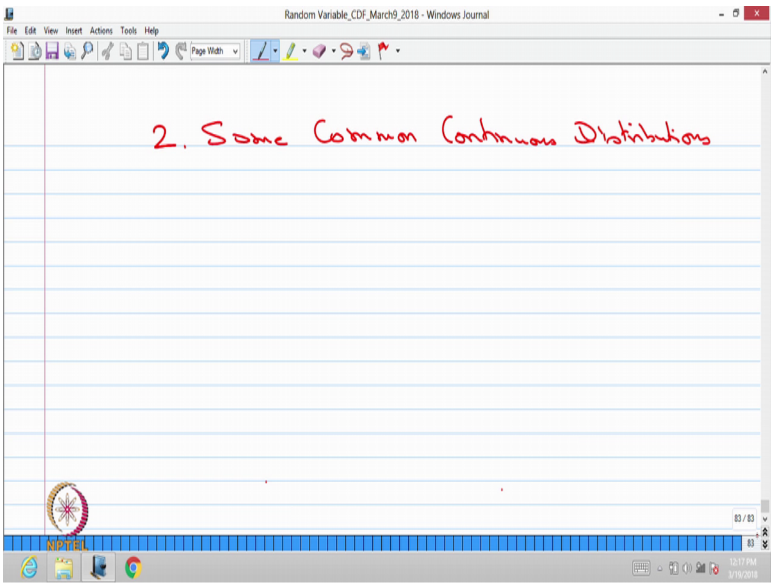
Introduction to Probability Theory and Stochastic Processes
Prof. S. Dharmaraja
Department of Mathematics
Indian Institute of Technology, Delhi

Lecture - 20

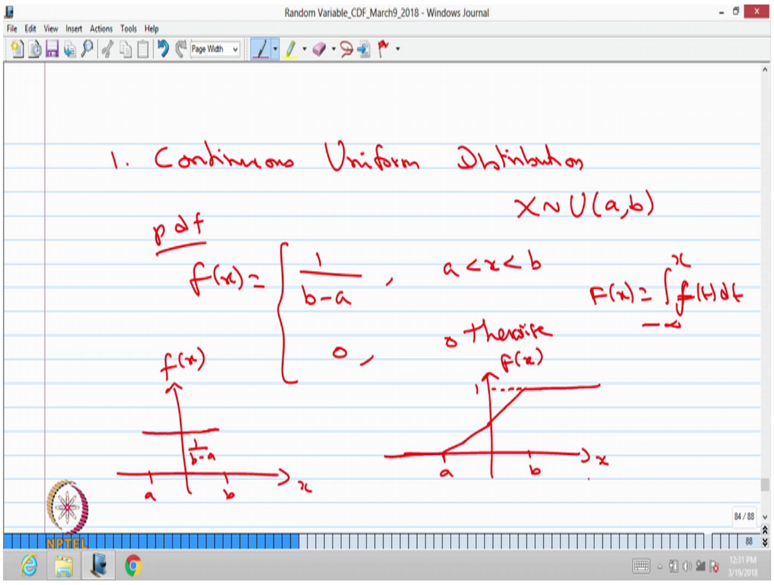
So, in the last class we have discussed some common discrete distributions in Standard Distributions module. Now, we will move into the some common continuous distributions that means a sum continuous type random variable occur more frequently whenever we come across different problems in the probability. Therefore, we introduce a word called common continuous distributions. That means, a I am going to discuss a few or some important continuous type random variable whose the probability mass function, then what is the cdf of those continuous type random variable, then what a what is the mean, variance, mgf it if it exist, then characteristic function and so on.

(Refer Slide Time: 01:00)



So, the title of these a lecture is, some common continuous distributions; in these we are going to discuss a few continuous type random variables which are occur a very frequently in the different problems of probability.

(Refer Slide Time: 01:23)



The first one that is continuous uniform distribution: a random variable X which is of a continuous time whose a probability density function is of the form f of x is 1 divided by b minus a, where X lies between a to b otherwise it is 0 where a comma b is in the real line. So, whenever any continuous type random variable whose probability density function is of the form f of x is 1 divided by b minus a as a probability density function which is greater than 0, otherwise it is 0 then we call it is a continuous uniform distribution. You can draw the probability density function to visualize.

