Discrete-Time Markov Chains and Poisson Processes

Professor Ayon Ganguly

Department of Mathematics

Indian Institute of Technology, Guwahati

Module: Hitting Time and Strong Markov Property

Lecture 12

Strong Markov Property

Welcome to the 12^{th} lecture of the course Discrete-Time Markov Chain and Poisson Processes. Today we are going to talk about what is called strong Markov property. Earlier just to recall that we have talked about what is called Markov property. What was Markov property just recall that and then we will go for what is called a strong Markov property. And we will then discuss what is the main difference between Markov property and strong Markov property. And finally, we are going to see 1 example where the strong markev property can be used to solve the problem.

Let us recall that the Markov property was as follows that if I have a Markov chain X_n if I have a some fixed time m, and I condition at the time m the state is i then after the time m the Markov chain is starting afresh. Markov chain is starting afresh from the state i. That means that if I have the time horizon like that the time is 0 here then 1 and so, on so, forth at some time I have the m^{th} time. So, m is a fixed time and if I write the state here. Then state is i here, then the Markov property tells us that well this part is independent of this part. Future is kind of independent and the Markov chain is starting from here. So, I just rescale this time to be 0 and so, on so, forth, then $m+1$ may be mapped to 1 and so, on so, forth. And that all the things that transition probability matrix remain same only thing is that initial distribution now is δ_i that means, I am starting from i and then the Markov chain is approaching to the time is approaching. That is basically the idea of the Markov property. Now, let us see in this case what I have pointed out this m is a fixed time. This m is a prefix time actually. So, it is not a random time, but it is a prefix time like m could be 1 m could be 2 m could be 3 does not matter, but this is not a random time, better word to give you is the non-random time. Now, natural question is that what will happen if instead of a non-random time if I take a random time. And strong Markov property talk about that of course, I cannot talk about any, any random time I have to talk about a particular kind of time which is called stopping time. So, I will discuss that but just for example, suppose I take that instead of m maybe I take the hitting time to the state i. Because if at T^i know the state is i because this is the hitting time that is the first time I enter into the state i starting from some non-i state. So, if I start from state i then T^i will be 0 that we have discussed earlier. So, I know at the at T^i the state is i. So, the state is known at the point T^i but T^i itself a random time. So, now if I replace this m with T^i So, that means if I have a time horizon like that, and some T^i is here, now, can I tell the same thing that this part is independent of this part or those kind of thing can we tell, can we tell that the Markov can start afresh from the state i onwards. Can we tell these things?. This question is answered by what is called strong Markov property. And as I mentioned that well, here instead of m, I cannot take any random time, but I have to take a specific type of random time. And that random time is called stopping time.

What is stopping time?. Stopping time of a Markov chain X_n is defined as follows. A random variability T which is of course time in our case random variability taking values in the set $0, 1, 2, \cdots, \infty$. So, T can take value 0, T can take value 1, T can take value 2 so, on so, forth. So, all non-negative integer value it can take along with it also can take infinity value also just like hitting time. Hitting time can take value 0, take value 1, take value 2, and so, on so, forth all are non-negative integer value it can take along with it can also take the infinitive value if starting from state j if I do not hit the state i then the hitting time of the state i will be infinity that we have discussed. So, T is a random variable which basically take values in the set is called a stopping time of a Markov chain X_n if following happens. So, what happens?. If the event $\{T = n\}$ is completely determined from the value of X_0 to X_n. So, this n is same as this n. So, if I know what are the values of time point 0, time point 1, time point 2 and so, on so, forth time point n, I can tell what is the where $\{T = n\}$ has occurred or not, I do not need to know any information of the future only up to n, I have to know to tell that where the $\{T = n\}$ has occurred or not. That means that $\{T = n\}$ this event can be completely specified or that that event depends only on the random variables $X_0, X_1, X_2, \cdots, X_n$. Only on this random variable this depends on that. And note that this n is same as this n and this particular thing has to be true for all possible non-negative values of n, 0, 1, 2 whatever the values of n this has to be true. Now, intuitively speaking what does this basically mean?. This means that if I say to stop at T, T is a random time, then if you see the process up to the n you can tell whether you have to stop at n or not. If you see observed value of the Markov chain up to the point n you can tell whether you have to stop at n or not. So, that means observing the observed value of the Markov chain I can tell when I have to stop if you if I am asked to stop at that time point T.

