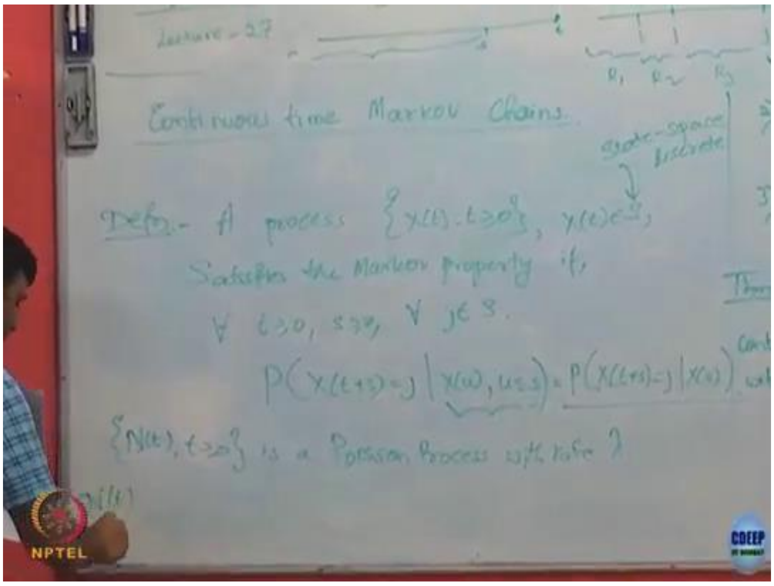
**Introduction to Stochastic Processes**
**Professor Manjesh Hanawal**
**Industrial Engineering & Operations Research**
**Indian Institute of Technology, Bombay**
**Lecture 47**
**Introduction to Continuous Time Markov Chains**

Now, we want to move to continuous time Markov chains. So, so far, we studied discrete time Markov chains when you studied discrete time, time was discrete, like we allowed, n equals to 1, n equals 2 like this, we only look for the change in the Markov chain at specific discrete times.

And we had allowed ourselves to look only into some countable set of states and that is why we called it a chain because we have certain states going from one state to another states like it is kind of forming a chain there. So, that is why we studied so far discrete time Markov chains.

But now I may do not want to look at the things at particular time instances, that maybe I may want to look at my process continuously. And now, I want to now study, what is the Markov property on such a continuous time. We have defined what is a Markov property, what is a Markov chain when my time is discrete, now we want to go and see how we are going to deal with the case where my time is not in discretely indexed, but it is a continuous random variable.

(Refer Slide Time: 1:44)




We already know what is a continuous random process, continuous random process is nothing but a sequence of random variables indexed continuously at every time. Now, we want to understand, what is continuous time Markov chain. Okay, so definition, so as usual we are going to always assume that this state space S is discrete. So, what it is saying? The probability that you are interested in knowing your process takes value state j at time t plus s given your process all the way up to s, what is this is telling?

So, this condition here tells that you have been told what is your process till time s and now we are interested in knowing okay what is the probability that in further t seconds, further t more time that is s plus t that you are going to take state j, what is this probability? Okay, so, like this your process is taking value over continuously this is s and let us say this is t. So, your process is taking continuously value and you have been given what is your process all the way up to time t.

And now from this you want to know what is the value your process is going to take at time s plus t here. This process, this probability only depends on the value your process has taken at time s I really did not care about all this before time s if you are going to look something beyond s I only note to know what has happened at time s and this is what this is telling this is same as what we did it in a Markov chain.

So, there we have said that okay the future depends only on the current state and I do not care about my past, here this is also saying the same thing okay past I do not care just tell me what is going to be happening at the current, but now it is just extension to the continuous time version. Here t and s both are continuous valued.

