Performance Evaluation of Computer Systems

Prof. Krishna Moorthy Sivalingam

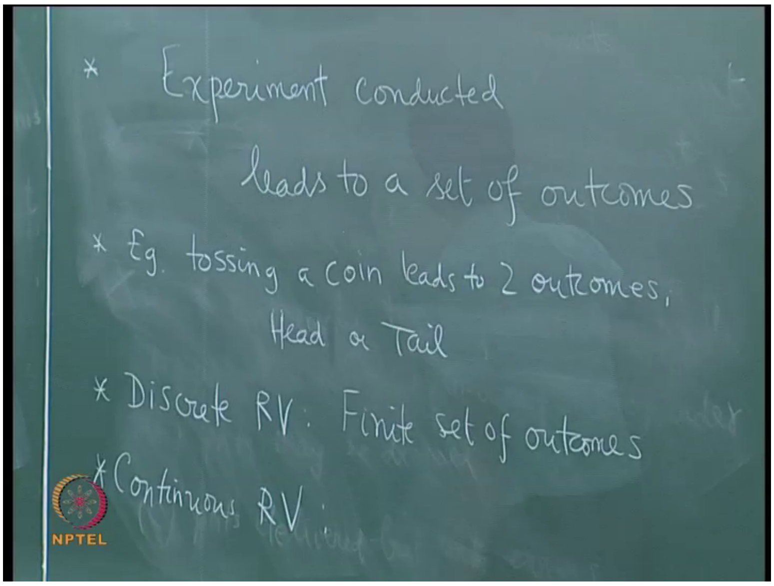
Department of Computer Science and Engineering

Indian Institute of Technology, Madras

Lecture No. # 05

Random Variables and probability distributions

(Refer Slide Time: 00:10)



So, we look at commonly used distributions. So, how do you, what is the random variable?

(( ))

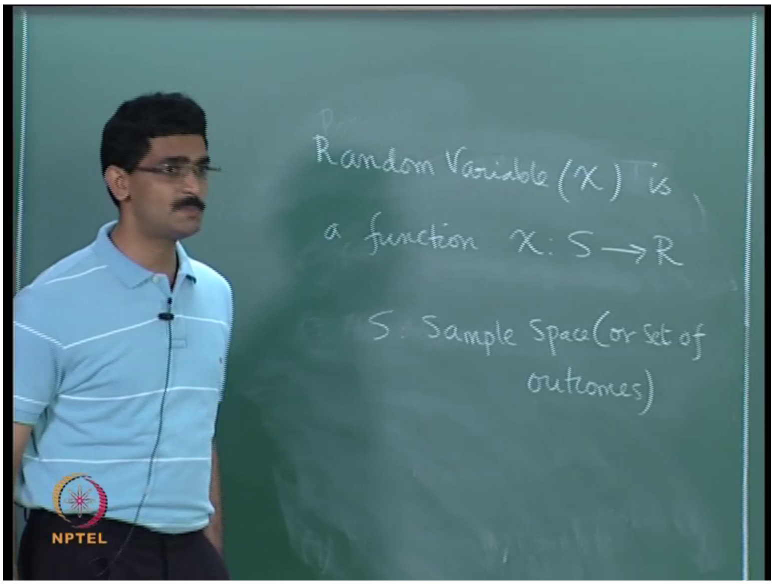
It is a mapping from sample space to real numbers good, this is in your (( ))

So, you normally think of when we design a system, we conduct experiments on the system; so you have an experiment and experiment will have a set of possible outcomes. So, that is where we started. So, you have an experiment conducted these two are set of outcomes. Simple experiment is tossing a coin, and the two outcomes are heads or tails. So, in that case, your set of outcomes is just 2 or if you repeat the experiment thrice, toss a coin 3 times, and then you count the number of heads; that is your outcome are again 8 possible outcomes there or may be multiple possible outcomes are there. So, you have this notion of a set of outcomes; for example, some of these, I have to write because this is also for NPTEL. So, if I have actually it is not heads or tails its head or tail only one kind of that solve. If you role a dice with 6 sides outcome is 1 to 6, then what is the discrete random variable, continuous random variables? The set of outcomes is countably finite, then you have a discrete random variable, where your random variable mapping, where only have from subset of will be subset of several each of these outcomes we call an event is going to be map to some corresponding probability.

So, if you are measuring in this case, if I toss a coin 10 times and I count the number of heads that are you going to have from 0 to 10? That is yours all possible outcomes for this particular tossing of a coin number of times, if I look at system where I measure delay, average delay of a packet that is no longer countable, it can be any value from small value smallest possible relay to highest possible relay depends upon the number of buffers number of other packets in the system and so on; that for that becomes a continuous random variable. So, you have two classes discrete and continuous, which by definition one is discrete one is continuous. So, I cannot define any further.

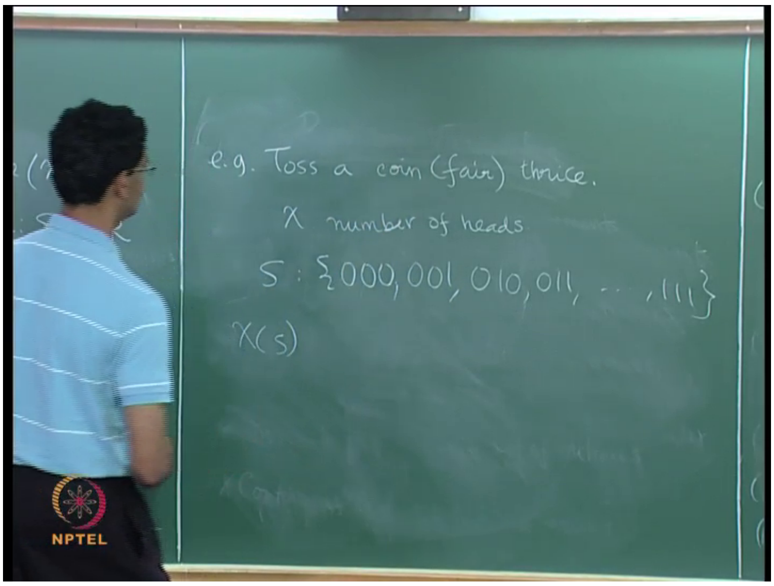
And when I say finite and fledging there, because it is countably finite, for example, number of packets coming per unit time number of packets number of customers arriving to a system per unit time is technically discrete random variable, but then it could be 0 through infinity. Therefore, it is not exactly finite, but you know what we talk about integer set of... So, you look at discrete variables first and then later on look at some other common continuous distributions.

