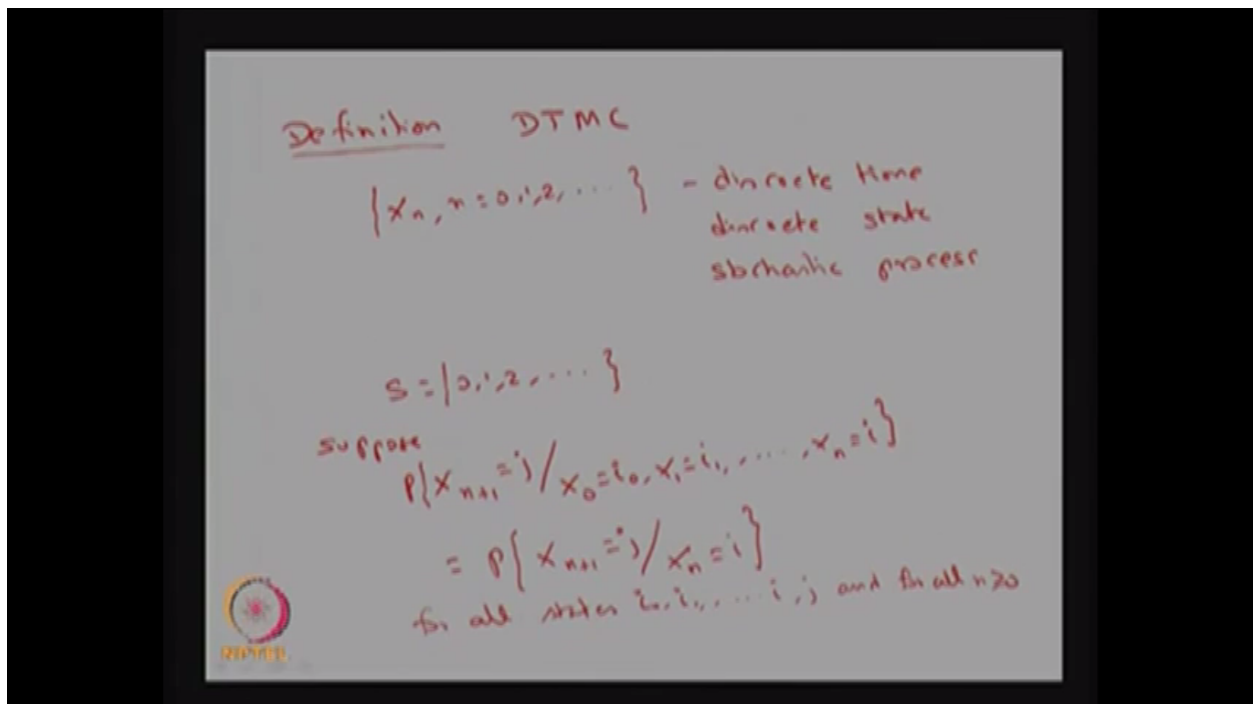


Discrete time Markov chain I am going to give the formal definition of a discrete time Markov chain. Formal definition of discrete time Markov chain in notation we in short we call it as a DTMC.

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 non-negative integers S is equal to 0, 1, 2, and so on and that's otherwise measured you can mention you can always assume that the state space S consists of the element 0, 1, 2, and so on even if it takes other values also you can always make a one-to-one correspondence and make the state space is going to be S is equal to 0, 1, 2 and so on.