Suppose, a is a negative and b is a positive and the probability density function is between the interval a to b is 1 divided by b minus a. So, you complete 1 divided by b minus a. So, you draw, so this height you think of one divided by b minus a otherwise; that means, a from minus infinity to a the probability density function value 0. Similarly, from b to infinity the probability density value is 0, only between the interval a to b the values 1 divided by b minus a.

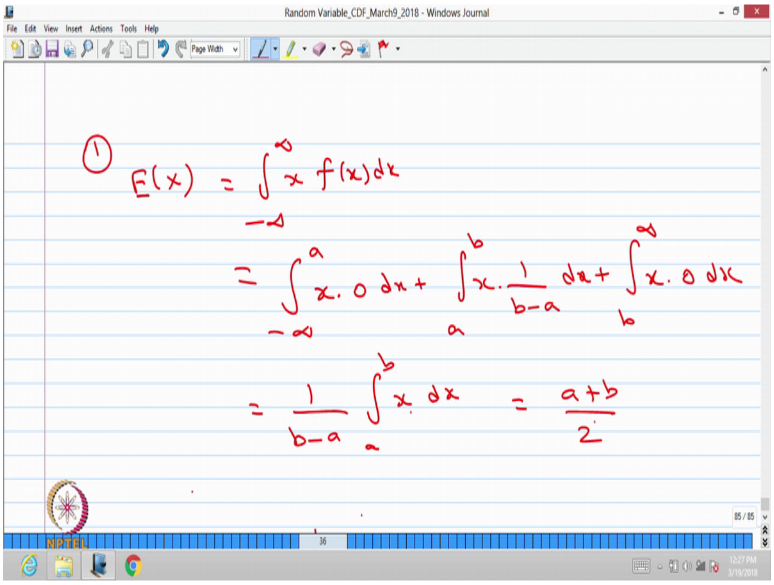
Why it is called a uniform, because the probability density is a uniform it is a constant it is not a function of x between the interval a to b since the density is a constant between the interval, and since it is the probability density function the integration has to be one therefore, that values 1 divided by b minus a. Therefore, this continuous type random variable is called a uniform distributed random variable. Earlier we have discuss the discrete a uniform distribution; that means, that is at the probability mass function is uniform are same in all n distinct points. Here it is a continues uniform distribution; that means, the density function between the interval is a constant which is same as one divided by length of the interval therefore, it is called a continuous uniform distribution.

If you draw the cdf of these, the cdf of continuous uniform distribution till a the value is going to be 0 at a till b you find out the integration capital F of x is integration from minus infinity to x, f of t dt, small f of t dt as probability density function if you substitute and find out the integration from minus infinity to a the probability density function is 0 therefore, the value 0. Whereas, a from a to b it is a your integrator from a to x 1 divided by b minus a dt. So, if you integrate you will get the a slanting line till b at the point b it becomes 1.

So, this type of random variable is call it as a continuous uniform distribution, the way I have explain to through the data of the cdf and the probability mass function of a discrete type random variable the same way by seeing the cdf and the probability density function one can conclude if the data has a cdf is 0 till some point after that it is a slanting line. And at some point it becomes a some constant value then it remains constant you can normalize it make it is a similar to the cdf of a continuous type then one can conclude this data follows a continuous uniform distribution.

Similarly, if you draw the histogram and the histogram is a between some interval it is a constant and all other value it is 0 then you can visualize that data follows a continuous type uniform distribution. So, here a and b or constant, lies between minus infinity to infinity and the probability density function is 1 divided by b minus a that is very important.