Let us see some couple of examples. First example is basically the hitting time just recall that hitting time we have defined as $\inf\{n \geq 0 : X_n \in A\}$, $A = \{j\}$. This was the hitting time. And now, you say that if suppose I have observed $X_0 = j$, that means basically nothing but that I can easily say that $T_j = 0$. These two are actually equivalence sets. So, if I observed $X_0 = j$, I can tell that $T_j = 0$. So, $T_j = 0$ is completely specified by it depends only on $X_0 = j$ does not depend on what are the future values are. So, if I take in this case, I taken $n = 0$ here. And so, $T_j = 0$ is same as $X_0 = j$. So, that means, this is completely depend on these this particular event and so, this one holds true. Now, what will happen if I take $n \geq 1$. Now, you see that $T_j = n$ when we say $T_j = n$ occurred that means, starting from some non-j state, I hit that j state for the first time at that time step n. Then we say this happens and that I can write in this way that $X_0 \neq j, X_1 \neq j, X_2 \neq j, \cdots, X_{n-1} \neq j, X_n = j$. Again, you see that if I know the values of $X_0, X_1, \cdots, X_n$ then I can tell that where the $T_j = n$ has occurred or not been occurred. That means that in this case, seeing this quantity I can tell whether $T_j = n$ has occurred or not and that is why this T_j which is the hitting time is also a stopping time. So, hitting time to the state j is always the stopping time for that Markov chain.

Now, let us take another example, which is not a stopping time. In this case, you see that what I have taken is that suppose I take this definition clearly this also be can take values in the set $\{0, 1, \cdots, \infty\}$ in this particular set. Now, see in this case what happens is that, if I have to see whether L^i equals to n or not then basically starting from 0, I have to know what happens in this part. Whether the chain vistis the state i here in the future after n or not, if the chain does not visit i then $L^i = n$ if the chain visits i after n then it will not be n it will be more than n. So, to say whether this has occurred or not I have to know what are the values of the Markov chain in the future to n. That basically means that this particular event now does not depend on only in this part. But it also depends on this part also. I have to know whether the Markov chain has visited or will visit the state i or not after n. So, that is why this L^i in general is not a stopping time is it is not important. Let us take another example which is intuitively make very sense suppose I am in a train I am moving this direction and there are 3 stations station A station B and station C. What I want?. I want to visit the station B. I want to actually leave the train get out of the train at the station B. This is my destination and I want to leave the train at the station B. Now, when the train has started from somewhere here and it moving and for the first time I am going here and for the first time I am going to the station B in this in this particular rail route and I do not know that when station B will come. So, what I am going to do?. I am going to ask my co passengers that I will tell him that, please help me and I want to get out of the train at station B please tell me when station B will come. They just normal reply which come to us is that well, you look for station A and station B is the next station to station A. So, when station A will come you just go to the door of the train, and then it will come to eventually it will come to station B you just get out of this train. You just get out of the train that is the normal answer we get from our co passengers. Now, if that information is given to me that station B is next to station A that is fine with me, I look for station A when station A will come I will go to that door and then I will get to get out of the train at the station B. But suppose some co passengers still told me that station B is followed by station C, then there is a problem that is not a useful information to me. I cannot get out of the train at station B if I have the information that station B is followed by station C if I have only this information that I cannot get out at the station B. And the idea of the stopping time exactly that; that if I know up to this I can know whether I have station B has reached or not. If I can do this, then that information works as a stopping time. But if I cannot tell whether the station B is coming based on only this information. Then that will not that information will not work as a stopping time just a second information, second information tells that all stations B followed by station C from that I cannot get out of the train at station B because I have to know the future to understand whether that the next station is station B or not. And that is not at all useful information to me. But if I am told that station B is next to station A that is a useful information to me. And that information can work at a stopping time. Because using that I can stop at station B, but the other information that is this station B is followed by station C using that I cannot stop at station B and so, this is not a stopping time information. So, this information I feel that may be helpful to understand what is the stopping time and now again visit these two examples that in the first case given the information obtained I can tell whether $T^j = n$ has occurred or not. So, that was that first example is T^j is a stopping time, but in the next example, L^i is not a stopping time because given up to the information n, I cannot tell whether L^i equals to n has occurred or not.

