infty} w \times w f_W(w) \mathrm{d}w \\
&= E(W^2) \\
&= \mathrm{Var}(W) + E(W)^2 \\
&= E(W)\left(1 + \frac{\mathrm{Var}(W)}{E(W)^2}\right)
\end{aligned}$$

Defining $\eta = \frac{\sigma_Y}{E[Y]}$ gives the result.

- This is by assumption [Actually it is provable by measure theory, but we would not do this here] as we are arriving randomly.
- We have:

$$E[X] = \int_{-\infty}^{\infty} E[X|Y] w f_W(w) \mathrm{d}w = \int_{-\infty}^{\infty} \frac{w}{2} \times w f_W(w) \mathrm{d}w = \frac{E[Y]}{2}$$

□

Now look at the last expression. If we did not have size bias, we would expect to wait $\frac{E[W]}{2}$, as we would randomly arrive into an interval of length W. But because of size bias, we actually have:

$$\frac{E[W]}{2}\left(1+\eta^2\right)$$

So the variability provides the bias here.

Going back to our question, we have a waiting time mean of $\frac{1}{\lambda} + T_0$, and:

$$\eta^2 = \frac{Var(W)}{E(W)^2} = \frac{\frac{1}{\lambda^2}}{\frac{1}{\lambda} + T_0}$$

So our question asks about $E[X]$, which is:

$$E[X] = \frac{1}{2}\left(\frac{1}{\lambda} + T_0\right)\left(\frac{1+\lambda+T_0\lambda^2}{\lambda+T_0\lambda^2}\right) = \frac{1}{2}\left(\frac{1}{\lambda^2} + \frac{1}{\lambda} + T_0\right)$$

□

2 Function of Random Variables

2.1 Quick Review of Basics

Let's say we have two random variables X, Y with pdfs $f_X(x)$ and $f_Y(y)$. We would ask the following questions:

1 What is the PDF of $X+Y$, assuming X, Y are independent?

Answer. Now define $W = X + Y$. We look at $f_W(w)$. What is this? For $W = w$, we need $X = a$ and $Y = w - a$, and it could happen for any such a. So we have:

$$f_W(w) = \int_{-\infty}^{\infty} f_X(a) f_Y(w-a) \mathrm{d}a$$

□

2 **What is the PDF of $W = X + Y$, assuming $X, Y \sim U[0,1]$?**

Answer. We apply the formula above:

$$f_W(w) = \int_{-\infty}^{\infty} f_X(a) f_Y(w-a) \mathrm{d}a$$

Now if you think about uniform random variables, you will realize we need to separate this integral into two pieces, one piece when $w > 1$ and one when $w \leq 1$.

For $w \leq 1$, we have:

$$f_W(w) = \int_0^w 1 \mathrm{d}a = w$$

For $w > 1$, we have:

$$f_W(w) = \int_{w-1}^1 1 \mathrm{d}a = 2 - w$$

So in total we have:

$$f_W(w) = \begin{cases} w & w \leq 1 \\ 2 - w & w > 1 \end{cases}$$

□

2.1.1 Review: Central Limit Theorem

This is the most famous theorem in statistics. Period.

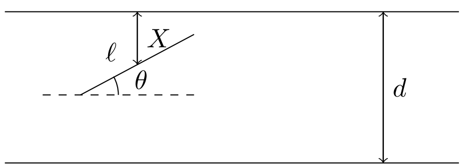
Theorem. Assume we have iid random variables $X_1, X_2, \cdots$, each distributed with mean μ and variance $\sigma^2 < \infty$. Then we have:

$$\frac{\sum_{i=1}^n X_i - n\mu}{\sqrt{n}} \to N(0,1) \qquad n \to \infty$$

2.2 Buffon's Needle Experiment

Consider the following problem:

You have a flag with stripes of length d, and assume you have a needle of length $l < d$. If you randomly drop a needle needle e on the floor, what's the probability that the needle would be contained in exactly 1 stripe?



The trick here is to consider variables X: the distance from the center of the needle to the closest line of stripe, and θ, the acute angle the needle makes with any stripe.

