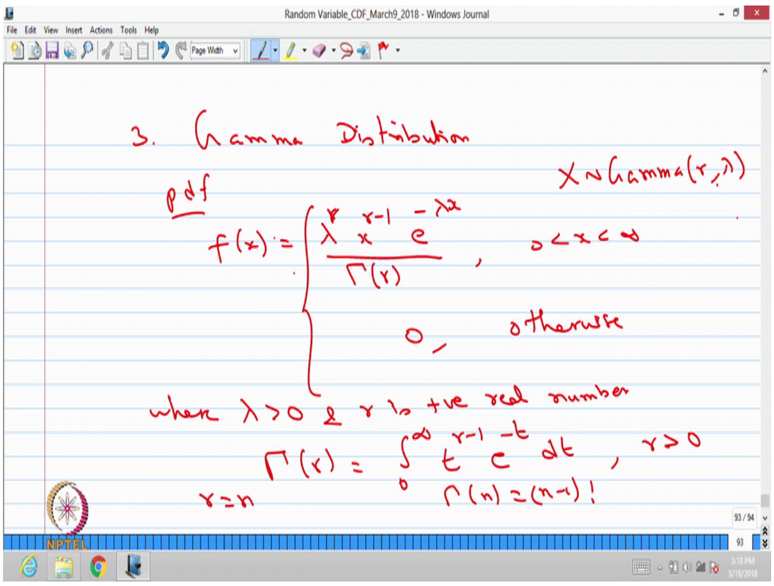
Introduction to Probability Theory and Stochastic Processes
Prof. S. Dharmaraja
Department of Mathematics
Indian Institute of Technology, Delhi

Lecture – 21

So, we have a discussed continuous uniform distribution and exponential distribution as the some common continuous type distribution. Now, we are moving into third one that is a gamma distribution, this is also the common continuous distribution.

(Refer Slide Time: 00:15)

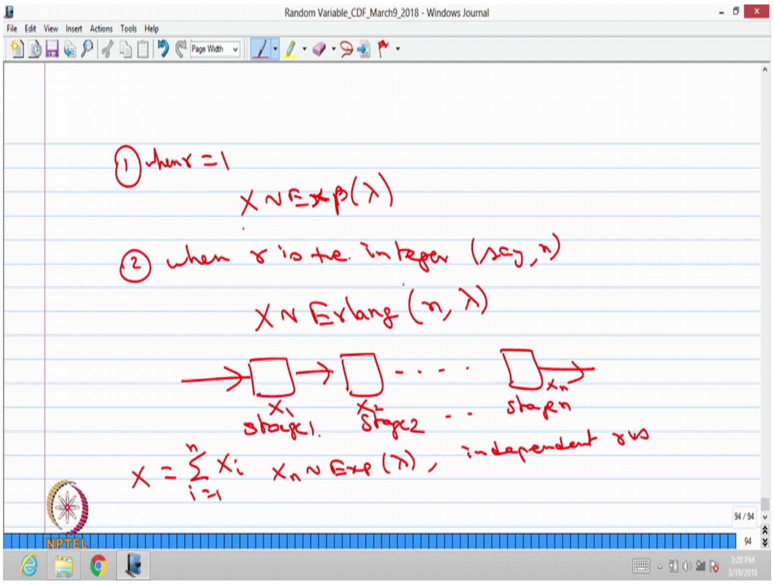


A random variable X is said to be gamma distributed if the probability density function is of the form f of X the takes a value lambda power r X power r minus 1, e power minus lambda X divided by gamma of r when X lies between 0 to infinity otherwise it is 0. Here lambda is strictly greater than 0 and r is a positive real number.

Whenever a continuous type random variable whose probability density function of this form we call it as a gamma distributed random variable. I have to define what is a gamma of r also. Gamma of r is nothing but integration 0 to infinity t power r minus 1 e power minus t dt. So, this is a way the a gamma of r is defined. r can be positive integers also, when r is going to be a positive integer label it as a n then gamma of n that is n minus 1 factorial. In general r is a real positive number and lambda is a again a real positive number which is greater than 0.