So, with that idea of stopping time let us move and let us see now what is strong Markov property. Again recall what was the Markov property. In case of the Markov property we know that if I have some time m and the state is i, I write time here I write state at the above. Then basically this part and this part kind of independent and Markov chain start for start afresh from here. Now, our question was that if I replace m with some stopping type and that is given here. What is tells that well, if I have some Markov chain of some time horizon and T is a stopping time, and if I know that the state at the time point T is i then you see that what happens that information is given in this theorem. It tells that well, if X_n be a Markov chain, with transition probability matrix P and some initial distribution does not matter what is the initial distribution. And T be a stoppings time for the Markov chain. Of course, for this particular Markov chain then conditional on $T < \infty$ and $X_T = i$. $X_T = i$ means basically state at the time when T is i. And I am taking $T < \infty$ if I do not take $T < \infty$ then X_T does not make much sense because T_∞ does not make any sense.

So, this condition I have to take to make this quantity to be a meaningful quantity that X_T is a meaningful thing that I take that T is finite. Stopping time is a finite one. So, it says that if conditional on $T < \infty$ and $X_T = i$, $\{X_{T+n}\}_{n \geq 0}$, is nothing but the Markov chain in this part. $\{X_{T+n}\}_{n \geq 0}$ it is X_T then X_{T+1} then X_{T+2} and so on so forth. So, this is nothing but the Markov chain in this part only is the Markov chain. So, this sub part of the original Markov chain is again a Markov chain, with the same transition probability matrix and initial distribution is δ_i so, as if from i it is again starting afresh. You see that the conclusion here and conclusion here are same. That means, even if I replace m with some stopping time the conclusion remains same. And then also I have the independence is also there. So, now this part is independent of this part. So, X_0, X_1, X_T these random variables are independent of X_T, X_{T+1}, X_{T+2} and so on so forth under the condition that $T < \infty$ and $X_T = i$. So, strong Markov property says that even in the Markov property if I replace the m with some stopping time, the conclusion of the Markov property still holds true. And why we use the strong word here because it is stronger in the sense that instead of a fixed time, now I can take the stopping time also. Let us go through this once again that it says that if I have a Markov chain with transition probability matrix P and whatever initial distribution does not matter, I have a random variable T, which is a stopping time for that Markov chain X_n. Then if I condition with respect to $T < \infty$ and $X_T = i$ that means T is a finite time and at the time point T the state is i if I condition with respect to all these things, then the next part of the Markov chain is independent of the previous part and the sub part of the Markov chain again is a Markov chain having the same transition probability matrix P. And the initial distribution is δ_i because at the point T, I know that the state is i. So, starting from state i, Markov chain process backups and proceed in time in future and the sub part that from X_T to or X_T onwards that part of the Markov chain is basically again a Markov chain having the same transition probability matrix P and it start from i that means the initial distribution is δ_i and the independence is also there that the that X_0 to X_T so, that is basically the Markov chain in this part is independent of Markov chain in this part. So, that happens.

