Communication Networks
Prof. Goutam Das
G.S. Sanyal School of Telecommunication
Indian Institute of Technology, Kharagpur

Module - 04
Queuing Theory
Lecture - 16
Arrival and Service Process

Hi. So, in the previous lecture which is lecture 15 we started discussing Queuing Theory. So, we have started understanding in a particular queue or a queueing system what are the things that are there. So, that is something we have discussed and then we have also started discussing how a queue interacts with the other events that are external to the queue.

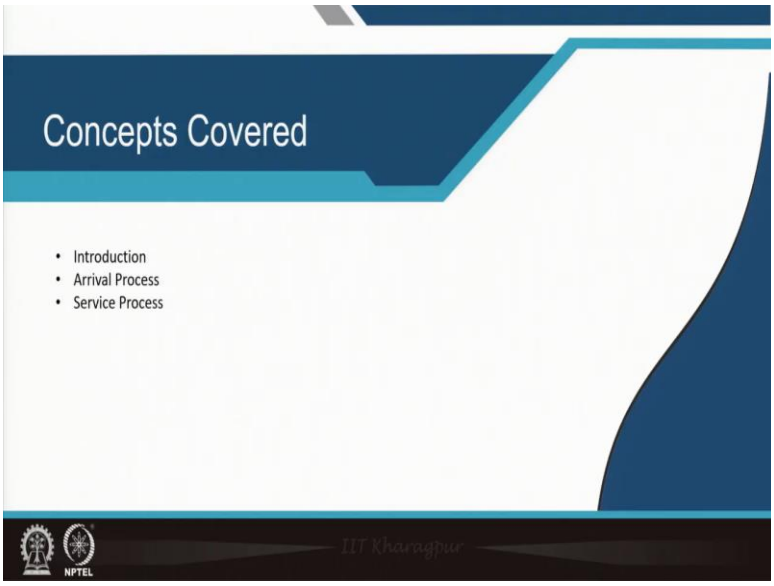
But what influences the queue, such as the arrival or the duration of the call or the duration of whatever service you are giving so those things? And we have also started discussing from there how you get a feel of the associated random process that goes through the queue. So, this is something we have started discussing ok.

So, basically what today will try to do? So, we have already identified that there are two things that are important in queuing, which are external to the queue, which is the arrival and service process. So, today we will try to characterize some of the very well-known arrival and service processes with their very means remarked properties, which have been used in deriving mathematical tools for queuing.

And you will see that it is not only just a mathematical beauty. So, it has a huge application because of all the telephony networks that we have been discussing like trunk calls and all trunk switch analysis we are trying to do. So, it has been seen that the arrival process follows the one I will be talking about which is the Poisson process, and the service process also will be talking about the exponential service process.

So, these things have been measured and it has been observed over the years that these things follow those kinds of distribution. So, it is not that it is just a mathematical thing that will be discussed these have huge practical applications and most of the trunk switches these days also are designed by the principle of the formula theory that will be developing now, ok?

(Refer Slide Time: 02:36)

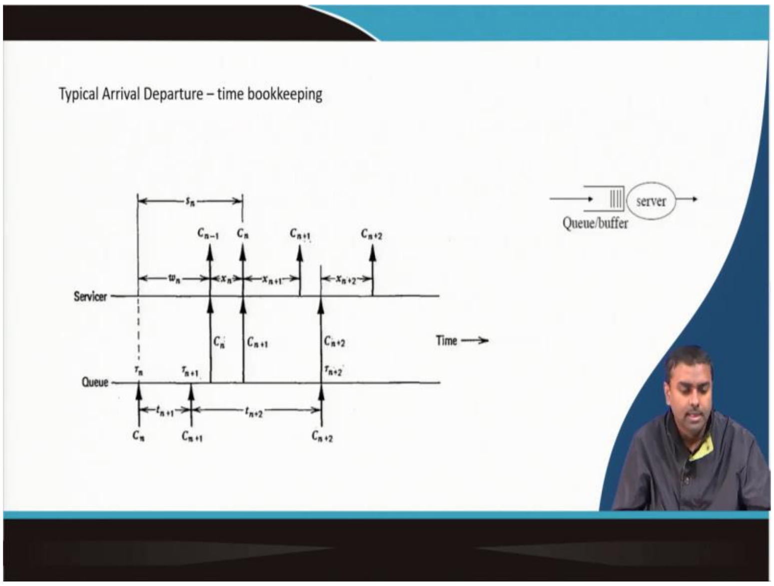


So, we have given some introduction as you can see now will start discussing the arrival and service process ok.

