Discrete – Time Markov Chain and Poisson Processes
Professor. Subhamay Saha
Department of Mathematics,
Indian Institute of Technology, Guwahati
Lecture 17
Lecture: Stationary Distribution I

So, hello everyone. Welcome to the 17th lecture of the course, Discrete-time Markov Chains and Poisson Processes. So, in last class we saw simple random walks and we saw that for simple symmetric random walk, all states are recurrent while for asymmetric simple random walk all states are transient. So, if $p = q$, then all states are recurrent and if $p \neq q$ then all states are transient. So, today we will start with a new topic called stationary distributions which is a very important topic. So, as we proceed along the topic you will realize why I am saying that this is a very important topic.

So, let us start. So, first we start with a definition. So, a row vector π, so it is a row vector where π_i, i belonging to the state space of non-negative entries is said to be a stationary or an invariant measure for a Markov Chain with transition probability matrix π_j, if it satisfies this condition. What is it? That $\pi_j = \sum_{i \in S} \pi_i p_{ij}$, for all $j \in S$. So, what is a stationary or an invariant measure? So, it is a row vector and what is the length of that vector? The length of the vector is equal to the cardinality of the state space. So, if it is a finite state space then it is a finite vector and if it is an infinite state space then it is an infinite vector.

So, it is a row vector where each entry is non-negative that means greater than or equal to 0 and it like the each entry of that row vector should satisfy this condition that π_j should be equal to $\sum_{i \in S} \pi_i p_{ij}$, for all $j \in S$ or in other words this should hold for each entry in the row vector. Now, when you see it in this way it may not be completely clear but let us see more closely. So, suppose this is the vector so this is $[\pi_1, \pi_2, \ldots, \pi_n]$. Say, let me for just illustration purpose let me think of it as a finite vector, it is saying now if you look at this it is just saying this should be equal to $[\pi_1, \pi_2, \ldots, \pi_n]P$.

So, for example if you look at π_1, what is it saying? That π_1 should be $\sum_{i \in S} pi_1 p_{1j}$, sorry so $\pi_1 = \sum_{i \in S} \pi_i p_{i1}$. So, p_{i1}, so you are going this and then you are going along. So, this is the pij, where the summation is over i that means you are, so remember P is again a matrix. So, this π_j equal to $\sum_{i \in S} \pi_i p_{ij}$. If you write it in the matrix form this just becomes $\pi P = \pi$.

So, in the matrix notation 1 can be written as πP equal to π. So, you know the matrix multiplication. So, you write π as a row vector, if you have to write this in the matrix form this is just π equal to πP. So, in matrix notation this condition for stationary or invariant measure becomes πP equal to π. So, this matrix notation is much easier to see or much easier to follow. Now, once you have πP equal to π, now you have πP equal to π. Now say, if I multiply both sides by P then what do I get? I get $\pi P^2 = \pi P$. So, I am multiplying both sides or left multiplying rather both sides by P, so $\pi P^2 = \pi P$ but we again know $\pi P = \pi$. So, now continuing in this way, it is easy to see that for all $n \geq 1$, $\pi P^n = \pi$ or in other words, so if I have to write in this way it is true that $\pi_j = \sum_{i \in S} \pi_i p_{ij}^{(n)}$. So, this is the entry of the nth transition matrix. So, π_j is So, it is not just, this condition $\pi_j = \sum_{i \in S} \pi_i p_{ij}$, it also says that it should also satisfy $\pi_j = \sum_{i \in S} \pi_i p_{ij}^{(n)}$ and the reason I have explained you here. So, this was stationary measure but you see the title of the slide says stationary distribution. So, what is stationary distribution? So, a stationary measure π is called a stationary distribution, if the sum of the entries is equal to 1 i.e., $\sum_{i \in S} \pi_i = 1$ or in other words this π_i is actually the PMF of some discrete distribution.

So, a stationary measure, so for stationary measure the only conditions are the entries has to be non-negative and the this condition has to be satisfied this π_j or in matrix notation it has to be $\pi P = \pi$ now a stationary measure π is called a stationary distribution, if this additional condition is also satisfied that the sum of the entries is equal to 1 or in other words it is actually a probability mass function of some discrete probability distribution. So, that is the definition of a stationary distribution.