Now, let us proceed to an example where I can use the strong Markov property to solve this example. What is this example?, the example start with the same idea of that birth-death process. Only thing is that here this probability of birth remains same for instead from i to $i+1$ or $i-1$ to i remain same. Whereas probability of death also remains same. It does not depend on either state i. So, it is also can be thing of a generalization of the Gamblers Ruin problem because in case of the Gamblers Ruin problem, I do have some terminating you know state here. But in this case, we are not assuming any terminating state as it is going on in time. I mean it is a cycle I have infinite states. So, the state can be any non-negative integer in this case. In this case, what is our aim?. Our aim is to obtain this probability. What is this probability you know this is nothing but the conditional probability that starting from 1 what is the probability that I hit 0 for the first time in n time step. I hit 0 for the first time in n steps that probability we try to find out. We recall that again this is nothing but the hitting time to 0. So, it is starting from 1 what is the probability that first time I am entering into the state 0 in the n^{th} step that probability will try to find out for all values of n. So, clearly that $P_1(T^0 = 0) = 0$ because starting from 1 in the 0 step I cannot be in 0. So, that probability has to be 1. Then I am looking for what is the probability that $T^0 = 1$. That I am looking for what is the probability that $T^0 = 2$ starting from 1 and I am going on going in. That is basically the thing we try to find out in this particular example. And as I said, we are going to use strong Markov property to solve this problem. Now, to find out these probabilities, we are going to talk about what is called probability degenerating function. What is the probability generating function by definition the probability generating function is this. For any random variable X, if X is a discrete random variable, then I can talk about the probability generating function of X. And by definition this is nothing but for $0 \leq s < 1$. This $\phi(s)$ is basically nothing but $E(s^X)$. I try to find out this expectation for any s in the in the interval 0 and 1 if 0 inclusive 1 excluded I try to find out this expectation. Now, clearly this expectation will be a function of s this particular function is called probability generating function of a random variable. Now, in this case, because I am talking about the conditional probability, I have to talk about the conditional expectation in this case. And if I write this one it basically nothing but $\sum_{n=1}^{\infty} s^n P_1(T^0 = n)$ and that can be written in this case also that it is nothing but s^X into multiplied by $X = x$ and I have to take the sum over the support of the discrete random variable X. So, same thing we are doing here in this case support is basically anything from 1 to infinity. I am taking the summation over this particular thing. I am getting the summation that s^n probability under $X_0 = 1$, $T^0 = n$ that sum I have to take and that was going to give me that probability generating function. What is the benefit of having probability generating function?. There are many benefits of the probability generating function. Probability generating function is a very very useful function in probability and the idea is that probability for discrete random variable if I know the probability generating function, I know the distribution of that particular discrete random variable. How?. The idea is very simple from this part, you see that in this expression, the coefficient of s^n is nothing but the probability that $\{T^0 = n\}$ under $X_0 = 1$. That means that coefficient of s^n is probability of $\{T^0 = n\}$ given $X_0 = 1$. So, that means, if I can calculate this particular function, then if I can collect the coefficient of s^n in this expression of $\phi(s)$, then I know all these probabilities. Not only that, there is many benefits also. This is the first point, second point is that if two discrete random variables X and Y have same probability generating function then X has same distribution as Y. So, 2 at discrete random variables X and Y have same probability generating function, then I can tell that X and Y have same distribution. So, that means that if X has this distribution these probabilities then Y also satisfy the same probabilities. So, if I can show that two random variable have same probability generating function, then that two random variables have same distributions. So, probability of $\{X \leq x\}$ is same as probability of $\{Y \leq x\}$ for all values of x. So, that is basically the idea. Now, another point I should mention here is that, what will happen, if I take s goes to 1 in $\phi(s)$. So, what is the limit s increases to in 1 of $\phi(s)$. Again, you can see it from here. So, that basically mean that as if I assume that I can take the limit inside of the sum, I can write it s increases to 1, s^n this limit is only on this part because the rest of the part does not depend on does not depend on s. So, this part is 1. So, this is nothing but $\sum_{n=1}^{\infty} P_1(T^0 = n)$ which is same as $P_1(T^0 < \infty)$ that I am summing over to $(T^0 = 1), (T^0 = 2), (T^0 = 3)$ so on so, forth. I am summing the probability of these events. That is nothing but $P_1(T^0 < \infty)$. Taking this limit actually from here I can also find out what is the probability that in finite time, I will hit state 0 starting from 1. I can taking once I can find this particular thing that probability generating function of T_0 taking the limit as s goes to 1, I can also find what is the probability of hitting 0 in finite time starting from 1. If this probability is strictly less than 1 then there is a possibility that in finite time I hit 0 starting from 1 that basically means that I will not actually hit 0. And if this quantity is 1 then I can tell that starting from 1 for sure I will hit the time 0. And note that in this particular case, once I hit 0, I will be in 0 forever because this is observing state from 0, I can go to 0 only with probability 1. And finally, one more point I should point out here that is basically that if I take the derivative of ϕ. Suppose I take that derivative with respect to s of ϕ again if I assume that the derivative can be taken inside the summation. So, that turns out to be n equals to 1 to infinity derivative of $s^n P_1(T^0 = n)$. So, this turns out to be $\sum_{n=1}^{\infty} ns^{n-1} P_1(T^0 = n)$. Now, if I take limit s increases to 1 of $\phi^{'}$ which is basically nothing but this $\phi^{'}$ is and again I if I assume that limit can be taken inside the inside the summation sign that actually tells me that it is nothing but $\sum_{n=1}^{\infty} nP_1(T^0 = n)$. Because this quantity actually goes to 1 when s goes to 1. So, now, what is this?. This is nothing but expectation of T^0 starting from 1. So, few things that if I take the derivative and take the limit then I get the expectation and if I take the limit without taking the derivative that will give me what is the probability that in finite time I will hit 0 starting from 1. So, these are the some things we can look into the some characteristics of the Markov chain based on the probability generating function of T^0. And so, our aim here is first to calculate this probability generating function and as

