EECS126 Course Notes [Spring 2021]

Patrick Yin

Updated April 1, 2021

Contents

1 Note

These course notes are my notes from EECS 126 : Probability and Random Processes. The course is linked here. These course notes are in progress.

2 Probability Basics

In progress...

3 Information Theory

In progress...

4 Discrete Time Markov Chains (DTMCs)

In progress...

5 Poisson Processes

5.1 Construction

A Poisson Process is an example of a counting process. A counting process $(N_t)_{t\geq 0}$ is a non-decreasing continuous-time integer-valued random process, which has right continuous sample paths.

Definition 5.1 (Poisson Process)**.** A rate-λ Poisson Process (i.e. PP(λ)) is a counting process with i.i.d inter-arrival times $S_i \overset{\text{IID}}{\sim} \text{Exp}(\lambda)$. Equivalently, a counting process is PP(λ) iff $N_0 = 0$, $N_t - N_s \sim \text{Poisson}(\lambda(t-s))$ for $0 \leq s \leq t$, and $(N_t)_{t\geq 0}$ has independent increments.

To elaborate on this, we will define T_i to be the arrival times, so $T_i = \min\{t \geq 0 : N_t \geq i\}$, which is the time of ith arrival. We also define the inter-arrival time, $S_i = T_i - T_{i-1}$, for $i \geq 1$.

Theorem 1. *If $(N_t)_{t\geq 0}$ is a PP(λ), then for $t \geq 0$, $N_t \sim$ Poisson(λt). I.e. $Pr\{N_t = n\} = \frac{e^{-\lambda t}(\lambda t)^n}{n!}$*

Proof.

$$
\begin{aligned}
Pr\{N_t = n\} &= Pr\{T_n \leq t < T_{n+1}\} \\
&= \mathbb{E}[\mathbb{1}_{\{T_n \leq t\}}\mathbb{1}_{\{t \leq T_n + S_{n+1}\}}] \\
&= \int f_{T_n}(s)\mathbb{1}_{\{s\leq t\}}\mathbb{E}[\mathbb{1}_{\{t\leq s+S_{n+1}\}}]ds \\
&= \int_0^t f_{T_n}(s)\mathbb{E}[\mathbb{1}_{\{t-s\leq S_{n+1}\}}]ds \\
&= \int_0^t f_{T_n}(s)e^{-\lambda(t-s)}ds \\
&= \int_0^t \frac{\lambda e^{-\lambda s}(\lambda s)^{n-1}}{(n-1)!}e^{-\lambda(t-s)}ds \text{ } (f_{T_n}(s) \text{ is Erlang}) \\
&= \frac{\lambda^n e^{-\lambda t}}{(n-1)!}\int_0^t s^{n-1}ds \\
&= \frac{(\lambda t)^n e^{-\lambda t}}{n!}
\end{aligned}
$$

□

Remark. By the memoryless property of Exp(λ), if $(N_t)_{t\geq 0} \sim$ PP(λ), then $(N_{t+s} - N_s)_{t\geq 0} \sim$ PP(λ) for all $s \geq 0$. Moreover, $(N_{t+s} - N_s)_{t\geq 0}$ is independent of $(N_\tau)_{0\leq\tau\leq s}$. In particular, Poisson Processes have independent and stationary increments. If $t_0 < ... < t_k$, then $(N_{t_1} - N_{t_0}), ..., (N_{t_k} - N_{t_{k-1}})$ are independent and $(N_{t_i} - N_{t_{i-1}}) \sim \text{Poisson}(\lambda(t_i - t_{i-1}))$ for all i.

5.2 Conditional Distribution of Arrivals

Theorem 2. *Conditioned on* $\{N_t = n\}$, $(T_1, ..., T_n) \stackrel{d}{=} (U_{(0)}, ..., U_{(n)})$ *where* $U_{(i)}$ *are the order statistics of n Uniform(0,t) random variables.*

In other words, given n arrivals occurred up to time t, the arrival times look like i.i.d Unif(0, t) random variables in distribution.