As we are throwing it randomly, we assume $X \in U[0, \frac{d}{2}]$ and $\theta \in U[0, \frac{\pi}{2}]$. Now we want to look at the probability that the needle stays within a stripe. That means, the vertical distance of the needle, $\frac{l}{2}\sin\theta$ has to be less than X. So the region is $X < \frac{l}{2}\sin\theta$. Then we integrate:

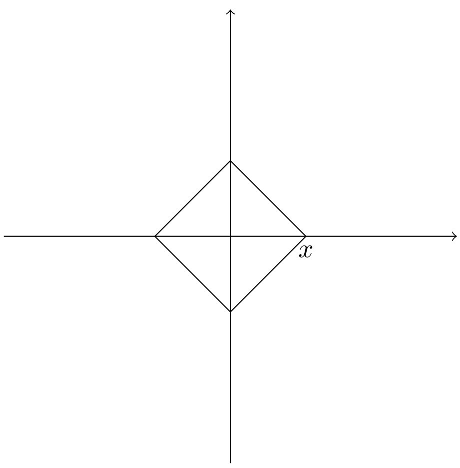
$$
\begin{aligned}
P &= \int_0^{\frac{\pi}{2}} \int_0^{\frac{l}{2}\sin\theta} \frac{4}{d\pi} \mathrm{d}x \mathrm{d}\theta \\
&= \int_0^{\frac{\pi}{2}} \frac{2l}{d\pi} \sin\theta \mathrm{d}\theta = \frac{2l}{d\pi}
\end{aligned}
$$

Note. Don't try to do the θ integral first - no luck there.

Note. Why is the $d < l$ assumption important? Hint, the devil is in the integral (for θ).

3 Geometrical Probability

Consider the following graph:



Consider a uniform random variable dropping in this diamond area of size $2x^2$. Consider its coordinates (X_1, Y_1) as random variables. Now we drop another one as (X_2, Y_2). What will be expected value of its Manhattan distance, or $E[D]$ where $D = |X_1 - X_2| + |Y_1 - Y_2|$?

Answer. First we half the amount of work by realizing X and Y are exactly the same so we just need to consider $|X_1 - X_2|$. We would do this the Mathematical way [There is a easier way if we just want the expectation, but why?]:

Let $D = k$, then we have $X_1 = a$, and $X_2 = a + k$ or $X_2 = a - k$. now lets do the case of $a + k$ first:

$$
\begin{aligned}
&\int_{-x}^{x-k} f_X(a) f_X(a+k) \mathrm{d}a \\
&= \int_{0}^{x-k} f_X(a) f_X(a+k) \mathrm{d}a + \int_{-k}^{0} f_X(a) f_X(a+k) \mathrm{d}a + \int_{-x}^{-k} f_X(a) f_X(a+k) \mathrm{d}a \\
&= \int_{0}^{x-k} \frac{(x-a)(x-a-k)}{x^4} \mathrm{d}a + \int_{-k}^{0} \frac{(x+a)(x-a-k)}{x^4} \mathrm{d}a + \int_{-x}^{-k} \frac{(x+a)(x+a+k)}{x^4} \mathrm{d}a \\
&= \frac{1}{x^4} \left(\left[\frac{(a-x)^3}{3} + k \frac{(a-x)^2}{2} \right]_0^{x-k} + \left[x^2 a - \frac{a^3}{3} - k \frac{(a+x)^2}{2} \right]_{-k}^{0} + \left[\frac{(a+x)^3}{3} + k \frac{(a+x)^2}{2} \right]_{-x}^{-k} \right) \\
&= \frac{1}{x^4} \left(-\frac{k^3}{3} + \frac{k^3}{2} + \frac{x^3}{3} - \frac{kx^2}{2} + kx^2 - \frac{k^3}{3} - \frac{kx^2}{2} + k \frac{(x-k)^2}{2} + \frac{(x-k)^3}{3} + k \frac{(x-k)^2}{2} \right) \\
&= \frac{1}{x^4} \left(-\frac{k^3}{6} + \frac{x^3}{3} + k(x-k)^2 + \frac{(x-k)^3}{3} \right)
\end{aligned}
$$

Then we similarly do it for the $-k$ one:

$$
\begin{aligned}
&\int_{-x+k}^{x} f_X(a) f_X(a-k) \mathrm{d}a \\
&= \int_{k}^{x} f_X(a) f_X(a-k) \mathrm{d}a + \int_{0}^{k} f_X(a) f_X(a-k) \mathrm{d}a + \int_{-x+k}^{0} f_X(a) f_X(a-k) \mathrm{d}a \\
&= \int_{k}^{x} \frac{(x-a)(x-a+k)}{x^4} \mathrm{d}a + \int_{0}^{k} \frac{(x+a)(x-a+k)}{x^4} \mathrm{d}a + \int_{-x+k}^{0} \frac{(x+a)(x+a-k)}{x^4} \mathrm{d}a \\
&= \frac{1}{x^4} \left(\left[\frac{(a-x)^3}{3} - k \frac{(a-x)^2}{2} \right]_k^{x} + \left[x^2 a - \frac{a^3}{3} + k \frac{(a+x)^2}{2} \right]_{0}^{k} + \left[\frac{(a+x)^3}{3} - k \frac{(a+x)^2}{2} \right]_{-x+k}^{0} \right) \\
&= \frac{1}{x^4} \left(\frac{(x-k)^3}{3} + k \frac{(x-k)^2}{2} + x^2 k - \frac{k^3}{3} + k \frac{(x+k)^2}{2} - k \frac{x^2}{2} + \frac{x^3}{3} - \frac{kx^2}{2} - \frac{k^3}{3} + \frac{k^3}{2} \right) \\
&= \frac{1}{x^4} \left(\frac{5k^3}{6} + \frac{x^3}{3} + kx^2 + \frac{(x-k)^3}{3} \right)
\end{aligned}
$$

Then the addition gives:

$$
f_D(k)\mathrm{d}k = \frac{1}{x^4} \left(\frac{2k^3}{3} + \frac{2x^3}{3} + kx^2 + k(x-k)^2 + \frac{2(x-k)^3}{3} \right) = \frac{1}{x^4} \left(\frac{4x^3}{3} + k^3 + 2xk^2 \right)
$$

For $k > x$, we have:

$$\begin{aligned}
f_D(k)\mathrm{d}k &= 2\int_{-x}^{x-k} f_X(a) f_X(a+k)\mathrm{d}a \\
&= \int_{-x}^{x-k} \frac{2(x+a)(x-a-k)}{x^4}\mathrm{d}a \\
&= \frac{2}{x^4}\left[x^2 a - \frac{a^3}{3} - k\frac{(a+x)^2}{2}\right]_{-x}^{x-k} \\
&= \frac{2}{x^4}\left[(2x-k)x^2 - \frac{(x-k)^3}{3} - \frac{x^3}{3} + k\frac{(2x-k)^2}{2}\right] \\
&= \frac{2}{x^4}\left[\frac{4x^3}{3} - 3xk^2 + 2x^2k + \frac{5k^3}{6}\right]
\end{aligned}$$

Then we can finally calculate the expected value:

$$\begin{aligned}
E[D] &= \frac{1}{x^4}\left(\int_0^x k\left[\frac{5k^3}{6} + \frac{x^3}{3} + kx^2 + \frac{(x-k)^3}{3}\right]\mathrm{d}k + \int_x^{2x} k\left[\frac{4x^3}{3} - 3xk^2 + 2x^2k + \frac{5k^3}{6}\right]\mathrm{d}k\right) \\
&= \frac{1}{x^4}\left(\frac{x^5}{6} + \frac{x^5}{6} + \frac{x^5}{60} + 2x^5 - \frac{45x^5}{4} + \frac{14}{3}x^5 + \frac{31}{6}x^5\right) \\
&= x\left(\frac{10+10+1+120+310+280-675}{60}\right) \\
&= \frac{14}{15}x
\end{aligned}$$

□

4 Queueing Theory

4.1 Review of Markov Chains

Note. Because of my own reservations, I have decided to present this material in a more formal (but still intuitive) way to aid understanding.

We start by reviewing the definition of a discrete time Markov Chain.

Definition. The process is called an *discrete-time Markov Chain* with state space S iff for all $x_0, x_n \in S$ we have:

$$P(X_n = x_n | X_0 = x_0, \cdots X_{n-1} = x_{n-1}) = P(X_n = x_n | X_{n-1} = x_{n-1})$$

Basically it means that any future state only depends on the last known state.

In the similar spirit, we define a continuous-time one:

Definition. The process X_t is called a *continuous-time Markov Chain* if for all $x_1, \cdots x_n \in S$ and all times $0 \leq t_1 \leq t_2, \cdots \leq t_n$ we have:

$$\mathbb{P}(X_{t_n} = x_n | X_{t_{n-1}} = x_{n-1} \cdots X_{t_1} = x_1) = \mathbb{P}(X_{t_n} = x_n | X_{t_{n-1}} = X_{n-1})$$

A *homogeneous* Markov chain is one where the righthand side only depends on $t_n - t_{n-1}$.

Similarly to discrete-time Markov chain, we define $P(t)_{xy} = P(X_t = y | X_0 = x)$ and call it the *transition semigroup* of this Markov Chain.

