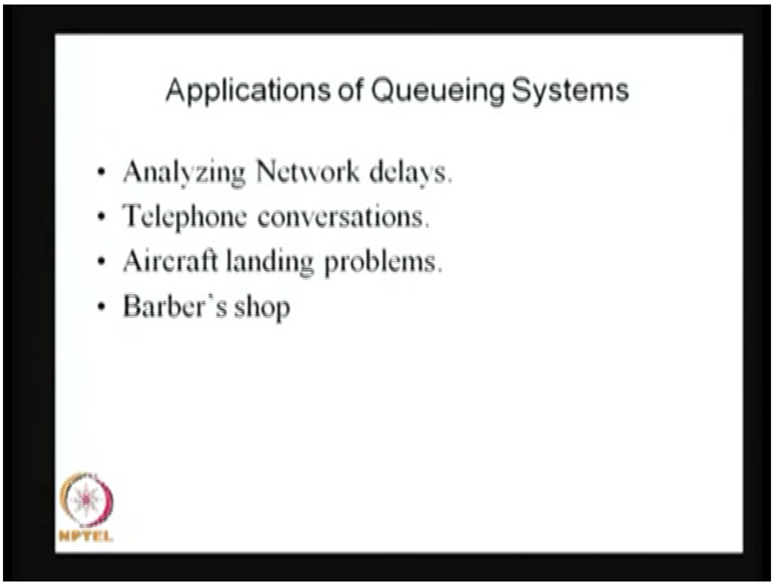
Introduction to Probability Theory and Stochastic Processes
Prof. S. Dharmaraja
Department of Mathematics
Indian Institute of Technology, Delhi

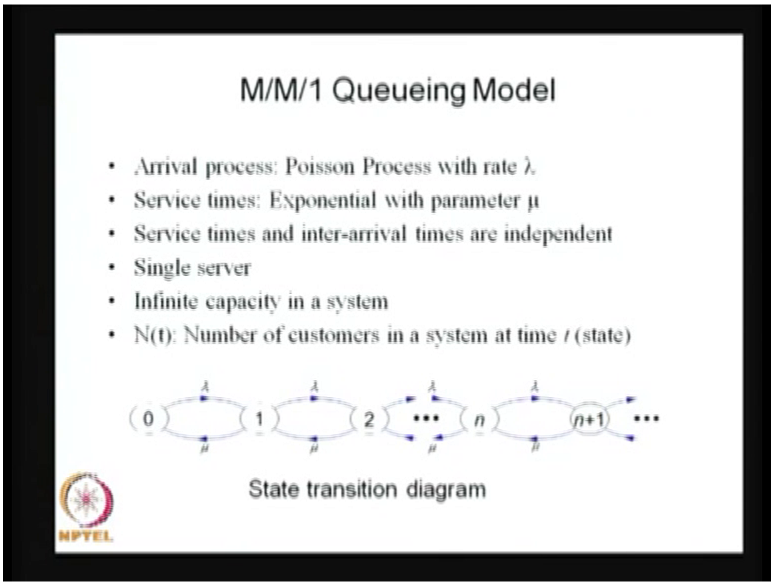
Lecture - 86

There are many applications of queueing system we are going to discuss the abstract queueing system in the further lecture.

(Refer Slide Time: 00:02)



(Refer Slide Time: 00:11)



The easiest or the simplest queueing model that is, a Markovian queueing model that is MM1 queueing model. Later, we are going to relate with the birth death process also. In the MM1queueing model the inter arrival time is exponentially distributed. As I discuss the Poisson process in the previous lecture, whenever you have a arrival follows a Poisson process. Then the inter arrival time follows a exponential distribution and or independent also.

So, here the 1st information is a arrival process follows a Poisson process with the intensity or rate lambda; that means, a the inter arrival times are independent and each one is exponentially distributed with the parameter lambda.

The 2nd information that is service time service times are exponentially distributed with the parameter mu. And the service times are independent for each customers and that is also independent with the arrival process; that means, there is no dependency over the arrival pattern with the service pattern, service times and the inter arrival times are independent. Then the 3rd information only 1 server in the system that is a queueing system in which only 1 server and the 4th information is missing; that means, it is a default it is infinite capacity model infinite capacity model.