(Refer Slide Time: 04:46)



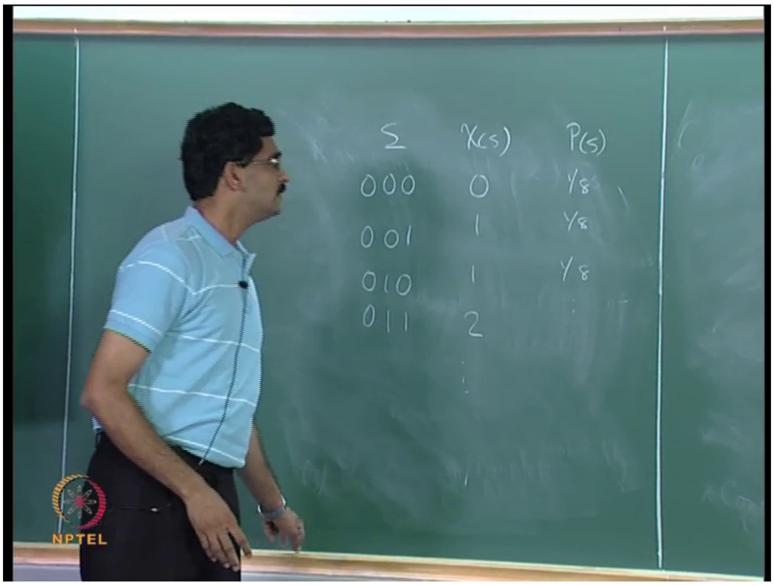
So, a random variable say call this x… (No audio from 04:55 to 05:09) is the function that maps your domain, which is the sample space the set of possible outcomes to real number which is the probability of each of those outcomes actually happen then the experiment, takes place. So, S is your sample space which x basically associates every sample or every possible outcome the probability of whether probability of that particular outcome happening when the experiments...

(Refer Slide Time: 06:05)



So, we will just look at one quick example it is for example, you toss a coin and let us assume as a fair coin thrice, let x be the number of heads. The number of times that heads appears; in each experiment has actually 3 trials within the same experiment. So, what are the possible values what is the sample space here… and the corresponding x of S is will be 0, 1, 1, 2. So, it is not corresponding value of S y.

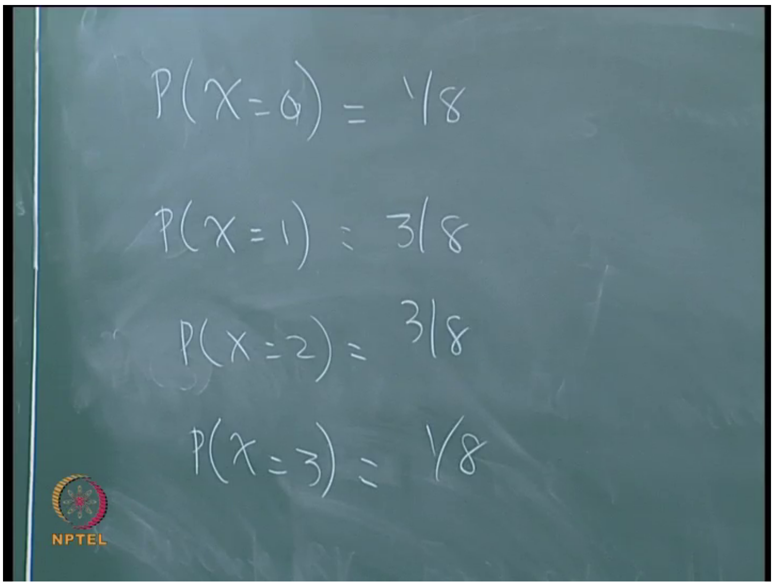
(Refer Slide Time: 07:29)



This is where I will have to redo this fine.

Sir, between1, 2, 3 and thereafter the probability of s as 1 by 8, 2 by 8…

(Refer Slide Time: 08:37)

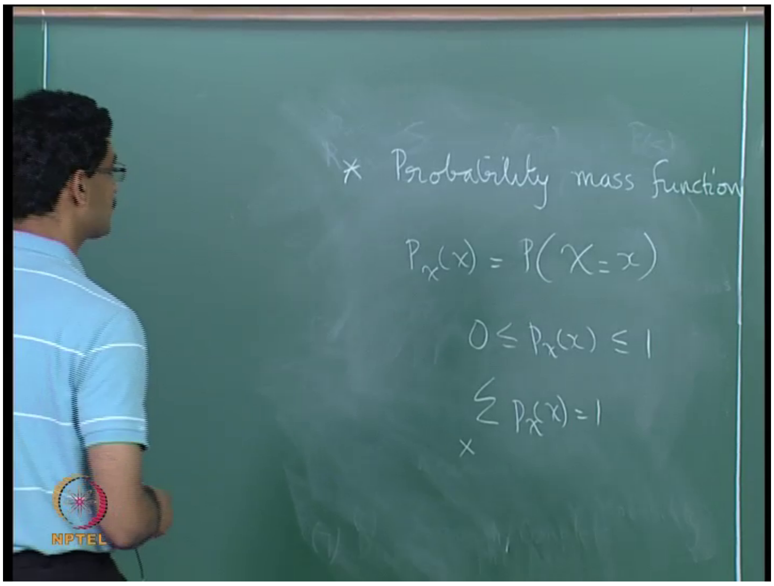


So, actually favor to say probability that X equals to 0 would be…

We can compute this is no 8. So, that is your mapping function that maps from this is your mapping function that maps from each that maps each outcome that the corresponding probability and then we have distribution for the random variable itself. Let us get into the few more definitions 6…

