Discrete-Time Markov Chains and Poisson Processes

Professor Ayon Ganguly

Department of Mathematics

Indian Institute of Technology, Guwahati

Module: Hitting Times and Strong Markov Property

Lecture 11

Hitting Time III

Welcome to the 11^{th} lecture of this course, Discrete-Time Markov chain and poison processes. Recall that in the last couple of lectures in basically three lectures, we have talked about that hitting probability. Now, in this current lecture will go into talk about the mean time to hit a collection of states.

Just recall again that T^A we have defined to be $\inf\{n \geq 0 : X_n \in A\}$, where $A \subseteq S$. Then basically in the previous three lectures we have talked about hitting probability which we denoted by h_i, which is $P_i(T^A < \infty)$ that we try to find out. Now, in this lecture, we are going to talk about the expectation, the expectation of T^A starting from i. So, this is nothing but $E(T^A|X_0 = i)$. We will talk about that, and again this one is also very very important quantity in practice, because the meaning of the expectation is basically mean, it is the mean time or expected time to hit the set A. So, starting from either expected or mean time taken by the Markov chain to hit A is given by this quantity, and the notation we are going to use k in this case. In case of the probability, we have used h notation, in case of expectation we will use k notation. Now, let us first discuss what that signify. In the example, we discussed in the previous lecture that we talk about the extinct probability. That what is the probability that starting with i individual the population finally extinct. That probability talked about. So, in this case a natural question is that, and another point I should mention here is that, we have pointed out that when $\sum \gamma_i = \infty$, the population is sure to extinct. Now, natural question comes to the mind is that well, what is if suppose that a particular population that $\sum \gamma_i$ is actually equal to infinity. So, that particular population will extinct, no matter from where it starts. Now, the question is that, on average how much time it will going to take to become extinct. That question is a very natural question. And in that kind of scenario, I have to find out the expectation of T^A and I start from i. So, basically nothing but the conditional expectation that starting from i, what is the average time, or mean time, or expected time to hit the set A. And in this case, we pointed out that T^A can take value $0, 1, 2, \cdots ,$. And finally, it can also take value infinity if the set is $\emptyset$. T^A can take these values. Now, naturally that means that this when I try to find out this expectation, the formula of the expectation is nothing but value multiplied by probability because it is a kind of a discrete random variable. So, that means I can write this one summation, $0 \cdot P_i(T^A = 0) + 1 \cdot P(T^A = 1) + \cdots + \infty \cdot P_i(T^A = \infty)$ so, the when I talk about the $E_i(T^A)$, I can write it $0 \cdot P_i(T^A = 0) + 1 \cdot P(T^A = 1) + \cdots + \infty \cdot P_i(T^A = \infty)$. So, this way I can write. Now, this term will not going to contribute anything, because I have 0 multiplied by, so, this term is 0. So, I have not written this term here, I have started writing from this 1 multiplied by corresponding probability, then 2 multiplied by the probability that $T^A = 2$, and so on so forth that part is written here, that A for n greater than equals to 1, $\sum_{n \geq 1} nP_i(T^A = n)$. So, this part has written here. Now, the rest of the part is that infinity multiplied by $\infty \cdot P_i(T^A = \infty)$. So, this expectation can be calculated using this formula. Now, one thing is for sure from here that if there is a positive probability for this, if this quantity is strictly greater than 0, then this expectation has to be infinity, because one infinity is there, it will make everything infinity. But in case so, if this probability is 0, this implies that $k_i^A = \infty$. If this probability is equals to 0, there are two possibilities in this case, one possibility is that $k_i^A < \infty$, and other possibilities that $k_i^A = \infty$. So, when this probability is 0, I have both the possibilities that, k_i^A can be finite, k_i^A can be infinite, that depends on whatever, what is this probability?. What is this summation give me?. And when this one has some positive probability, then definitely that $k_i^A = \infty$. So, that means basically we have the scenario like that k_i^A has to be positive because T^A is always positive. So, expectation of T^A is always positive, but there are two possibilities now, one is that k_i^A can be finite, k_i^A can be infinite. If I have a positive probability that $T^A = \infty$, then expectation has to be infinite there is no other option. But if T^A is always finite, with probability 1, then there we may have two possibilities, that expectation may be finite, expectation may be infinite. Both the possibilities is possible, that completely depends on what kind of probability it assigns to define values of n, that $T^A = n$, and what are the probability that $P_i(T^A = n)$. It completely depends on that. So, this is also called the expected hitting time to the state A, and informally we will write this way. So, they are almost same as before.