I mentioned that once I get that by collecting the coefficient of this I can actually answer our main question not only that, we can also answer some other questions as a sub product.

Let us move on to the next slide. Our aim now is to calculate $\phi(s)$. To calculate $\phi(s)$, I can proceed as follows; suppose I start from state 2 note that this is basically that starting from state 1 what is the expectation of this quantity or the conditional on a conditional on the fact that I am starting from state 1. But to find out it suppose the chain is starting from state 2 what happens let us see that see, if I start from state 2 and if I have to hit state 0 then I have to move to state 1 there is no other option directly from 2, I cannot go to 0, if I want to go from 2 to 0, I have to go via 1 there is no other option. That is why what I am doing is that suppose this is the time 0 and the state is 1 when I have to go from 2 to 0 as I mentioned, I have to go via 1 that means starting from 2 first I have to hit 1 and suppose for the first time I am hitting 1 at the time T^1 that is our standard notation. So, here the state is 1. From the state 2, I start at time T^0 at the time T^1, I hit state 1 for the first time. So, 1 has not occurred here the first time 1 occurred here. Then it goes on and finally at that time point T^0, I hit the state 0 suppose this difference time is basically denoted by $\tilde{T}^0$, this is the difference between T^0 and T^1. So, that means I can now write $T^0 = T^1 + \tilde{T}^0$. This time is T^0 to T^1 and then T^1 to T^0 which is the time $\tilde{T}^0$. So, I can write in this manner. So, using the strong Markov property at the time T^1 note that T^1 is stopping time we have already proved so, I can use the strong Markov property at that time T^1 and that is what basically I am doing here. And then what I can tell is that is starting from state 2, the conditional distribution of I mean conditioning or the fact that $T^1 < \infty$ that T^0 can be written as $T^1 + \tilde{T}^0$, where $\tilde{T}^0$ is the time taken by the chain to hit 0 after T^1. That means basically starting from state 2 at the time 0, I first hit the state 1 at the time T^1 and then, I finally hit state 0 at the time T^0 and in between time it noted by $\tilde{T}$. And not that when I am writing here I am implicitly assumed that $T^1 < \infty$ if this is not true, then this beaker backup does not make any sense because T^1 goes to infinity. Strong Markov property also tells us that this part of the Markov chain is independent of this part of the Markov chain. The future part that this part this is basically a Markov chain start afresh at the time 1. Not only that, this part is independent of this part. Now, T^1 depends only on this part. And this $\tilde{T}^0$ depends only on this part because $\tilde{T}^0$ as the time that starting from 1 hitting to T_0 and that does not have any effect from here. That this says that $\tilde{T}^0$ and T^1 are independent. Because T^0 depends only on this because this is the stopping time. So, this has to depends only on this and $\tilde{T}^0$ depends only on this part because this is the additional time taken by the chain to hit the state 0 after T^1. That also does not depend does not have any effect of the chain here. This part of the chain has no effect of this part of the chain these two parts are independent. $\tilde{T}^0$ only depend on this part. So, that is why