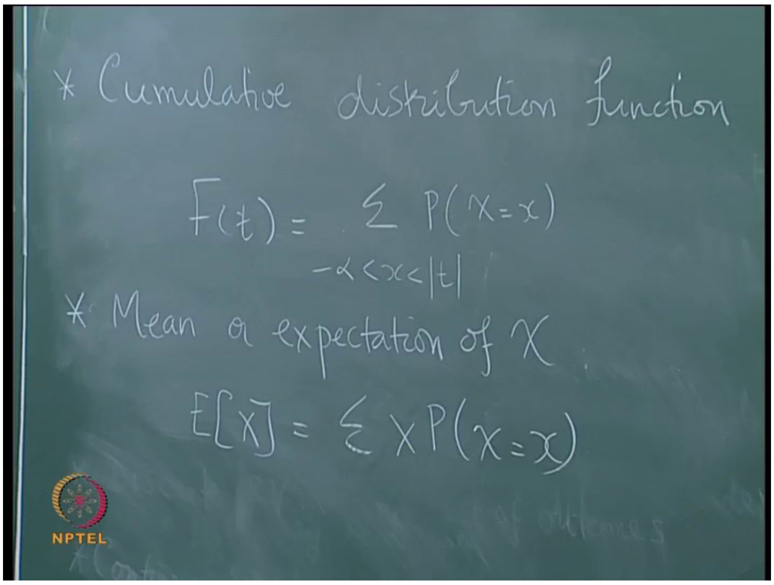
Make sure that (( ))

(Refer Slide Time: 10:06)



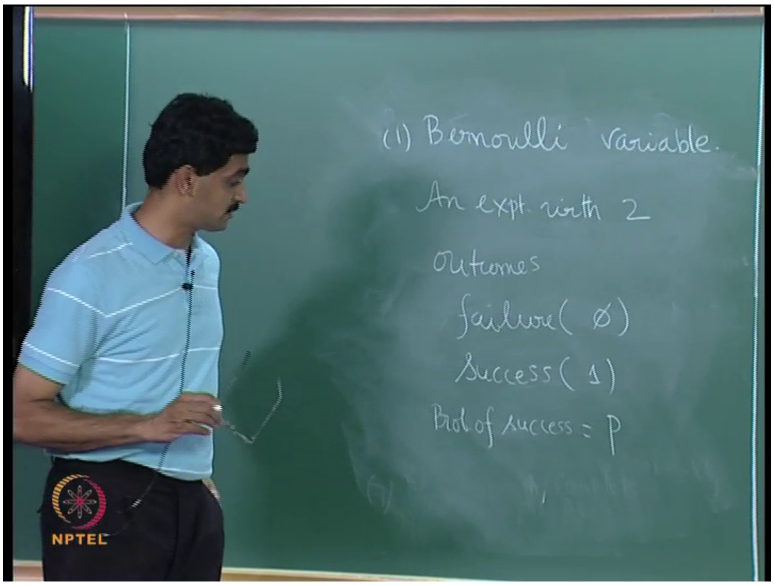
So, probability math's function is use the some of these notations are from trivet's book, which is little bit different from here, but for now will just use this is the probability that. So, variable x takes and the value x. So, the mass function defines the probability that the random variable takes a given value x and that we will see examples of some of these as we along.

(Refer Slide Time: 11:28)



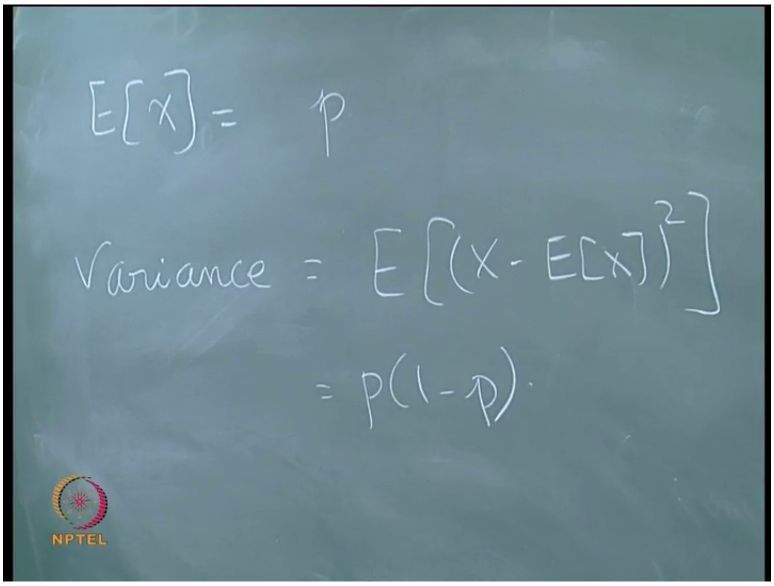
The new also have the cumulative. So, f of x is (( ))… So, then you have your mean or expectation. So, the mean is again x into I should be again this I have to redo this part. Value of x in upper case lower case is not looking good on the same thing, with this basic set of definitions. We now look at some of the common variables at many of you probably remember from your other codes as the first.

(Refer Slide Time: 13:46)



So, the first variable let you see often used see the Bernoulli variable. So, this is in general any experiment with us two outcomes. So, where I experiment with two outcomes and which we will denote as failure value 0 or success the value is 1 and the probability of success is cannot specified. So, this is not your fair find task this is simply some problem probability of an experiment succeeding or not succeeding. So, when you send a packet on the network the packets either gets delivered or it gets blocked or if you just look at delivery as dropping then the probability of successful delivery this speed or dropping is 1 minus speed . So, this would be I got it.

(Refer Slide Time: 15:25)



So, what is the expectation for this Bernoulli variable? What is the expected value?

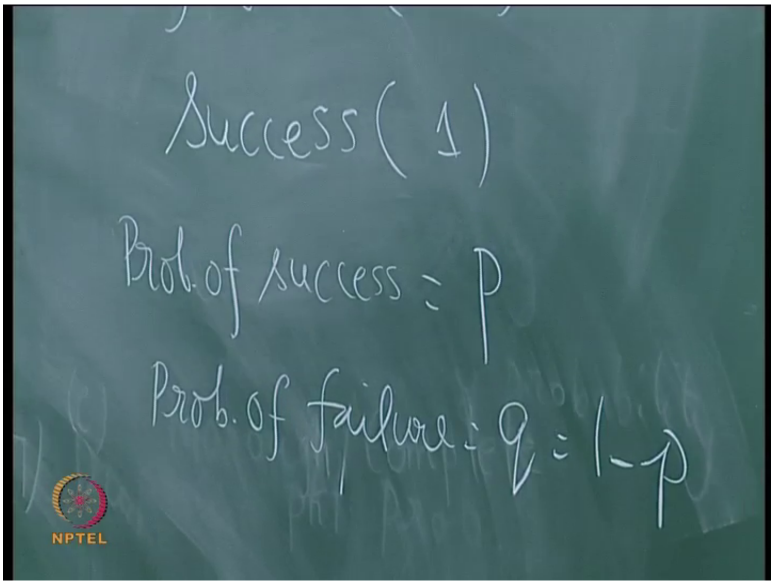
(No audio from 15:29 to 15:46)

Then we have not defined variant got the variance defined as could you define variance?