Proof. Let $0 = t_0 \le t_1 \le ... \le t_n \le t$, then

$$\begin{aligned} f_{T_1 T_2 \dots T_n | N_t}(t_1 \dots t_n | n) &= \frac{Pr\{N_t = n | T_1 = t_1, ..., T_n = t_n\}}{Pr\{N_t = n\}} f_{T_1 \dots T_n}(t_1 \dots t_n) \\ &= \frac{Pr\{N_t - N_{t_n} = 0\}}{Pr\{N_t = n\}} \prod_{i=1}^{n} f_{S_i}(t_i - t_{i-1}) \\ &= \frac{e^{-\lambda(t-t_n)}}{e^{-\lambda t}\frac{(\lambda t)^n}{n!}} \prod_{i=1}^{n} \lambda e^{-\lambda(t_i - t_{i-1})} \\ &= \frac{n!}{t^n} \text{ (density of uniform random order statistics)} \end{aligned}$$

□

5.3 Merging

Theorem 3. *If* $(N_{1,t}) \sim PP(\lambda_1)$ *and* $(N_{2,t}) \sim PP(\lambda_2)$ *are independent, then* $(N_{1,t} + N_{2,t}) \sim PP(\lambda_1 + \lambda_2)$.

Proof. We will show that the sum of the two independent Poisson Processes satisfies the three properties of a PP:

1. $N_{1,0} + N_{2,0} = 0 + 0 = 0$
2. For $0 \le s \le t$,

$$\begin{aligned} (N_{1,t} + N_{2,t}) - (N_{1,s} + N_{2,s}) &= (N_{1,t} - N_{1,s}) + (N_{2,t} - N_{2,s}) \\ &\stackrel{d}{=} \text{Poisson}(\lambda_1(t-s)) * \text{Poisson}(\lambda_2(t-s)) \\ &= \text{Poisson}((\lambda_1 + \lambda_2)(t-s)) \end{aligned}$$

3. $(N_{1,t} + N_{2,t})_{t \ge 0}$ has independent increments since $(N_{1,t})_{t \ge 0}$, $(N_{2,t})_{t \ge 0}$ has independent increments.

□

5.4 Splitting (a.k.a Thinning)

Theorem 4. *Let* $p_1, ..., p_k$ *be non-negative such that* $\sum_{i=1}^{k} p_i = 1$ *and* $(N_t)_{t \ge 0}$ *be a PP(*λ*). Mark each arrival with label "i" with probability* p_i*, independently of all other arrivals so that* $(N_{i,t})_{t \ge 0}$ *be the process that counts arrivals marked with "i". Then* $(N_{i,t})_{t \ge 0}$*, for* $i = 1, ..., k$*, are independent Poisson Processes with respective rates* $p_i \lambda$ *for* $i = 1, ..., k$.

Proof. We will only prove for $k = 2$. This is sufficient because we can simply do induction to get $k > 2$. For $k = 2$, we let $p_1 = p$ and $p_2 = 1 - p$.

$$\begin{aligned}
Pr\{N_{1,t} = n, N_{2,t} = m\} &= Pr\{N_{1,t} = n, N_{2,t} = m, N_t = n + m\} \\
&= Pr\{_{1,t} = n, N_{2,t} = m | N_t = n + m\} Pr\{N_t = n + m\} \\
&= \binom{n+m}{n} p^n (1-p)^m e^{-\lambda t} \frac{(\lambda t)^{n+m}}{(n+m)!} \\
&= e^{(-p\lambda)t} \frac{(p\lambda t)^n}{n!} e^{(-(1-p)\lambda)t} \frac{((1-p)\lambda t)^m}{m!} \\
&= \text{Poisson}(p\lambda t)\text{Poisson}((1-p)\lambda t)
\end{aligned}$$

$\square$

5.5 Random Incidence Paradox

Consider $(N_t)_{t \geq 0} \sim PP(\lambda)$ and pick a random time t_0. What is the expected length of the inter-arrival interval in which t_0 falls?

Say it falls between T_i and T_{i+1}. Then the length of the inter-arrival interval is $L = (t_0 - T_i) + (T_{i+1} - t_0)$. We know that $T_{i+1} - t_0 \sim \text{Exp}(\lambda)$ by the memoryless property of the exponential distribution. We also know that

$$Pr(t_0 - T_i > s) = Pr(\text{no arrivals in } (t_0 - s,\, s)) = Pr(N_{t_0} - N_{t_0 - s} = 0) = e^{-\lambda s}$$

so $t_0 - T_i \sim \text{Exp}(\lambda)$. By linearity of expectation, $\mathbb{E}[L] = \frac{2}{\lambda}$. If we arrive at a random time, we are more likely to land in a long interval.

6 Continuous Time Markov Chains (CTMCs)

6.1 Construction

Intuitively, a CTMC is a markov chain where we need to wait for $Exp(\lambda)$ time before transitioning to the next state.