Now, our interest is to find out the behavior of a queueing system or the behavior of number of customers in the system at any time t. Therefore, you can define a random variable n of t that is nothing but the number of customers in the system at time t. Therefore, this is going to follow form a stochastic process over the t.

Since, the inter arrival time is a exponentially distributed and the service times are exponentially distributed, the memory less property is going to be satisfied throughout all the time. Therefore, this Stochastic process there is a Discrete state continuous time stochastic process satisfying the Markov property. Therefore, this is a Markov process since, a inter arrival time is exponentially distributed and the service time is exponentially distributed and both are independent.

And the service time is also independent for each customers, therefore, this stochastic process satisfies the memory less property at all time points. Therefore, this discrete state because the possible values of n of t since, it is a number of customers the possible values are 0, 1, 2 and so on countably infinite. Therefore, it is a discrete state and you are observing the system over the time. Therefore, it is a continuous time therefore; this stochastic process is a discrete state, continuous time, stochastic process satisfying the Markov property based on these assumptions.

Therefore, n of t is a Markov process since the state space is a discrete. Therefore, this is a Markov chain therefore, this is a continuous time Markov chain therefore, n of t is a CTMC. So, one can write the state transition diagram for this CTMC; that means, the possible states are 0, 1, 2 and so on. So, this will form a notes and you try to find out, what is the rate in which the system is moving from one state to other state.

Since it is a MM1queue model queueing model therefore, whenever the system is in the whenever the system is in the state 0 by the inter arrival time which is exponentially distributed the number of customers in the system will be incremented by 1. Therefore, that rate will be lambda or the system moving from the state 0 to 1 it spends exponentially distributed amount of time here before moving into the state 1.

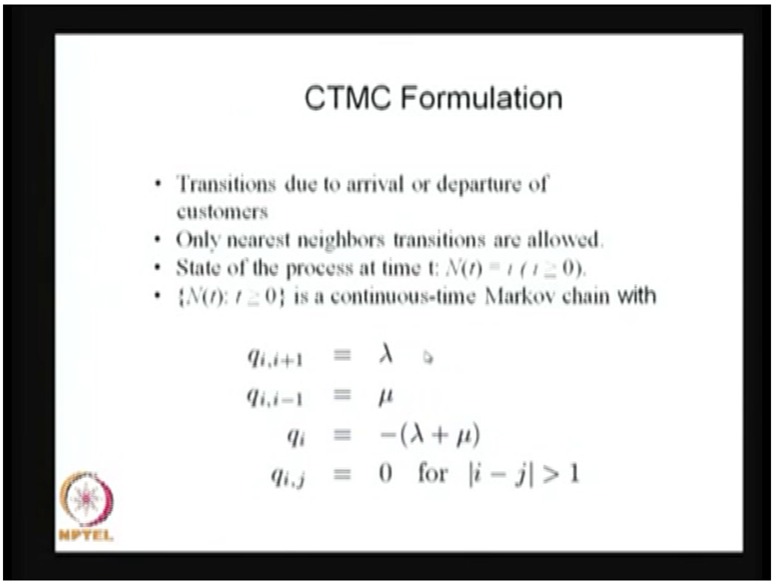
Once the system come to the state 1 either one more arrival is possible or the customer who is under service then service could have been finished. Therefore, the service time is a exponentially distributed with the parameter mu. Therefore, the system goes from the state 1 to 0 with the parameter mu. Similarly, from 1 to 2 because of the inter arrival time is exponentially distributed with the parameter lambda therefore, this is lambda.

Since, the arrival follows a Poisson process in a very small interval of time; only one customer is possible with the probability lambda times delta t and so on. Therefore, there is no way the system goes from one state to jump into more than 1 state that is not possible forward. So, only one step forward is possible because of the arrival process follows a Poisson process and the since we have only one servers in the system the system also decremented by only one level below

Therefore, this is going to form a birth death process. The reason for this CTMC going to be a birth death process because of that arrival process follows a Poisson process. So, whatever the assumptions we have it for the Poisson process that is going to be satisfied. And since, we have only one server in the system and he does the service for only one customer at a time. After finishing that server after finishing the customer service, then, it move into the next service immediately and so on if the customers are available in the queue.