(Refer Slide Time: 02:46)

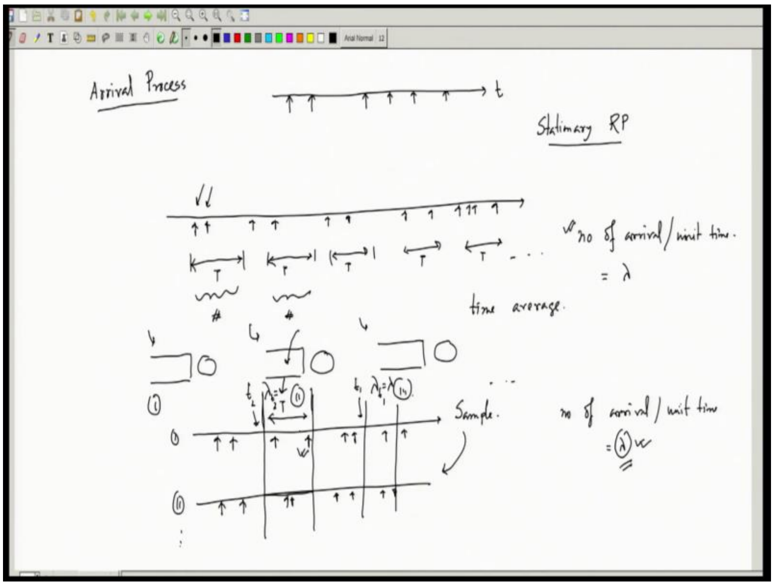
So, it is popularly known as Markovian queues and we will see with single servers and multiple servers that something we will see later on.

(Refer Slide Time: 02:56)



But we have also started discussing a typical queuing system where you have a queue and server and then we have started discussing the timeline of how customers arrive, how they interact with the queues what kind of series of events occurs, and what the associated randomness that is involved in this process. And then try to see, what are the basic things we can talk about during this arrival process ok? So, that will be the most important part that will be discussed in the beginning ok.

(Refer Slide Time: 03:43)



So, let us try to see what we are trying to discuss about a process called; so, the arrival process is nothing but in time if you see so, in the timeline. So, it is either packet or customer or call whatever it is they in the time frame they do arrive and generally, they are random because each customer is arriving at the queue, of course, it is to a particular service center switching center or queue.

So, it is it is a random arrival that will be happening and this is the account of these random arrivals with their timestamps ok. So, this is described by it is a time series; that means, different time instances where particular arrivals arrival number 1, arrival number 2, 3, and so on have happened.

So, its account of these things or bookkeeping of these things ok that is arrival. And because it as you can see is random so, it varies with time, the pattern does not repeat. So, it varies with time in a random fashion. So, it is an associated random process ok or it is described by an associated random process. So, what will try to do we try to characterize if we can describe the associated randomness?

And that random process if you can characterize through some mathematical description, but before describing the randomness you all know a little bit about random process. So, here will give a very brief introduction to the random process that we are talking about. So, over here the random process will be a stationary random process ok. So, what do I mean by stationary random process; that means whatever statistical description will be given that statistical description does not change with time.

So; that means, if you take all the statistical properties let us say you take the average you take the second moment, or from the second moment you derive the standard deviation so, these things do not change over time. So, at whichever time is supposed for a queue particular queue there are arrivals or let us say for a trunk switch there are customers who are arriving.