Then the probability density function of the formal lambda power r X power the r minus 1 e power minus lambda X divided by gamma of r; that means, if you integrate between 0 to infinity of this f of X that is going to be same as the denominator is a independent of r therefore, you will have a lambda power r gamma of r can be outside. And if you integrate that quantity is nothing but a gamma of r divided by lambda power r therefore, it cancels out you get the value 1. That means, lambda power r divided by gamma r that is a normalizing constant because that makes the whole integration is 1.

(Refer Slide Time: 03:17)



One more observation, when r is equal to one you can see the probability density function, when r is equal to one this f of X will be simplified into lambda times e power minus lambda x, where X is lies between 0 to infinity otherwise 0.

If you recall this is same as that the probability density function of expansion distribution therefore, when r is equal to 1 this becomes exponential distribution with the parameter lambda. That is the notation Exp exponential distribution with the parameter lambda. If you know the value of lambda and r you know the probability density function of gamma distribution therefore, both r and lambda are the parameters. So, a notation we can say X follows gamma with the parameters r gamma lambda.

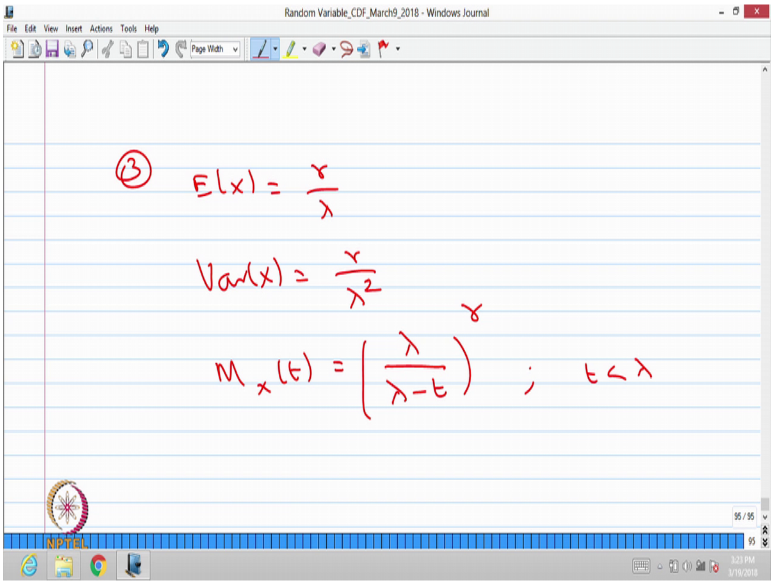
Again when r is a positive integer say n there is a another name for this gamma distribution that is called Erlang distribution, with the parameters n gamma lambda. I am just replacing a r by n when r is a positive integer. You can think of a very a simple example of a distribution the time taken by the system taking different stages.

So, this is a stage one, this is a stage 2, like this is a stage n. Suppose a some units spend the time from all the n stages that you denoted by the letter X in which stage one time spending is X 1, time spending in a stage 2 is X 2, and similarly the time spending in the n'th stage is X n and each one follows a exponential distribution with a parameter lambda; And the time taken in each step each stages are independent. Then the total time spent has a summation of X i's, i is equal to 1 to n this follows a Erlang distribution with a parameters n come over lambda.

For example, somebody entering into the hospital in which the time spending getting the admission in the hospital that takes an exponential distribution with the parameter lambda. Then going to the particular doctor then the a the time taken spending with the doctor that follows exponential distribution with the parameter lambda, the same patient is going to the stores collecting the medicine that also follows exponential distribution.