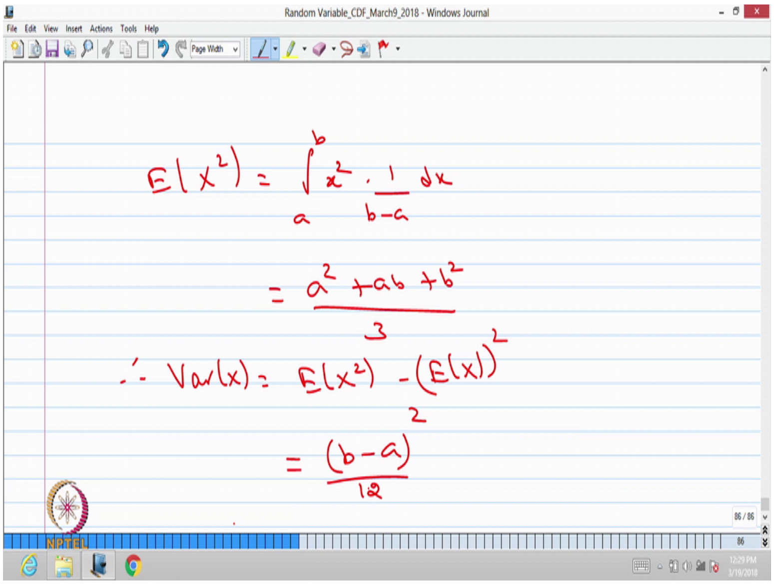
(Refer Slide Time: 06:30)



One can go for finding mean variance and mgf characteristic function and so on for this standard continuous type random variable. So, the mean for the random variable f x that is since it is a continuous type it is a minus infinity to infinity x times f of x, dx. We are doing this calculation with the assumption that the mean exists the assumption that in absolute sensor this integration is a finite quantity without absolute a we are finding the integration that is the value of a expectation. This is same as minus infinity to a x times 0 plus a to b x times the probability density function is one divided by b minus a dx plus b to infinity x times 0 dx the probability density function is 0 between minus infinity to a as well as b to infinity. Therefore, this is nothing but 1 divided by b minus a integration a to b x times dx.

So, one can simplify and you can get the answer for mean and variance. This is same as a plus b divided by 2. The mean of continuous uniform distribution is a the interval is a to b, then addition of the time interval divided by 2 that is going to be the mean. That a intuitively also one can say if it is a uniform then adding those endpoints divided by 2 that is a going to be the middle point. So, whenever it is a uniform distribution of the continuous type a plus b by 2 that is going to the mean that is same as average.

(Refer Slide Time: 08:42)

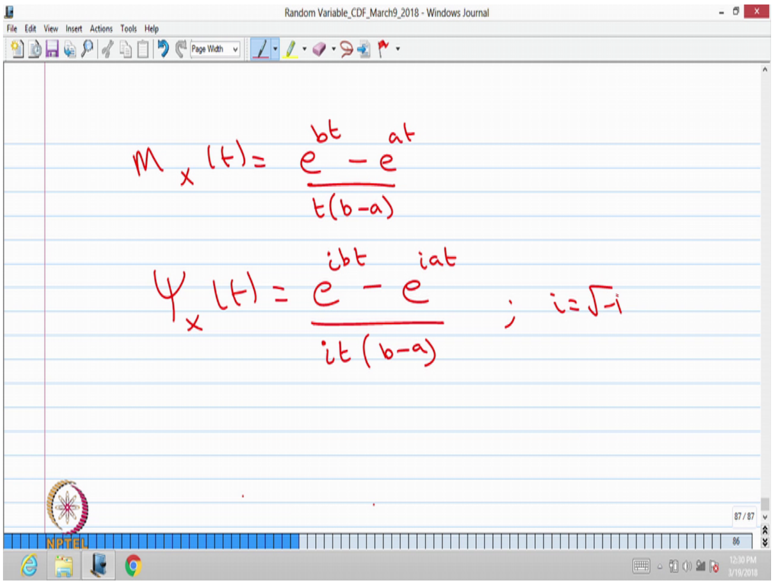


Similarly, one can find the variance, to find out the variance first you should find out what is expectation of X square. That is the same way a to b x square one divided by b minus a dx.

If you simplify this integration you will get a square plus a b plus b square divided by 3 therefore, the variance of X is E X square minus E of X the whole square. This is not the only way you can go for variance of X is equal to expectation of X minus b the whole square. So, you can compute that expectation also or we can find the E of X square then you can use this formula then you can get the variance. So, this is going to be if you substitute the value of E of X square that is a square plus a b plus b square by 3 and E of X is a, a plus b divided by 2 and the whole square. After simplification you can get b minus a the whole square divided by 2.