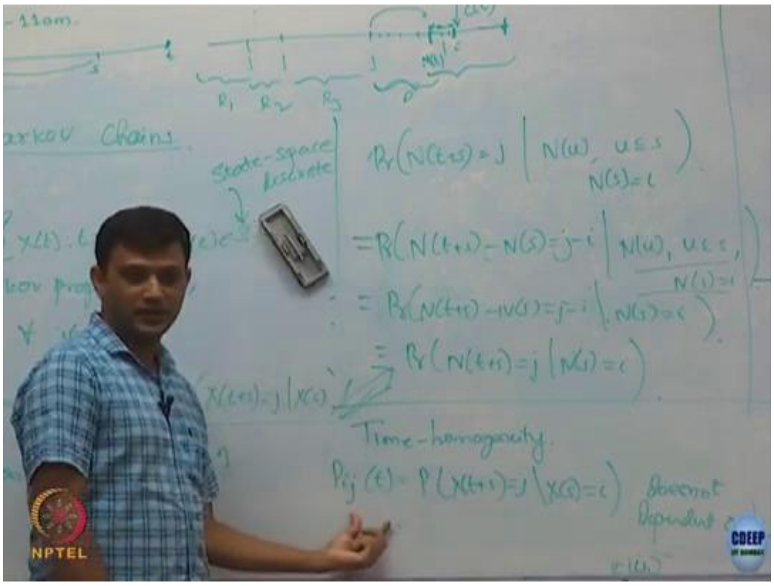
So, we have already seen some processes continuously index what is such a one process? Okay, renewal process where what are Mt, Mt was continuously evolving and what would you say that Mt is distributed how? We just said that it has, it is distributed Poisson with rate lambda t like what was, so what would you say? Poisson process but for general process (())(7:03) say anything like if you have a renewal process ui with certain mean value what is Mt is going to look like.

So, fine so, we had specifically mentioned about my Poisson process by the way is Poisson process is a continuous time process, how did you define a Poisson process? Poisson process is a counting process we said and it going to do the number of arrivals or number of some events happening till time t.

So, there also it was a sequence of continuously index random variable. Okay fine, let us say now if this Poisson process is a CTMC. Okay, Poisson process is a continuous time random process, we want to see whether it is a, it satisfies this definition. Let us say Nt is a Poisson process with rate lambda. So, we have, we know that at any time t this Nt is what, Nt is the number of arrivals, or lives that has occurred in the interval 0 to t.

So, all these values are integer valued numbers N of t. This is a counting process so every time it is going to count on many times things have happened in the interval 0 to t. So, N of t is going to take value 0, 1, 2, 3 like that.

(Refer Slide Time: 9:08)




Now, whether, if I am going to ask the question okay number of renewals, number of arrivals till time t plus s given till this, does it depend only on N u N of s not all u less or equals to s and therefore specifically let us assume that N of s is equals to some i. So, when I said this, N of u, given, given N of u, u less or equals to s means I have described you my process all the way up to s.

When I said, I have described my process still s I have told you, what are the states that has been taken by this process till time s, so let us say at time s, it has taken a particular state i. And now I want to see that whether this probability it only depends on what is the state taken as s and does not depend on what has happened before it. Okay, is that true?

So, let us say, anyway that is true, right? Like, because what has happened till s and I am just counting, so my counting accumulates, after that point I am only going to see what more has come. So, let us write that in a formal way. This is intuitive idea, but we have to bring in the definition of my Poisson process to argue this.

So, instead of asking this question Nt plus s going to j with N of s equals to i I am asking that okay u the number of increment that has happened between the interval t plus s to s is j minus i. So, it is the same thing, like instead of going directly from i to j at time t plus 1 it is I am asking that between s and t plus s interval j minus i renewals are happened and with this we had done. What is the property of a Poisson process?

Poisson process has independent increment. So, what is this is saying this so far this is saying that okay this is like the number of arrivals in the interval 0 to s is i and this is a number of arrivals that has happened between s to t plus s. So, the interval here is s to t plus 1 and here interval is 0 to s and these are independent event, this is nothing but probability equals to j minus i and here i so this is I have assumed that N of s is already knew to, known to me.

So, you already see that you have eliminated dependency of this probability on all the things that has happened before time s. And this you can just know go and do the reverse process and this is saying okay, it is equal j one, yeah this is still Ns equals to i and then we are done. So, by definition of the Poisson process, you already have this condition.

So, what is the crucial point you use? You wanted to make sure that the increments in disjoint intervals are independent. If they are not, you do not know it. So, okay, let us say something has happened in this interval, and the amount of the increment that is going to happen in the next some intervals is going to depend and what has happened in the previous interval, then they are not independent.