Normal the variance is also p (( )) p to…

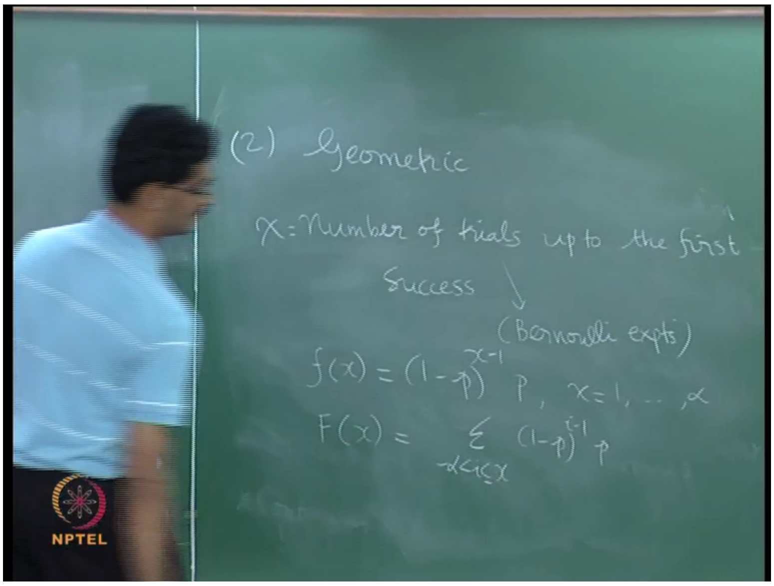
Yes is that variance definition. So, this again we will not spending deriving that it is P into 1 minus P.

(Refer Slide Time: 16:47)



And sometimes we also define probability is defined as q which is nothing, but 1 minus P. So, depending on where these are needed will use this system q.

(Refer Slide Time: 17:07)



So, what is the next possible extension of Bernoulli if I do repeated trials in what are the other variables? Said you can derive from Bernoulli or one is a geometric variable the geometric variable is used when you repeat Bernoulli some number of times repeated x times n times then so on until the for success. So, that is what you like to the represent to the geometric variable. So, this represents the number of trials first success (( )) whereby trial we mean trials Bernoulli experiment is what we try to do here.

So, what is the corresponding math's function probability math's function for this? So, I have x minus 1 failures and finally, one success. Therefore, this is 1 minus P, minus 1 then finally, one success and x can take values from 1 to infinity, anywhere need at least 1. What is f of x you are c d f give these vary of summation still starts coming and what is the summation of this?

1 minus T whole bar

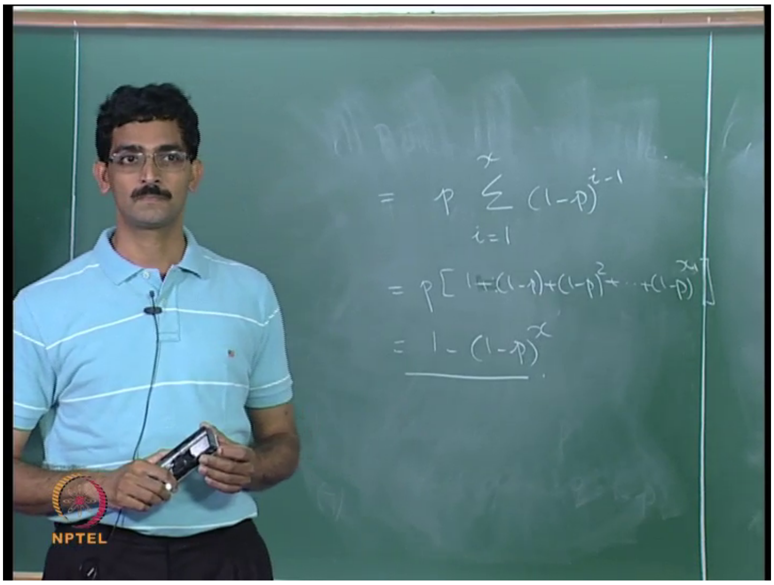
1 minus

1 minus T whole bar

Sure you are guessing or…
(No audio from 20:16 to 20:32)

Well will skip the minus infinity part at (( )). So, will at least x that have less than or equal to x.

(Refer Slide Time: 20:44)



So, we put as you it…
Does not… what is the expression for that?
What is the geometric series.

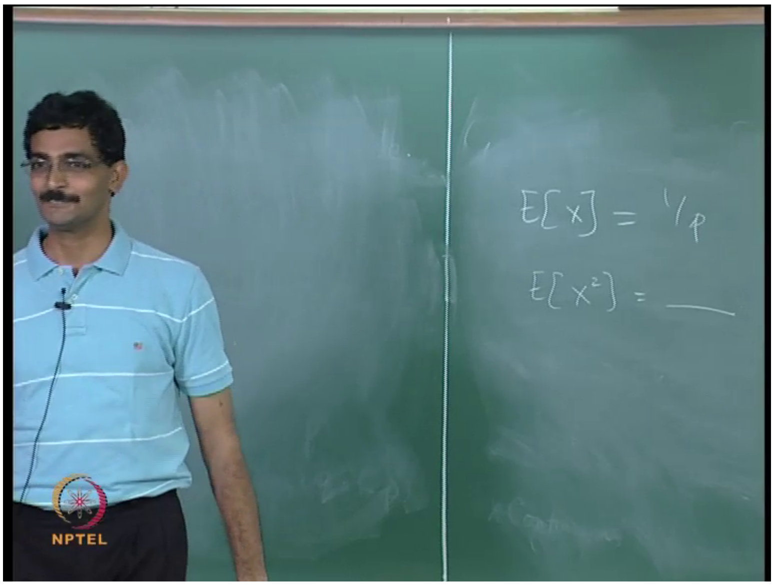
It is 1 minus all the first experiment values. So, it is 1 minus (( )).

