Discrete – Time Markov Chain and Poisson Processes
Professor. Subhamay Saha
Department of Mathematics,
Indian Institute of Technology, Guwahati
Lecture 23
Lecture: Some Problems II

Hello everyone, welcome to the 23rd lecture of the course Discrete-Time Markov Chains and Poisson Processes. So, in the last lecture, we solved two problems. One was about finding stationary distributions of a finite state Markov chain and the other was that random walk on a tetrahedron, and finding the mean return time. So, one important concept that we learned in last lecture is about this doubly stochastic matrix. And that fact is that, if you have a finite state irreducible Markov chain, which has a doubly stochastic transition probability matrix then the unique stationary distribution is very special. And it is of the form for each component is $\frac{1}{n}$, where n is the size of the state space. So, in today's class also we will use this fact in one of the problems. So, today we will solve few more problems, so, let us start.

So, let us read the problem carefully. So, such problems are like so, when you have a word problem, then first you need to understand it properly, you need to formulate it mathematically and then solve it. So, these word problems are slightly difficult problems. So, that is why I am trying to solve a few in today's lecture. So, each morning an individual leaves his house and goes for a run. He is equally likely to leave either from his front or back door. So, there are these two doors, front door and back door. Now, first thing is the individual can, it is equally likely that, so both are $\frac{1}{2}$. So, he is equally likely to leave either from front or back door. Now, upon leaving the house or while leaving the house, he chooses a pair of running shoes. So the individual has multiple pairs of running shoes. So, he chooses one of them again, uniformly. So upon leaving the house, he chooses a pair of running shoes or goes running barefooted, if there are no shoes at the door from which he departed. Now, so if, so he has say some so actually we will see that he has four pairs of running shoes. Now, first thing that every morning he decides to choose a door uniformly, either he, so he chooses a door with equal probability, so either back door or front door. Now, depending on how many shoes are there on that particular door which he chooses, he will choose a shoe from the available pair of shoes. And if there is no pair of shoes in that particular door, then he will go out running barefoot. Now, on his return, now when he comes back from running; again, he is equally likely to enter and leave his running shoes either from the back door or front door. Now, when he is returning again, either he will enter to the back door or the front door, whichever door he enters, he will leave the pair of shoes there, provided he had a pair of shoes if he went barefooted, then there is no question of leaving the shoe. Now, if he owns a pair of a total of four pairs of running shoes, what is the long run proportion of time he runs barefooted? So, what is the long run proportion of time he runs barefooted? So again, whenever you see this terminology long run proportion of time, you should immediately start thinking of stationary distribution. Because, we saw this result that this long run proportion that of that, the long run proportion that a Markov Chain is in a particular state is equal to the stationary distribution. So again, this problem is asking you about long run proportion of time he runs barefooted, so you should think of stationary distribution. But, before thinking of stationary distribution, first you need to formulate this entire problem mathematically.

So, let X_n be the number of pairs of shoes at the front door on the nth morning. So, what I am defining my Markov chain to be? So, X_n is the number of pairs of shoes on the front door on the n-th morning. So, here there is a symmetry between front door and back door, because he, while going out also, he chooses them equal, with equal probability, while coming back also he chooses them with equal probability. So, we will just consider the number of pairs of shoes in one particular door and you will see that will be sufficient. So, let X_n be the number of pairs of shoes at the front door on the nth morning. Again, if we define our X_n like this, it is easy to see that the, that this is a Markov chain. So, if the number of shoes on the nth morning is say i, then you can easily then the probability then number of shoes that will be there in the next morning in order to say in order to say something about that, or in order to state the probabilities you only need to know the current state and nothing from the past. So again, the X_n defined in this way is a Markov chain. Then, X_n is a Markov chain with this transition probability matrix. So here I need to explain this a bit. So, I will explain a few states others you will, you can do it in a similar fashion. So, suppose it is the first thing, it says that it goes from 0 to 0 with probability $\frac{3}{4}$ why is that? Now, 0 means in the front door, suppose there are no shoes. Now, what is the probability that in the, so suppose on the nth morning in the front door, there are no pairs of running shoes. Now, what is the probability? Then the next time also there will be no pair of running shoes. Now, in how many ways is that possible? You need to just carefully count it. Now, so front door, so the individual leaves from that door. Now he is now running barefooted. Now, if he comes back from that door, then, so in the front door, there are zero shoes. So, now first, which door the individual chooses to go out? Suppose if the individual chooses to go out from the front door. Now, if he chooses the front door, now there are no running shoes, so he will