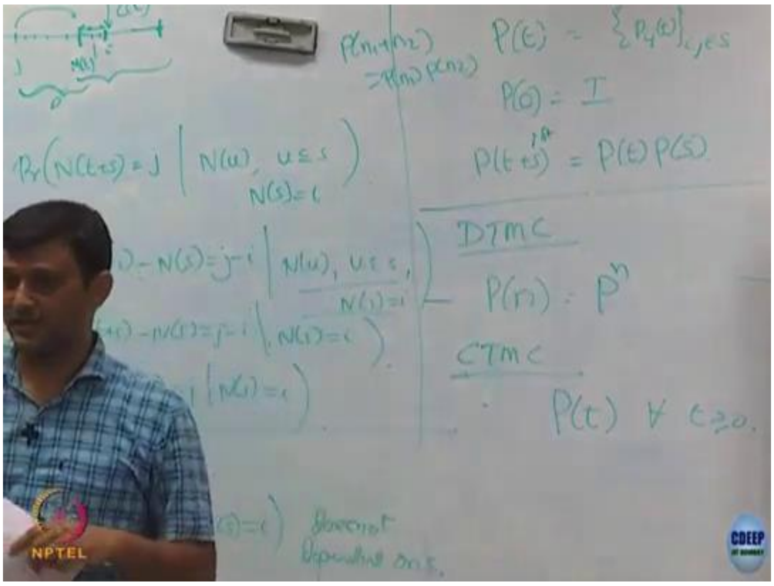
Then in that case, I actually needed to know what has happened before. Before I can say what is the probability that this is going to happen in this intervals. So, there is no Markov property there. But in the Poisson process, given that like this in the increments in disjoint intervals are independent this allows us to make sure that I do not need to know before what happened to say what is happening in this current interval.

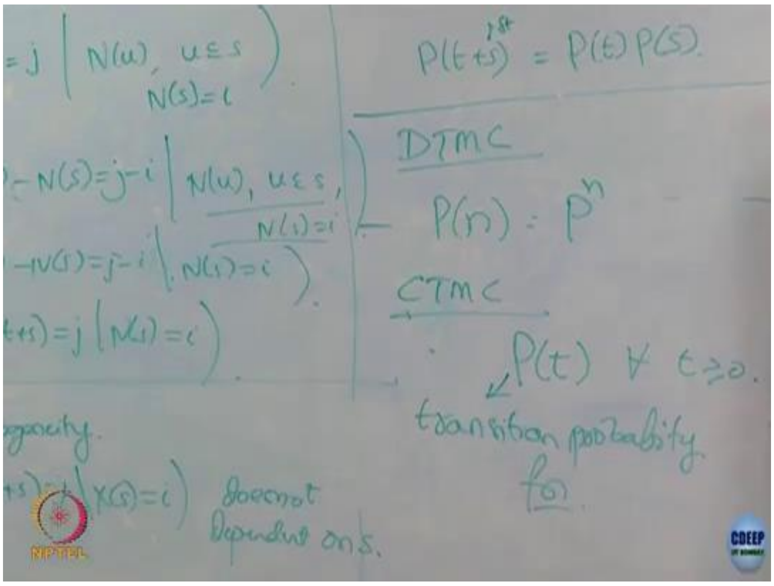
So, that is fine. Now, so now most of the things now we are going to do next is very similar to what we did for the DTMC like what we and we started with DTMC, we defined. What is my transition probability matrix? How I define, what is my one step transition probability matrix? Then we said how I am going to write my joint final time distributions in terms of my this transition probability matrix?

So, let us see how we can do all the things for my continuous time Markov chains. So, first of all, we are going to, for simplicity we are going to assume time homogeneity like the other like in the DTMC. Then going to assume that so, you are asking the question, time s were in state i and what is the probability that you go to state j at time t plus s.

So, here you are looking for interval s to s plus t. So, its length is simply t I am going to assume that this probabilities only depend on that length of the interval, not exactly from the point where you started. So, it does not matter where you are going to start, what matter is in this probability is, what is the length of that duration, this is we are just going to call it time homogeneity, it does not matter which s you are going to put as long as the length of that interval is same, I am going to it is going to be governed by the same probability.

(Refer Slide Time: 16:24)




Now, we are going to denote by this P of t equals to Pij t. So, all this collection of probabilities for every state i and j for a given t I am going to denote it by P of t. So, is my P of t is a stochastic matrix? My P of, capital P of t, it is going to stochastic matrix all the row is going to sum to one. And we are going to like in the earlier case we are going to define P of 0 to be I identity matrix.

