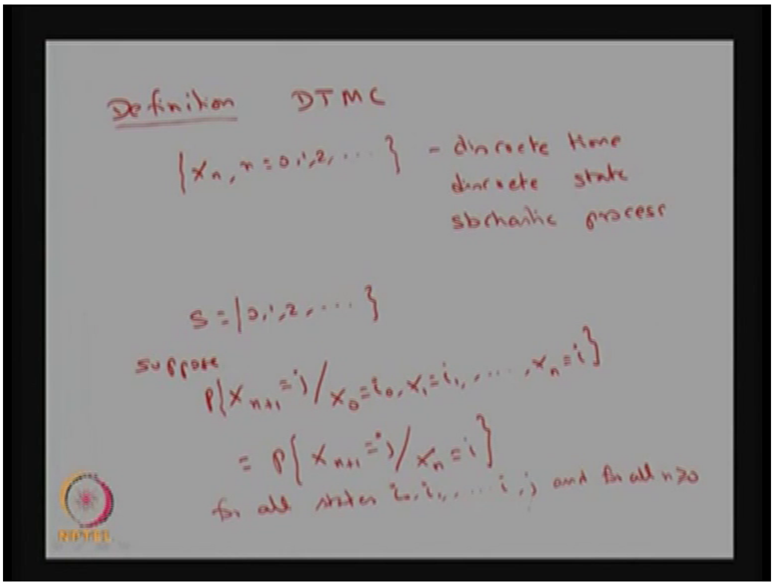
Introduction to Probability Theory and Stochastic Processes
Prof. S. Dharmaraja
Department of Mathematics
Indian Institute of Technology, Delhi

Lecture – 61

(Refer Slide Time: 00:12)



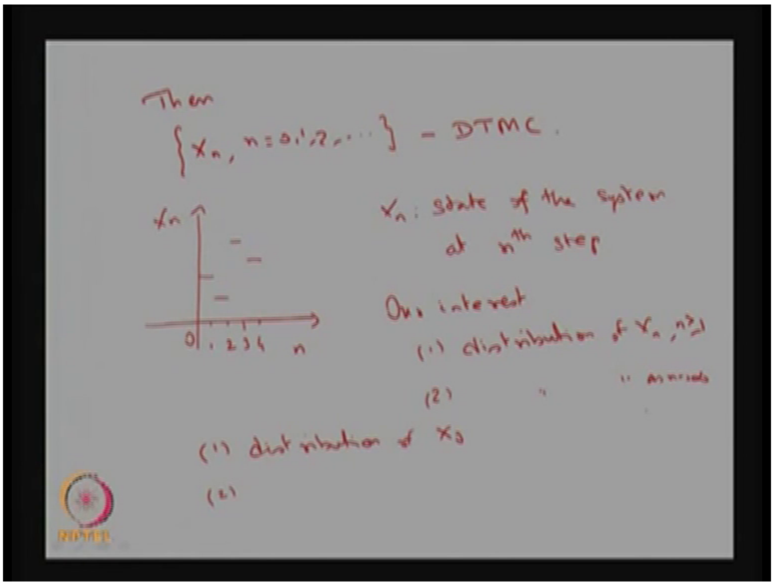
Discrete Time Markov Chain; I am going to give the formal definition of a Discrete Time Markov Chain, formal definition of Discrete Time Markov Chain in notation. We in short we call it as a DTMC. Consider a discrete time, discrete state stochastic process. Consider here discrete time, this is a discrete time, discrete state stochastic process assume that X n takes a finite or countably.

Countable number of a possible values unless otherwise mentioned the set of possible values will be denoted by the set of nonnegative integers S is equal to 0, 1, 2 and so on. Unless otherwise measured, you can mentioned, you can always assume that the state space S consists of the element 0, 1, 2 and so on. Even if we takes a other values also, you can always make a 1 to 1 correspondence and make the state space is going to be S is equal to 0 1 2 and so on. Suppose, the probability of the X n plus 1 will be taking the value j given that, it was taking the value X naught is equal to i naught X 1 was i 1.

And so on and X n was i that probability is same as the probability of X n plus 1 will be j. Given that X n was i for all states for all states i naught whatever be the value of i naught i i and j and also for all n greater than or equal to 0. If this property is satisfied by for all states, i naught i 1 i comma j as well as the for all n greater than or equal to 0.

Then this stochastic process that is a discrete time, discrete state, stochastic process is going to be known as a Discrete Time Markov Chain. So, basically this is the Markov property and the Markov property is satisfied by all the states, as well as all the random variables. So, if this Markov property is satisfied by any stochastic process then it is called a Markov process. And since it is the time space is discrete and the parameter space is discrete. Therefore, it is called a Discrete Time Markov Chain then we see.

