Discrete – Time Markov Chain and Poisson Processes
Professor. Subhamay Saha
Department of Mathematics,
Indian Institute of Technology, Guwahati
Lecture 21
Lecture: Limit Theorems II

Hello everyone, welcome to the 21st lecture of the course Discrete-time Markov Chains and Poisson Processes. So, in the last lecture, we saw that if you have, if X_n is an irreducible positive recurrent Markov chain, then starting from any state i, the long run proportion of time, the chain visits state $i$or the long run proportion of time the chain spends in state i that is equal to π_i which is the unique stationary distribution, where π_i is equal to $\frac{1}{E_i(T_i)}$. So, we saw that stationary distribution appears as a limit that showed that, okay. Stationary distribution is indeed important because I said that it is a very important thing, we have already seen many important properties of stationary distribution. In today's lecture, we will see some more. And another thing that we saw in previous lecture was this complete characterization of states in terms of this p ii k like, when it is transient, recurrent, positive recurrent, and null recurrent. So, if $\sum_{k=0}^{n} p_{ii}^{(k)} < \infty$, then it is transient if this quantity is equal to infinity, then it is recurrent. Now, in order to further sub classify it into positive recurrent and null recurrent, we need to look at this quantity which is $\frac{1}{n+1}\sum_{k=0}^{n} p_{ii}^{(k)}$, if the limit of this is equal to 0, then it is null recurrent and if it is positive, then it will be positive recurrent. So, today we will see some more limiting theorems, so this the first theorem that this long run proportion is π_i that was 1 of the limiting theorems, so this module is about limiting theorem. So, today we will see 1 more very important limiting theorem, so let us start.

So, we saw that, okay, if it is positive or null recurrent, so we saw this kind of a thing that $\frac{1}{n+1}\sum_{k=0}^{n} p_{ii}^{(k)}$, so say if i is, so this converges to 0, if i is null recurrent and this converges to $\frac{1}{E_i(T_i)}$, if it is positive recurrent. So, again you see that we already saw that this long run proportion is equal to the station distribution, but again, what is this Ei Ti? Now, since it is positive recurrent suppose we are assuming reducibility as well then this is precisely the stationary distribution. So, we see even the limit of this thing $\frac{1}{n+1}\sum_{k=0}^{n} p_{ii}^{(k)}$ that is equal to the stationary distribution. The i-th component of the stationary distribution if the state or if the chain is irreducible and positive recurrent. Now, for a state i what about this limit? Remember what is this, so it is starting from i, $X_n = i$, does this have a limit? So, as $\lim_{n\to\infty}$ does this converge to something. Remember, if you recall 1 of the theorems that we saw in the module of stationary distributions, we saw that if you have a finite state Markov chain then this $p_{ij}^{(n)}$, the limit of that if it exists we called it as limiting distribution and if it exists that has to be the stationary distribution. So, basically this you start from a state and then probability that you add Xn equal to j as n goes to infinity does this converge to something that theorem said that if this limit exists then that has to be equal to the stationary distribution, but it did not say anything about when or whether the limit exist or not. So, here we are now trying to investigate that part whether that limit exist or not, but here we instead of i and j we are just looking at $i = j$. Again, we will see, finally when we see the result, we will see everything for ij, but initially, first let us concentrate on $p_{ii}^{(n)}$. So, whether this starting from i, $P(X_n = i)$ if this as $\lim_{n\to\infty}$ this limit exists. In order to answer that question, first let us look at an example, so we consider the Markov chain with the following transition probability matrix. So, the Markov chain has two states 0 and 1 and the Markov chain is actually very simple, so if this is 0 this is 1, it just goes from 0 to 1 with probability 1 and it goes from 1 to 0 with probability 1. So, it is a extremely simple Markov chain and it is easy to see that it is irreducible because you can go from any state to any state and if you just try to calculate, you will see that this has a unique stationary distribution given by half comma half. So you can easily check that. Now as soon as it has a unique stationary distribution or just, because it has a stationary distribution, it is irreducible, so all states are