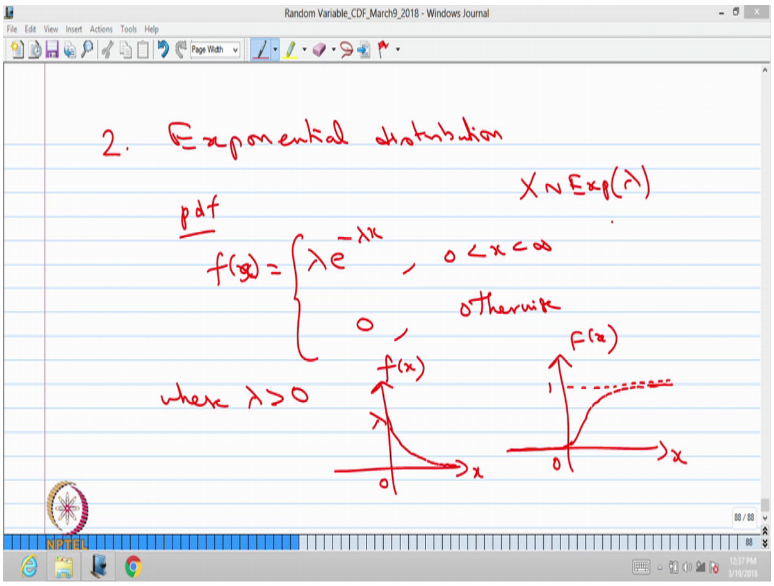
This a very important result the mean of a continuous uniform distribution between the interval a to b is a plus b divided by 2 and the variance is a b minus a if the whole square divided by 2.

(Refer Slide Time: 10:16)



One can get the moment generating function for continuous uniform distribution also. I am not going for the derivation and directly giving the result e power b t minus e power a t divided by t times b minus a. Since, you know the mgf finding the characteristic function is by replace t by i times t. Therefore, the characteristic function of a continuous uniform distribution is e power i times b t minus e power i times a t divided by i times t multiplied by b minus a, where i is nothing but square root of minus 1 complex.

(Refer Slide Time: 11:09)



Now, who will move into the next distribution that is the exponential distribution? So, before that in notation the continuous uniform distribution is call it as a X follows capital U with the interval a comma b. So, whenever we give a open interval a comma b with a capital U. That means, a the random variable X follows a continuous uniform between the interval a to b whereas, a for discrete uniform distribution we say X follows capital U within bracket X 1 comma X 2 comma so on till X n. If we supply the n points with a capital U that means, a that random variable is a discrete uniform distributed.

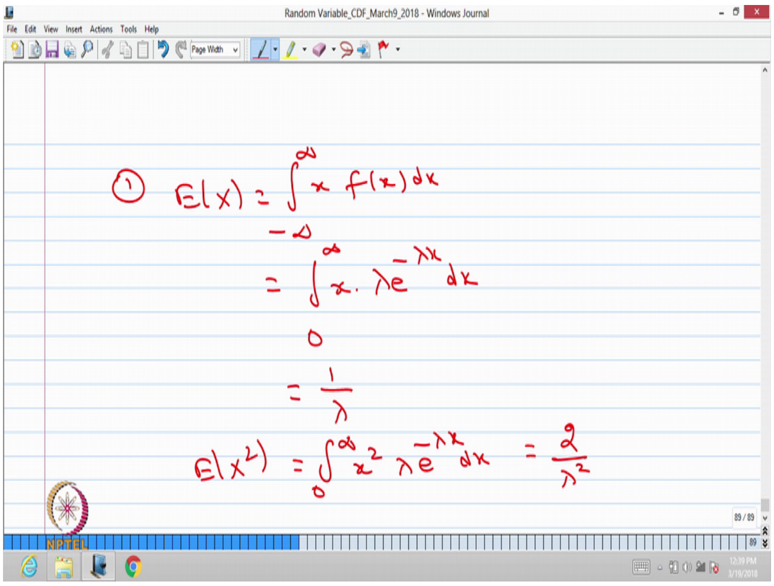
Now, you are moving to the second one that is exponential distribution. A continuous type random variable is said to be exponential distributed random variable when the probability density function of that random variable is of the form lambda times e power minus lambda x, where X is lies between 0 to infinity. Otherwise it is 0 where the lambda is strictly great than 0 then only it becomes a probability density function because this is greater is equal to 0 and integration from minus infinity to infinity is going to be 1, because minus infinity to 0 the probability density function is 0 and integration from 0 to infinity lambda times e power minus lambda x that is going to be 1 when lambda is greater than 0.