This is the...

Because (( ))

So, this is 1… This will be come to the formula for that then 1 minus a r divided by 1 minus depending on. So, you r terms here or x terms in this case it will come to finally what you are saying?

(Refer Slide Time: 22:59)

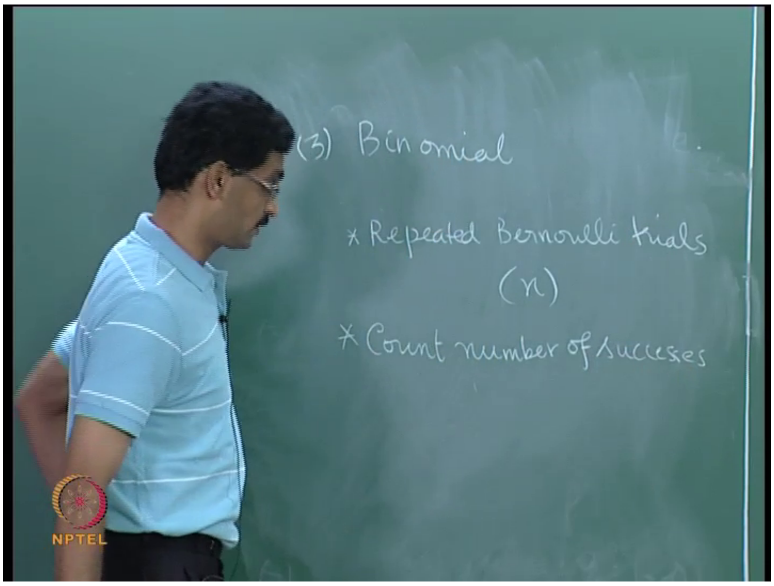


But that is what 1 by p write it that if you are not satisfied if you go back to your basics. When you use these often in when you to deal networks in packet transmissions. So, that is now repeating the Bernoulli's trials until we get the first success. So, that is your geometric series than when you find these commonly used in the context of systems networks. What real of example do you have for geometric variable?

Number of free trans…

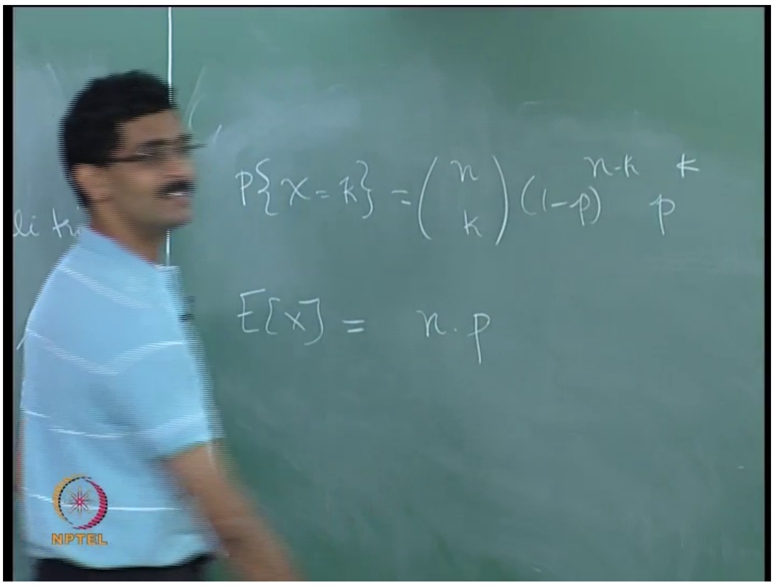
Number of free transmission it is the packet and it fails probability is 1 minus p if you are free transmit. So, finally, the average number of retransmissions need at the number of retransmissions in the case of repeated there in the case of link that is got errors and so on. That is what geometric series will try to represent it you know number of retransmissions needed. So, therefore, there if you are is it 1 hour p number of transmissions, total number of transmissions 1 minus p. So, back my basics and derive this again yes if p is 1 and can we need 1 transmission not convinced of this go back again can derive that.

(Refer Slide Time: 25:20)



Then your third variable is your binomial where again I do repeat binomial trial Bernoulli's trials. So, will have n such trials. So, n trials and then we count the number of successes. So, x is the number of successes when I conduct the experiment jacket or n times.

(Refer Slide Time: 26:33)

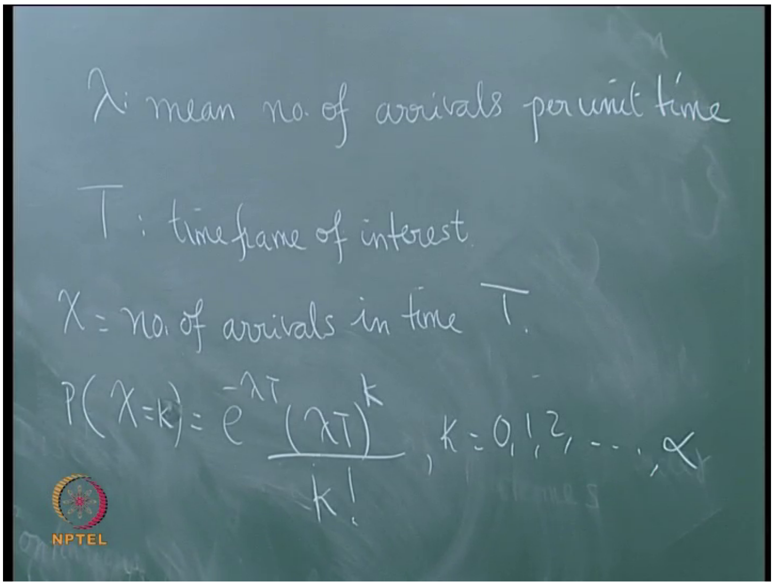


So, the probability that x takes value k…

(No audio from 26:36 to 27:06)

This is what we did with the first example when we have trial toss the coin thrice in this case I am conducting n times a this is your classic binomial and what is the expected value can be. So, again we derived if you want to, but since a when trials and it seems reasonable that p be the probability of success every time should be n p.