positive recurrent. So it is a irreducible positive recurrent Markov chain with these stationary distributions. So, in that sense it is a nice Markov chain, now let us see for such a Markov chain, if whether this limit exists or not. Now, here you see, since like it is a very the dynamics of the Markov chain is very simple it goes from 0 to 1, 1 to 0. So, starting from 0 the probability that it will be at 0 in odd number of steps that is 0 because it goes from 0 to 1, 1 to 0, so it is at 0 at all even steps. So, $p_{00}^{(2n-1)} = 0$ and $p_{00}^{(2n)} = 1$ for all $n \geq 1$. Now, that is very simple from the dynamics of this Markov chain, but that tells you that this limit does not exist, because what is this like if you look at, so this if you look at, so this $p_{00}^{(n)}$ is nothing but just a real sequence and how does that sequence look like so it is like in or so $p_{00}^{(n)}$ is. So, if I just look at from say the first so odd place you have 0 and even place you have 1. So, it is a sequence of 0 and 1, so obviously this does not have a limit because the lim soup of this is 1 and limit as $n \to \infty$ of this is 0. So, you know like this kind of n, so it is kind of an alternating sequence, so the limit does not exist. So, you see for such a simple Markov chain which has all these properties of irreducibility, positive recurrence, unique stationary distribution still this limiting distribution does not exist. But, now the question is okay then is there some issue with this Markov chain?

Now, in order to understand that first we need to learn a definition, the definition of a period. Now, what is period of a state? So period of a state i is defined by the greatest common divisor of all integers n greater than or equal to 1 for which this $p_{ii}^{(n)} > 0$, i. e., what is d_i, d_i is the gcd of all $n \geq 1$ such that $p_{ii}^{(n)} > 0$, if this is a non-empty set, if this is an empty set then we define the gcd to be equal to 0, define the period to be equal to 0. Now, what is greatest common divisor? Again this is something you should know from high school. So, given a set of numbers, what is the greatest common divisor or the highest common factor, that it is also called the highest common factor, so what is that? It is basically, so great again so its greatest common divisor, so the name itself will tell you what it is. So, first of all it has to be a common divisor. So, if you are given a set of numbers, so in this case the set of numbers is the set of all n for which this $p_{ii}^{(n)} > 0$. So,I look at this set of numbers, now the greatest common divisor is a number which is first of all a common divisor, that means it should divide each number in that set, it is a common divisor and it is the greatest common divisor that means, if you have any other common divisor that will be less than or equal to that. Say for example, If I look at, if my set is 4 and 8, so if my set is, so here it is suppose my set is 4 and 8 then 1 is a common divisor because 1 divides everything, 2 is a common divisor, 4 is a common divisor, but what is the greatest common divisor, 4 is the greatest common device. But in this case, this set can be an infinite set. So, actually, so when this is non empty, this actually will be an infinite set. So, but again it does not matter, it is the same thing. The, so the greatest common divisor of a given set of numbers finite or infinite does not matter is a number that number, which is a divisor of all the numbers in that set, so it is a common divisor and if there is any other common divisor that has to be less than or equal to gcd, it is the greatest common divisor also this is called the highest common factor. So, what is this period? The period is the greatest common divisor of this set of numbers, what are the numbers? So it is the collection of all those $n \geq 1$ for which $p_{ii}^{(n)} > 0$. That means starting from i you can be at I, in the n-th step with positive probability. If n is such that for which this probability is positive then you take that n. So, you look at the collection of all such ns that gives you a set of numbers, the greatest common divisor of that set of numbers is called the period of that particular state i. And if this set is empty, then you call it, like then you just define the period to be 0. The important case is this first case when it is non-empty right. Further, if i is called aperiodic, if the period is 1. Now, let us see some examples to understand this concept of period better. So again consider the Markov chain, so this is basically the example which we saw in