go out running barefooted. Now, while coming back, now since he is running barefooted, while coming back, it does not matter whether he comes back to the front door or the back door, the number of shoes in the front door will not increase, because he went out barefooted. So while coming back, whether he chooses the front door or the back door, it does not matter, because the number of shoes will not increase, because he had no shoes on. So, there is no point, there is no question that he will leave a shoe, while returning, because he did not have any shoe on. So, if he chooses the front door, because there were no shoes on the front door. So it does not matter whether he comes back via the front door or the back door the number of, the next day the number of shoes in the front door will be zero. So, here there are two possibilities. So front door, going through front door, coming back through front door, going through front door, coming back through front door, in both these cases it is the possibility is that it will remain at 0. And what are these possibilities choosing the front door, that probability is $\frac{1}{2}$; and then coming back through front door that probability is half that is one situation. Another situation is choosing the front door with probability $\frac{1}{2}$, and then returning via the back door that probability is also $\frac{1}{2}$. Now, another situation, now so this is if the individual chooses the front door. Now, if the individual chooses the back door, now there are also all four pairs of running shoes are now on the back door, because, so the total number of pairs is four, so in the front door, it is 0, then in the back door there are four, because the sum should be equal to 4. So, now if the individual chooses the back door, now he will choose one particular pair of running shoes with probability $\frac{1}{4}$, because there are 4, so he will choose that with probability $\frac{1}{4}$. So again, it does not matter so any shoe, so again, because the number of shoes does not matter. As long as there is a shoe, he will wear a shoe and wear, and go out running wearing that shoe. So choosing a shoe does not matter. So, if the probability of choosing the back door is $\frac{1}{2}$. Now, if he returns again via the back door, then the number of shoes on the front door will not change. But, if he leaves from the back door, but returns from the front door, then now the number of shoes on the front door will be one pair. So that, so if the individual leaves from the back door and then returns from the back door in that case also the number of shoes in the front door will remain 0. So, you remain from 0 to 0. So again, that probability is half times $\frac{1}{2}$, sorry, $\frac{1}{2} \times \frac{1}{2}$. So, if you sum it up, it gives you $\frac{3}{4}$ that is this particular $\frac{3}{4}$ and 0 to 1. So, if he goes via the back door and returns through the front door, then again that probability is choosing the back door probability is half. Then, returning to the front door that probability half you get $\frac{1}{4}$. So, you get $\frac{1}{4}$, so that is this one way. So, 0 to 0 that probability is $\frac{3}{4}$ and 0 to 1, that probability is $\frac{1}{4}$. Now, suppose if we look at say, 1 to 0, so again, I will not explain each of them, but I am just giving you the idea, so that you can do it for other states. Say, suppose 1 to 0 what is the probability of going from 1 to 0? Now, again, first, suppose he chooses the front door, now from the front door there is a shoe, so he will wear that and go out. Now, if he again return through that front door itself, so then the number of shoes will remain one itself. So, that is the only way it can remain from 1 to 1, no sorry. So, he, so that is one way that the number of shoes in the front door will remain one itself, so again let me repeat. So, there are now one pair of shoes in the front door. Now, suppose if the individual chooses the front door, goes via the front door and returns via the front door that then the number of shoes in the front door will remain 1, so that probability is again $\frac{1}{2} \times \frac{1}{2}$. That is one way that it will remain starting from 1, it will remain at one. So, if the number of shoes on the nth morning is 1, then the number of shoes on the n plus oneth is morning will also be 1. Now, if he, so again, so he has chosen now the front door; but if he returns through the back door, now it will become 1 to 0, that is 1. So in this scenario, it becomes 1 to 0. Now, let us suppose the individual chooses the back door, chooses the back door, returns via the back door. So, now there was one shoe on the front door, so it will remain as it is. So, that is another way where it will remain from 1 to 1 how? Chooses the back door to go out, and then chooses the back door to come back. So again, so in this situation also it will remain, so that is half times $\frac{1}{2} \times \frac{1}{2} + \frac{1}{2} \times \frac{1}{2}$, you get $\frac{1}{2}$. So, this is this $\frac{1}{2}$, so 1 to 1 goes via the, so chooses the front door, comes back via the front door, chooses the back door, comes back via the front door, then it will remain 1 to 1. Goes via the front door, returns via the back door, it will become 1 to 0, goes via the back door, returns via the front door, then it will become 1 to 2. So, that is these two $\frac{1}{4}$ and $\frac{1}{2}$. So, from one, it can either go to 0 or it will remain at 1, or it will go to 2 and I have explained all the possibilities. So 1 to 1, the probability is $\frac{1}{2}$, 1 to 0 the probability is $\frac{1}{4}$. So, he goes via the front door and returns via the back door that is when it will become 1 to 0. And 1 to 2 that probability is also $\frac{1}{4}$ why? So, goes via the back door and then returns via the front