One followed by the other and all are independent then the total time spending in the hospital by entering into the hospital, spending time in the admission, spending time with the doctor, spending time with the getting the medicine, then the total time that is going to be Erlang distribution with the parameters 3 come over lambda. As long as a each exponential distribution has a same parameter all are independent and the total time is a sum of all the stages of time then it is going to be considered as the Erlang distribution with the parameters n come over lambda, or you can call it as a gamma distribution with the parameters n gamma lambda also both are one and the same.

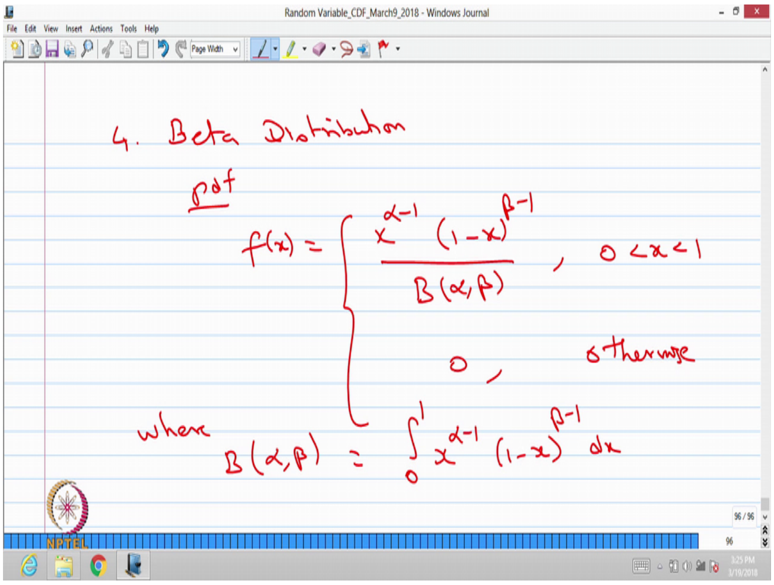
(Refer Slide Time: 07:27)



We can go for finding mean and variance if you find the mean for this gamma distribution the simple calculation will give the result r divided by lambda. You can verify when r is equal to 1 it becomes a exponential distribution and you know that the mean of exponential distribution is one divided by lambda. Similarly, if you do the simple calculation you can get the variance of X is r divided by lambda square.

So, here also one can verify with the exponential distribution. You can find the mgf of gamma distribution that is lambda divided by lambda minus t the whole power r and here the t is lesser than lambda. So, this is the mean variance and mgf of gamma distribution.

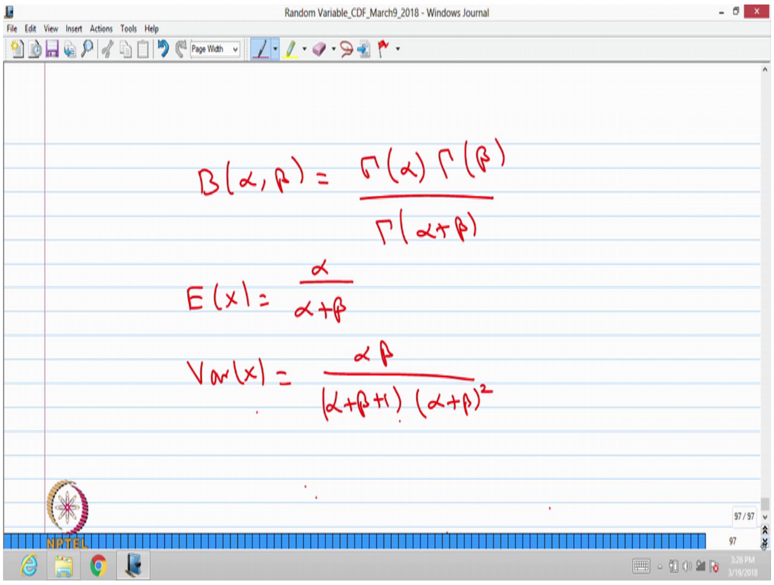
(Refer Slide Time: 08:33)



Now, will move into the next standard distribution that is called the beta distribution; Number 4, beta distribution. A continuous type random variable X whose probability density function is of the form f of x is x power alpha minus 1, 1 minus X power beta minus 1 divided by the beta function of alpha comma beta and this is a valid when X lies between 0 to 1 otherwise it is 0, where the beta function is beta function of alpha comma beta that is nothing but integration from 0 to 1 X power alpha minus 1 one minus X power beta minus 1 dx.