(Refer Slide Time: 27:37)



There is also negative binomial which do not use a whole lap. So, we leave it at that when we come to the classifies on process which is use the whole lock mainly because its usably for lots of derivations its mathematically attractable and therefore, people tend to like this Poisson variable and so on.

(No audio from 28:04 to 28:24)

So, we will define the Poisson slightly differently. So, if lambda these is the parameters is the represents the main number of arrivals. So, I have an arrival process am sitting on counting the number of arrivals or the number of packets coming per second all the number of request to cash per unit time and so on.

So, it is a counting process that keeps track of the number of arrivals in a unit time and lambda characterization when in each of these you will find there is a parameter character. In this case n is the number of trials feast probability of success. So, likewise here this is the mean number of arrivals that is either empirical measured or as just assumed to exist. And then you look at sometime T sometime frame T where T is the time frame of interest.

(No audio from 29:24 to 29:34)

This is a T is usually considered to be very small, but for just assume that T some time frame that is interested there and the Poisson process represents the, what we want to do this find the probability of x number of packets arriving in this time T.

So, x is the number of arrivals in time T (( ))…

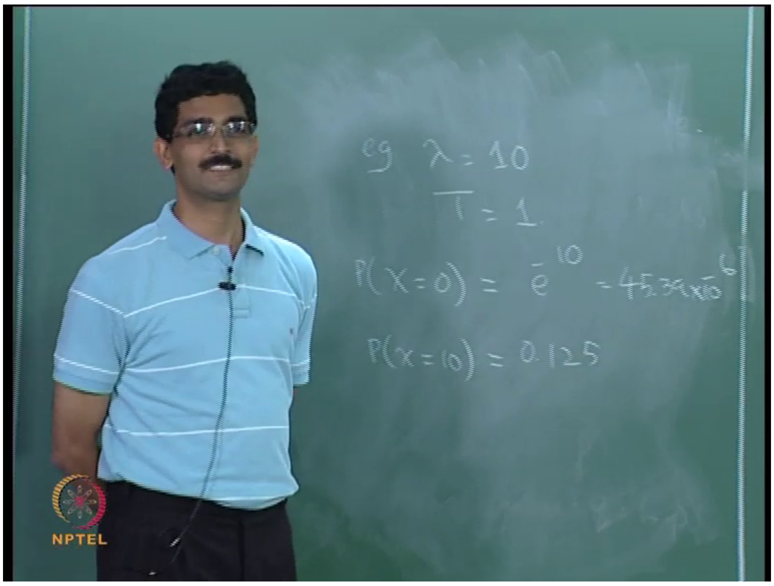
(No audio from 29:53 to 30:40)

Where k gives goes from…

(No audio from 30:45 to 31:05)

So, T power minus lambda T lambda T it part that verify k factorial that is the probability of having k arrivals in this mean time T. If T is 1 this is the possible of lambda into lambda to the power k. So, lets us this is probably the more important one and what where it couple of more examples.

(Refer Slide Time: 31:29)



So, will say that say lambda is 10 and T is 1. So, many times you want to find out the probability of x number of packets to coming for unit time. So, what is the probability that, no packet will come to the system in 1 time unit. Even though lambda is 10 there are still non 0 probabilities that will be no packet coming in this. You retrieve window of time 1. So, what is that value e power minus 10? Very very small value next time we have calculate as there 6

0 4 students of their need the calculator 45.39 into 10 power minus 6 negligibly small. So, that is your and if we look at probability of 10 packets coming in unit time will be close to 1…

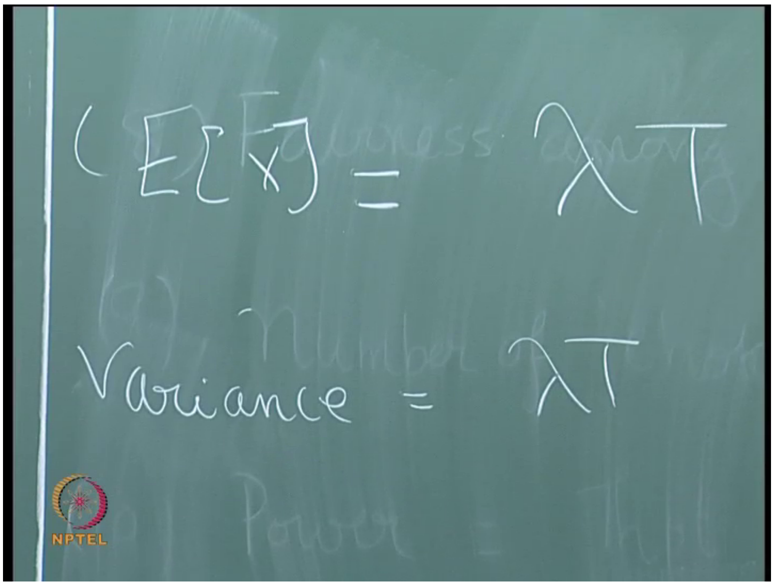
(No audio from 32:57 to 33:37)

Anyone calculate point.
(No audio from 33:40 to 33:57)

125 point 125.

125, yes as the way that is. So, even though that are 10 mean arrivals actually probability of getting 10 arrivals in unit time is point 125. So, we can actually plot this and see with respect to k for the system how these plots of...

(Refer Slide Time: 34:33)



And the expected value of X is lamda. That is known that if you want you can go back in derive that. So, expected value of X is lambda which is basically number of packet per unit time that is what the definition of metric S of X. Now, the interesting thing is like to calculate the variance, now the system. So, what is the variance of Poisson distribution?