the previous, so example 34 and example 35 is the same Markov chain, in 34 we saw that this Markov chain does not have the limiting distribution or in other words $\lim_{n\to\infty} p_{ii}^{(n)}$ does not exist. Now, let us see that what is the period of say 0, now we have already seen that $p_{00}^{(2n-1)} = 0$ and $p_{00}^{(2n)} = 1$. So, now what is this set? Set of all $n \geq 1$, where $p_{ii}^{(n)} > 0$, so this is precisely the set $\{2, 4, 6, \ldots\}$, which is the set of all even numbers. Now, what is the greatest common divisor? So, 2 is a common divisor, so basically 1 and 2 are common divisors of set of all even numbers because if you take anything greater than 2, that will not divide 2. So, 1 and 2 are the greatest, are the common divisors of the set of all even numbers among them, 2 is the greatest, so d_0 is equal to 2. Similarly, see for even 1 also the same thing is true that $p_{11}^{(2n-1)} = 0$ and $p_{11}^{(2n)} = 1$, because the 1 also you go to from 1 to 0 and then you come back to 1. So, starting from 1, you can be, you are at 1 at even steps, so even by the same argument, you can also check that d_1 is also equal to 2. So, here we see that this 1 thing is there that here it has period 2. So, is that a problem because of which this $p_{ii}^{(n)}$, the limit of that did not exist, it can be because we saw why it did not exist, because the, if we looked at this sequence of $p_{ii}^{(n)}$, it was this alternating 0 1 0 1, so maybe there is some connection with this periodicity thing. But that, we will see slightly later, but before that, let us seeonemore example.

So now, we look at this S, so the state space is the set of all integers. So, $0 \pm 1 \pm 2$ and so on and suppose you starting from I, so again you go to, from i to $i-1$ with probability a, you go to i from, sorry, so this is i, this is $i+1$, and say this is $i-1$. So, you go from i to $i-1$ with probability b and you remain at i with probability c, where $a+b+c=1$, where $a, b, c > 0$. Now, fix a state i, now, suppose that $p_{ii}^{(1)}$, so again, what is $p_{ii}^{(1)}$? That is just simply pii, $p_{ii}^{(1)}$ is p_{ii}. So, that from here that is equal to c, so if that is strictly greater than 0, that means if I look at the set, the set of all $n \geq 1$, $p_{ii}^{(1)} > 0$, so then 1 belongs to that set. So 1 belongs to the set of all n for which $p_{ii}^{(1)} > 0$, because $p_{ii}^{(1)}$ which is nothing but just p_{ii} that is c, which is strictly greater than 0. So, if $c > 0$, 1 belongs to this set. Now, as soon as 1 belongs to this set, the only possible common divisor is 1, because if you take anything greater than 1 it will not divide 1. Remember, common divisor means it has to divide every number in that set and as 1 is in the set anything greater than 1 cannot be divisor. So, 1 is the only common divisor and hence it is the greatest common divisor. So, in this case the di is equal to 1. So, or in other words, if c is greater than 0, then i is what is called aperiodic. So, in this case, i is aperiodic. Now, that is if $c > 0$, if $c = 0$ that means what you can go from i to $i-1$ and i to $i+1$. So again, so this is just simply like a simple random walk, and in a simple random walk, we saw that if you are starting from a state you can come back to that state only in even number of steps, you cannot be come back to that state in odd number of steps. So, again for any i this is equal to 0 and now this is positive. Again not, like the previous example, it is not 1, but it is strictly greater than 0. So, starting from state i, there is a positive probability that you come back to i in any even number of steps. For example, if you have to come back in two steps, say there are two ways of doing that either you go from i to $i+1$ and then come back or you go from i to $i-1$ and then come back. So, all these for any even step this probability is strictly greater than 0. So, now if I look at, again the set, set of all n greater than or equal to 1 where $p_{ii}^{(1)} > 0$ that is again the set of all even positive even integers. So, $2m, m \geq 1$. Now, we have already seen this in the previous example, the only common divisors are 1 and 2 so 2 has to be the greatest common divisor, so in this case d_i is equal to 2. So, depending on whether $c > 0$ or $c = 0$ we get 2 kinds of things. So, if $c > 0$, then the period is 1 and if $c = 0$ then period is 2. And now again, you see, so here you see, I have done it for any state i, so it is true for all states. Now, even in this previous example also, you saw that the period is same for both the states.