door, so then that probability is $\frac{1}{4}$. So, in this situation, it will become 1 to 2. So remaining also you can easily check. But now, the important thing is again, if you look at this transition probability matrix, again, this is doubly stochastic. You can see, so the column sums are also one, row sum anyway will be 1, but the column sum is also 1. So again, using that useful fact, we get now here the size of the state space is 5, so the stationary distribution is 1 by 5 for each component. Now, what is the long run proportion of times he runs barefooted? Now, if the number of shoes is 0 in the front door, and he chooses the front door then he will run barefooted. And what is the long run proportion that the number of shoes in the front door is 0? That is equal to the first component of this stationary distribution. Because you have seen that the long run distribution or long run proportion is actually equal to the stationary distribution. So here you are interested in the long run proportion of times there are 0 shoes, 0 pair of running shoes in the front door. So that should be π_0, which is equal to $\frac{1}{5}$ and you choose the front door with probability $\frac{1}{2}$, so that is $\frac{1}{2}\pi_0$. Now, another possibility is all four shoes are on the front door and the individual chooses the back door, so, in that situation also he will run barefooted. So, what is the long run proportion of times that the front door has four shoes? Again, that should be equal to π_4, which is equal to $\frac{1}{5}$, so it is $\frac{1}{2}\pi_4$. Again, the probability of choosing the back door is $\frac{1}{2}$ and all four shoes in the front door. So, in that situation also the individual will run barefooted. And what is the long run proportion of times that the front door has four shoes? Again it should be equal to π_4 from the results that we have learned till now. So, it is $\frac{1}{2}\pi_0 + \frac{1}{2}\pi_4$, which gives you $\frac{1}{5}$, because both π_0 and π_4 are $\frac{1}{5}$ each. So, finally the answer is $\frac{1}{5}$. So, here the, so there are many things about this problem. First of all, like this kind of word problems if the transition probability matrix is not straightaway given to you. So from the wording, you first need to figure out what the transition probability matrix is. You need to write down the transition probability matrix properly; that is the first challenge. And then if you are asked questions like long run proportions, you need to find the stationary distribution. So, if you have to find the stationary distribution, the first approach is you write down the equations and try solving it. But in some special cases, you do not have to do all that work. And that special case is if the matrix is doubly stochastic, and which is the case for this particular example. In that case, you know what the stationary distribution is without doing any calculation. So, we got that the stationary distribution is $\frac{1}{5}$ each and now we get the long run proportion. So that is this problem.

So moving on, again, we have a word problem. So again, let us read this carefully. In a good weather here, the number of storms is Poisson distributed with mean 1. So, you know what Poisson distribution is. So it is a discrete random variable, and you know it is probability mass function. But anyway, we will not need all those things in this particular problem. So, it is a Poisson distribution with mean 1, you know that for a Poisson distribution what the mean is, that is the λ. Poisson(λ) means the mean is λ. So that is the, so here the parameter is 1. But anyway, as I said, like you will see in this particular problem, you will not need all those things only thing you will need the mean. But in general, if you are given the mean of the Poisson distribution, you can write it. Write down its probability mass function, because the mean gives you the parameter of the Poisson distribution. And in a bad year, it is Poisson distributed with mean 1. So in a good year, the number of storms is Poisson distributed with mean 1 and in a bad year, it is Poisson distributed with mean 3. Assume that any year's weather condition depends on the past years only through the previous year condition. So, this is the thing which will give you, allow you to model this situation using a Markov chain. Now, if a good year is equally likely to be followed, either by a good year or a bad year and that a bad year is twice as likely to be followed by a bad year, as by a good year. Find the long run average number of storms per year, find the long run average number of storms per year. So, as soon as you see this word long run, again you should start thinking about stationary distribution. So, how do we solve this problem? Again, let $X_n = 0$ if the n-th year is bad and $X_n = 1$ if the n-th year is good. So, we call $X_n = 0$, we define $X_n = 0$, if the n-th year is bad and we say it is equal to 1, if the n-th year is good. Then, clearly X_n is a Markov Chain. That is because of this given hypothesis that the weather condition depends only, any years weather condition depends on the past years, only through the previous year condition. That is why X_n is a Markov chain with transition probability matrix. Now 0 is bad, now it says that a bad year is twice as likely to be followed by a bad year as by a good year. So, 0 to 0 the probability, so 0 to 0, probability is $\frac{2}{3}$. Because it is twice as likely that a bad year will be followed by a bad year than a good year, and so 0 to 1 is $\frac{1}{3}$. But, if it is a good year, so if, so 1 means good year; if a good year is equally likely to be followed by a good year or a bad year. So, 1 to 0 is $\frac{1}{2}$, 1 to 1 is also $\frac{1}{2}$. So, if it is a good year, then the next year will be good with probability $\frac{1}{2}$; and