Therefore, the system goes to the 1 state below by only one move only it would not move from 2 to 0 or 3 to 1 and so on therefore, this CTMC is a birth death process.

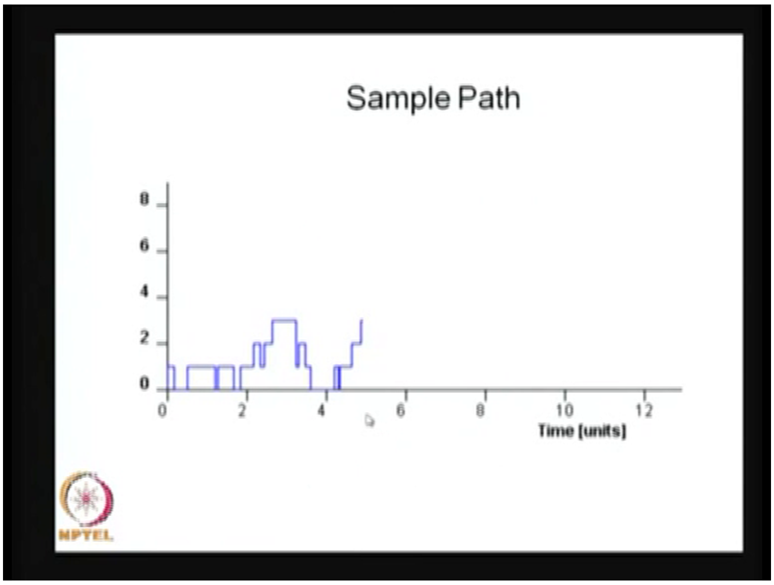
(Refer Slide Time: 07:05)



Therefore, I am connecting the CTMC with the MM1queue in particular the CTMC the birth death process. Because, of the transitions due to arrival or a departure of a customer and only nearest neighbor's transitions are allowed.

Because, of the assumptions which you have made therefore, this is going to a continuous time Markov chain with the rate in which the system moves from the state i to i plus 1 that, rate is lambda. And the system moves from the state i to i minus 1 that rate is mu and all other rates are going to be 0 other than the diagonal element. And these rates also constraint not the state dependent rates. Therefore, this is a birth death process with the birth rates lambda under death rates mu.

(Refer Slide Time: 08:08)

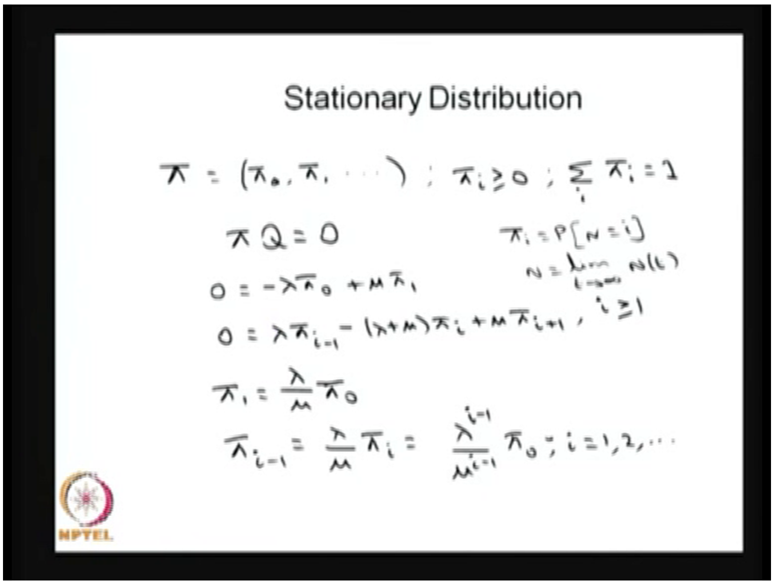


So, this is a sample path suppose, at time 0 1 customer in the system then, services over then the 2nd customer enter into the system. Now, the number of customers in the system is 1 and so on so; that means, this duration is the service time for the first customer and from this point to this point, that is a inter arrival time of the 2nd customer entering to the system and from this time point to this time point, that is the service time for the second customer which is independent of the service time for the first customer.