Definition 6.1 (CTMC)**.** Let S be a countable state space. A CTMC is defined in terms of a rate matrix Q satisfying $[Q]_{ij} \geq 0$ for $i \neq j$, $i, j \in S$ and $\sum_{j \in S}[Q]_{ij} = 0$ for all $i \in S$. Specifically, the transition rate for state i is $q_i := [Q]_{ii} = -\sum_{j \neq i}[Q]_{ij}$. We also have $[Q]_{ij} = q_i p_{ij}$ such that $\sum_{j \in S} p_{ij} = 1$ where $p_{ii} = 0$ and $p_{ij} \geq 0$. p_{ij} are the transition probabilities for an associated DTMC called the jump chain.

A CTMC with rate matrix Q works as followed:

1. Start with $X_0 = i$.
2. Hold for $\text{Exp}(q_i)$ amount of time, then jump to state $j \in S$ with probability p_{ij} where $j \in S$.
3. Repeatedly apply the previous line above at next states (starting at state j).

We can equivalently define CTMCs by their jump rates q_{ij}. On entering state i, consider independent random variables $T_j \sim \text{Exp}(q_{ij})$ for $j \in S \backslash \{i\}$ and jump to state $j^* = \text{argmin}_{j \in S}(T_j : j \in S)$ at time T_{j^*}. This valid due to the splitting property of Poisson Processes.

Remark. This is called a markov chain by the memoryless property of the exponential distribution:

$$Pr(X_{t+\tau} = j | X_t = i, X_s = i_s, 0 \leq s < t) = Pr(X_{t+\tau} | X_t = i)$$

6.2 Stationary Distributions

Definition 6.2 (CTMC Stationarity)**.** A probability vector π is (without considering pathological cases) a stationary distribution for a CTMC with rate matrix Q if $\pi Q = 0$. This called the rate conservation principle. This is equivalent to $\pi_j q_j = \sum_{i \in S} \pi_i q_{ij}$ for all $j \in S$. In other words, assuming that $Pr(X_t = i) = \pi_i$, the rate at which transitions are made out of j is equal to the rate at which transitions are made into j.

6.3 Classification of States

Similar to DTMCs, we can classify the states.

- We say i and j communicate (i.e. $i \leftrightarrow j$) iff $i \leftrightarrow j$ is a jump chain iff we can travel $i \to j$ and back.

- Classes in CTMC are same as those in associated jump chain.
- State j is transient if, given $X_0 = j$, $(X_t)_{t\geq 0}$ re-enters state j finitely many times with probability one. State j is recurrent otherwise.
- For a recurrent state j, define $T_j = \min\{t \geq 0 : X_t = j \text{ and } X_s \neq j \text{ for some s } < \text{ t } \}$.
- State j is positive recurrent if $\mathbb{E}[T_j|X_{0\text{J}}] = \infty$.
- State j is null recurrent if $\mathbb{E}[T_j|X_0 = j] = \infty$.
- Transience/Positive Recurrence/Null Recurrence are class properties
- There is no concept of periodicity.

6.4 Big Theorem

Theorem 5. *We define $P_{ij}^t := Pr(X_t = j|X_0 = i)$ and $m_j := \mathbb{E}[T_j|X_0 = j]$. For an irreducible CTMC, exactly one of the following is true:*

1. *Either all states are transient or all states are null recurrent. In this case, no stationary distribution exits, and $\lim_{t\to\infty} P_{ij}^t = 0$ for all $i, j \in S$.*
2. *All states are positive recurrent. In this case, a unique stationary distribution exits and satisfies $\pi_j = \frac{1}{m_j q_j} = \lim_{t\to\infty} P_{ij}^t$ for all $i, j \in S$.*

Remark. Stationary distribution in CTMC is not the same as the stationary distribution in the jump chain. Generally speaking, $\widetilde{\pi}_j = \frac{\pi_j q_j}{\sum_{i\in S} q_i \pi_i}$ given that $\sum_{i\in S} q_i \pi_i < \infty$ where $\widetilde{\pi}_j$ is the stationary distribution of the jump chain.

6.5 Examples

6.5.1 M/M/s queue

Customers arrive to a system with s servers according to PP(λ). If a server is available, the arrival is immediately serviced, which takes $\overset{IID}{\sim}$ Exp(μ). If no server is available, the arrival waits until one becomes available. Let $(X_t)_{t\geq 0}$ denote the number of customers in system at time $t \geq 0$. We can model this with $q_{n,n+1} = \lambda$ and $q_{n,n-1} = \begin{cases} n\mu & 1 \leq n \leq s \\ s\mu & n > s \end{cases}$.