Now, is there something special? Yes, like transience, recurrence, positive recurrence, null recurrence, even period is a class property. So, if i communicates with j then d_i is equal to d_j. In both the previous examples that we saw, it is very easy to check that both the Markov chains are irreducible. So every state has the same period, in the first example, it was both 0 and 1 had period 2, in the second example, if $c > 0$, then all states have period 1 and if$c = 0$, then all states have period 2. So, even period is a class property. So, if i communicates with j, then d_i the period of i is equal to period of j. So, a small thing, let me just so here, so basically, so instead, of this we should write it as d_i equal to d_j that is the notation that we are using not $d(i)$, but d_i. So, if i is equal to j, if i and j communicate then d_i is equal to d_j, so if i and j is in the same communicating class, then they have the same period, so states in the same communicating class have the same period or in other words periodicity just like recurrence, transience, positive recurrence, null recurrence is again a class property. So, if X_n is irreducible, then all states have the same period, in particular if all states have period 1, we say that the Markov chain is aperiodic just like if all states are recurrent. we say the Markov chain is recurrent or if all states are transient, we say that the Markov chain is transient. Similarly, if it is an irreducible Markov chain, and all states have period 1, then obviously all states will have the same period. Now, if it is equal to 1, then we say that the Markov chain is aperiodic. Now comes the important theorem, which says that let Xn be an irreducible positive recurrent, and aperiodic Markov chain, then for any initial distribution μ, so that means what is μ, so μ is like $P(X_0 = i)$ is μ_i, that is the meaning of any initial distribution μ, then $\lim_{n\to\infty} P_\mu$. Again, P_μ means just this, that you are starting from the initial distribution, $P_\mu(Xn = i)$, this exists and is equal to π_i for all i, where this π_i, $i \in S$, is the unique stationary distribution. Again, since it is irreducible and positive recurrent, we know that the stationary distribution exists and it is equal to the, it is, it exists and it is unique and it is equal to $\frac{1}{E_i(T_i)}$. So again, this quantity or the limiting distribution is the stationary the unique stationary distribution. Now see, this is not again we have already seen this at least for in the case of finite state Markov chains, that if the limiting distribution exists, then it has to be stationary distribution. We have already seen this result, the question was when the limiting distribution exists. So, here, we give a condition on the Markov chain under which the limiting distribution exists. So namely, if it is irreducible, positive recurrent and aperiodic, then for any initial distribution mu, so no matter what initial distribution you start with this limiting distribution the by limiting distribution, I mean this quantity that starting from any initial distribution the probability that at the nth step, you will be in state i as $n \to \infty$, that converges to π_i, where π_i is $\frac{1}{E_i(T_i)}$,because this is a irreducible positive recurrent Markov chain, so it has a unique station distribution π where π_i is $E_i(T_i)$. So, in particular suppose my μ is δ_i, so that means if I say start from i and then if I look at, say starting from i what is the $P(X_n = j)$ this as $n \to \infty$ will converge to π_j. So here, I am looking at the special case, where my initial distribution is δ_i or I start with from i with probability 1. So since, this is true for any initial distribution, I can obviously choose my mu to be δ_i, so for any two states i, j, $\lim_{n\to\infty} p_{ij}^{(n)} = \pi_j$. So, in the theorem in the module of stationary distributions, we saw that for, if it is a finite state Markov chain, then if the station limiting distribution exist that has to be stationary distribution, but what this theorem is telling you is giving you a condition or a set of conditions rather, under which this limiting distribution exist. So that this is equal to π_j is actually not surprising that, this is equal to the stationary distribution is not surprising. But the important thing is this, what so the main content or crux of this theorem is condition under which the limiting distribution exists. So, the first example which we saw namely this example, so what was the issue with this, the issue with this example was that it had period 2, because we have seen this here, that this Markov chain has period 2. So, this is not aperiodic, hence that is why, so for the limiting distribution 2 exist there this aperiodicity is needed if it is not aperiodic the limiting distribution