next year will be bad with probability $\frac{1}{2}$. But if it is a bad year, then the next year will be a bad year with probability $\frac{2}{3}$ because bad year will follow a bad year that is twice likely. And if the current year is bad, the next year will be good with probability $\frac{1}{3}$. So, we have written down the transition probability matrix. So, we have formulated the problem as a Markov chain and we have written down the transition probability matrix. So, we have completed the first challenge or we have overcome the first hurdle. Now again, since the question is about long run average number of storms, first let us find the stationary distribution. Now here, you need to do the calculation, so I have checked it. Now, so this is a small exercise for you just check that this. So again, it is easy to see that this is an irreducible Markov chain. And since it is a finite state Markov chain, irreducible means all states are positive. So, it will have a unique stationary distribution and you can check that the unique stationary distribution is given by this $(\frac{3}{5}, \frac{2}{5})$. Also, notice that this is an aperiodic Markov Chain. Why? Because p sorry, because $p_{00} > 0$, and similarly, $p_{11} > 0$. So in that, so remember, what was the period? So period was the GCD of all n, such that $p_{ii}^{(n)} > 0$. Since the set contains 1, so the greatest common divisor has to be 1, because anything bigger than that will not divide 1. So, we have already seen this, if the ones state probability. So, if $p_{ii}^{(1)} > 0$, then i has to be aperiodic. So, this is, and this is an irreducible Markov chain, so all states have the same period. So, this is an irreducible positive recurrent, aperiodic Markov chain. So, which, so it is which we also call as the ergodic Markov chain. So, this is an irreducible positive recurrent a periodic Markov chain, irreducible you can easily check the finite state Markov chain. So, all states are positive recurrent and it is aperiodic, because both $p_{00} > 0$ and $p_{11} > 0$. So, this is aperiodic and this is the unique stationary distribution. Now, we have seen that if it is an ergodic Markov chain that if it is irreducible, positive recurrent and aperiodic then in the long run the probability that you will find the Markov chain in state i that probabilities is π_i. No matter from what initial distribution you start with that $P_\mu(X_n = i) = \pi_i$; or it is not actually equal to π_i. So, $\lim_{n\to\infty} P_\mu(X_n = i) = \pi_i$. That means in the long run, the probability that you will find the chain in state i that is equal to π_i, no matter what the initial distribution was. And this is because the chain is recording or in other words, it is irreducible, positive recurrent and aperiodic. So, thus what is the long run average number of storms per year? So in the long run, the probability that it will be a bad year. So remember, 0 is bad, the probability that it will be bad year is $\frac{3}{5}$, because π_0 is $\frac{3}{5}$. And in a bad year what is the mean number of, remember you have been asked the long run average number of storms, that means the mean number of storms. So, in the long run the probability that it will be a bad year that is $\frac{3}{5}$, that is because the first component of this stationary distribution is $\frac{3}{5}$. And if it is a bad year, then the mean number of storms is 3. So it is again as I said, so the only thing we are interested in is in the mean number of storms, so it is Poisson distributed. If the question was something else, then the exact distribution would have mattered, but here only thing we are interested in is the average number of storms, and that is 3. So if it is a bad year, then the average number of storms is 3. So, the probability that it will be a bad year is $\frac{3}{5}$ in the long run times the average, so it is $\frac{3}{5} \cdot 3$. And in the long run, the probability that it is a good year is $\frac{2}{5}$ and in a good year, the mean number is 1. So, you get $\frac{3}{5} \cdot 3 + \frac{2}{5} \cdot 1 = \frac{11}{5}$. So, the long run, average number of storms per year is $\frac{3}{5} \cdot 3 + \frac{2}{5} \cdot 1 = \frac{11}{5}$. And why is that? Because here we are using the fact that in the long run, the probability that you will find a Markov chain in state i is equal to π_i provided the Markov chain is ergodic or it is irreducible, positive recurrent and aperiodic. All these three conditions are true for the particular example that we are looking at. So in the long run, the probability that it will be a good year, or it will be a bad year is $\frac{3}{5}$. And if it is a bad year, then you are given that the mean number of storms is Poisson distributed, then the number of storms is Poisson distributed with mean 3. So, here the exact is, at least for the given problem, the exact distribution is not important. The important thing is the mean, so the mean is 3. And in the long run, the probability that it will be a good year is $\frac{2}{5}$ and in a good year, the mean number of storms is 1. So, you get $\frac{3}{5} \cdot 3 + \frac{2}{5} \cdot 1 = \frac{11}{5}$. So, here we use this last theorem, which said that if you have an irreducible, positive recurrent, aperiodic Markov chain, then in the long run the probability of finding the Markov chain in a particular state is equal to that component. So, the probability that you find the chain in state i will be given by the ith component of the unique stationary distribution. Since, it is irreducible and positive recurrent, so the stationary distribution is unique and it will, so this quantity, the long run probability of finding this Markov chain in state i is given by π_i. So, using that we get this long run average number of storms per year. So we will stop here today. Thank you all.