At time t, we are going to assume that in 0 times, you are going to be just falling back in your own state not you are going to move to any other state. So, that is why only diagonals there will be 1, other are going to be 0. And then this we are going to skip but we can verify that if I have transition probability matrix for time t plus s this can be written as P of t and P of s.

So, this is a similar way what we did for the DTMC where we could split my transition probability matrix for a sum of two discrete times into that of individual times. So, we are just doing the same thing here so it follows the same steps there just skipping this. Now, here is the contrast, let us contrast in our DTMC we had written P of n, what is P of n, P of n there in the, in the case of DTMC is the n step transition probability matrix that we had written in terms of one step transition probability matrix by raising it to the power n.

So, now if such a similar analogy holds for CTMC. So, what is the meaning of n here? When steps, when you are going to take t here, what is the interpretation in CTMC, it is like interval of length t. There will be in interested in transition at every instance. When you are looking at n you had ever looked at n times steps, but in time even in some interval 0 to t, you will be interested in transition at every instant that means it has infinitely many uncountably many transitions possible in that.

So, that does not make sense, like I am going to write P of t to P to P for this P to the power capital T or like. So, even it is not clear like how to define a one cup, one step transition probability here like my transition can happen at any instance starting from 0. So, that makes does not make sense to represent my transition probabilities matrix here, in terms of one step transition probability matrix.

So, earlier in a DTMC for a time homogeneous case it made sense to give my n step transition probability everything in terms of one step transition property, but in, when you move to DTMC, sorry CTMC, you have to specify my transition probability matrix for every t. So, in CTMC I need to know this P of t for all t greater than or equals to t

Student: (())(20:34) P to the power.

Professor: We will see that, why it has to be of that form not anything else. So, P of t plus 1 we have so we are saying that it is going to split. So, is that exponential is the only one which will have this property. Is that the case?

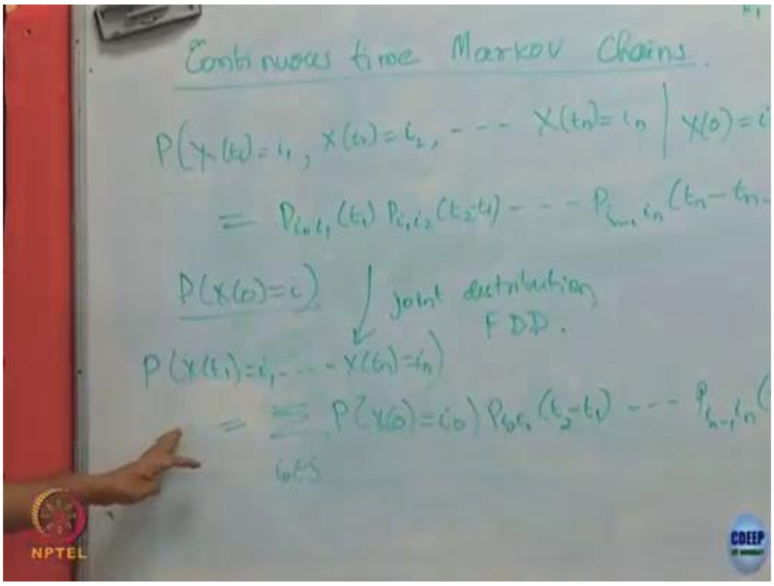
Student: t plus delta.

Professor: t plus delta t whatever like you just define S as delta t. So, we will see bit in a couple of minutes that it is indeed true that this distribution is going to be exponential distribution but not in on this what in this like here it is clear but it will sit in a different setting and then we will again revisit it but whether this P of t has to be exponential again.

So, I said I have skipped this you have to go and use the same steps as we did it in DTMC case okay, there also we had split, there we did it for any n 1 plus n 2, we had said this is going to be P of n1 into P of n2. You have to just do it like. So, here just like split this probabilities conditioning on two parts, and just apply the Markov property that you are going to get, okay fine.

So, this Pt here that is why we are going to call it as transition probability function. Because it has to be specified for every t. So, every t it is a matrix. We are just going to call it as transition property function.