may not exist. So, if the Markov chain is irreducible, positive recurrent, plus aperiodic, then the limiting distribution exists and it is equal to the stationary distribution. So you see again like why stationary distribution is important, so what this basically tells you is that it says that, so you start from any, so, we have seen that if you start from a stationary distribution, you remain there that means the distribution of each X_n is equal to that stationary distribution. But now, if you start from any initial distribution, what this theorem is telling you is that as $n \to \infty$ $P(X_n = i)$ that is equal to π_i, so even if you do not start from the initial distribution if your chain is irreducible, positive recurrent and aperiodic, then basically the chain moves towards initial, moves towards the stationary distribution. So that is why some books also use the term equilibrium distribution for stationary distribution, because every system wants to go towards the equilibrium. So, if you, if the chain is irreducible, positive recurrent, plus a periodic then the, no matter what initial distribution you start with the chain will eventually move towards the stationary distribution that is not a very precise mathematical statement. The precise mathematical statement is this that $\lim_{n\to\infty} P_\mu(Xn = i) = \pi_i$ that means the limiting distribution is the stationary distribution. And this theorem is telling you something more, so remember that theorem, in the module for stationary distribution, the theorem was only for the, statement of the theorem was only for finite state Markov chain, state that if it is a finite state Markov chain, then if the limiting distribution exists, then it has to be stationary distribution, but what this theorem is telling, okay, you do not need actually finite state Markov chain. If it is, if the limiting distribution exists and the limiting, again if the if even if it is a infinite state Markov chain the and provided it is irreducible, positive recurrent and aperiodic then the limiting distribution exists and it is equal to the stationary distribution. So, even if you do not start from the stationary distribution, you eventually move towards the stationary distribution. So, that is why it is also sometimes called an equilibrium distribution, so that is because every system wants to move towards equilibrium and if you have an irreducible, positive recurrent, aperiodic Markov chain, then its distribution moves towards the stationary distribution. Now, I will end with 1 just small remark, that this kind of a Markov chain if it is irreducible, positive recurrent and aperiodic some books also use the term for such a Markov chain they use the term ergodic. So, if in a book you see that X_n, let X_n be an ergodic Markov chain, that means it is irreducible, positive recurrent and aperiodic. So, ergodic is a common term used for if a Markov chain has all these three properties, that it is irreducible, positive recurrent and aperiodic then such a Markov chain is also called aperiodic. So, I just wanted to mention this because in some book, you might encounter the book on Markov chains, discrete time Markov chains, you might encounter this terminology ergodic and this is what it means irreducible, positive recurrent and aperiodic. So, what we have seen today is two more important properties of stationary distribution again not two more, actually like one more that it is actually the limiting distribution. So, we already knew it for the case of finite state Markov chains, but there we did not know when the limiting distribution actually exists.

So, today we saw a criteria under which the limiting distribution exist and also, so once the limiting distribution exists it is equal to the unique stationary distribution provided, the chain is irreducible, positive recurrent and aperiodic. So, you see, so suppose someone asked you, okay, if you look at a chain of, so long time from now, so that is the meaning of so in real life what is the meaning of $\lim_{n\to\infty}$? Suppose the chain starts from now, the current time and someone asked you, so if you now look at the chain, a long time from now what is the probability that you will, that the chain will be in state i. Then the answer is if the chain is irreducible, positive recurrent and aperiodic, then in the long run the probability that you will find the chain in state i is equal to π_i, that is the way you should interpret this mathematical theorem that or the mathematical result that the limiting distribution is π_i. So, in the long run, if you look at the chain, the probability that you will find the chain in state $i$is equal to π_i, where π_i is the unique stationary distribution. So, we also know what this $\pi_i = \frac{1}{E_i(T_i)}$. So, we will stop here today. Thank you all.