# Stats 150, Fall 2019
Lecture 17, Thursday, 10/31/2019

To motivate analogies between discrete and continuous Markov chains, we'll look at one of this week's homework problems.

## 1 Remarks about the Homework

From this week's homework, recall the cute problem regarding a repairman who completes tasks according to an exponential ($\mu$) distribution. However, he is interrupted by a Poisson point process ($\lambda$).

Recall that in the past few lectures in constructing our Poisson Point Process, we considered more than just Poisson points on a line (we instead consider points on a strip). We relate our present problem to our first picture of a Poisson point process with rate 1 per area.
We remove the vertical bars as they were only parts of our construction. That is, the unit squares are only an artifact of our first construction. We found that any figure of unit area follows this Poisson point process. Now we want to mark the $\mu$ points differently and we project downwards.

By our construction (and familiar properties) of Poisson point processes, we have a few facts to take from our picture.

- The $\times\times\times\times\cdots\times\times$ marks on the timeline are a Poisson point process with rate $\lambda$. We drew the picture to make it so, projecting a $1\times\lambda$ rectangle down to the timeline with width 1.
- The $\otimes\otimes\otimes\cdots\otimes\otimes$ similarly are a Poisson point process with rate $\mu$.
- The $\times\times\cdots$ and $\otimes\otimes\cdots$ points are independent because in the planar Poisson point process, the areas constructed from each of the $\lambda$ and $\mu$ strips will be disjoint. Precisely, these strips are $[0,\infty)\times[0,\lambda]$ and $[0,\infty)\times[\lambda,\lambda+\mu]$.
- Moreover, consider the discrete sequence of marks $\times\times\times\times\times\times\otimes\times\times\times\times\otimes$, by throwing away the timing and only looking at the sequence. These are independent Bernoulli($\times/\otimes$) trials, where $\mathbb{P}(\times)=\frac{\lambda}{\lambda+\mu}$ and $\mathbb{P}(\otimes)=\frac{\mu}{\lambda+\mu}$.
- In the continuous model, let $T_\times$ be the continuous time to the first $\times$, so that $T_x\sim$ Exponential($\lambda$). Let $T_\otimes$ be the continuous time to the first $\otimes$, so that $T_\otimes(\mu)$. This is a race between independent exponentials (as they are disjoint parts of a PPP). It turns out
$$\mathbb{P}(T_\times<T_\otimes)=\frac{\lambda}{\lambda+\mu}.$$
In better form, let $T_\lambda$ and $T'_\mu$ be two independent Exponential($\lambda$) and Exponential($\mu$) variables. Then $\mathbb{P}(T_\lambda<T'_\mu)=\frac{\lambda}{\lambda+\mu}$, and the event $(T_\lambda<T'_\mu)$ is independent of $T_\lambda\wedge T'_\mu:=\min(T_\lambda,T'_\mu)$.

Now, bringing these ideas back to our homework problem, given that the repairman is interrupted 6 times, the waiting time until the first completion is Poisson ($\mu$). Pitman gives that this picture of the Poisson point process makes this question a lot more trivial than it may seem. This spacing is Exponential($\lambda+\mu$). This is a race between exponentials, so given that the repairman finishes the task, he must have done so at a faster rate.

We may ask, what happens in this model if interruptions are a renewal process? Immediately, we do know that our Poisson tricks does not work in this case.

## 2 Further Remarks

Moreover, looking at $T_\otimes$ as the time to first $\otimes$ as $T_\otimes = \mathcal{E}_1 + \mathcal{E} + \cdots + \mathcal{E}_N$, where the $\mathcal{E}_i$ are the Exponential$(\lambda + \mu)$ spacings before the points either $\times$ or $\otimes$. We know

$$\mathbb{P}(N = n) = \mathbb{P}(\overbrace{\times \times \times \cdots \times}^{n-1} \otimes) = \left(\frac{\lambda}{\lambda + \mu}\right)^{n-1} \left(\frac{\mu}{\lambda + \mu}\right),$$

which implies $N$ is Geometric$\left(\frac{\mu}{\lambda+\mu}\right)$. In this model, we have that $N$ is independent of $\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3, \ldots$. We see then that the Exponential$(\mu)$ time $T_\otimes$ is represented in this picture as the sum of a Geometric $\left(\frac{\mu}{\lambda+\mu}\right)$ number of copies of iid Exponential$(\lambda+\mu)$. Therefore, the poisson strip model offers an explanation or proof of the basic fact that if $\mathcal{E}_1, \mathcal{E}_2, \ldots$ are iid Exponential$(\nu)$ and $N$ is independent of these with Geometric$(p)$ on $\{1, 2, 3, \ldots\}$, then

$$\implies \mathcal{E}_1 + \mathcal{E}_2 + \cdots + \mathcal{E}_N \stackrel{d}{=} \text{Exponential}\,(\nu p)\,,$$

where $\nu = \lambda + \mu$, and $p = \frac{\mu}{\lambda+\mu}$.