Now the question is, why am I calling it stationary? From where this terminology stationary is coming? So, this following theorem justifies the use of the terminology stationary. Now let us see what is it saying, so it says let X_n be a Markov chain with transition probability matrix P and initial distribution π, initial distribution π means $P(X_0 = i) = \pi_i$ for $i \in S$, the state space. Now, if further if π_i s a stationary distribution that means anyway, so you know this, for this π all entries since this is an initial distribution so this is a probability mass function so all entries are non-negative. The sum is equal to 1, so when you say that if π is a stationary distribution what you mean is that it also satisfies this condition that $\pi_j = \sum_{i \in S} \pi_i p_{ij}$ or in other words πP equal to π. So, if the initial distribution also satisfies that condition that $\pi P = \pi$ or in other words if it is a stationary distribution then $P(X_n = j) = \pi_j$ for all $n \geq 0$ and for all $j \in S$.

That means what? So, if you start with a stationary distribution that means if your initial distribution is a stationary distribution, then up to that for all n the distributions of X_n is also that same distribution. That is why, it gets fixed, that is why you use the terminology or that is the reason for using the terminology stationary. So, if you start from a stationary distribution, you remain there. So, if $P(X_0 = i) = \pi_i$ where π_i is a stationary distribution then $P(X_n = i) = \pi_i$ for all $n \geq 1$. So, if you start with the stationary distribution then you remain there. That is why you use this term stationary because stationary means which does not move or you are fixed there. That is why this terminology stationary. Now the proof is pretty simple. So, we will prove by induction. So, for induction first there is a basis step, you need to put prove it for $n = 0$ but you have already assumed that π is an initial distribution. So, from definition $P(X_0 = j) = \pi_j$ because that is the definition of initial distribution. So, this condition is true for $n = 0$ that is the basis step. Now we assume that $P(X_n = j) = \pi_j$ for all j, for some n. So, that is how induction works. So, there is a basis step where you prove it for $n = 0$ then you assume it for n and you prove it for $n + 1$, that is how induction goes. So, we have now assumed it for n, that is the induction hypothesis. So, we are assuming that $P(X_n = j) = \pi_j$. Now by Markov, sorry, now we want to show that $P(X_{n+1} = j) = \pi_j$. So, now by Markov probability, $P(X_{n+1} = j)$ is nothing but, so first we use here what is called the law of total probability so we condition on the previous jump, on the previous state so $P(X_{n+1} = j|X_n = i)P(X_n = i)$ and you sum over all i in the state space. So, this is nothing but just law of total probability.

But now, we will use Markov property, we know that X_n is a Markov chain. So, $P(X_{n+1} = j|X_n = i)$, is nothing but just p_{ij} this small p_{ij} because that is the transition probability matrix of the Markov chain. So, this part probability is nothing but p_{ij} and this by induction hypothesis, since you have assumed that $P(X_n = j) = \pi_j$ for all $j \in S$, so this will be equal to π_i . So, this should be equal to π_i . So, we get this but now we know that π is a stationary distribution. So, now from the definition of stationary distribution this should be equal to π_j . So, hence you have shown that probability $P(X_{n+1} = j) = \pi_j$ and this is true for all $j \in S$. So, that completes the proof and so by induction we are done that $P(X_{n+1} = j) = \pi_j$ for all $n \geq 0$.

So, again so, for $n = 0$, it follows from the definition of initial distribution, then we assume it for n and then we prove it for $n+1$ and for that we just condition on X_n and then since this is the Markov chain, so we know what $P(X_{n+1} = j|X_n = i)$, is just simply given by the p_{ij}, then of the transition probability matrix p_{ij} and since by induction hypothesis $P(X_n = i) = \pi_i$, so and now finally we use the definition of stationary distribution to get that this is equal to π_j . So, that completes the proof. So, if you start from a stationary distribution, you remain there hence the terminology stationary.