$\tilde{T}^0$ is independent of T^1. Moreover, a very important thing is that the distribution of $\tilde{T}^0$ is same as starting from 1 hitting to 0. Why?. The reason is that I have already reached one here and strong Markov property says that the Markov chain starts afresh here only thing is that is starting from 1. So, clearly the distribution of T^0 will be same as if I am starting from 1 and I am at hitting to the state 0 that is basically $\tilde{T}^0$. So, that distribution of $\tilde{T}^0$ will be same as that distribution of T^0 when I know $X_0 = 1$. So, this basically to solve these information I can use that two information basically here that 1 is $\tilde{T}^0$ is independent of T^1 under the condition that this quantity is less than infinity. And the distribution of $\tilde{T}^0$ is same as the distribution of T^0 when I am starting from 1 these two things now we are going to use and we are going to look into this. So, I have $E_2(s^{T^0})$ and that very can be written as $E_2(s^{T^1+\tilde{T}^0})$. Now, because these two are independent, I can write it as $E_2(s^{T^1} \cdot s^{\tilde{T}^0})$. Now, I will use the independence that this part is independent of this part and if I we know that if X and Y are independent then expectation of XY is same as expectation of X into expectation of Y. Now, in this case if I take $s^{T^1} = X$ and $s^{\tilde{T}^0} = Y$ then these two random variables are independent. So, loosely speaking I can write it is $E_2(s^{T^1})$ multiplied by now, when I reach T^1 now the Markov chain starts from 1 so, $E_2(s^{\tilde{T}^0})$ that way I can agree to write. And if I write in this particular manner what I get is that you see that whatever is the definition of this thing this particular quantity we have denoted by $\phi(s)$ is basically starting from 1, s^{T^0}. So, this part is clearly $\phi(s)$. Now, what about this part?. Now, look into this Markov chain once again. This $\phi(s)$ is basically nothing but starting from 1 what is the expectation of s^{T^0} that is basically the thing. So, starting from 1 the expectation of s^{T^0} that we have denoted by $\phi(s)$. Now, you see that starting from 1 hitting to the state 0 in n step is basically same as starting from 2 hitting 1 in n step. Why?. The idea behind that is that here the I mean these this particular Markov chain this probability of going right is p and probability of going left is q. So, this p and q is fixed throughout. So, it does not matter whether I am trying to find out the probability distribution of T^0 starting from 1 or whether I am trying to find out probability distribution of it T_{i-1} starting from i whether I am moving from this to this or whether I am moving from this to this that distribution will remain same because that problem for this Markov chain the probability of going towards left is p does not matter from which state I am starting similarly going to as left is q does not matter from which state I am starting. So, that is why this thing is true and this is true for all possible values of n. If this is true, then basically if you just go back then I can write these quantities same as $\sum_{n=1}^{\infty} s^n P_2(T^1 = n)$ and this 1 is nothing but $E_2(s^{T^1})$. So, starting from 2, $E_2(s^{T^1})$ is also $\phi(s)$ and you will see that we got exactly that the same thing here. So, this quantity is also $\phi(s)$ and then this says that this one is same as this thing and then we got that starting from 2, expectation of s^{T^0} which is T^0 is basically hitting time to 0 is same as $\phi^2(s)$ square if $\phi(s)$ is nothing but the expectation of starting from 1, the conditional expectation of s^{T^0}. So, we got this expression, now, we try to use this expression in the further analysis.