Let us move into example, to see how we can calculate this kind of expected hitting time. This is one of the example that we have already discussed. And that in this case basically we have three states from 1, I can only go to 1 in one step. So, from 1, I will be in 1 always, from 4 also in one step I can go into 4 itself. So, from 4, the chain will remain in 4 forever. From 2, there is two possibilities, in one step I can go to 3, or in one step I can go to 4, both of have that half-half probabilities. And from 3 also the scenario are same, that from 3, I can go to 4, or go to 2 in one step and that one step transition probabilities from 3 are half and half respectively. With that now, my question is that, starting from 2, what is the mean time taken by the chain to get observed into 1 or 4. In this case what we try to find out, if I take $A = \{1, 4\}$, we try to find out $K_2^A = E_2(T^A)$. That is basically we try to find out so, that is basically nothing but $E_2(T^A|X_0 = 2)$. So, to do this one basically we define this one k_i. So basically, this power thing I am not writing to make this notation written simple. k_i which is nothing but the expected hitting time to the set A. Before that, just few points I should mention in this case. First of all, that in this case, I have three communicating classes and it is very easy to see that 1 does not communicate with 2, because from 1, I cannot go to 2. So, 1 and 2 does not communicate. And 1 does not communicate with 3, 1 does not communicate with 4, that for sure. So, 1 is a separate class. Similarly, 4 is a separate class 4 does not communicate, once I am in 4, I cannot go to any other states. So, 1 and 4 are two separate communicating classes. Regarding 2 and 3 they communicate, because from 2 in one step there is a positive probability go to 3. Similarly, from 3 there is a positive probability to go to 2 in one step. So, of course, these 2 and 3 communicate. So, I have 3 communicating classes in this case, among them this one is closed, this one is also closed, but this one is not closed. Why this one is closed?. Because once I am in 1, I cannot go out of 1. Onec I am in 4, I cannot go out of 4. So, these two are close classes, whereas, these 2 and 3 are not close classes, because from 2 there is a possibility to go to 1. And once I go 1 that means, I am going outside this particular communicating class. So, that means this is not a close class. And so, that means that when I am talking about 1 and 4, I can talk about the absorption because once I go there, it will, the chain will remain in 1 or 4 forever, it will not come out from that. With that let us proceed now. We define k_i and in this case what is K_1?. K_1 is nothing but the expected time of hitting A starting from $X_0 = 1$ i.e., $k_1 = E(T^A|X_0 = 1)$. Now, if you are in 1 then basically T^A equals to 0. if $X_0 = 1$, or $X_0 = 4$ that is for sure, because, if I am already in 1 or 4 that means I am already in A. So, $T^A = 0$, and if T^A is identically equals to 0, that means that K_1 and K_4 has to be 0. And intuitively it also makes sense because if I am here that means in zero time, I will be in 1. So, the expectation is also 0. Similarly, the same thing happens if I am here also. So, that is why that K_1 and K_4 are 0 in this case.