Now, this customer arrival process does not change or the statistical property of that process does not change with time; that means, if you take a certain time window let us say 1 second or let us say 1 hour. So, 1 hour in 1 day if you take and then try to evaluate how many customers are arriving at that 1 hour and then if you keep on moving and then try to see the average value of it, it will not be actually if you take average over ensemble it will not be changing over time.

Let us try to understand this ensemble average a little bit. So, two kinds of averages can happen. So, let us say I have a time window, where different kinds of arrivals are happening so, whatever it is. Now, suppose I have fixed a time window called T ok over this T I will be taking multiple samples and the way we derive the statistical properties is we take average means; we take actually in multiple windows we take samples and over that we infer the statistical property.

So, if you wish to take the average we take multiple such instances of T windows or windows of T length we calculate or count how many arrivals are happening so, that is probably counting. Then we will try to see what the average number of counts that are happening ok. Now, this averaging we can also do this in two different ways let us try to understand what these two different ways are that will give you a basic understanding of the random process.

So, even though you have been taught a random process probably in your earlier courses, I will still give a brief idea about it. So, what will be happening over here as you can see if you just keep on observing we can observe over multiple time instances, ok? So, basically you have recorded this particular arrival for a very long duration of time and then for different T over the time for the same arrival history that you have recorded.

If you take and then do an average over all these time windows so, basically you keep on doing that and then you take an average over all these windows and then try to give over T time how many arrivals are happening, or from there we can divide this by T so, per unit time how many arrivals are happening or it is called the arrival rate. So, the number of arrivals per unit of time is ok.

So, you will be evaluating the count. So, count you will be getting everywhere and then those counts you will average over the number of instances, and then you divide by T so, that per unit time number of arrivals you will be getting. So, this is a very important parameter which is called lambda. Now, all we have to see is whether this average is the average we are talking about. So, this is one average of course, but this is called the time average.

What are the other averages that are there? The other average is called the ensemble average. What do I mean by this ensemble average? So, the ensemble average is something like this. So, you prepare these kinds of queues as identical queues multiple such queues an infinite number of queues. So, you will be preparing all identical queues for which you know identical kinds of arrivals are happening ok?

So, say with the same property, with the same kind of customers it makes arrival ok and for each of this queue. So, let us say this is queue 1, this is queue 2, this is queue 3, and so on. So, for each of these queues, you record the time ok and you see the arrival and of course, for each one of them, there will be different, different cases ok? They are identically prepared; which means, statistically identical or the customers have similar kinds of properties.

But when they will arrive that is they depending on their whim right? So, you do that you record for 1 for 2 and so on. Then you choose a particular time window of duration T. Now, you observe how many arrivals over here in the 1st case this is now called sample one of the sample, you take another sample the 2nd one over this T again you see how many arrivals are happening and then you go across this is called the ensemble go across the ensemble. At the same frozen time, you go across the ensemble earlier we were doing.

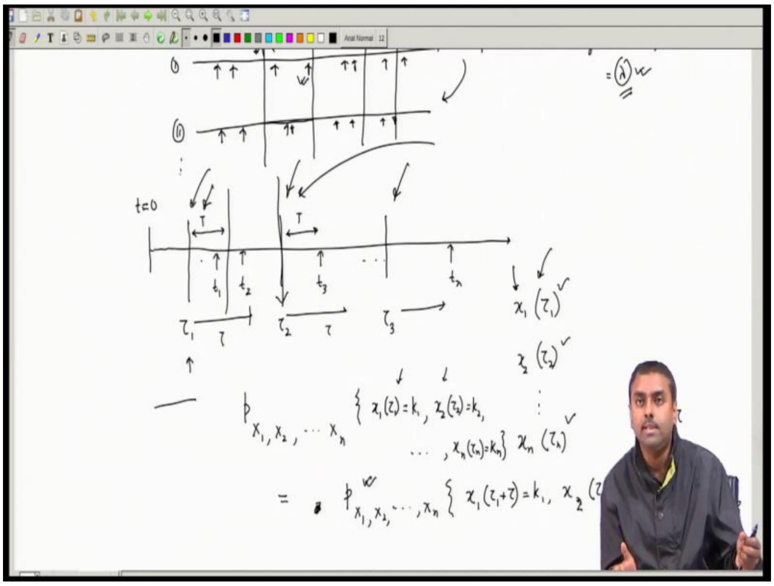
At the same frozen let us say sample we were going across the time so that was time average. Over here if we do the same thing again we try to calculate the number of arrivals per unit time which is called the lambda if we calculate. So, that will be the new lambda, which is the ensemble lambda this is where we are interested means this is what we are interested in. Now, what we should get is something like this.