Lamda

So, in this case E of x is there this particular case that…

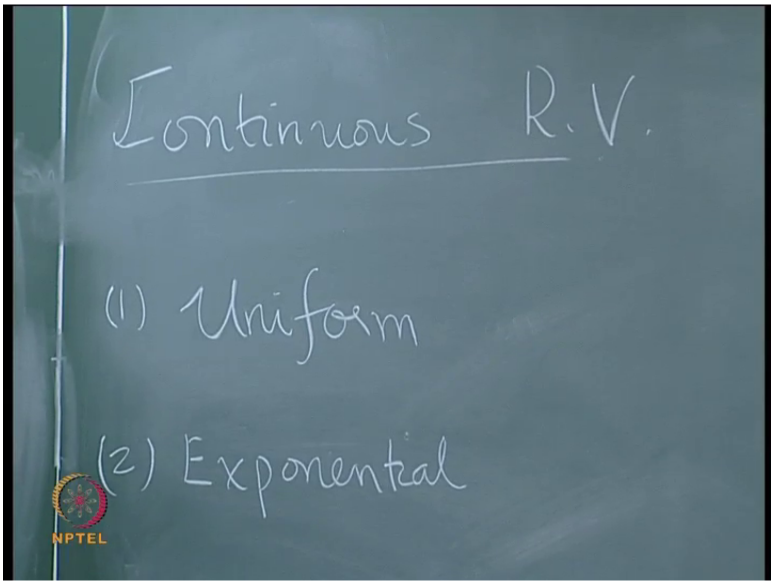
Variance also lambda T

That is we derive that look like to guess it. So, not this is the derivation for this. So, variance is also lambda T.

(No audio from 35:37 to 36:01)

So, that is the Poisson distribution we will visit this with respect to exponential later on. There are lot of other variables like the negative binomial beta and all that which know. We do not need to look at that is the set of discrete random variables.

(Refer Slide Time: 36:17)



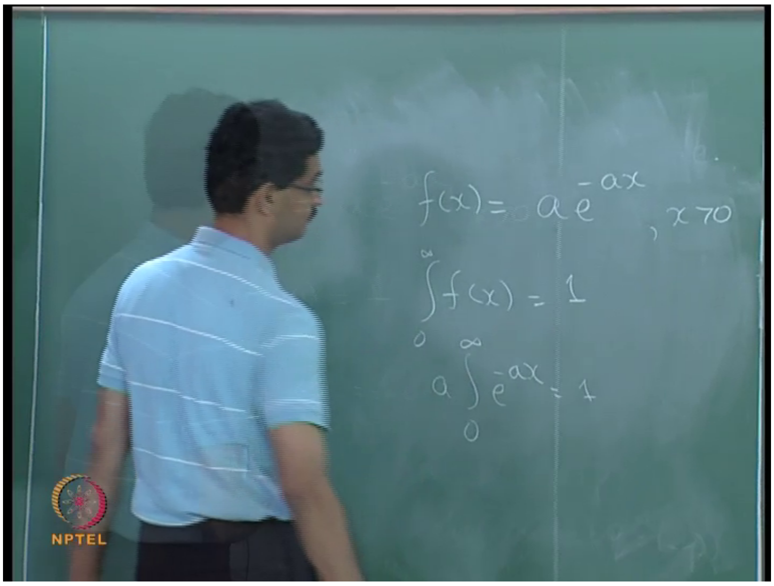
Now will go to get my correctly.
(No audio from 36:18 to 36:42)
So, for the continuous random variable again you have the p d f and c d f probability distribution density function and they are c d f. So, look at couple of things will start at that point. So, first is the uniform distribution which is also there in the discrete case, but it is also if the uniform distribution which will look at later than the more interesting on is your exponential distribution. So, exponential again is very convenient for analysis and we can prove lot of nice properties with respect to exponential. So, we look at that the definition for exponential you will which some might have seen and the other codes.

So, what is the definition for exponential? Remember what is for exponential variable?

(( )) lambda E power lambda and…

I will not lambda that use lambda as a little bit confusing. Alpha mu we generally used mu, because we fine that as we it is 1 hour mu.

(Refer Slide Time: 38:05)

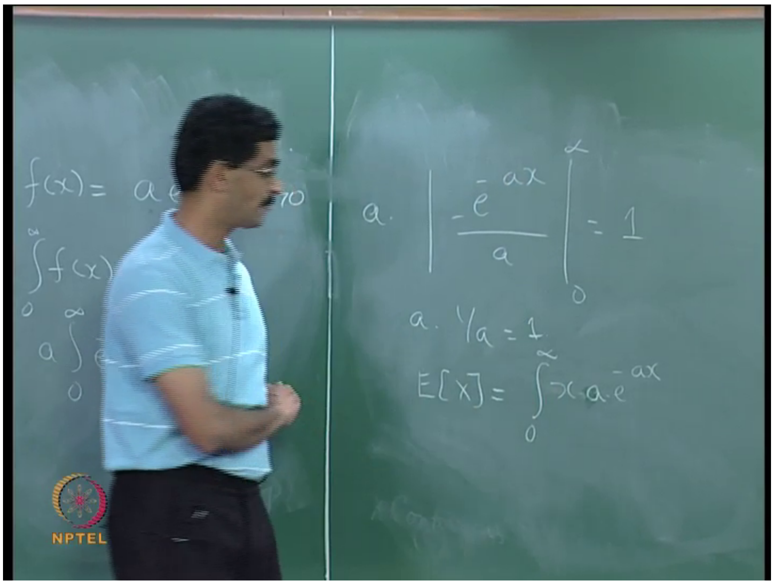


So, I will just use a e power minus a x if you 1 alpha is simply good and the only condition is that x is possible. Because that looking a time between used the use as to modal the time for example, inter arrival time between 2 packets what you modal that that is the random variable and when it describes it is an exponential variable we have some nice properties. So, here a represents them some parameter you just like lambda represents the mean number of arrivals we will see what a represent. So, what is a represent?

Inverse of the mu.

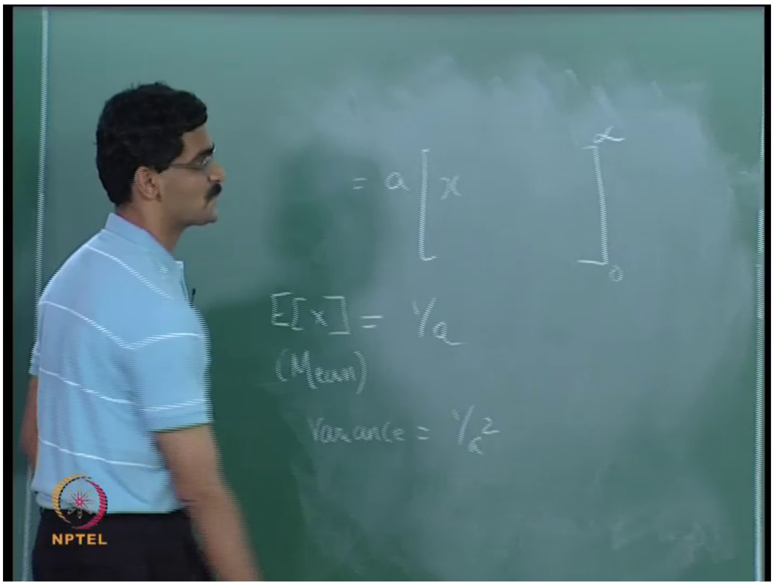
So, mean of this is 1 of a. So, I guess with some of you know some of you do not. So, let us figure this out so, that is accepted. Some of all the f of x integrates from 0 to infinity the goes to 1. So, therefore, that integration this will remember what is we have taken that. So, be a long times to serrate my deviation.
(No audio from 39:50 to 40:10)
So, therefore, what is it work out this a is 1 by for that is good. So, 1, that is…