Let us proceed and what we have till now is that we have already understood what is the value of K_1 which is 0, what is the value of K_4 which is also 0. Now, we have another two states, which is state 2 and state 3. We are looking for whether we can able to find out what is the value of K_2, or what is the value of K_3?. Of course, in this case I have four states. So, I have four k values, among of these four k values, 2 we already find out that K_1 has to be 0, K_4 has to be 0. Now, we are looking for what is the value of K_2 and what is the value of K_3?. How we can find out let us see and the idea here again nothing but use of the system of linear equations. What we do look into this K_2. The K_2 by definition is nothing but expectation of $E(T^A|X_0 = 2)$. Starting from 2, what is the expected time of hitting either state 1 or state 4?. Just A is basically our state 1 or state 4. Starting from 2, what is the expected time of hitting state 1 or state 4?. Now, what we are going to do here is that, we are going to condition with respect to next transition. So, we start with two and in the next transition, I can move from 2 to 1, or 2 to 3. From 2 in one step I can move to 1, or I can move to 3. So, I am going to include that particular thing here. So, I write it $E(T^A|X_0 = 2)$, next transition $X_1 = 1$, then multiplied by $P(X_1 = 1|X_0 = 2)$, this is one case, and another case would be $E(T^A|X_0 = 2, X_1 = 3)P(X_1 = 3|X_0 = 2)$, i.e., $k_2 = E(T^A|X_0 = 2|X_0 = 2, x_1 = 1)P(X_1 = 1|X_0 = 2) + E(T^A|X_0 = 2, X_1 = 3)P(X_1 = 3|X_0 2)$. So, from 2 in this particular case I can move to any of the two states either 1 or 3. From 2, first step if I take to 1, I make the condition with respect to this, and then multiplied by the corresponding conditional probability. Similarly, the conditional expectation that initially I am starting from 2, next transition I am taken to 3. Under that that I try to find out what is the expected value of T^A and then I try to find out that should be multiplied by what is the probability that starting from 2, I go to 3 in one step. And these two probabilities are known to us that both the probabilities are half. Now, let us talk about this expectation. Just see that this expectation, I can write as $1 + E(T^A|X_0 = 1)$. The first part of the expectation I can write in this particular form, why?. The idea is again the same, that from 2 if I move to 1, then Markov chain can forget the previous and it can start from time 1 afresh, only thing is that the initial distribution will be δ_1. So, using the Markov property that at time 0, I start from state 2 at time 1, I am moving to state 1. So, using the Markov property at the time point 1, I can completely forget this one, and I can start from 1. I can start from one Markov chain actually start afresh from the state 1. So, I can write this one that under the condition that $X_0 = 1$ the initial state is 1, what is the $E(T^A|X_0 = 1) + 1$. Why this one is coming?. Because from the state to 2 when I am moving to state 1, I have already taken a one step, so that one step will come into the expectation. So, basically, when I am taking the step, this is the expected value starting from 1, what is the expected value of hitting the states in A, plus I have already taken one step to move from 2 to 1, so, that 1 has come there. So, these two conditional probabilities can be written in this form it is 1 plus conditional expectation of T_A, given $X_0 = 1$ multiplied by this conditional probability, and this conditional probabilities given to me to be half, that in one step, I can move from 2 to 1 with probability half. Similarly, the next one will be 1 plus expected value of T_A, under the condition $X_0 = 3$ multiplied by half. Because, again this probability is half, and this one again starting from starting at the time 0, I am in state 2, at the time 1, I am in state 1. So, again, Markov chain can start refresh from the time 1, and I can completely forget about that, only thing is that, it will basically start from the initial state 3. So, that is why this expectation come into the picture, plus 1 because I have already taken one step to move from 2 to 3. So, I get this one. Now, notice that this particular thing is nothing but K_1, and this particular thing is nothing but K_3. So, I can write now, K_2 is equals to half time 1 plus K_1, plus half times 1 plus K_3. So, I have this linear system, if I just simplify it, it turns out to be $K_2 = 1 + \frac{1}{2} \times k_1 + \frac{1}{2} \times k_3$. Similarly, K_3 can be written from 3, I can move to 2 or 4 in one step. So, when I am writing K_3, it will be the probability of moving from 3 to 2 which is half, plus multiplied by 1 because this one step I have taken plus K_2. 1 plus K_2, and plus half into 1 plus K_4. Because the probability of moving from 3 to 4 in one step is half. And that should be multiplied with 1 plus K_4, because one step I can take from 3 to 4, that one step comes here, plus K_4 is the expected value of T^A given $X_0 = 4$. And so, we now have two system of linear equations. And we can solve this. Basically, we have now a system of linear equation, which involves two equations and 4 unknowns, K_1, K_2, K_3, and K_4. Among of these 4 unknowns, two of them we have already find out. So, if I plug in these two values K_1 and K_4 here, I will finally have a system of linear equation, which having two equations involving two unknowns. We have to solve that, and once we can solve it, I am done. In this case solving is simple, because we got $K_3 = 1$, if I just replace the K_3, so first thing is that this term will not contribute this term is 0, because $K_1 = 0$, this time will not contribute because $K_4 = 0$. So, what I get is $K_3 = 1 + \frac{1}{2}k_2$, if I replace this K_3 with $1 + \frac{1}{2}k_2$, what I get is K_2 is equals to this quantity. And now I have a equation in K_2 only, I can easily solve, and if we solve, we will get that K_2 turns out to be 2. So, this example tells us that on an average, in two unit of time, starting from 2, I will hit the state 1, or state 4, the chain will be absorbed in the state 1 or in state 4. On an average in two units of time if the chain starts from state 2. That is the idea of solving this kind of problem. And the basic idea is almost same as that of the finding out the hitting probability starting from some given state, but the only crucial point keep in mind this 1, this 1 plays a role in case of the expectation, but when we talk about the probability, this 1 terms do not come there. So, keep this thing in mind sometimes we forget to add 1 in case of the expectation sometimes we add extra 1 in case of the probability, and that is not correct. So, keep in mind when I am talking about the hitting probability calculation of hitting probability will not involve 1, but calculation of the expected time of hitting will involve this 1 because one step when I am taking that need to be added to the expected value, but that need not to be added to the probability.