## 3 Analogies Between Discrete and Continuous Time Markov Chains

- Suppose $Y_n$ for $n = 0, 1, 2, \ldots$ is a discrete time Markov chain with transition matrix $P$ and $Y_0 = y$ for some $y$ in state space $S$. Let

$$H := \min\{n \geq 1 : Y_n \neq y\}$$

We notice that by definition, $H_y \in \{1, 2, 3, \ldots\}$. Then except in trivial cases with $P(y, y) = 1$, we have

$$\mathbb{P}(H_y < \infty) = 1.$$

Then assuming this is not absorbing (that is, $P(y, y) < 1$), we have that

$$\mathbb{P}(H_y = n) = [P(y, y)]^{n-1}(1 - P(y, y)), \quad \text{for } n = 1, 2, \ldots$$

where this bracketed factor denotes the power of the single element and not the matrix power. That is, $H_y \sim$ Geometric$(1 - P(y, y))$.

Now, let $J_y := Y_{H_y}$, which we know for sure is not equal to $y$. Then we may look into some other state $z$ which has

$$\mathbb{P}(J_y = z) = \frac{P(y, z)}{1 - P(y, y)},$$

which holds for $z \neq y$. So $H_y$ and $J_y$ are independent.
Bringing this foward into continuous time, take $(X_t, t \geq 0)$ to be a continuous time Markov chain with the semigroup of transition matrices $(P_t, t \geq 0)$ where $P_t = e^{Qt}$. Recall from last lecture that we found

$$Q = \lim_{t \downarrow 0} \frac{(P_t - I)}{t},$$

which is the matrix of transition rates. We may ask similar things as before in terms of holding times. Let $y$ be a state. We write

$$H_y := \inf\{t > 0 : X_t \neq y\}$$

The idea here is that we are holding for a while and then jumping to another state. Then we have the following analogies from discrete time that carry over to this continuous time discussion.
- In continuous time, we want to think about $(1 - P(y, y))$ in a small amount of time. Now because this converges to 0, we want to normalize (divide by) $t$. Looking at the formula for $Q$ above, we notice that

$$H_y \sim \text{Exponential}(\lambda_y = -Q(y, y)).$$

- Let $J_y := X_{H_y}$, which is the state that we jump to in the picture above. From the row sums argument, we have:

$$\mathbb{P}(J_y = z) = \frac{Q(y, z)}{\lambda_y}.$$

This is partitioned in a way that is completely analogous to before.

- Moreover, $H_y$ and $J_y$ are independent.

# 4 Corollary

We have a very simple way of simulating a continuous time Markov chain. If we start at $y$, for each state $z \neq y$, we have an exponential variable $\mathcal{E}_{yz}$ with rate $Q(y, z) \geq 0$. We assume these are variables are independent, and let

$$H_y := \min_z \mathcal{E}_{yz}, \quad J_y := \arg\min_z \mathcal{E}_{yz}$$

Then $(X_t, t \geq 0)$ iid and $X_0 = y$ is constructed as $X_t = y$ for $0 \leq t < H_y$ and $X_{H_y} = J_y \neq y$. Then given $J_y = z$, we iterate. It is a fact that in the race of exponentials, the time to completing the race is independent of the end of the race.

# 5 Examples

A very simple example is the two-state Markov chain, which is also known as the alternating exponential renewal process. Let these states be 0 and 1 with:

$$Q = \begin{bmatrix} & 0 & 1 \\ 0 & 1 - Q(0,1) & Q(0,1) \\ 1 & Q(1,0) & 1 - Q(1,0) \end{bmatrix} =: \begin{bmatrix} -\lambda & \lambda \\ \mu & -\mu \end{bmatrix}$$

The picture illustrates why this is also called an alternating exponential renewal process, as we are simply flipping between the two rates. Now with our same terminology from before,

$$P_t(x, y) = [\text{Exponential}(tQ)]\,(x, y),$$

which is given by a similar formula in Durett. We have done this in our homework in the discrete case. There is a similarly simple formula which can be found via the differential equations quite easily. Now, we may ask what happens in the limit. Recall that a large rate means that an event happens fast. Imagine the case where $\lambda$ is large and $\mu$ is small. Then this gives a picture: In terms of renewal theory, we have an iid cycle every time we drop back down to 0. Appealing to the Renewal Reward Theorem, we compare the mean time for the above state against the mean time for the cycle:

$$\lim_{t\to\infty} P_t(x,1) = \frac{1/\mu}{1/\lambda + 1/\mu} = \frac{\lambda}{\lambda+\mu}.$$

Similarly, for $P_t(x,0)$, we have $\frac{\mu}{\lambda+\mu}$.

Lecture ends here.