(Refer Slide Time: 03:38)



This stochastic process X n takes value n starting with 0, 1, 2 and so on, is the Discrete Time Markov Chain. We can just to have a look of how the sample path look like for different value of n and the y axis is X n. Suppose at n is equal to 0 it started with some value at x is equal to n is equal to 1.

It would have been the different value and n is equal to 2 it would have been the different value. So, these values are the either it could be a finite value or countably infinite number of values. Therefore, the state space is going to be discrete and the parameter space is going to be discrete. So, like that it is a taking a different value over there. So, this is going to be the sample path or trace off the stochastic process X n. Suppose, you assume that X n is the state of the system at nth step or nth time point.

And this X n satisfies the DVSS, this Markov property then, the stochastic process is going to be call it as a Discrete Time Markov Chain. And our interest will be suppose, the stochastic process satisfies the Markov property, our interest will be to know the 2 things. 1 is, what is the distribution of X n for n is greater than or equal to 1 you know where the system starts. So, X naught you know your interest will be what could be the distribution of X n.

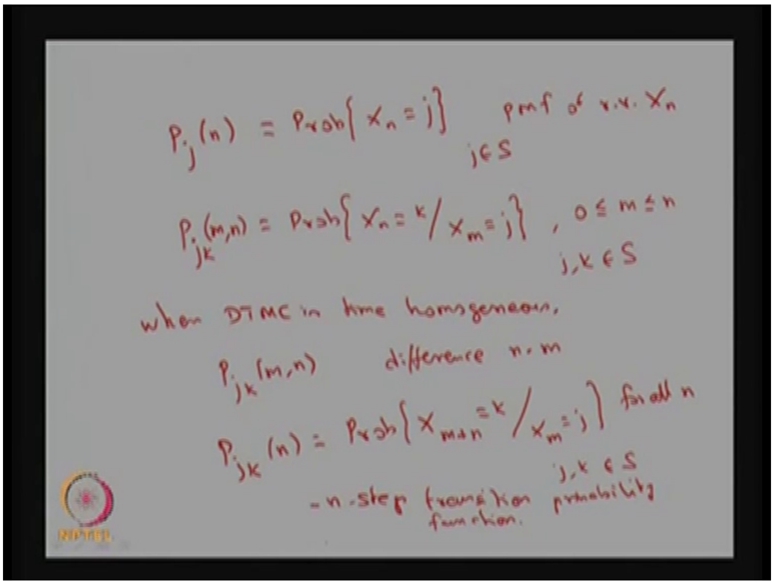
That is nothing but, what is the probability that the X n will be in some state j? And also what could be the distribution of X n as a n tends to infinity, as n tends to infinity or our interest will be finding out the distribution of the X n. So, at any finite n as well as the n tends to infinity that will be of our interest.

To compute this, you need 2 things 1 is you need, what is the distribution of X naught? That is the initial distribution vector where the system starts at the 0th step. What is the distribution of X 0 and also 2nd things of your interest will be; what is the transition distribution? Or how the transition takes place? What is the distribution of a transition from any nth step to n plus 1th step for all n?

So, if you know the two things the initial distribution vector as well as the distribution of the transition. From nth step to n plus 1th step using these 2 quantity. You can find out what is the distribution of X n for any n as well as you can find out the distribution of X n as n tends to infinity.

For that, I am going to define few conditional probability distribution as well as the marginal distribution. For the random variable X n, through that we are going to find out the distribution of X n for any n as well as n tends to infinity. So, the 1st one I am going to define.

(Refer Slide Time: 07:53)



The probability mass function as the P suffix j of n that is nothing but, what is the probability that X n takes the value j? So, this is the probability mass function of the random variable X n. What is the probability that a X n takes the value j that, I am going to denote it as that P suffix j of n? Where here, the j is belonging to the state space capital S; this is the probability mass function of the random variable X n. Similarly, I am going to define the conditional probability mass function as a P suffix j k of n. That is nothing but, what is the probability that the X n takes the value k? Given that, X sorry I need a 2 suffixes to index for here.