If you supply the value of lambda you are known with the exponential distribution therefore, lambda is the parameter. So, we can use the notation X follows the Exp within bracket lambda whenever we write Exp with in bracket lambda; that means, a that one is exponential distributed random variable with the parameter lambda. One can draw the probability density function, the probability density function starts at lambda at 0 and it will be keep going down and down and it becomes 0 at infinity. Since, this is the probability density function area below this curve from 0 to infinity that is going to be one when lambda is greater than 0 the probability density function will be touching asymptotically 0 at infinity and the area below that curve that is going to be one between the interval 0 infinity. And the cdf it is 0 till 0 and keep increasing and it becomes asymptotically it touches 1 at infinity. So, this is cdf.

The same interpretation if the data has a cdf of this form then you can conclude that data follows a exponential distribution it is a non-linear whereas, a uniform distribution as a slanting line it is a first order in x. Whereas, this one is a non-linear and then probability density function for a uniform distribution is a constant between the interval whereas, a here it is a function of x it is lambda times e power minus lambda x. There are some books they use the word a negative exponential distribution, but here we use the word exponential distribution; that means, the probability density function is lambda times e power minus lambda x. There are some books they use the parameter is 1 by lambda instead of lambda, whether we use 1 by lambda or lambda does not matter at the end of the day weather we will compute all other moments everything is going to be a function of parameters. So, you should remember whether you write lambda times e power minus lambda x are if the reciprocal form of lambda.

So, in this course I am using consistently lambda times e power minus lambda x that is the probability density function of exponential distribution with a parameter lambda that is a notation x follows Exp with the parameter lambda.

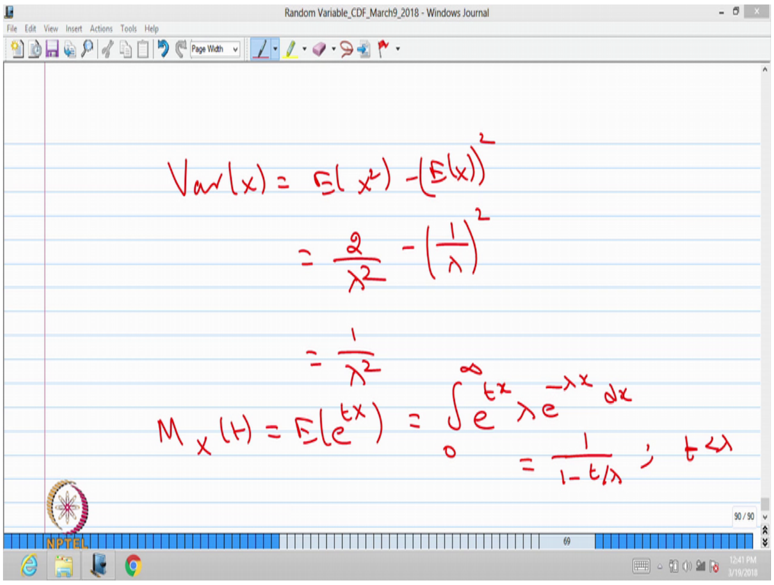
(Refer Slide Time: 16:19)



One can get me provided it exist it is same as the minus infinity to infinity x times f of x dx. This is same as the since the probability density function is greater than 0 between the interval 0 to infinity. So, you can directly write a 0 to infinity x times lambda times e power minus lambda x dx. If you simplify this integration you will get the answer that is one by lambda.