(Refer Slide Time: 22:27)




Now, you can verify that simply I am going to write this final dimensional distributions for this. So, we will see that this can be just like after applying the Markov properties 0, i 1, t1, i1, i 2, t2 Pi minus 1, and if you further write it as x naught equals to i, this is your initial distributions. The same thing you could write it as unconditional. So, this is a conditional joint distribution so you are looking at your CTMC taking value i1 at time t 1 i2 at time t2 and in at time tn, given that initially started with state i naught.

You could simply split it like this, just apply the Markov property repeatedly then this probability is nothing but probability that you go from state i 0 and i1 in time t1 and then probability that you go to from state i1 and i2 in time t2 minus t this, but this is like a conditional distribution, but now, if you are interested in this simply the joint distribution and which we call finite dimensional distribution FDD.

So, now, for this you could just bring in this unconditional or your initial value. So, if you know your initial value you could write this. So, we always said like to complete the characterization of your process you need to give full description of your finite dimensional distributions. So, to give finite dimensional distribution for the DTMC what all the things we said we needed, we said we need a one step transition probability matrix and your initial distributions.

So, here also what all the things you need to do a complete distribution, description of your process? You need to give your initial distribution and then you need to give your transition probability function here, okay. So, so far what we have been doing is we have a continuous time Markov chain. We have described about Markov property and how to write down its finite dimensional distributions.

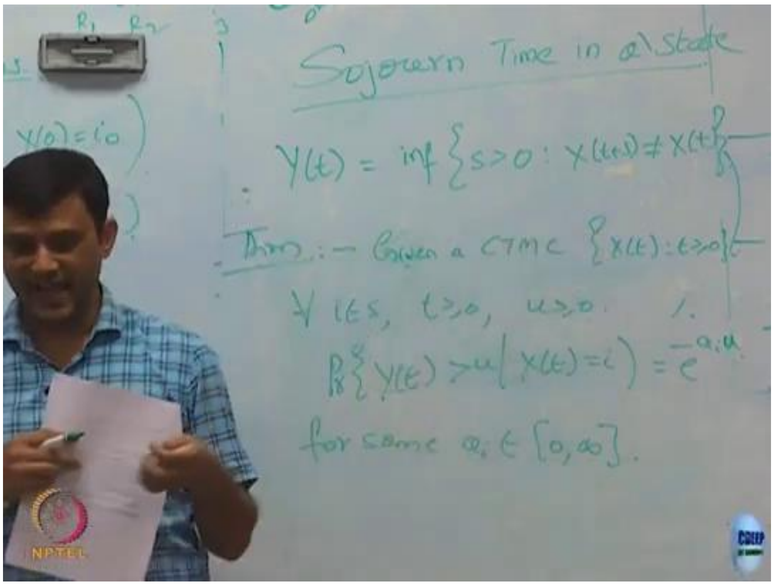
Now, in a CTMC once you are going to hit a state you are now you are continuously watching it right at every instant you are watching you are random process, you want to know how much time my process is going to stay in that state before it jumps to a new state. Like if it is a CTMC it must be the case that okay it hit some state maybe state in that state for some time and then went to a new state and share it on that state for some time and then move to another state.

Or it may be such a very, very shaky random process that it is not staying in any state for any good amount of time it came to one state immediately switched to another state and it can be very, very like unstable thing like it is not spending any time in any state for a longer time. But you want to capture this. So, this is anyway continuous time process like I am watching it continuously.

So, the discrete times you only worried about your Markov chain at particular instances. But here I have to now completely describe how it is behaving it every time instance, maybe one way to describe that is just give alternative characterization by saying that if I hit this state, if I am at this state at particular time, I am going to stay in this state for this much more time.

Maybe that is a random value, but if you can do the distribution, maybe you can alternatively characterize your continuous time Markov chain.

(Refer Slide Time: 28:39)




So, that is those are all can be given to you at sojourn time in a state. So, sojourn time is the amount of time you are going to spend in a state before you move out of it. So, we are going to define at any time t. So, can you pass it what is happening here? So, I am defining this function Y of t for any given time t, what I am looking is I am looking, let us say X of t is any state at time t.

Now, I am looking for t of s so t of s will take some state like, this is the state taken at time t of s. I am looking for s from time t, in which it is going to take a time, going to take a step other than Xt I what I am looking it, I am going to look for s, which is the smallest among all of them.