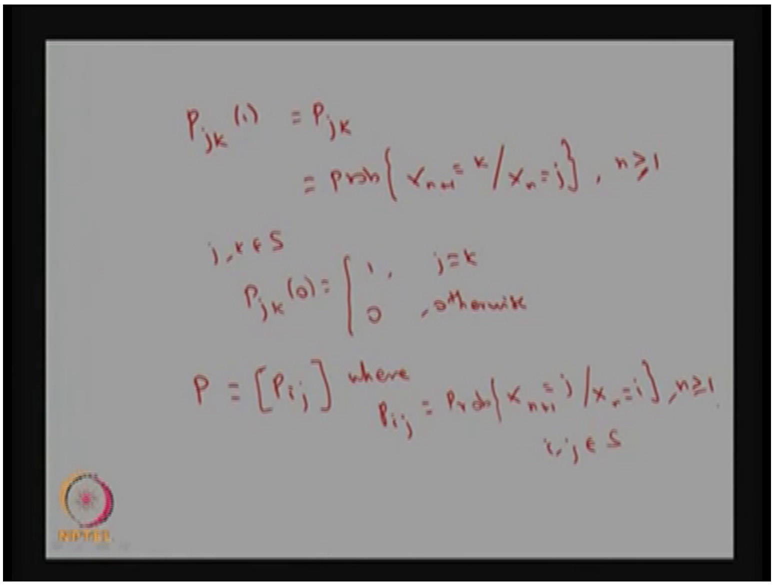
With the 2 variables m comma n, what is the probability that X n will be the state k given that X m was j ? Obviously, the m is lies between 0 to n whatever m and every n and the j comma k is belonging to capital S. So, this is a conditional probability distribution of the random variable X n with the Xm and the mth step. The system was in the state j and the nth step the system is in the state k. And this is a conditional probability with the 2 arguments m comma n. So, this is the probability that the system makes a transition from the state j at step m to the state k at step n.

And this is called a transition probability function of the Discrete Time Markov Chain. When the DTMC is a time homogeneous, this is very important, when DTMC is a time homogeneous; that means, it satisfies the time invariant property; that means, the P j k of m comma n depends only on the time difference n minus m.

Whenever, the DTMC is a time homogeneous; that means, in the time invariant. So, the actual time is not a matter only the time difference is the importance. Therefore, this is going to be depends only on the time difference n minus m. In this case, I do not want the two arguments m comma n I can go for writing P j k of n that is nothing but. What is the probability that the m plus nth step the system will be in the state k? Given that the mth step it is in the state, it was in the state j; for all n and here j comma k belonging to S. So, the m does not matter only the interval or the interval length of step n is matter so.

That means, if the system was in the state j and the it is a transition into the state k in n steps because the DTMC is the time homogeneous. So, the X m to X m plus n it is valid for all m for all n we are finding out for the n step transition this is called n step because the DTMC is a time homogeneous and this is called a n step transition probability function. This is the n step transition probability function. Using this, we can define the 1 step transition probability that is denoted by.

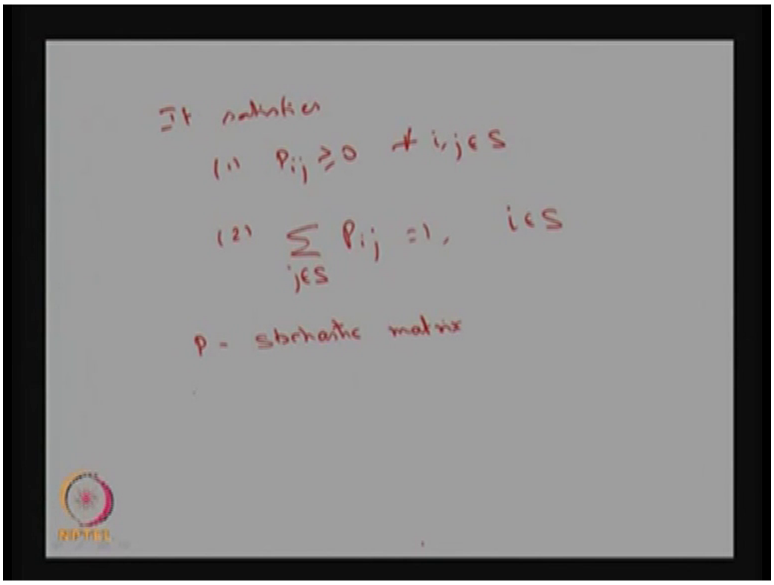
(Refer Slide Time: 12:46)



P suffix jk of 1 or we can avoid the bracket 1 also. So, you can write it as the P j oh P j k that is nothing but, what is the probability that the X n plus 1 is equal to k. Given that X n is equal to j for all n greater than or equal to 1; obviously, for j comma k belonging to S if you find out the 0 step transition probability that, values is going to be 1 for j equal to k. Otherwise, it is going to be 0. This one step transition probability I can make it in the matrix form as the capital P is the matrix. And that consists of P ij. Where the Pi j is nothing but 1 step transition probability matt elements of X n plus 1 is equal to j.

