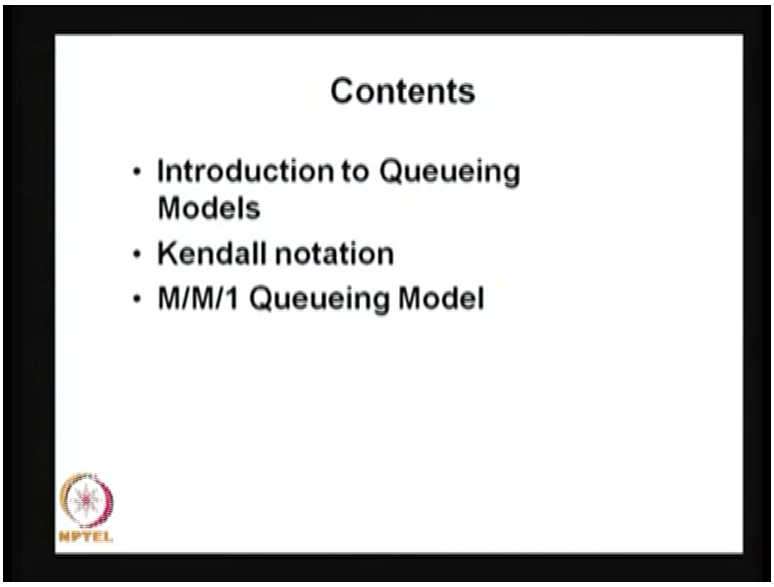
Introduction to Probability Theory and Stochastic Processes
Prof. S. Dharmaraja
Department of Mathematics
Indian Institute of Technology, Delhi

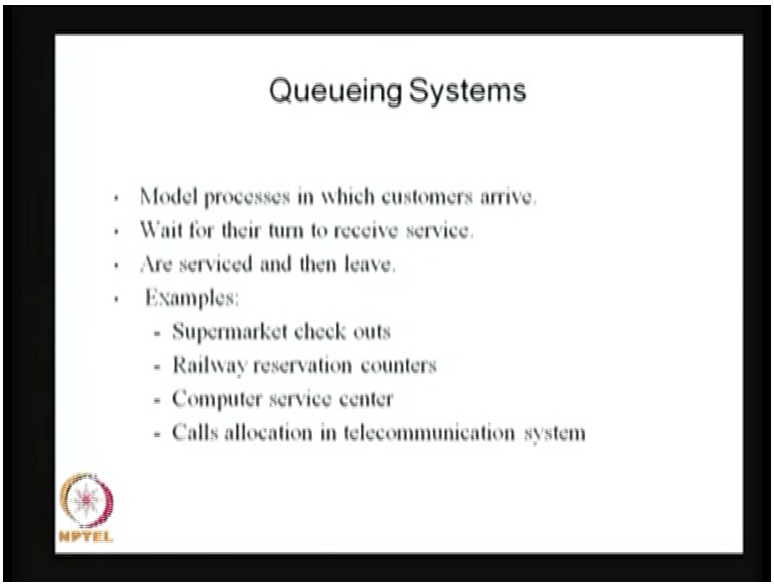
Module – 12
Simple Markovian Queueing Models
Lecture - 85

(Refer Slide Time: 00:18)



In this talk, I am going to discuss the queueing models. So, for that, I am going to give the introduction to the queueing models. Then, I am going to discuss the Kendall notation then followed by that the simplest queueing model M M 1 queue will be discussed. And this is going to be the applications of continuous time Markov chain in queueing models. So, in this lecture, I am going to discuss only the simplest queueing model M M 1 queues.