So, I am going to look for the time that yeah, so the time I am going to leave, I am going to stay in this state, right. And basically, I am basically here when I check this condition I am looking at when I actually left my state. So, when I said t, I do not care in which state I am whatever that state is Xt that is like let some random state. What I am just looked this how long I stayed in that state. And I am just looking at the minimum amount of time that is required.

And this could be a random quantity, right? Because both X of t and X of t plus s themselves are a random quantity. And this is what we are going to call it as it sojourn time. So, sojourn time will be basically the minimum amount of time you are going to spend on a particular on a state at that time at a time t.

So, this is defined for every t. So, you look for a process at some time t, you just see when is the earliest that it is going to change its state that is going to give you a sojourn time. Now, we come to this theorem, what it is saying is okay take any CTMC let us focus on a particular state i and at time t. You have given that at time t you are in a particular state i.

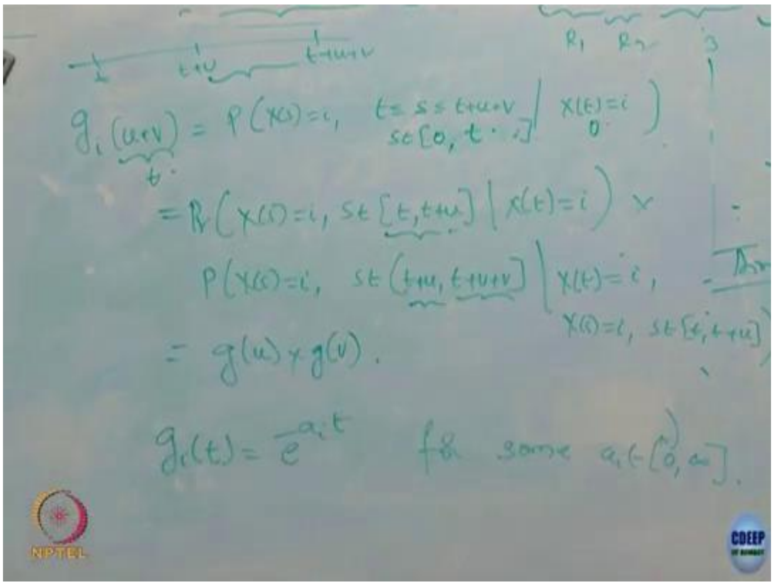
Now, what have been asked is the probability that Yt is greater than u what is Yt greater than u means the probability that is going to stay in that state i for at least u more minutes okay at least u more minutes going to be e to the power ai u for some ai which is between, which is a positive real number.

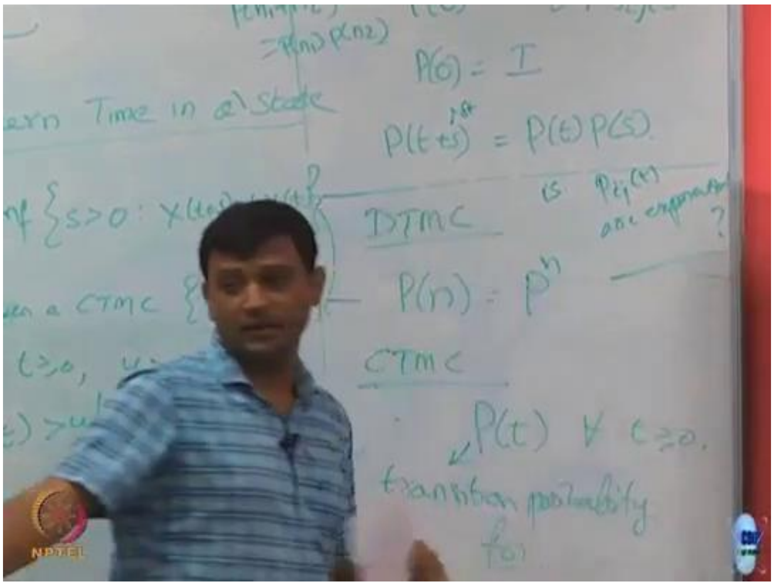
So, what is the distribution of Yt condition on Xt equals to i then it is going to be exponentially distributed with some ai which is function of that state which your process has taken at time t so, if you realize that so, okay fine so, this ai let us say for every i I have this ai, once you know what is your state space, you have for all i's you know ai's. Let say a time arbitrary time t you realize that you are in a particular state i.