Given that X n is equal to i. Here, i comma j belonging to the state space S we should remember that, the state space S is consist of finite elements or countably infinite number of elements. Accordingly, this matrix is going to be either when S is going to be a finite elements. Then the P matrix is going to be a square matrix. Since, the P ij is the one step transition probability of a system moving from the state i to j in 1 step. And since it is a time homogeneous this is valid for all n, this is valid for all n greater than or equal to 1 and this satisfies.

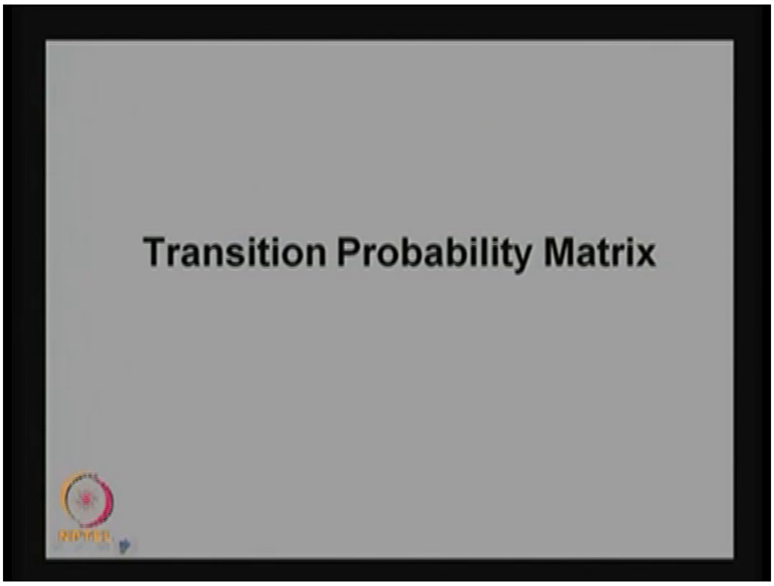
(Refer Slide Time: 14:56)



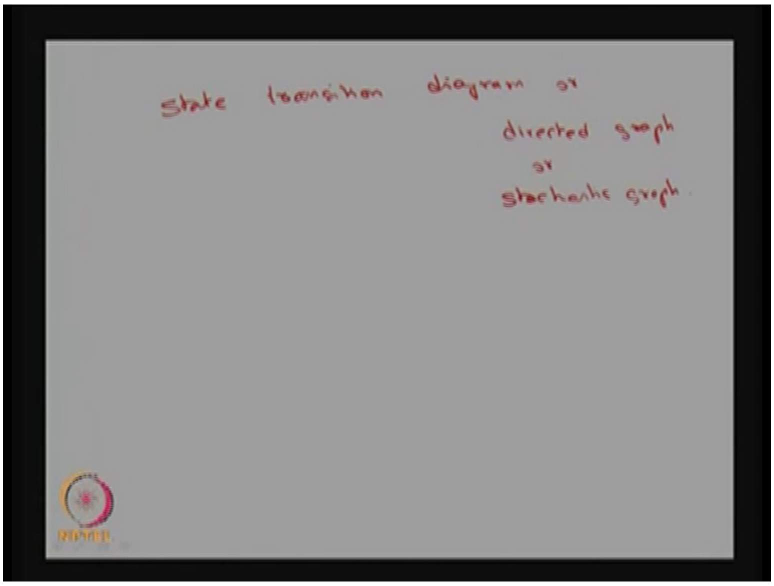
The 1 step transition probability matrix satisfies 2 properties. The each entity will be greater than or equal to 0; for all i comma j belonging to S because, these are all only the conditional probability of system moving from the state i to j in 1 step. Therefore, either it will be a 0 or greater than 0 for all possible values of i comma j. The second condition, if you make the summation over j for fixed i then, that that is going to be 1. i belonging to S; that means, the row sum is going to be 1 because, it is a conditional probability of system moving from 1 state to all other states. If you add all the other possible probabilities then that is going to be 1.

And since, this 1 step transition probability matrix satisfies these 2 properties and this matrix is P is known as a stochastic matrix. Because of satisfying these 2 conditions the matrix 1 step transition matrix is also called a stochastic matrix. Now, I am going to explain: what is the pictorial way of viewing the 1 step transition probability matrix or the stochastic matrix that is provided by state.

(Refer Slide Time: 16:48)



(Refer Slide Time: 16:53)



Transition diagram or the other word it is called a directed graph the DTMC can be viewed as a directed graph. Such that, the state space S is the set of vertices or nodes and the transition probabilities that is a 1 step transition probabilities or the weights of the directed arcs between these vertices or nodes. Since, the weights are positive and the sum of the arc weights from each node is unity, this directed graph is also called a stochastic graph.