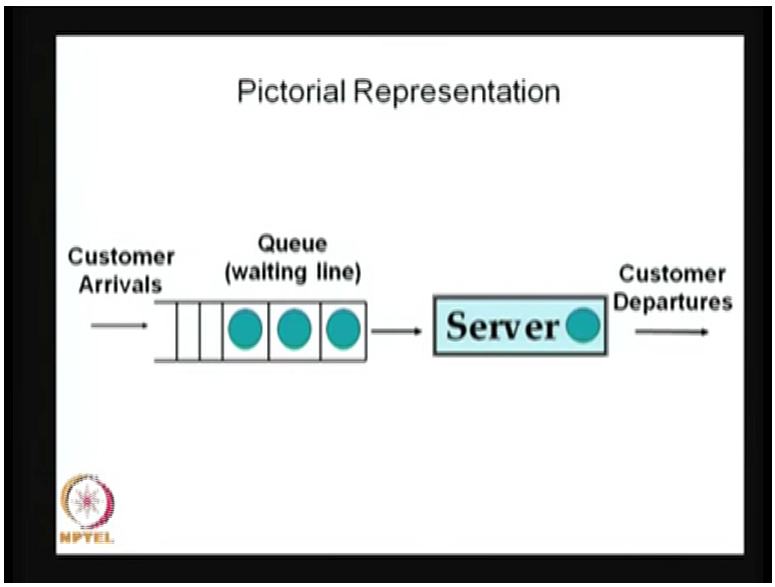
(Refer Slide Time: 00:55)



So, how one can define the queueing system? You can see many examples in which whenever, you go to the supermarket to get some items or you see the railway station counters or you can see the computer service center, many PCs are there and printers and so on.

So, how the queueing system is created and also you can see the examples in the calls allocation in telecommunication systems. In all those examples, you can see something is getting served and leave the system.

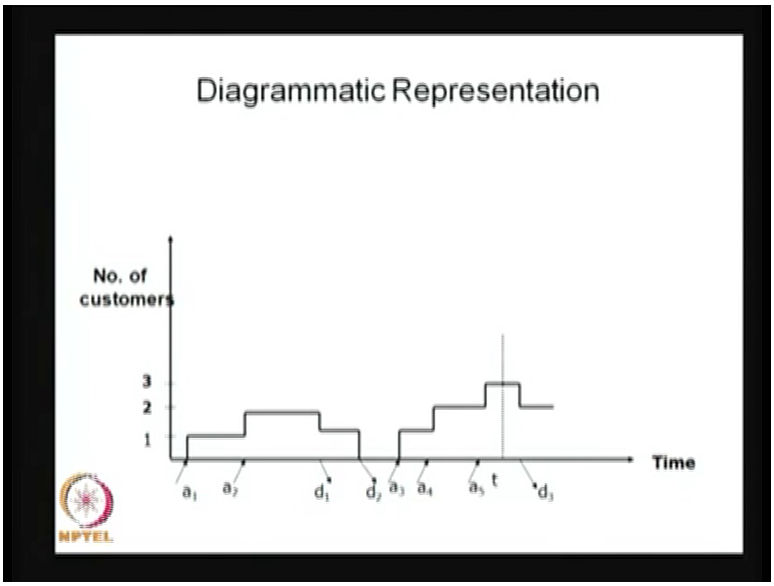
(Refer Slide Time: 01:38)



We can give the queueing system we can represent the queueing system in a pictorial form. Some customers are coming into the system and waiting for their service. Once the service is over, then they departure from the system.

So, this is the way one can visualize the queueing system in a pictorial form.

(Refer Slide Time: 02:10)