Another point into looking into what do we get, that starting from 2 the conditional expectation of s^{T^0} is same as $\phi^2(s)$. Now, look into $\phi(s)$ from a different perspective. By definition $\phi(s)$ is that now what I am doing?. I am using the Markov property not the strong Markov property because the time is fixed now. And what I am doing is that I am here I am starting from 1 and from 1, I can either go to 2 or I can go to 0 in 1 step. I am conditioning with respect to the first transition, transition at the time 1. So, from 1, I can move to 2 and that probability is p so, I can write this 1 is p times starting from 1 next transition is 2 then conditional expectation of s^{T^0} plus from 1, I can move to 0. So, given $X_0 = 1$, $X_1 = 0$, I try to find out the conditional expectation of s^{T^0} multiplied by the probability which is q. Now, look into this expectation the p is as it is here. Now, look into this expectation. This expectation means that $X_0 = 1$, $X_1 = 2$ and then under this condition I try to find out what is the expectation of s^{T^0}. Now, when I move to 2 that means, that Markov property now tells me that at the time point 1 the Markov chain is starting afresh and at the time point 1 the state is 2. So, this expectation can be now written as expectation that starting from 2 it is nothing but s^{1+T^0}. I reached 1, 2 now, from 2, I have to hit state 0, but 1 step I have already taken. So, this 1 has come here given $X_0 = 2$ so, Markov chain has a start from here afresh just forget about this 0 forget about 0 time point and as if time point started from here then rescale the time point from $0, 1, \cdots$. So, this expectation turns out to be that and then s^1 comes out here that s comes out here the rest of the part is written here. Now, when $X_1 = 0$, that means in 1 step it is 0. So, that means $T^0 = 1$ in this case that means that it will be nothing but a starting from 1, that expectation is s because there is no randomness there the expectation is s. So, finally we get this and this basically tells me because beforehand we got that that $E_2(s^{T^0}) = \phi^2(s)$, I write that $\phi(s)$ is same as $ps\phi^2(s) + qs$. So, that means ϕ satisfies this particular condition just instead of writing $\phi(s)$. Now, I am just writing ϕ if I do this is turns out to be this. This is $ps\phi^2 + qs - \phi = 0$. Now, let us solve for ϕ if I solve for ϕ using because this is a quadratic equation in ϕ if I solve I get this particular solution note that we have 2 possibilities here that positive sign or negative sign. Now, you see that $\phi \leq 1$. Why?. Because ϕ is nothing but starting from 1 s^{T^0} for $0 \leq s < 1$ that is what basically we have taken. Now, if $0 \leq s < 1$ and T^0 only takes the positive integer values. So, that s^{T^0} always lies between 0 and 1. So, clearly when I take the expectation of that that also lies between 0 and 1. So, I can also put 0 in this side. Now if I take the positive sign here. Then what happens?. What happens $\phi(s)$ is same as $\frac{1+\sqrt{1-4pqs^2}}{2ps}$ or $\frac{1-\sqrt{1-4pqs^2}}{2ps}$. Now if I take s

goes towards 0 what happens that see that numerator actually converges to 2 because this quantity goes to 0 this is 1 plus 1 so, converges to 2, but the denominator converges to 0. So, this whole quantity converges to infinity and that cannot happen because ϕ has to lies between 0 and 1. So, the positive sign I cannot take so, only possibility is that I have to take the negative sign here and finally I have $\phi(s) = \frac{1-\sqrt{1-4pqs^2}}{2ps}$. And if now you take this quantity s decreases to 0 here you will get $\phi(s)$ converges to 0 you will get that you can check this one. You have to use the L'Hospital rule that is it. That means I have to take the negative sign here and final expression of $\phi(s)$ given by this for s lies between 0 and 1.

Now, if I just expand this particular square root in terms of power series, I have this particular expression. And now, if you just collect the coefficient of s coefficient of s squared so, on so, forth, you will get it turns out to be ps, then s^2 is not there. So, $0 \times s^2 + pqs^2 + pq^2sq$, then you will have $0 \times s^4$ and it goes on to that way. So, now, if I collect that the coefficient of s will be nothing but probability that $T^0 = 1$ under the condition I am starting from 1. So, that clearly says that $P_1(T^0 = 1) = p$ because it is the coefficient of s here that is p then $P_1(T^0 = 2) = 0$, because the coefficient of s^2 is 0, then $P_1(T^0 = 3) = pq^2$, the coefficient of s to the power 3 is pq^2. Then $P_1(T^0 = 4) = 0$ and so, on it is it is going on. So, this way we can find out the expression of that all the probabilities of the form $P_1(T^0 = n)$ for different values of n, $n = 1, 2, \cdots$, these quantities now we know. And so, using moment generating function and then using strong Markov property as well as Markov property, we are actually able to solve this problem. Now, another important thing we can look into this probability and as I pointed out as s goes to 1, ϕ converges to this probability. So, let us try to find out this limit and when I try to find out this limit that is nothing but this s is basically 1, I can put this turns out to be that now, if you put that $p + q = 1$ if you use this relation, then this particular quantity can be written in this form I again left is an exercise so, you just remove this q with $1 - p$ plus this q with $1 - p$ here you will get this one. Just check it you will get this one quantity. What does this tells is basically tells that if $p < q$ which is basically equivalent to say that $p \leq \frac{1}{2}$ that means $p \leq q$ starting from 1 it is sure that we will hit 0. So, if a population follows this kind of evaluation of the Markov chain, then if the population start starts from 1 it is sure that population will extend if $p < \frac{1}{2}$ notice that p is basically nothing but probability of birth. And if $p > \frac{1}{2}$ then there is a positive probability in the finite time will hit 0. But there is also a positive probability that starting from 1 the chain will not hit 0. So, now, there are two possibilities are possible. Those both the things are possible that in finite time I may hit in finite time I may not hit and both the cases we have some positive probabilities.