Now in the lectures we talked about transition rates from one state i to another state j. This is formally defined as the Q-matrix.

Definition. let S be a countable set. Then a *Q-matrix* on S is a matrix that satisfies $0 \leq -q_{ii} < \infty$, $q_{ij} \geq 0$ for all $I \neq j$, and $\sum_j q_{ij} = 0$ for all i.

We then define $q_i = -q_{ii}$ and given a Q-matrix we define the *jump stochastic matrix* as followed:

For $q_x \neq 0$:

$$p_{xy} = \frac{q_{xy}}{q_{xx}} \qquad p_{xx} = 0$$

If $q_x = 0$, then $p_{xy} = \mathbf{1}_{\{x=y\}}$.

So the Q-matrix encodes the rates of the transition while the P-matrix (*jump stochastic matrix*) encodes the probability of the embedded markov chain jumping from one state to another.

Now we can finally formally define a continuous-time Markov chain:

Definition. A Markov process X with $X_0 = \lambda$, generator Q, a Q-matrix, is a stochastic process with jump chain $Y_n = X_{J_n}$ being a discrete time MarkoV chain with $Y_0 = \lambda$ and transition matrix P. Moreover, conditional on $Y_0, \cdots Y_n$, the holding times $S_i \sim \exp(q_{Y_{i-1}})$.

4.2 Small Diversion to Recurrence/Transience

Note. This section is not taught in class but I thought this would be very helpful in achieving a greater understanding to queueing theory.

In our class, we explained the concept of *steady-state equilibrium*, but it turns out that not all Markovian chains have this! To *truly* understand steady-state equilibrium, we need to dive into the concept of recurrence/transience:

Definition. For a Markovian process X_t, we call a state *recurrent* if $P_x(\{t : X_t = x\}$ is unbounded$) = 1$. It is *transient* if this probability is 0. The state is *positively recurrent* if it is recurrent and the mean return time of the markov process X_t to this state, given one starts there, is finite.

Basically, a recurrent state is one which the markov chain returns to an infinite number of times, and transient state only a finite number of times before never returning. [With Probability 1]

Next up we introduce steady-state equilibrium, or in its math jargon, *invariant distributions.*

Definition. let Q be the generator of a continuous time Markov chain and let λ be a measure. It is *invariant* if $\lambda Q = 0$. It is an *invariant distribution* if $\sum_i \lambda_i = 1$.

Decoding this, $\lambda Q = 0$ are the *flow balance equations* we talked about in class (as to balance the probability flux going into and out of a state in steady-state), written in the fancy matrix form. Think of λ_i as the probability of being in such state in steady-state.

Then we have an important theorem:

Theorem. Let X be an irreducible continuous time Markov chain with generator Q. Then:

Some state is positive recurrent $\Leftrightarrow$ Every state is positive recurrent $\Leftrightarrow$ X is not explosive and has invariant distribution $\lambda(x) = (q_x m_x)^{-1}$, where $m_x = E_x[T_x]$, the mean expected time to return to state x.

What this tells us is that a Markov process has a steady-state only if it returns to every state infinitely often and has a finite average return time.

Finally, we have the following theorem that cements what we learned in class, about steady state being a long-term average:

Theorem. Let X be irreducible, non-explosive, and continuous Markov with generator Q. Suppose λ is invariant, then:

$$p_{xy}(t) \to \lambda(y) \quad \text{as} \quad t \to \infty$$

That means, as $t \to \infty$, the probability of staying in any state converges to the invariant distribution, or the steady state!

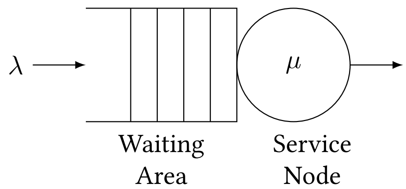
Note. Non-explosive is a *extremely* technical Markov process concept which can be basically summed up as: The process has 0 probability of changing states infinitely often in finite times, so everything is nice.

4.3 Queues: $M/M/1$ Queue and $M/M/\infty$ Queue

We have people coming in. We then have k servers serving the people (and after that they leave). This, obviously is called a *queue*. We characterize queues in the following way:

$$\underbrace{M}_{\text{The method customers arrive in queue}} \quad / \quad \underbrace{G}_{\text{The method customers are being served}} \quad / \quad \underbrace{k}_{\text{Number of Servers}}$$

where M means Markovian (or Memoryless, depending on who you ask), G means general. Typically $k = 1$ or ∞.