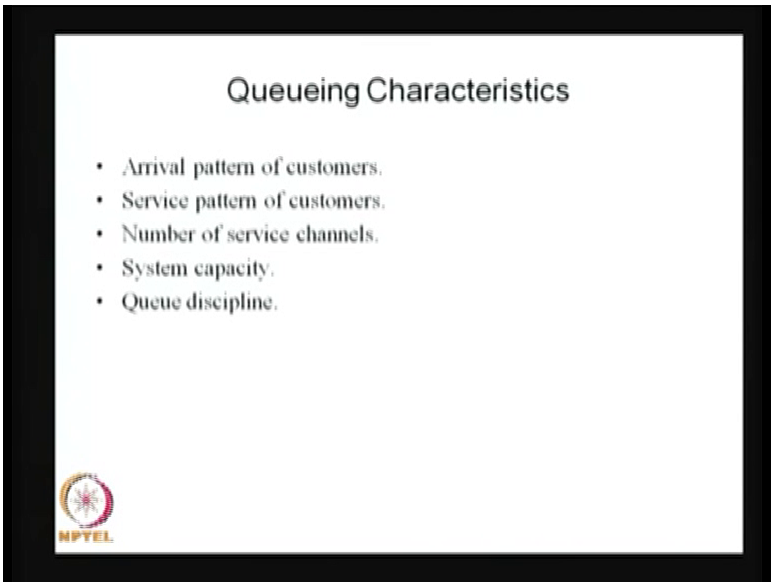
This is a diagrammatic representation and the x axis is the time and y axis is the number of customers in the system. Suppose, at time a 1, the first customer entered into the system then the number of customers in the system is incremented by 1. The customer who entered the system is to getting the service.

During his service time, the next customer enters the system that with the time point a 2. Therefore, now the number of customers in the system is 2. Going on at this time point, the first customer service is over. So, he departure from the system, that is d 1; the time point in which the first customer service is over.

Now, the number of customers in the system is 1. The time point t suffix 2, the second customer service also gets over. Now the number of customers in the system is 0. The third customer enter at the time point a 3. So, during this interval, the system was empty. So, like that, the system is a keep increasing whenever one customer entering to the system and decreasing by one whenever the service is completed.

So, this is the diagrammatic representation of a any queueing system. Here, I made the assumption it is a very simplest one only one customer entering into the system and only once customer is getting served and leave the system and so on. So, this is the simple way of a simple diagrammatic representation of the queueing system.

(Refer Slide Time: 04:10)



So, to define the queueing system, you need a few important characteristics. Using that, one can easily frame the queueing system. So, for that, you need the first information that is a arrival pattern of customers; how the customers are entering into the system, how frequently, whether the customers are coming in a very constant interval of time or in a random fashion. If it is constant, then we say the inter arrival time is a deterministic. If the customers are entering into the system with the inter arrival time, that is some random variable. Then, we should know what is the distribution of inter arrival time.

So, this information is needed to define the queueing system, the arrival pattern that includes whether it is a deterministic or probabilistic. If it is a deterministic, then what is the inter arrival time; that constant time. If it is a probabilistic, then, what is a distribution and so on.

Similarly, after the customer is entering into the system, you should know how the service takes place. Whether the service time for each customer who enter into the system is it a constant or a random. If it is a constant amount of service for each

customer, then what is a time, how much time it takes for each service, If it is a probabilistic, then what is a distribution of service time.

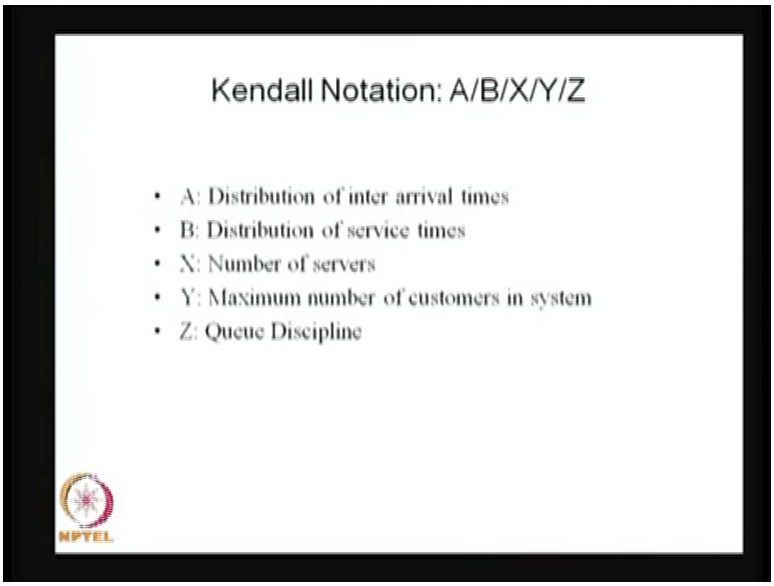
Then, the third important information or the characteristic is a number of servers in the system. How many service channels are available to do the service, whether you have only one server in the system or more than one or countably infinite numbers. So, according to that the queueing system may vary. So, the third information is number of number of servers in the system.