So, now moving on, now the following theorem so we prove yet another theorem about stationary distribution the following theorem says that the limiting distribution we will see what this limiting means, if it exists is a stationary distribution. So, let us see the theorem and again this theorem is true for when the state space is finite. So, let S be finite or in other words we are looking at a finite state Markov chain. Suppose that for some $i \in S$, $p_{ij}^{(n)} \to \pi_j$ as $n \to \infty$, for all $j \in S$. That is why, we are, so you see this π_j is a kind of a limiting distribution because why? It is the limit of this p_{ij} and that is why we are calling this as limiting, this π_j as limiting distribution. So, suppose such a limiting distribution exist, so again we are not saying that this always exists, the statement of the theorem says that suppose such a π_j exist such that $p_{ij}^{(n)}$ converges to π_j as $n \to \infty$, for all $j \in S$, then this limiting distribution π_j is a stationary distribution and again when S is finite.

So, this is important, I am not claiming this. This theorem does not claim it for S infinite. So, this is a statement about a finite state Markov chain which says that, if the limiting distribution and what is the meaning of limiting distribution? The limiting distribution meaning is this thing that there exists some $i \in S$ such that $p_{ij}^{(n)}$ so this is basically the ij-th entry of the n-th transition probability matrix or this is basically $P(X_n = j|X_0 = i)$, if that converges to π_j for all $j \in S$, then if I look at this vector π_j for $j \in S$ that is a stationary distribution. So, the first thing we need to show is that, so anyway so since each $p_{ij}^{(n)} \geq 0$ so it is easy to see that $\pi_j \geq 0$, for all j so the entries are non-negative. Now we show that the sum is 1 again that is very simple so $\sum_{j\in S} \pi_j = \sum_{j\in S} \lim_{n\to\infty} p_{ij}^{(n)}$ but now this is a finite sum so I can take the limit outside. So, it becomes $\lim_{n\to\infty} \sum_{j\in S} p_{ij}^{(n)}$ which is now we know that this is again a probability vector because this is a row in the n-th transition probability matrix. This is the row sum of the nth transition probability matrix and since that is a stochastic matrix or in other words the row sum is equal to 1 for the n-th transition probability matrix, we get that this is equal to 1. So, that gives us summation of π_j , $j \in S$ is equal to 1 and to pull this limit out of this summation, we use the fact that this is a finite sum because if it is an infinite sum this pulling in or pulling out of limit is not a trivial thing, it requires technical assumptions but since here this state space is finite so this is a finite sum so we can easily pull out limit or pull push the limit inside. So, the first condition that $\sum_{j\in S} \pi_j = 1$ is also satisfied. Now we will have to show that πP equal to π. For that again, π_j is equal to $\lim_{n\to\infty} \sum_{j\in S} p_{ij}^{(n)}$ which is equal to $\lim_{n\to\infty} \sum_{j\in S}$ this term. Now again so here we are basically using the Chapman Kolmogorov equation, so $p_{ij}^{(n)}$, I can write it as $\sum_{k\in S} p_{ik}^{(n-1)} p_{kj}$ so this is nothing but just Chapman Kolmogorov equation. But now, again since this is a finite sum I push the limit inside, so $\lim_{n\to\infty} \sum_{k\in S} p_{ik}^{(n-1)} p_{kj}$ but it does not matter whether it is $n-1$ or n, so if $\lim_{n\to\infty} p_{ik}^{(n)} = \pi_k$, so then $\lim_{n\to\infty} p_{ik}^{(n-1)}$ is also π_k is also so this is just simple like if X_n goes to something as $n \to \infty$ then X_{n-1} goes to the same limit as $n \to \infty$. So, this limit is equal to π_k because that is our assumption that π_k is a limiting distribution so you get this is equal to $\sum_{k\in S} p_{ik} p_{kj}$ which is precisely π equal to πP . So, we have shown all the three conditions, that all the entries are non-negative they sum to 1 and also they satisfy this πP equal to π.

So, we have shown that this limiting distribution if it exist is a stationary distribution. So, for a finite state space this actually gives you a way of finding a stationary distribution So, if you can find p like the entries of the nth transition matrix and you can take limit of each entry and if the limit exists that gives you a stationary distribution. So, this is a way of finding stationary distribution, but this is not the most easiest way because calculating this n-th finding this $p_{ij}^{(n)}$ is not always an easy task but anyway this is a method whether it is a useful method or not is a different thing but this is a method but also it is not just a method, it is telling you that if the limiting distribution exists, it is a stationary distribution.