The probability density function is defined when lambda is greater than 0 therefore, expectation of X mean for the exponential distribution is reciprocal of the parameter. One can get the variance for the variance we can compute the E of X of square first in the same way that is the 0 to infinity x square times lambda times e power minus lambda x dx. If you do the simplification you will get 2 by lambda square therefore, the variance of X is going to be E of X square minus E of X the whole square that is a 2 divided by lambda square minus we got a mean of the random variable is a 1 by lambda therefore this is 1 by lambda whole square therefore, you will get 1 by lambda square.

(Refer Slide Time: 17:34)

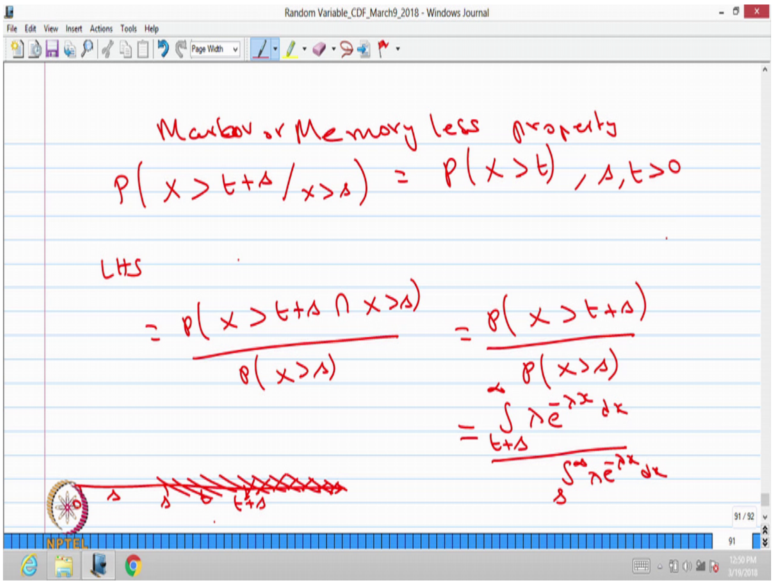


So, it is very important result the mean for exponential distribution with the parameter lambda is 1 by lambda and the variance of X is 1 by lambda square. You can get the mgf of exponential distribution that is expectation of E power tx, that is same as the integration from 0 to infinity e power t times x lambda times e power minus lambda x dx. And disintegration the expectation is going to be a finite quality whenever that is going to be less then lambda and the value is going to be 1 divided by 1 minus t divided by lambda.

So, this integration is going to give the value whenever the t is going to be less than lambda and the value is 1 minus t divided by lambda. So, the mgf exists between the interval from minus infinity to lambda and the value is exponentially whereas, lambda is strictly greater than 0. Since you know the mgf you can always get the characteristic function by replacing d by i times d.

Now, we will go to the one important property that is the probability of X greater than t plus x, given x is greater than s. What is that value?

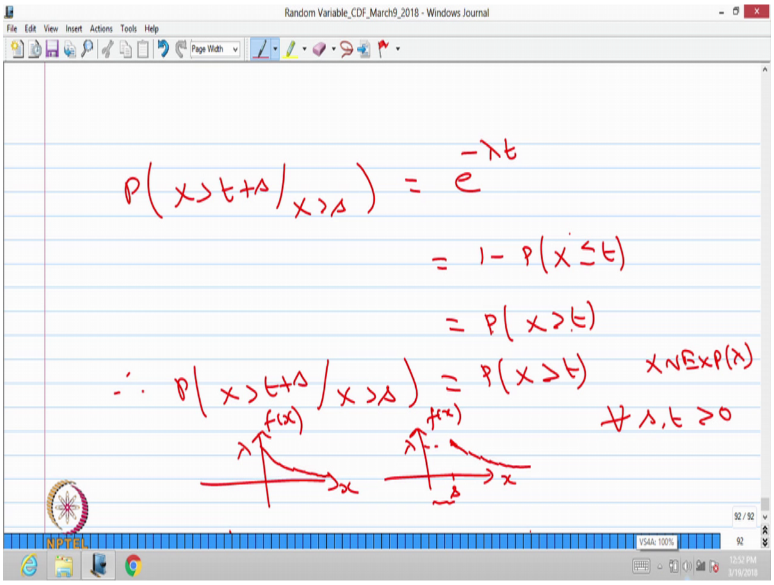
(Refer Slide Time: 19:32)