Now, the amount of time that it is going to continue to stay in that same state is now is exactly according going to happen according to this distribution, which is now exponentially distributed with parameter what this parameter ai. So, in a way what we are saying is, if your continuous time process has this Markov property, it must be the case that the amount of time it is going to stay in a particular state at any arbitrary time is also memoryless. Exponential distribution we know it is a memoryless property. So, what is Markov property? Markov property in a way like I do not care what has happened before it, what matters is my current state after it is only going to cover next thing.

In a way there, it is also given the current state it is it does not depend on anything which has happened before. So, for that in a way, like we would not be surprised with such a result that if you already know that you are going to be in a particular state at time t, it must be the case that the amount of time you are going to spend in that state is again exponentially distributed. So, fine so, let us quickly go through the arguments.

(Refer Slide Time: 35:33)






So, let me define these quantities here I am interested in probability what X of s equals to i and t. So, I am going to define this probability as that a given that I am in instead i at time t, that I continue to stay in that state i in the interval t to t plus u plus v. So, alternatively, this is nothing but say, okay, I am in this interval, I was already in state at time t, i will continue to stay in that state in this interval.

So, this one you can simply expand it again by chain rule. So, for this probability I just applied chain rule I split this interval t to t, u plus v to t to t plus u and then t plus u to t plus u plus v. So, I have just split my total time between t to t plus u plus v into this part t plus u and then now by definition what I am looking at like given x t equals to i I am now asking the question okay x of s that I am going to stay in the same state i in this interval starting from t to again t plus u.

According to this definition, this is nothing but g of u. Earlier it was for g of u, but now now it is only for the interval g of u. And what is this other part now? For this other part, I am going to again apply my Markov property, from Markov property I know that actually, I have already applied my Markov property and I have already know that my x of s is i in the interval t to t plus u and I do not care what has happened before that.

So, I already know at x of s I have not going to i I now I am asking the question, okay, but what is the probability that it is going to continue to stay in the state i again, and now what is the length of this interval? It is going to be v. So, it must be the case that this is g of v. So, this probability, I could split it like this. And as one of you said, like, it is only the property of the exponential distribution, that I could do this.

So, it must be the case that this probability here that I am going to stay so, okay, so what have we just showed is g of t i is equals to is equals to e to the power ai t for some ai i here is 1. So what we are basically said is, given that I am instead, t equals to i, the amount of time I am going to continue to stay in that is again exponentially distributed, and that is why I am going to get it but you see that.

Okay, fine, this is the distribution of the amount of times we are going to continue in the same stream we are looking at end but what is this here, what is this probabilities here and why we could split it like this?

So, what is Pt plus s it is the probability that given some state at time 0 that you are going to take some other state at time t plus s and that time going to express it as product of these two terms. Do you feel that it should be also like the only possibility that this can happen if all this my probabilities, so here their probabilities. They all going to be exponentially distributed so okay what is the definition of this, g i of t? Is like this is amount of time I am going to continue to stay in state i again for t more times.

Like this is, let us say this is going to, this i here is the your initial state, let us say, let us put this is t equals to 0 here and let me make this equals to t, then I will be just looking at 0 to t this interval and this is going to be this. So, the amount of time I have spent is going to be this much distribution, then, is it not exactly the thing what I am asking for? Okay, fine, so this is the step you need to check that exponential is the only distribution that can be split like this. That is what I said like from that property we are going to use this.

So, how we are going to do this? We have already not come across this. Okay, so fine. So, how we are going to do this, it requires a couple of steps. So, maybe like you guys need to figure it out how we are going to do this or we will just put it as an assignment question. Okay. Is it the case that is what we had we try to argue, right? What is the interpretation of a P there and what is the interpretation of g here?

So, okay, fine. So there are two things you need to convince that it is not really necessary that my each of this Pij here in this matrix has to be necessarily like exponential or like is it the case that that is the only way. So, now if you are going to use this argument that the sum has split if we could split it like this if exponential is the only possible distribution why is that not necessarily be here or it is the only case. Okay, so let us check that. And then here is Pij are exponential. Okay, so let us stop here.