The fourth information that is a system capacity, whether the capacity is a finite one or infinite capacity accordingly, the number of customers in the system may go maximum the finite capacity or it may infinite number of customers can be wait in the system to get the service. Therefore, the system capacity is also important characteristic.

The fourth one queueing discipline, when the customers entering into the system, whether they are getting served or whether they are placed in a first come first order or first come last serviced or random fashion or priority based and so on. So, the queueing discipline also important to know to know the how the queueing system is at any time to know that dynamics of number of customers in the system. And you should know how the queueing discipline is taken care.

Similarly, the service discipline also. How the service is also takes place during the picking the customers for the service. So, these are all the minimum important information to characterize the queueing system. One is a arrival pattern, second is a service pattern and the third is a number of server, the fourth is a capacity of the system and the service discipline or queueing discipline.

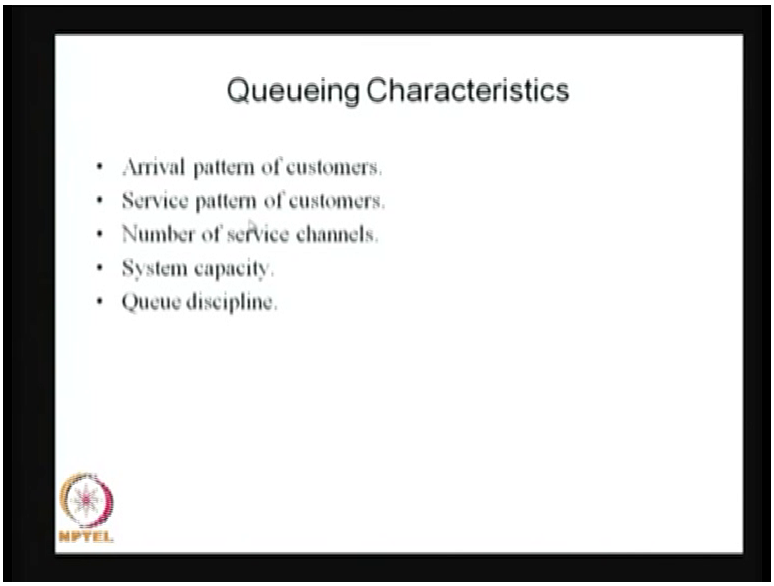
(Refer Slide Time: 07:53)



So, based on that, the Kendall made a notation and that notation is called the Kendall notation. The Kendall notation consists of capital A capital letter A slash capital B slash capital X slash capital Y and capital Z. So, the possible values we are going to assign for A B X Y Z accordingly, one can define the queueing system and each letter is corresponding to some important characteristic of the queueing system.

A denotes the arrival pattern information. Here the A denotes the distribution of inter arrival time the letter corresponding to the capital A. The second one B whatever the letters you are going to assign for the second one that denotes the distribution of the service time the way I have said the characteristic.

(Refer Slide Time: 09:01)