When we talk about stationarity so, we should say that the statistical properties of these ensembles do not vary over time. So, it does not matter which time instance I take is it over here or is it over here; it does not matter which time instance I take irrespective of that time instance you are always getting similar statistical property. One of the statistical properties is this lambda, which is the average arrival rate so, that must not vary.

So, if this lambda T is wherever you go, that will be irrespective of T. So, if this is T 1 and this is T 2 they will be all irrespective of so, suppose this is my t 2 and this is my t 1 that first instance. So, wherever you start this one if the lambda becomes irrespective of this one and the subsequent other statistical properties also means all the higher order moments also this is the first order moment this is the average we are taking.

So, all the moments if they are not changing, or rather we should say can be described by a distribution, which is invariant of time, if this happens then we call this a stationary random process. Now, you might be asking what distribution we are talking about which needs to be stationary. So, let us give a little bit of hints about the distribution that we are talking about over here ok? So, what we are trying to do is something like this.

(Refer Slide Time: 14:09)



So, basically whenever we have these particular things ok. So, let us say I have some particular things which have happened. So, this arrival has happened I have time equal to 0 and this is the first arrival that has happened ok? So, let us say at t 1 I have something then t 2 I have something, t 3 I have something ok.

So, whatever is happening now, every time instance I can measure some property ok let us say my property is at any time instance. So, these are the arrival which has happened, but now I will be defining or let us just differentiate it may be this is the arrival I will be now putting some instant tau 1, some instant tau 2, where I am probing the system some instance tau 3.

From there I will be assigning some T amount of time and I will be trying to see how many arrivals have happened. So, this number of arrivals that count that is the variable ok, which is a random variable. How many arrivals will be happening at Tau 1 that might be a function of Tau 1 ok, so that is possible.

So, let us say that I have let us say some tau 1 at time tau 1 I have some x 1 number of arrival ok. At tau 2 so, basically, I take this time T and I have another x 2 number of arrival and so on ok. So, x n some tau n number of arrival something like that, ok.

So, these things are my random variable now. So, x 1 is a variable that tells that at the time instance where I have frozen it as tau 1 what is the number of every arrival that has happened? Now, because there will be multiple such samples. So, everywhere things will be varying so, this particular value will be random. It can take any value I do not know which sample I will be picking accordingly the value will be dependent. So, it is a random thing.

So, every time instance you freeze you will be getting some random value ok. So, that might be x 1 x 2 x n only thing is that it is a random process because it also depends on that time instance ok? At Tau 1, what will be happening? At Tau 2, what will be happening? How many arrivals will be happening ok?

So, now, suppose I am trying to describe because it is random so, I need to describe a probability distribution function. So, what is that probability distribution function? So, I will be now describing a joint probability distribution function. So, which is p let us call those random things as X 1 X 2 X n and what I am trying to do is that this x 1 let us call it tau 1 is equal to some let us say k 1 x 2 tau 2 is equal to some k 2 x n tau n some k n, ok.

So, this joint distribution I am trying to evaluate. So, at a given tau 1, tau 2, tau 3 like this. I can like take an infinite number of time instances so, basically, my instances will be infinite. So, these infinite time instances whatever values I will be getting are joint distributions I am trying to calculate ok. It is a joint distribution at every instance you take them and it is a joint distribution; that means, that at tau 1 x 1 number is arriving.

So, let us call it maybe 1 or 2, 1 arriving at tau 2, 2 arriving at tau 3; 1 arriving something like that this all possible joint distribution I am trying to calculate, that what is the probability that this whole event will be happening. The joint distribution we are interested in is ok.