6.5.2 Birth Death Chain

Individuals give birth $\overset{\text{IID}}{\sim}$ $PP(\lambda)$ and have lifetimes $\overset{\text{IID}}{\sim}$ Exp(μ). Let X_t be the number of individuals in population at time t. We can model this with $q_{n,n+1} = n\lambda$ and $q_{n,n-1} = n\mu$. Note that this means $q_n = n(\lambda + \mu)$ so then $p_{n,n+1} = \frac{\lambda}{\lambda+\mu}$ and $p_{n,n-1} = \frac{\mu}{\lambda+\mu}$, so the DT jump chain is also a birth-death chain.

6.5.3 M/M/∞ queue

This is the same of the M/M/s queue problem except there are infinite servers. In this case $q_{n,n+1} = \lambda$ and $q_{n,n-1} = n\mu$. If we solve $\pi Q = 0$, we see that $\pi_n = \frac{e^{-\lambda/\mu}(\lambda/\mu)^n}{n!}$. By the Big Theorem, $X_t \xrightarrow{d}$ Poisson(λ/μ) where X_t is the number of people in the system at time t.

6.6 First Step Equations (FSE)

If $A \subseteq S$, define $T_A = \min_t\{t \geq 0 : X_t \in A\}$. We want to compute $\mathbb{E}[T_A|X_0 = i]$, so we will use FSEs to do this. Define $t_i := \mathbb{E}[T_A|X_0 = i]$ and $t_i = 0\ \forall i \in A$. Then we want to if $t_i = \mathbb{E}[\text{hold time}] + \sum p_{ij}t_j\ \forall i \in S$. Thus our FSEs are

$$t_i = 0 \quad \forall i \in A$$

$$t_i = \frac{1}{q_i} + \sum_{j \in S} p_{ij}t_j \quad \forall i \notin A$$

6.7 Uniformization

Uniformization is an approach to compute CTMC transition probabilities by simulating a DTMC.

6.7.1 Context

For context, consider a CTMC with transition rates $(q_i)_{i\in S}$ and assume $\exists M > 0$ s.t. $q_i \leq M \quad \forall i \in S$. We want to find P_t for some $t \geq 0$. Here,

$$\left[P_t\right]_{ij} := Pr(X_t = j|x_0 = i)$$

Markovity gives the Chapman-Kolmogorov Equations, which is $P^{s+t} = P^sP^t \quad \forall s, t \geq 0$. We can also show that for $h \approx 0$, $P^h \approx I + hQ + O(h)$ so

$$\begin{aligned}
P^{t+h} &= P^tP^h \\
&\approx P^t(I + hQ + O(h)) \\
\frac{P^{t+h} - P^t}{h} &= P^tQ + \frac{O(h)}{h} \\
\frac{\partial}{\partial t}P^t &= P^tQ \\
P^t &= e^{tQ} := \sum_{k\geq 0} \frac{(tQ)^k}{k!} \quad \forall t \geq 0
\end{aligned}$$

So we've found a way to compute P^t, but this becomes intractable for large state spaces. This is where uniformization comes in.

6.7.2 Construction

Let us define a uniformized DTMC which is a DTMC, given $\gamma \geq M$, with transition probabilities

$$p_{ij} = \frac{q_{ij}}{\gamma} \quad p_{ii} = 1 - \frac{q_i}{\gamma} \quad i, j \in S$$

If $P_u =$ transition matrix with uniformed DTMC, then $P_u = I + \frac{1}{\gamma}Q$. So observe $\pi P_u = \pi + \frac{1}{\gamma}\pi Q$. In other words, $\pi P_u = \pi \iff \pi Q = 0 \iff \pi$ is a stationary distribution for both the CTMC and uniformized DTMC. So we're beginning to see that the behavior of the two chains are similar. In fact, we can see that

$$P_u^n = (I + \frac{1}{\gamma}Q)^n \approx e^{\frac{n}{\gamma}Q}$$

So to estimate P^t, we can run the uniformized DTMC for $n \approx \gamma t$ steps because $P^t = e^{tQ} \approx e^{\frac{n}{\gamma}Q} \approx P_u^n$. Notice that with a larger γ, we get a better approximation.

7 Random Graphs

Definition 7.1 (Erdős–Rényi Random Graphs)**.** Fix $n \geq 1$ and $p \in [0, 1]$. A random graph G(n, p) is an undirected graph on n vertices, where each edge is placed independently with probability p.