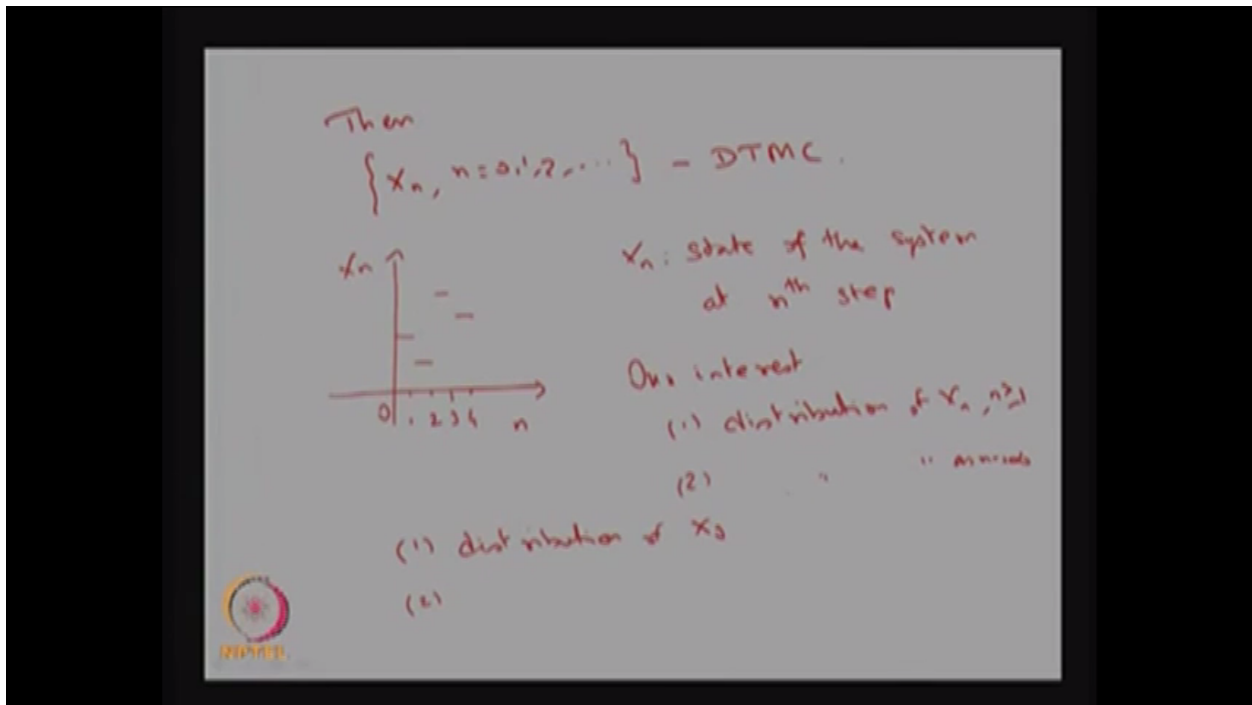
Suppose the probability of the X_{n+1} will be taking the value J given that it was taking the value X_n is equal to i X_1 was i_1 and so on and the X_n was i that probability is same as the probability of X_{n+1} will be j given that X_n was i for all states i j whatever be the value of i j , i_1 , i and j and also for all n greater than or equal to 0 if this property is satisfied by for all states i j , i_1 , i, j as well as 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 discrete time Markov chain.



So basically this is the Markov property and the Markov properties are satisfied by all the states as well as all the random variables. So if this Markov properties 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 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 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 the n is equal to 2 it would have been the different values. So these values are 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 taking a different value over the n . So this is going to be the sample path or trace of the stochastic process X_n .

Suppose you assume that X_n is the state of the system at n th step or n th time point and this X_n satisfies the previous 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 two things one is what is the distribution of X_n for n is greater or equal to 1. You know where the system starts so X_0 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 has a n tends to infinity. Our interest will be finding out the distribution of the X_n . So at any finite n as well as n tends to infinity that will be of our interest. To compute this you need two things. One is you need what is the distribution of X_0 that is the initial distribution vector where the system starts at the 0th step what is the distribution of X_0 and also second 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 n th step two $n+1$ th step for all n . So if you know the two things the initial distribution vector as well as the distribution of the transition from n th step to the $n+1$ th step using these two 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'm 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 first one I am going to define the probability mass function as the P suffix j of n that is nothing but what is the probability that the X_n takes a value j . So this is the probability mass function of the random variable X_n . What is the probability that X_n takes a value j that I'm going to denote it as the P suffix j of n where here the j is belonging to the state space 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 P suffix jk of n that is nothing but what is the probability that the X_n takes a value K given that, sorry I need two suffixes, two index for here with the two variables m, n what is the probability that the X_n will be the state k given that X_m was j . Obviously they m is lies between 0 to n . Whatever m and every n and the j, k is belonging to S . So this is a conditional probability distribution of the random variable X_n with the X_m and at the m th step the system was in the state j and the n th step the system is in the state k and this is a conditional probability with the two arguments given m, n . So this is a 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_{jk} of m, n depends only on the time difference n minus m whenever the DTMC is a time homogeneous that means is in the time invariant so the actual time is not a matter only the time difference is the importance. Therefore, this is going to depend only on the time difference n minus m . In this case I don't want the two arguments m, n I can go for writing P_{jk} of n that is nothing but what is the probability that the m plus n th step the system will be in the state k given that the m th step it is in the state it was in the state j for all n and here j, 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 that the system was in the state j and it is a transition into the state k in n steps because the DTMC is a time homogeneous. So the X_m to X_{m+n} it is valid for all n . 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 n step transition probability function. This is the n step transition probability function.