Now, let us move and put this idea in terms of a theorem, and the theorem again looks almost same as that of the hitting probability case only change is that again that 1. Thus theorem goes like that the vector of mean hitting times k^A which is basically nothing but collection of all $k_i^{'}$ is the minimum non-negative solution of the system of linear equations this. And this system of equations written exactly that way if I am already in A, then the expected time is 0, if I am outside A, then the expected time can be written in this form. So, what I can write that K_i^A, I am talking about $i \notin A$ in this case. That I can write as well, if I can go from i to some j, then in one step, then this p_{ij} will come multiplied by, what is the time to go from j to A. What is the expected time from going from 1 to A, $K_j^A + 1$ that will come. Now, this j can be anything in the state space, so, that sum will come now, because we know that $\sum_{j\in S} p_{ij} = 1$ so, that means that I can write this one as $\sum_{j\in S} p_{ij}k_j^A + \sum_{j\in S} p_{ij}$. Now, this term is basically 1 so, it turns out to be 1 plus $\sum_{j\in S} p_{ij}k_j^A$. Now, you see that I have also listed here one thing and that is basically nothing but if this $j \in S$. Now, this S can be written in terms of two sets. one is A and one is a A^c, that means, this sum can be written as $1 + \sum_{j\in A} p_{ij}k_j^A + \sum_{j\notin A} p_{ij}k_j^A$. Now, when $j \in A$ this is basically 0 that is given here. If $i \in A$, $K_{ij} = 0$ similarly, if $j \in A$, $K_{ij} = 0$, so, this term will not contribute anything. So, finally, I have K_i^A equals to 1, plus the sum that is what basically written here. Again, keep in mind that when I am solving it finally, I have to look into minimal nonnegative solution. And that notion of the minimality is same as before, that is basically if $x_i, i \in S$ another non-negative solution in this case, then x_i has to be greater than equals to K_i^A, that is the definition of the minimality it in this case. So, that is the idea put in terms of the theorem, and this minimal non-negative solution is important when we have infinite number of solution of this particular system of equations.

Let us discuss another example, where we can use this theorem to find out the required quantity we are interested in. So, that in this case, it goes like that it is a standard snake ladder game, only thing is that in this case it is a 9 cross 9 square, we start from here and our aim is to reach 9 it is the finishing point. We start from 1 our aim is to reach 9, and the game goes like that, I am tossing a fair coin, if head comes then we will advance one step. So, if head comes, we will advance one step. So, when I am when in 1. if I toss head come, I go to 2. If tail comes, I advanced two step. So, if I am in 1, if tail comes, I go to 3. So, from 1, I toss a coin, if head comes, I go 2, if tail comes, I go to 3, then the standard procedure applies that if I am in a foot of a ladder, I will I will climb out to the head of the ladder. And similarly, if I am land in a head of a snake, I will slide down to that tail of that snake. So, that way the standard game will continue only thing is that I am tossing a coin and based on the head or tail, I am taking one step forward or two steps forward.