So, now let us move ahead. Now we will see some examples of how to calculate stationary distribution for certain Markov Chains. So, we start with our first example. So, consider this two state Markov Chain with transition probability matrix this one. So, you go from 0 to 0 with probability $1-\alpha$, 0 to 1, with probability α, 1 to 0, with probability β and 1 to 1 with probability $1-\beta$. So, this is a Markov Chain whose states space is consists of only two states, 0 and 1 and this is the transition probability matrix. So, we will try to calculate its stationary distribution. So, let (π_0, π_1) be a stationary distribution. Now if it has to be a stationary distribution then it needs to satisfy. So, remember so these are the two conditions which you get from πP equal to π and this is the condition that it should sum to 1 and also, we need to find a solution where each π_i is greater than or equal to 0. But now, so this is just a simple system of equations so I am not solving it here, but you can solve it easily. You solve it and see that you get π_0, you get a unique solution $\pi_0 = \frac{\beta}{\alpha+\beta}$ and $\pi_1 = \frac{\alpha}{\alpha+\beta}$. So, for this particular example, we see that stationary distribution exists and is unique. So, you see given a Markov Chain a priority is not clear that whether a stationary distribution will exist or not. That is an existential question, and another question is, if it exists how many such stationary distributions can exist. So, there are two questions one of existence and the other of uniqueness. So, the first question is given a Markov Chain whether stationary distribution exists or not. That is the first question, and the second question is, if it exists how many such stationary distributions are possible. So, for this particular example, that we just saw we saw that there exists a unique stationary distribution.

So, now moving ahead. Now again we consider another two state Markov Chain which is a very simple Markov Chain, with transition probability matrix $\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix}$. So, this is basically the identity matrix. So, again it is a very simple Markov Chain. Again, there are two states, so call those states as 0 and 1. So, then again it is a very trivial kind of a Markov Chain, from 0 you stay in 0 and from 1 you stay in 1. So, here both 0 and 1 are what are called absorbing states. So, again this is a Markov Chain, so whenever you see a transition probability matrix you should just try to find out, what how many communicating classes it has? What are those communicating classes? So, here it has basically two classes, two communicating classes and both are recurrent. So, both are absorbing states, so if you start from 1, you remain at 1 or if you start from 0 you remain at 0. So, two classes both are recurrent but anyway here we are interested in finding the stationary distribution but see, since this is a identity matrix, if I take any $(a, 1-a)$, where $a \in [0,1]$, then that is a stationary distribution because this $\pi P = \pi$ S satisfies trivially. Why? Because this is an identity matrix, if you multiply something with the identity matrix you get back that. So, that is why this πP equal to π is trivially satisfied. So, you take any m but you also need the entries to be non-negative so you take any vector of size 2, of length 2 any row vector of length 2, this $\pi P = \pi$ will be satisfied but also you need the entries to be non-negative and they should sum to 1 so that is why if you take any row vector of this form $(a, 1-a)$, where $a \in [0,1]$ then that is a stationary distribution. So, for this particular example, there are infinitely many stationary distributions. So, in the previous example, we saw that stationary distribution exist and is unique but in the second example we saw that again stationary distribution exist but there are infinitely many stationary distributions. So, we saw these two examples tells you that it is not a very trivial matter. So, in both cases, we will see one more example where we will see that stationary distribution does exist. So, all three cases are actually possible, that it will exist and unique, exist but infinitely many does not exist. So, all three cases are possible till now we have seen only two examples, one where stationary distribution exist and is unique and second example where stationary distribution exists but there are infinitely many stationary distributions. So, let me here also mention one thing, so you see that if there are more than one stationary distributions. It is not possible that there will be only finitely many stationary distributions. Why? Because if π_1 and π_2 are two stationary distributions. Now if you take any convex combination, what is the meaning of convex combination? That means you take lambda π_1 plus 1 minus, sorry, let me just rewrite this again. So, say if π_1 and π_2 are two stationary distributions, then if I look at lambda pi, plus 1 minus lambda π_1 plus 1 minus lambda π_2 where lambda belongs to 01. So, if there exist more than one stationary distribution and then if I take all possible convex linear combinations, what is the meaning of convex linear combination? That means combination of this form lambda π_1 plus 1 minus lambda π_2 for any lambda in this closed interval 0 and 1. So, for lambda equal to 0, then it gives you π_1 and for lambda equal to 1 it gives you π_2 but if you take lambda in between 0 and 1, you get some different vector and all you can easily check, see now it is easy to see that again the entries will be non-negative. The sum of the entries will be equal to 1 and also, so if I look at, since you know the linearity of matrix multiplication. So, basically what I am saying is that $\lambda\pi_1 + (1-\lambda)\pi_2$, if I multiply this with P, you can see because you know how matrix multiplication works. So, this again will just become $\lambda\pi_1 + (1-\lambda)\pi_2$. Just if you follow the rules of matrix multiplication and if you also use the fact that both π_1 and π_2 are stationary distribution. So, either there will exist a unique stationary distribution or there will exist infinitely many stationary distributions. So, it is not possible that there will exist a finitely many stationary distributions. So, either it is one or infinite and what is the reason? If there exist more than one, so if just there exists two from those two, you can create infinitely many just by taking convex linear combinations as I have shown you. So, so again there are, so three possibilities either no, there exists no stationary distribution, so 0 or there exist unique stationary distribution 1 or infinitely $\{0, 1, \ldots \infty\}$. So, finitely many stationary distributions is not possible because if there is more than one taking convex linear combination you can create infinitely many. So, till now I have not given you an example where stationary distribution does not exist. So, now let us move to that example.