Now, these joint distributions must be stationary what does that mean, that means even if I shift my time instance; that means, these joint distributions I calculate like this X 1 X 2

X n at let us say x 1 tau 1 plus some amount of value I wish to put tau equals to k 1. Sorry I should not write equal to now, what I will be saying is that should be making it equal to tau 2 plus tau.

So, everywhere I am shifting this tau 1 by some amount of tau, tau 2 also I am shifting by some amount of tau ok. So, the time instance their relative differences will remain the same only thing is that the whole frame will be shifted by the amount of tau. So, these things in this joint distribution must be equal to the previous one for every value of tau ok, this must be true.

Once this happens; that means, the joint distribution is basically wherever you take your time the joint distribution remains the same ok. So, if the joint distribution remains the same all other lower marginal distribution will remain the same. So, therefore, if I just freeze one particular time instance I do not take this joint distribution only tau 1 I take I do not have tau 1, tau 2, tau 3, tau. If I take only tau 1 which is the marginal distribution from the joint distribution I can derive that ok by successively integrating ok.

So, if I take any marginal distribution at any time instance tau 1 ok or tau 2 or tau 3 for instance. So, there if I do; that means, I am freezing the whole time or concentrating the whole time into one-time instance, and then I take the distribution. So, that is when you will be getting the first moment ok. That is when as you can now see all these time instances just get concentrated over here.

And that is what we were doing for average and we were saying that this average irrespective of where I take will remain the same ok. So, once I do this complex joint distribution of infinite time instances if that is stationary then the subsequent marginal distribution will remain stationary. So, therefore, in the first moment, and second-moment autocorrelation everything will remain stationary.

So, this is the random process that we are discussing about. For arrival when we say mathematically we are trying to give a description this is what we are talking about. You might say practically let us take the call arrival right? So, that call arrival might not have this property you might say you might argue that call arrival probably might not have this property.

I agree probably over a day it might not be because there are busy hours there are let us say not so, busy hours. So, let us say in an office area if you are trying to see the call arrival. So, in the daytime it will be very busy because everybody is in the office, they are trying to use they are trying to make calls. So, at that point probably that office area will be very busy in the morning.

So, where the average amount of calls coming per unit time will be much higher than in the evening or maybe late in the night or early in the morning right that might happen. But over the days if you try to see it at the same time instance it will remain stationary.

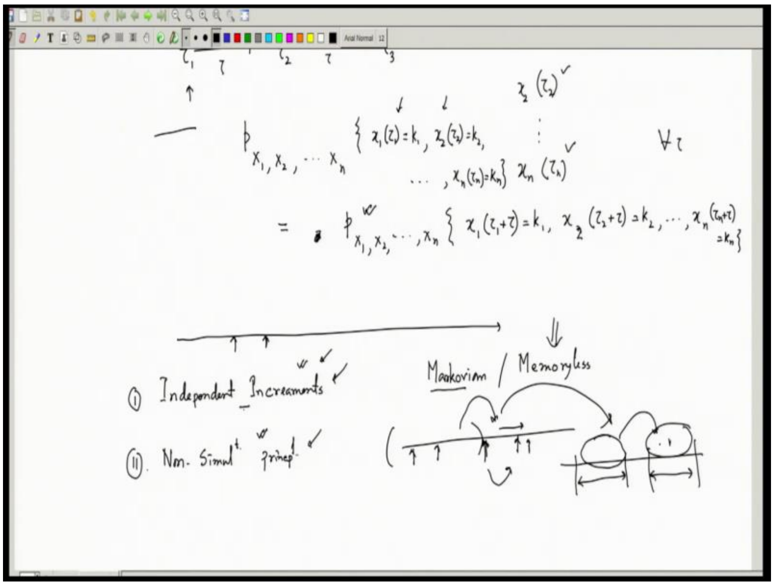
So, an almost similar kind of average rate will be happening if you are not building up new buildings over that office area. So, as long as the other things remain the same, you will be seeing that probably the average arrival rate will remain similar, ok? Not only the average arrival rate the other properties also will remain similar. So, if we have a corresponding stochastic process ok.