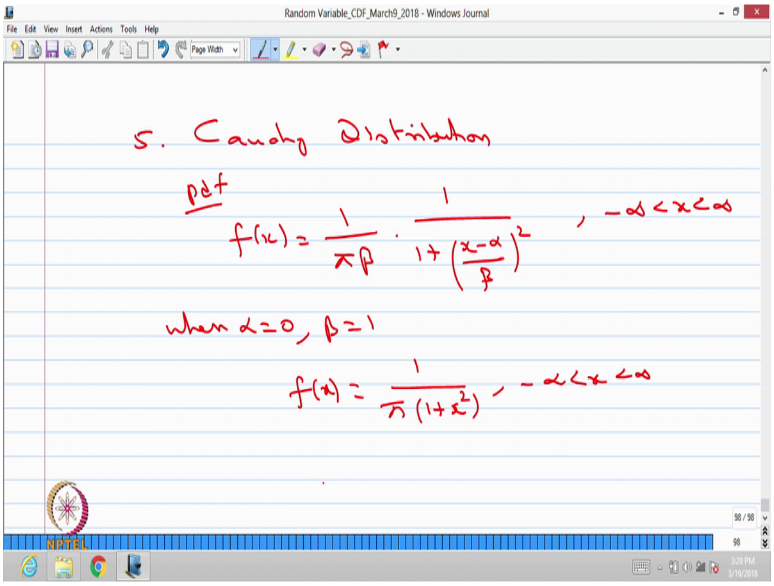
So, if you integrate between the interval 0 to 1 of this function you will get the beta function and the probability density function the denominator beta function. That means, if you integrate it is going to be 1 that means, the beta function is going to be the normalizing constant for the beta distribution.

(Refer Slide Time: 10:12)



One can write beta function in terms of gamma function also that is a gamma of alpha gamma of beta divided by gamma of alpha plus beta. For this a random variable also one can get mean that is a alpha divided by alpha plus beta and you can get the variance in the same way that is alpha beta divided by alpha plus beta plus 1 times alpha plus beta power.

(Refer Slide Time: 10:58)