So, again see now we consider an infinite state Markov Chain. So, consider the Markov Chain with the following transition diagram, $p + q = 1, 0 < p < 1$. So, again here there are two parameters p and q but this sum up to 1 and p is strictly between 0 and 1. So, this gives you the transition diagram so from 0 you always go to 1 with probability 1 but from 1 you go to 0 with probability q and 2 with probability p, similarly from 2, you go to 1 with probability p and 3 with probability q, from 3 you go to 2 with probability q and 4 with probability p and so on. Anyway, this is the transition diagram. Now let π_i be a stationary distribution, now if it is a stationary distribution so it should satisfy that $\pi P = \pi$. Now let us light write that one by one so $\pi_0 = \pi_1 q$. So, if you again so this is an infinite transition matrix. So, basically if you look at this, how was it so it was π_j is equal to sum over, sorry not sum over j, sum over i, $\sum_{i \in S} \pi_i p_{ij}$. So, this is what so this is sum over all i, see, if $p_{ij} = 0$ that will not contribute in the sum. So, this is basically sum over all i from where you can go to j. So, now let us look at here so 0, so from where you can go to 0? You can go to 0 only from 1. So, $\pi_0 = \pi_1 q$ which is this. So, the sum is only over those i's from where you can go to j. So, if so if I have to write it for 0, then you can go to 0 only from 1, so you get $\pi_0 = \pi_1 q$. Now π_1 , if it is π_1 now you can go to 1 from 0 and from 2. From 1 you go with probability 1 and from 2 you go with probability p. So, $\pi_1 = \pi_0 + \pi_2 p$. But $\pi_0 = \pi_1 q$. So, I replace that, so, $\pi_1 = pi_1 q + \pi_2 p$. Now if this $\pi_1 q$, I bring it to this side then what do I get, so $\pi_1(1-q)$. So, if I bring it to this side, then I get $\pi_1(1-q)$which gives me $\pi_1 p$ but and the other thing remaining on the right hand side is $\pi_2 p$. But you know that $0 < p < 1$. So, in or in other words it is not equal to 0, so you can cancel it and you get $\pi_1 = \pi_2$. Now, again similarly if you do like this, you will actually get $\pi_2 = \pi_3$, $\pi_3 = \pi_4$, $\pi_4 = \pi_5$ and so on. But now you see, this is not possible. Why? Because, if everything is same, so remember this is an infinite state Markov Chain. So, you forget π_0, first just look at π_1 . It says that your π_1 should be equal to π_2 , π_2 should be equal to π_3 . So, π_1 , π_2 , π_3 everything should be same and π_0 should be equal to π_1 but if everything is same. So, remember you need summation π_i to be equal to 1, i running from 0 to so, here the state space is 0 to the set of all positive or non-negative integers. So, this should be equal to 1 but if since this all this has to be equal then what is the problem? The problem is that so if this sum has to be equal to 1 then what you get? Now remember everything is same, now if you add a same thing say suppose, let us look at a general fact. So, if I am adding a same thing, say a_i, i running from 1 to ∞. So, in this case, if you are adding a same thing that is equal to infinity provided this thing is strictly greater than 0. So, that tells you all this π_1 , π_2 , π_3 , π_4 since they are equal and they should, their sum should be less than or equal to 1. So, if it is anything positive the sum will be infinity. So, they should all be equal to 0. So, that tells you that $\pi_1 = \pi_2$ equal all these should be equal to 0. But if $\pi_1 = 0$ that tells you that $\pi_0 = 0$ but then finally you end up with that all entries should be equal to 0. But even in that case, the sum will not be equal to 1 because sum of 0 is 0. So, if each π_i for $i = 1, 2, \ldots$, if it is positive, then the sum will be infinity. So, the sum each on entry cannot be positive so it has to be 0, remember it has to be non-negative. So, if it cannot be positive it has to be 0 but if it is 0 then it turns out that each entry will be equal to 0 but in that case again the sum will be 0 and not 1. So, you see in all cases it is not possible. So, you can solve this equation, but the problem is, it will not be a distribution. So, you can say so in this way you can say that if I look at this vector of all 0s. So, if I look at the vector of all 0s that is a stationary measure. So, that is a stationary measure but it will not be a stationary distribution because its sum is not equal to 1. So, its sum is not equal to 1 so it is not a stationary distribution. So, a stationary distribution so for this particular example a stationary, for this particular example stationary distribution does not exist. So, here we have given now all three examples, an example where stationary distribution exists and is unique, an example where stationary distribution exist but there are infinitely many stationary distributions and a third example where there does not exist a stationary distribution. But remember in this third example a stationary measure exists but you will see like, since we are working with probability.