Now, if we have understood that we can take fairly well this assumption that it is a stationary process that we are trying to describe. So, now, what can we say more about this stationary process? Can we try to derive this distribution this joint distribution or some distribution that describes those call arrival processes I get if I can get that then only I can go ahead and do some analysis.

Then only we can we can try to say ok for all kinds of random arrivals I can still say what will be my blocking probability given this system design. So, to system design; that means, the queue and server are designed from there we want to get a relationship concerning the input process ok.

Whatever the input random process we take from the system design can we now claim that this will be the blocking problem for this particular stationary input process so, that will be our target. Now, let us try to see if we can say those things. So, for those things, we might require some more property which is very realistic, but it fundamentally will give us some mathematical insight.

(Refer Slide Time: 24:35)



So, let us try to understand those properties. So, I will list down two properties actually that will be important for us. 1 is called independent increment and the 2nd thing is called nonsimultaneity principle, ok. So, these two things I will describe these fundamentally because once you understand these two things you will see that very nicely the whole distribution will be derived just by assuming these two things.

So, let us try to understand this independent increments and nonsimultaneity principle. So, I will start with the second one because that is the easier one. So, what is this nonsimultaneity principle? Non simultaneity principle says that over here what we are trying to do we are trying to account for events.

So, the nonsimultaneity principle says in this world this is true actually. In this world no two events; however, geographically separated they might be; however, distributed they might be no two events occur simultaneously, it is almost like the Paulis exclusion principle. So, it says a very fundamental property that in time no two events can occur simultaneously this is impossible this will not ever happen why do I say so, is this making sense?

So, we might be thinking that ok if I arrive it might be possible that I take one of my friends and we arrive together in a particular room. So, are we not simultaneously arriving? You might argue that way, but what I can say is that you know that 100-meter race people used to do this photo finish right.

So, before the photo finished was there people used to think that ok, there was probably a simultaneous arrival; that means, two persons have arrived at the same time at the touchline, and probably they are too should be both should be declared first or second, and so on.

But after the photo finished people started understanding that no maybe if I make my time smaller and smaller maybe a femtosecond interval if I try to observe then will be able to see that even between those two arrivals there are some; however, small it is there will be some differences. So, if you go into the minuscule of time intervals from femtosecond to nanosecond to so on you will always see there will be a difference between two arrivals.

So, that is the basic fundamental property that nothing no two events in this world can happen simultaneously. So, that is the nonsimultaneous principle ok, we will see the mathematical description of that in the next class. And then the other thing is called the independent increments this is also a very fundamental property of this kind of arrival they will be classified as a special kind of arrival.

This is called something we are enforcing this is not like the nonsimultaneity principle that always happens. This is something will be enforcing and this is something we will be calling Markovian property or memoryless property. So, it is called Markovian or memoryless property ok? Markovians came because of the person who has innovated these things.

So, Marko his name from his name, and it is memoryless you will later on see that this is something called where the history is erased these processes are special processes where the history means the future events. So, suppose I am staying at this all the future events if I try to predict because these are all probabilistic events. So, I can always try to predict from here I try I can try to assess or guess what is the probability that this will be happening so, that I can always do it.

So, when I am trying to predict my entire history about what has happened previously if this has some influence in my prediction then that is called the memory process, but if it does not have any influence; that means, only by seeing the present I will be able to predict what will be happening in future; and I do not have to know what has happened in the past. So, if it is not dependent on past events or past occurrences then this is called memoryless.

The corresponding associated property will be independent increments or independent increments if we define independent intervals then we will see these independent increments what does that mean, in a timed event, we can define two independent or nonoverlapping intervals.

Then we can again say this is the very generalized principle that that whatever is happening over here has nothing to do with or no dependency over whatever is happening over here and vice versa so, that is called the independent increments. So, these two things will be mathematically defined in a more detailed way in the next class, ok?

Thank you.