That means that if I now try to write this in terms of the transition probability matrix of a Markov chain, how we can write, look into that. First point is that, suppose I am in 1, I toss the coin I got a head. So, I start at 1, I got a head I move to 2, but it is true I will not stop I will climb the ladder, I will actually finally go to 7. Similarly, if tail is come, I will move to 3, but I will not stop at 3, I will climb the ladder and I will go to 5. That means that 2 and 3, these states I cannot stop anytime. So, we will not stop in the foot of the ladder. So, that means this 2 and this 3 these two steps I do not need to consider. Similarly, at the head of the snake, I will not stop if I land in the head of this snake, I have to slide down. So, similarly this 6 and this 8, I do not need to consider in the states. So, in the states I need to consider 1 then 4 then 5, then 7 and then 9. So, I have 1, 2, 3, 4, 5 states in this case and the states are 1, 4, 5, 7 and 9 where actually I can stay. So, if I write this one now, I let us try to find out the probabilities. So, from 1 can I move to 1?. No. If head or tail comes whatever comes I have to move out of 1, there is no point that I will be in 1. So, from 1 where I can go?. If head comes, I will go to 7. If tail comes, I will go to 5. So, from 1 only possibility is to go to 7, or go to 5. And the probability of going to 7 is basically nothing but probability of getting head, and because I am tossing a fair coin the probability of getting head is half. Similarly, probability of moving to 5 in one step starting from 1 is same as probability of getting a tail, and the probability of getting a tail is half. So that is why these two quantities are half, that from 1, I can go to 5 with $\frac{1}{2}$ probability in one step, from one I can go to 7 with $\frac{1}{2}$ probability in one step, the rest of the probabilities in this row has to be 0, because the row sum has to be 1. I am sorry, that will not be 2, that will be 4. Now from 4 where I can go?. Now, I am in 4. From 4 if head appears start 4 start at 4, if head appears, I move to 5 and when I move to 5, I have to stay at 5, if tail appears, I will move to 6 and when I go to 6, then I have to slide down. Finally, I have to come to 1. So, from 4, I can go only to 5 or to 1 and in this case, the probability of going to 1 in one step is $\frac{1}{2}$ similarly going to 5 is $\frac{1}{2}$. Similarly, I can write all of them, but once I reach 9, I will be 9 forever I will not that play in so means I will be 9 forever. So, from 9, I can only go to 9, so I have 1 here. So, in this way the matrix can be written.

And once the matrix can be written, I can again write the system of linear equations. And here also the idea same see that when I am talking about K_1 that from 1, I can only move to 5 and move to 7. So that means it will be 1 plus probably p_{15} times K_5 plus p_{17} times K_7 i.e., $k_1 = 1 + \frac{1}{2}k_5 + \frac{1}{2}k_7$. So, I can write this similarly, I can write complete equation here only thing K_9 will be 0, because 9 in this case A is my 9, that is we try to find out hitting time to 9 because I am trying to find out on an average how much time I need to complete the game. So, that is why I try to find out nothing but K_1 that is basically my aim and I define that K_i like this and I can write this way I have to solve this system of linear equation and if you solve it, you will get this. Now, the question is that how can you solve, the solving is very simple. In this case, just little bit of algebra you have to do because in this case, you will see that $K_9 = 0$, so that actually do not contribute. So, now this is a equation 7 and 4. Now, if you plug this 7 here, this expression here, it will turns out to be equation in 1, 5 and 4. Now these three will be equation in K_1, K_5, and K_4. So, 4, 3 equations I have and 3 unknowns I have, I can easily look for the solution. So, the substitution method we will help you in solving and if you solve finally, this result you will get. One point I should specify here, look in this case, I have a unique solution. So, in this particular case, that minimality I do not need to look for because I have only one solution. So automatically this is minimal. And if you solve it, you will find out this is a non-negative solution. There is no question about that all the $K_i^{\text{'s}}$ are greater than or equal to 0. That is good. So finally, this is the solution. We are done. What we get is that on an average, about 6.625 tosses I have to do to complete the game. With that, I stop. Thank you for listening.