And this is the time point 2nd customer enter and this is a time point in which the 3rd customer enters. Therefore, the inter arrival time is from this point to this point and so on. So, this is the dynamics of a number of customers in the system over the time. Therefore, this stochastic process is a Discrete State Continuous Time stochastic process at showing the Markov property. Therefore, this is a continuous time Markov chain.

So, later I am going to simulate the MM1 queueing model using some simulation technique.

(Refer Slide Time: 09:31)



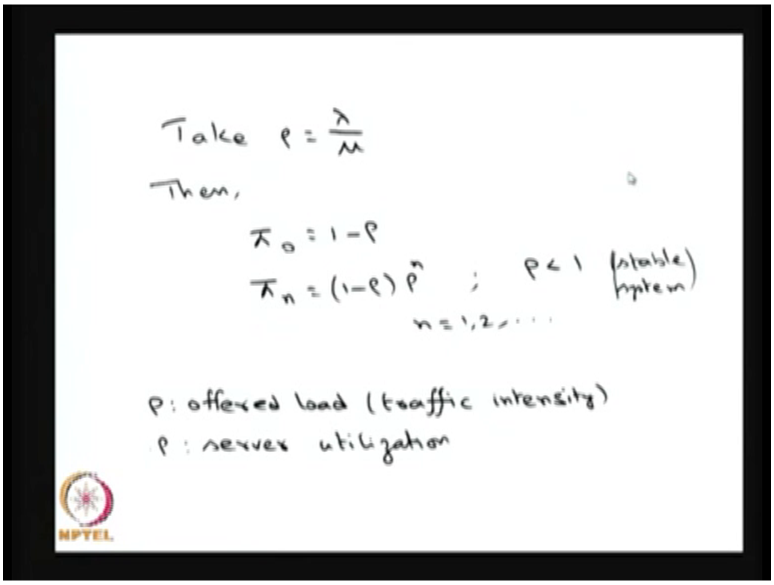
So, the conclusion is the under the underlying stochastic process for the MM1queueing model is a birth death process the n of t is a stochastic process. So, this stochastic process is a birth death process therefore, now we are going to discuss the stationary distribution time dependent probabilities and so on.

So, how to find the stationary distribution? Solve pi Q is equal to 0 pi is the vector consists of a pi is where pi is are nothing but what is the probability that, n customers in the system. What is the probability that i customers in the system in a long run? So, that long run is defined in this way the n of t is a stochastic process as a t tends to infinity. The number of customers in the system in a long run that is going to be the n and pi is nothing but a probability that n, i customers in the system in a longer run.

So now we are going to solve a pi Q is equal to 0 with the normalized equation summation of pi is equal to 1. So, once you frame the equation you will get a pi 1 in terms of pi naught and the pi i minus 1 in terms of first pi i then substitute recursively you will get in terms of pi naught.

So, since it is a homogeneous equation you will get all pi is in terms of pi naught. So, use the normalizing equation summation of pi i is equal to 1.

(Refer Slide Time: 11:08)



You will get pi naught so the pi naught is equal to 1 minus rho where rho is lambda by mu. And since, I am relating this stochastic process with a birth death process with the infinite capacity. If you recall the stationary distribution exists as long as the denominator of a pi naught that series converges. So, that will converge only if a lambda by mu is less than 1 if lambda by mu is greater than or equal to 1. Then, that denominator diverges accordingly you would not get the stationary distribution.

So, to have a stationary distribution you need a rho has to be less than 1 that also you can intuitively say whenever the system is stable that is corresponding to rho is less than 1. In that, you will have a stationary distribution; that means, a in a longer run this is a proportion of the time the system will be empty and the pi n is nothing but the n customers in the system in a longer run that is a 1 minus rho times rho power n where rho is less than 1.

This rho can be visualized as the offered load also because the rho is nothing but the mean arrival rate and the mu is the mean service rate and this ratio will give the offered load. And 1 minus pi naught that is the probability that the system is nonempty and that is nothing but the server utilization server utilization is nothing but what is the probability that the server is busy? The server will be busy as long as a the system is not empty. So, the rho is the server utilization that can be obtained in the from this formula and in a longer run the server utilization is rho.