So, before filling up right hand side first we will compute this quantity then we will write down. So, let us start with the left hand side. If you compute this probability of X is greater than t plus s given X is greater than s that is nothing but probability of X is greater than t plus s intersection X is greater than s divided by probability of X is greater than s provided probability of X is greater s is greater than 0. You can explain this concept in a easy way just draw a line 0. Suppose you take the length s here. So, this point is s you take another length t therefore, this point is t plus s.

So, when I say X is greater than t plus s that means, you shade this point, when X is greater than s that means, a shade greater than s. Now, you look for what is the common portion of X is greater than t plus s with X is greater than s that is nothing but a greater than t plus s. So, therefore, this quantity is a probability of X is greater than t plus s, in the denominator it is probability of X is greater than s. Since, X follows exponential distribution probability of X is greater than t plus s means nothing but t plus s to infinity lambda times e power minus lambda x dx. And the denominator X is greater than s means integration from s to infinity lambda times e power minus lambda x dx, either you compute this integration and simplify or you can go for 1 minus of probability of X is less than or equal to t plus s and the denominator also 1 minus probability of X is less than or equal to s you can compute that integration then you can do the simplification.

(Refer Slide Time: 22:00)



So, the end of the simplification you will get the probability of X is greater than t plus s given X is greater than s that is same as the e power minus lambda t. The e power minus lambda t one can write that is a 1 minus the probability of X is less than or equal to t that is same as probability of X is greater than t you can directly also can write e power minus lambda t is same as a probability of X is greater than t or 1 minus probability of X is less than or equal to t. That means, a the probability of X is greater than t plus s given X is greater than s that is same as probability of X is greater than t whenever X follows exponential distribution with the probability of lambda this is true for all s and t greater than 0.

So, the left hand side is a conditional probability right hand side is the probability of X is greater than t and one more observation the left hand side involves s as well as t whereas, the right hand side involves only t which is free from s. That means, the information about the s is a disappearing the right hand side whereas, in the left side the conditional probability involves t as well as s. This is possible only when X follows exponential distribution.

That means, now I am concluding the probability of X greater than t plus s is given X is greater than s that is same as probability of X is greater than t this is for always comma t greater than 0 whenever X follows exponential distribution. This result we call it as a memory less property. That means, the random variable it is going to take the value given it is going to take the value more than s and the probability of getting the value more than t plus s that is same as the probability of getting the value more than t which is a not a function of s it is called the memory less property.

There is a another name for this property that is called Markov property. (Refer Time: 24:38) names for this property either you can call it as a memory less property or it is called the Markov property, m a r k o v, because of the information till it is not occurring the s is a disappear therefore it is called the memory less property. Not only this distribution satisfy is the memory less property there is a one more distribution also satisfies a memory less property that is a geometric distribution. When X is a geometrical distributed with the parameter p then probability of X is greater than m plus n given X is greater than m that is same as probability of X is greater than n where n and m are positive integers.

So, there are 2 distribution satisfies the memory less property one is exponential distribution which is of the continuous type, the other one is a discrete type that is geometric distribution. That means, the probability density function of the random variable x that is a with the probability density function start from lambda and it goes to 0 to infinity. If it does not take the value till s it is going to take the value more than s than the conditional probabilities again it is a probability of X is greater than t; that means, a from this point also the probability density function is going to be of the same form as the version. The probability density function will not change whatever the s you take given property of X is greater than s the condition probability of X greater than t plus s that is same as from del X is greater than t.

So, it has the same probability density function at every point in which you does not take the values. That means, this much memory is erased, that means, the interval from 0 to infinity that information is erased; that means, the memory is erased at every stage therefore, it is called a memory less property. If you choose a some other distribution in stop exponential distribution finding the conditional probability of left hand side you will get a function of s as well as t.

Therefore, no other distribution satisfies the memory less property, whereas the exponential distribution satisfies the memory less property. Similarly, geometric distribution satisfies the memory less property of discrete.