(Refer Slide Time: 39:30)



Therefore we know that will integrate up to 1. So, my E of X is x e power... So, what is the mean of this distribution? What is the mean value taken by x? So, again I guess I should have said that dx plus infinity and of course I forgot my dx here.

(Refer Slide Time: 41:14)



(No audio from 41:09 to 41:27)

So, what is x e a minus a x integrates to...

(No audio from 41:33 to 42:02)

What your u v (( ))

Which is u, which is v?

u is h sequential u

u is x

If you d become 1 and…

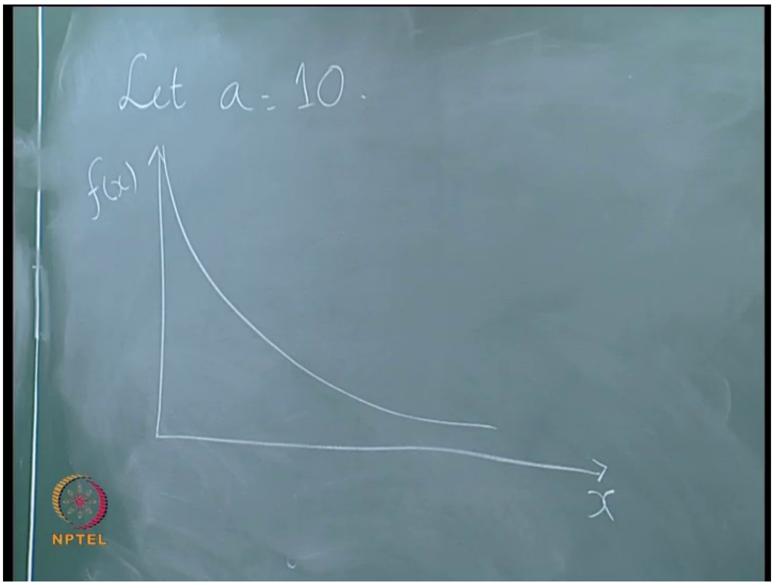
(No audio from 42:31 to 42:46)

Let you go through that the measure you later on an actually derivate it will come to for now a day it is for that is way your expected a represents inverse to the mean or mean is 1 about their. So, that for E x is I should not be saying E of x there this same mean…

(No audio from 43:12 to 43:24)

and the variance on over this for pictorially. What is this look like this?

(Refer Slide Time: 43:41)



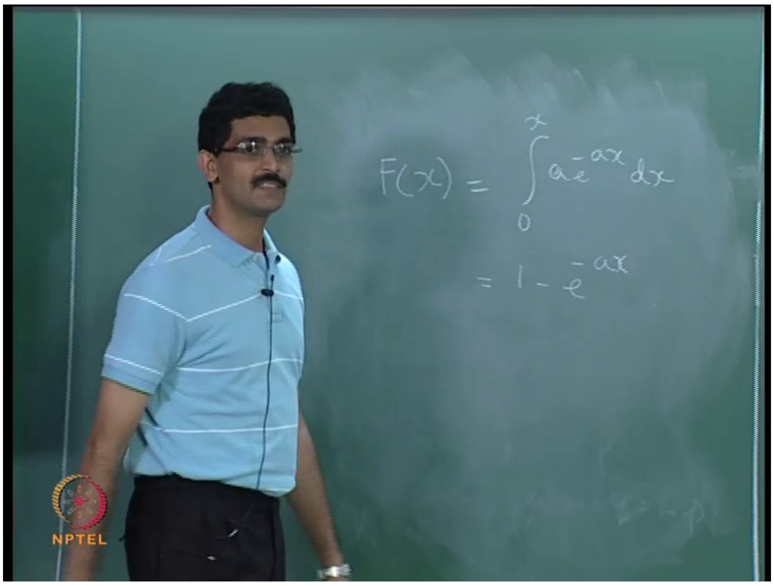
So, let a equals 10 which means you for per unit or the mean is 0.1 seconds. Whatever this unit is your graph look like I have to plot your f(x).

(No audio from 44:00 to 44:12)

So, which your starting value is x equal to 0 it is in some for... Now, will just leave it. So, it is your some…

(No audio from 44:33 to 44:44)
it is plot it is x goes to infinity, your value goes to 0.

(Refer Slide Time: 44:51)



Now more interestingly what is F (x)? Basically 0, the x the summation of all probabilities from 0 to this value x and that is equal to 1 minus e to the power minus a x; that is your F (x). So, we will come back on Friday, and look at why this is an important property trying to show this, so called the memory less nature of exponential one thing we have an talked about is property is both Poisson as well as exponential how they related with each other and how the memory less property is actually going to be useful when you do lot of analysis later on.

Because when we do Poisson then also we take…

Because the Poisson arrival tries that the inter arrival is exponential the exponential, these two are type together therefore, this Poisson is also memory less. So, it is inter expect inter arrival time memory less also be…

(( )) we do not depend on what kind that what time that…

So, what it means is we will have to formally state that will do the next time, but essentially if you have been waiting for the bus for the half an hour, you want to know what is the probability that the bus will come on the next 5 minutes. So, that the beginning of the system then we started waiting for this bus that same probability holds even after having waited for ever x number of for 1 hour and so on. That is why the vague that you waited for an hour, next know difference at all, the probability of having to wait for another 5 minutes is going to be same or which over pointing time, we look at in the system, that is where memory less concept and so on. That is what we try to use this to show.