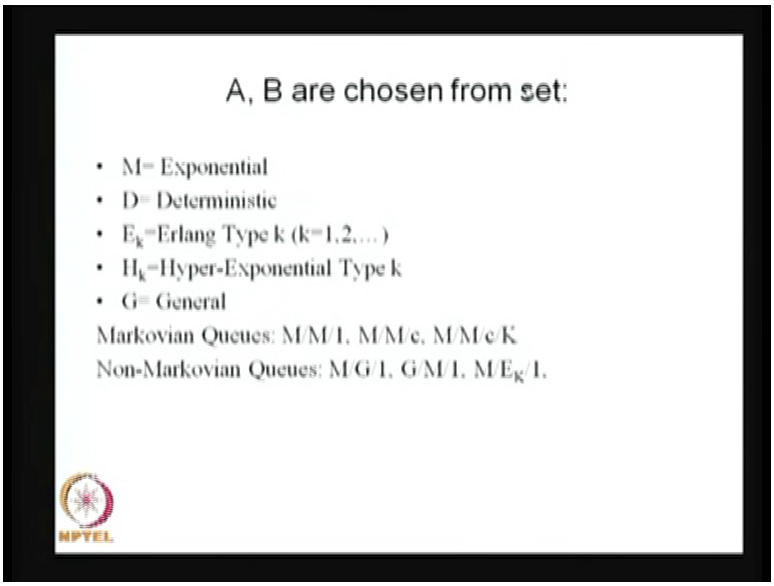


The first one is arrival pattern second one is service pattern and so on. The same way, we have given the Kendall notation.

So, the capital A is for the letter whatever the letter you are going to assign for capital A, that is for the distribution of inter arrival time and B is for the service time distribution. The third one X, whatever the number you are going to write, that is the number of servers in the system. The fourth one what is a capacity of the system. The fifth one what is the queueing discipline whether it is a first come first served, last come first served, priority, random and so on.

Now, I am going to give what are all the different possible values for these letters.

(Refer Slide Time: 09:57)



The first two A is for the distribution of inter arrival time, B is for the distribution of service time. The both can be chosen from these letters. If you write M in the first place; that means, the inter arrival time is exponentially distributed.

Even though it is exponentially distributed, we use the letter M because of exponential distribution satisfies the memory less property or Markovian property so to denote that we use the letter M. So, whenever you write M in the place of A or the second place B, then; that means, the inter arrival time is exponentially distributed or service time is exponentially distributed respectively.

Suppose, you write the letter D in the place of A or B; that means, that distribution is a deterministic; that means, it is going to take it is not a probabilistic. It is takes a constant amount of time whether you placed it in the first or second accordingly. So, it is going to be a constant amount of time going to take for the inter arrival time or service time whenever you place it in A or B respectively.

Similarly, if you use the letter E suffix k; that means, it is a Erlang distribution of type k or you can say Erlang distribution of a stage k that can be 1 2 and so on; that means, the inter arrival time is a Erlang distributed with the stage k if you place it in the first letter.

Similarly, H suffix k means a hyper exponential distribution of a type k, whenever you have a inter arrival time means a other than exponential deterministic and so on. So,

usually, other than exponential, you can use the letter G. G means a general distribution. General distribution is also it is a known distribution. The only thing is it is other than exponential distribution.

So, either you can use the letter M, D, E k, H k or G. So, G can be other than M itself in the visual or in general form, it is a known distribution that other than exponential, we use the letter called a G for general distribution. So, these are all the possible values for the A and B whereas, the third one is the number of servers in the system and the fourth one is a capacity of the system and the fifth one is the queueing discipline. The default discipline is a first come first served therefore, no need to write the fifth information.

And the sixth information is also there. What is a population of the customers who are entering into the system, the default population is infinite; that means, from infinite source the customers are entering into the system that is a sixth information. As long as we will not write as long as the system in which the population is infinite as well as the queueing discipline is first come first served, then we will not write.

So, we write only the first four information, that is a inter arrival time distribution. Second one is a service time distribution. The third one is number of servers and the fourth information is a capacity of the system.

So, in these examples, the inter arrival time and service time both are exponentially distributed by default, they are independent also. And the third letter denotes number of servers in the system so here only one server in the system. Here, C means ah, it can be greater than or equal to one that is a multi-server system and fourth letter K means, a capacity of the system. Suppose, we did not write the forth information here; that means, it is a infinite capacity system.

And this is also infinite capacity system and since the inter arrival time and the service time are exponentially distributed, this model is called the Markovian queues because, it satisfies the Markov property. Whereas, a non-Markovian queues either service time or the inter arrival time can be a non-exponential distribution or non-exponential distribution. In default, we can use a letter general distribution.

So, whenever G comes in the first place or the second place, then we use non or when we then we say, it is a non-exponent non Markovian queues and if the fourth letter is missing; that means, it is a infinite capacity system.