Now, another point we have discussed that again we can find out this expectation by taking the derivative first and then taking the limit and so, when we are looking into this one, you will note that I have to only look into this the reason being behind is that in the second case in this case, there is a positive probability $P_1(T^0 = \infty) = 1 - \frac{p}{q}$ if $q < p$ that basically mean that $p > \frac{1}{2}$. So, there is a positive probability of having this thing. So, that basically means that under the condition that $p < \frac{1}{2}$ under this condition if I try to find out what is the expectation of T^0 starting from 1 it has to be infinity because infinity times some positive probability will be there. So, that has to be infinity there is no concern. Only the interesting part is to look into this part. And so, we are only for try to find out the expectation under $p < \frac{1}{2}$ again which is equivalent to $p \leq \frac{1}{2}$. And so, what I have to do?. I have to first find out what is the $\phi^{'}(s)$ and if I can find out this I can take the limit s increases to 1 if I will take the limit. So, to do this one I start with this, this is nothing but same as what we got here. This is nothing but same as this that I have written and now I take the differentiation on both the sides of this equation with respect to s. So, if I do this, this turns out to be $2\phi(s)$ this part I keep as it is then the derivative of this function, which is basically $2 \times \phi\phi^{'}$ because ϕ is a function of s also then I take $p\phi^2$ is there and derivative of s that give me this minus derivative of ϕ and finally derivative of qs which is q. And which has to be equals to 1. Now, from here I can solve for $\phi\phi^{'}$ and it can be shown that $\phi^{'}$ is given by this expression. Now, take the limit as s increases to 1. Now, we just remembered that limit s increases to 1, $\phi(s)$ this quantity is 1 if $p \leq \frac{1}{2}$. This we have done here. This is 1. So, I just $\phi^2 = 1$, I replaced and $p + q = 1$ then this $\phi = 1$ and s is also 1. So, it turns out to be $\frac{1}{1-2p}$, which can be written in this particular form also. That basically, this is nothing but $1p$ you write in terms of $1 - q$. So, you get this particular expression. That means, we get that this expectation of T^0 when I am starting from 1 is same as $\frac{1}{q-p}$ for $p \leq q$. Let us look into this what does it mean?. This means that both the expression are fine. So, this means that in this case, what happened $p \leq \frac{1}{2}$. So, if p is close to 0, that means this quantity will be small. This expectation will be small, if p is close to 0. And that basically means the difference between q and p is quite high. On the other hand, if p is close to $\frac{1}{2}$, then this expectation will be large. And in fact, when $p = \frac{1}{2}$, this expectation is infinity. So, that means if the difference between p and q increases, that expectation actually decreases. So, that means if p is close to 0 that means, if the birth probability is very small, then that on an average, very quickly we will hit 0 starting from 1. And if the birth-probability is close to $\frac{1}{2}$, then obviously, we will hit 0. But the mean time to hit 0 starting from 1 is increasing. And if p is equals to $\frac{1}{2}$ or greater than $\frac{1}{2}$, then starting from 1, the mean time to hit 0 is actually infinity. So, with this interpretation, I stop here. Thank you for listening.