So, we want the measure to be actually, so we want this $\sum_{i=0}^{\infty} \pi_i = 1$ condition to hold. So, for this particular example the stationary distribution does not exist. Now you see, if you have noticed it for in the first two examples, it was finite state Markov Chain and in third example, I went to an infinite state Markov Chain to give an example of a Markov Chain where stationary distribution does not exist. So, why did not I give an example of a finite state Markov Chain, where a stationary distribution does not exist. The reason is that is actually not possible. That is our next theorem which tells you if $\{X_n\}_{n\geq 0}$ is a Markov Chain with finite state space then there exists at least one stationary distribution. So, if $\{X_n\}_{n\geq 0}$ is a Markov Chain with a finite state space then there exists at least one stationary distribution. So, if you have a Markov chain with a finite state space then the possibility that there does not exist a stationary distribution is not possible. So, either it will have 1 or infinitely many but 0 is not a possibility. So, only when you go to an infinite state Markov Chain all 3 possibilities are possible. That no stationary distribution, exact or unique stationary distribution and infinitely many stationary distributions.

So, for a finite state Markov Chain, there always exists at least one stationary distribution. Like in example two, we saw that uniqueness is not guaranteed there can be infinitely many stationary distributions, but existence is guaranteed. So, for a finite state Markov Chain existence of stationary distribution is guaranteed, but if you go to infinite state Markov Chain it is possible that a stationary distribution does not exist. So, in that case all three possibilities are true, 0, 1 as well as ∞.

So, this is the thing which I was telling you. So, remember, so this is the calculation which I did in the previous part. So, if $\pi_1 = c$, then $\pi_0 = cq$ and if $\pi_1 = \pi_2 = \pi_3 = \cdots = c$ so all these are same, so you get, so this sum cannot be 1, so in this case there does not exist a stationary distribution and if $\{X_n\}_{n\geq 0}$ is a Markov Chain with finite state space then there exists at least one stationary distribution. So, we will stop here today. Thank you all.