Definition. A $M/M/1$ queue, shown above, is one that customers arrive in a poisson process of rate $\lambda > 0$, and the single server service customer with service time exponential with parameter μ. Equivalently, $q_{i,i+1} = \lambda$, $q_{i,i-1} = \mu$.

Theorem. Let $\rho = \frac{\lambda}{\mu}$. Then X, a $M/M/1$ queue is transient iff $\rho > 1$, recurrent iff $\rho \leq 1$, and positive recurrent iff $\rho < 1$.

In the positive recurrence case X has equilibrium distribution $\pi(n) = (1-\rho)\rho^n$. The waiting time $W \sim \exp(\lambda - \mu)$ at equilibrium.

Note. Note here *waiting time* means waiting + service time! I know, math jargon is confusing.

Remark. We can solve for all positive recurrent results using the flow balance equations $q_{i,i+1}P_i = q_{i+1,i}P_{i+1}$ for all i, and here the Q-matrix is just $q_{i,i+1} = \lambda$ and $q_{i+1,i} = \mu$. Using these equations and the equation:

$$\sum_{i=0}^{\infty} P_i = 1$$

Gives you the solution to the equilibrium distribution, as the first set of equations can be rearranged to read:

$$P_n = \frac{\lambda}{\mu} P_{n-1} = \left(\frac{\lambda}{\mu}\right)^n P_0$$

And then use the sum equation to solve for P_0, then all others.

Once we have that $\pi(n) = (1-\rho)\rho^n$, we can calculate pdf of waiting time as:

$$\begin{aligned} f_W(w) &= \sum_{k=1}^{\infty} f_W(w|k)P(X_t = k) \\ &= \sum_{k=1}^{\infty}(1-\rho)\rho^k \frac{\mu^k}{(k-1)!} w^{k-1} e^{-\mu w} \\ &= e^{(\lambda-\mu)w} \end{aligned}$$

This calculation is basically done through conditioning on the number of people in the queue (including the guy being serviced). If you read the section above, then the transient result is also easy to understand: if $\rho > 1$, you have more people coming in than going out, so your queue length would go to infinity and never come back. The $\rho = 1$ case is just a technicality.

Definition. A $M/M/\infty$ queue is same as a $M/M/1$ queue but with infinitely many servers, so everyone gets served immediately.

Theorem. The queue length X_t is positive recurrent for all $\lambda, \mu > 0$. Furthermore the invariant distribution is Poisson with rate λ/μ.

Proof. The quantitative results follows immediately after solving the flow balance equation. $\square$

4.4 M/G/1 Queues

Now assume we have a $Poi(\lambda)$ arrival but only ϵ_i iid service times. As in, our arrival process is still Memoryless, but we have a General (G) distribution for our service times. What can we do? Turns out we can still find a Markov chain here, if we look at the Markov process at the end of each service just before the person leaves (here D_n denotes the nth departure from the system):

Proposition. The process $X(D_n)$ is a Markov chain with transition probability $p_{i,i+k-1} = p_{0k} = E[e^{-\lambda\epsilon}(\lambda\epsilon)^k/k!]$ for all $i > 0$, $k \geq 0$, and 0 otherwise.

Note. Basically, what this says is that the probability that at the next end of service, there would be $i + k - 1$ people in the system, given that there are i people in the system at the end of current service is:

$$E[e^{-\lambda\epsilon}(\lambda\epsilon)^k/k!]$$

Proof. If we look at the process between departures, there are only arrivals. The probability that k customers arrive (and then one leave) is Markov, and we know how to calculate that:

$$P(A_n = k) = E[P(A_n = k|\epsilon_n)] = E[e^{-\lambda\epsilon}(\lambda\epsilon)^k/k!]$$

Done. The 0 case accounts for the fact that no one can be leaving if no one is there. $\square$

4.5 Spatial Markov Queues

Similar to a time Markov chain, we can define a spatial Markov chain:

Definition. For a Borel measurable set $S \subset \mathbb{R}$, we define the Poisson process of intensity γ, $X(S)$, as:

$$P(X(S) = k) = \frac{(\gamma A(S))^k}{k!} e^{-\gamma A(S)}$$

Where $A(S)$ is the area of S, and $X(S) = k$ means there are k poisson events in the area of S.

It has the similar property to time poisson processes that given $X(S) = k$, the location of the k events are uniformly randomly distributed in S.