$$P_{ij}(n) = \text{Prob}\{X_n = j\}$$
 pmf of r.v. X_n
 $j \in S$

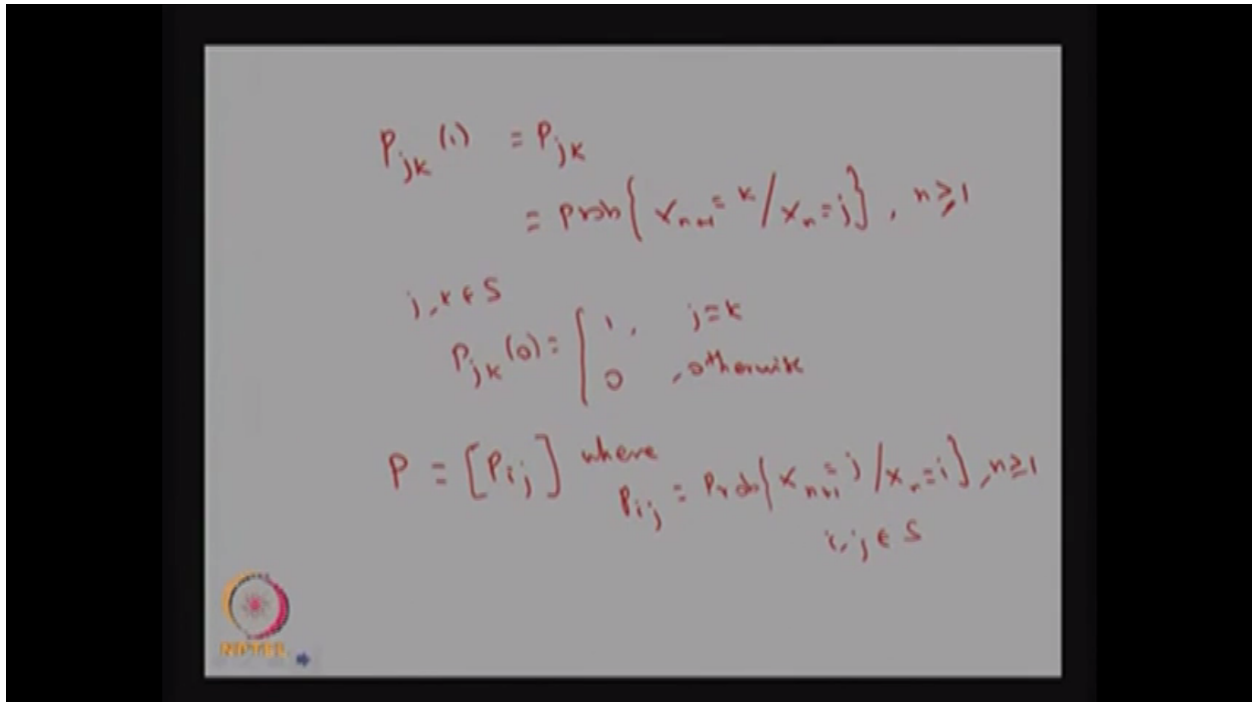
$$P_{jk}^{(m,n)} = \text{Prob}\{X_n = k / X_m = j\}$$
, $0 \leq m \leq n$
 $j, k \in S$

when DTMC is time homogeneous,

$P_{jk}^{(m,n)}$ difference $n-m$

$$P_{jk}^{(n)} = \text{Prob}\{X_{m+n} = k / X_m = j\}$$
 for all n
 $j, k \in S$
 - n -step transition probability.

Using this we can define the one step transition probability that is denoted by P suffix jk of 1 or we can avoid the bracket 1 also so you can write it as the P_{jk} that is nothing but what is the probability that the X_{n+1} is equal to k given that X_n is equal to j for all n greater or equal to 1. Obviously for j, 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 1 step transition probability I can make it in the matrix form as the P is the matrix and that consists of P_{ij} where the P_{ij} is nothing but one step transition probability elements of X_{n+1} is equal to j given that X_n is equal to i here i, j belonging to the state space S . You should remember that the state space S is consists 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 one step and since it is a time homogeneous this is valid for all n great or equal to 1 and this satisfies – the one-step transition probability matrix satisfies two properties. The each entity will be great or equal to 0 for all i, j belonging to S because these are all only the conditional probability of system moving from the state i to j in one step therefore either it will be a 0 or greater than 0 for all possible values of i, j .

The second condition if you make the summation over j for fixed i then 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 one state to all other states if you add all the other possible probabilities then that is going to be 1. Since this one-step transition probability matrix satisfies these two properties and this matrix is P is known as a stochastic matrix because of satisfying these two conditions the matrix one-step transition matrix is also called a stochastic matrix.

Now I am going to explain what is the pictorial way of viewing the one-step transition probability matrix or the stochastic matrix. That is provided by state 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 one-step transition probabilities or the weights of their 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.