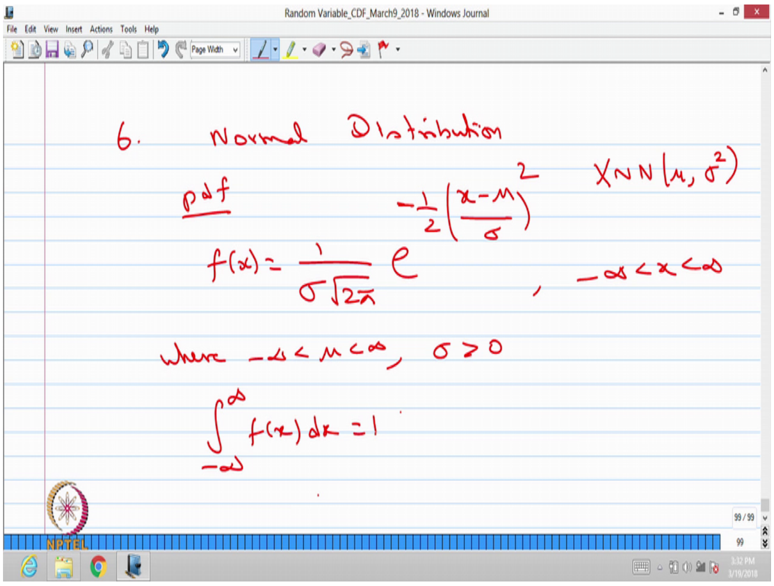


The next standard distribution that is number 5 that is Cauchy distribution; This is also continuous type and the probability density function of the form f of x that is 1 divided by pi times beta multiplied by 1 divided by 1 plus x minus alpha divided by beta the whole square, where x lies between minus infinity to infinity. Sometimes we see the special case that is when alpha is equal to 0 and beta is equal to 1, then the probability density function of the from 1 divided by pi times 1 plus x square.

So, this is a standard Cauchy distribution the importance of this Cauchy distribution the mean does not exist which we have given the proof while doing the mean for the random variable. Since the mean does not exist further moment does not exist therefore, the mgf does not exist is a very important distribution in which mean does not exist therefore, all the moment of order in this does not exist therefore, the mgf also does not exist.

Now, we will move into the next very important distribution in the probability course that is normal distribution.

(Refer Slide Time: 12:48)



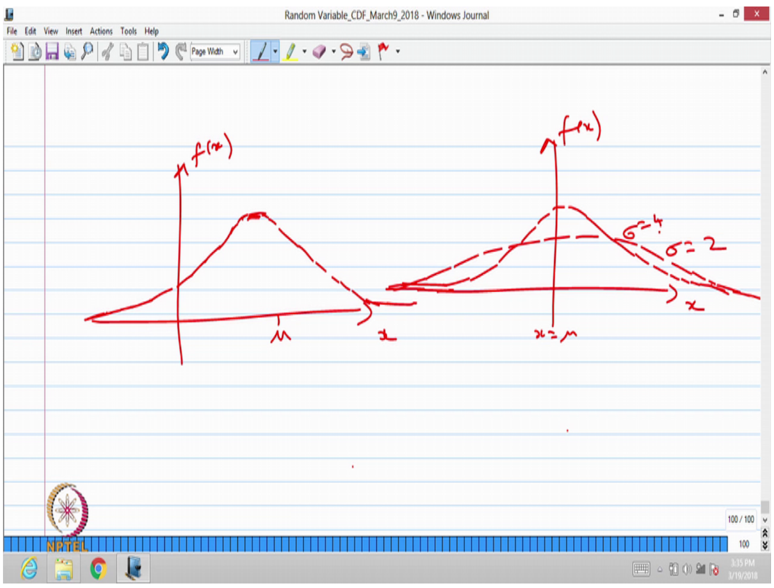
Why it is important? That when you are going to solve many more problems in the probability we will use the normal distribution quite a lot of time and also this a very important distribution because of some important result that is called the central limit theorem. Let me explain what is normal distribution and then will discuss the all the moments mgf and so on than finally, we give the central limit theorem also.

Here continuous type random variable is said to be normal distribution whenever the probability density function is of the form 1 divided by sigma times square root of 2 pi e power minus 1 by 2 times x minus mu divided by sigma the whole square, where x lies between minus infinity to infinity. This probability density function involves except mu and sigma therefore, one can define what is the range of mu, where mu can lies between minus infinity to infinity and the sigma is a positive real number. Once we know the value of mu in sigma we are known with the distribution of a normal distribution therefore, mu and sigma are called parameters.

So, in notation we use X follows a capital N mu comma sigma square we write. Why we write sigma square? That is because of the first parameter is nothing but the mean of the normal distribution and the second parameter sigma square is nothing but the variance of the normal distribution therefore, we write both the parameters mu and sigma square, where sigma is the positive square root of sigma square that is called as standard deviation of the normal distribution.

You can verify whether this is going to be a probability density function or not by integrating minus infinity to infinity f of x dx that is equal to; since it is a even function you can use the calculus in particular improper integral evaluate this integration and one can control this integration is going to be 1.

(Refer Slide Time: 15:27)



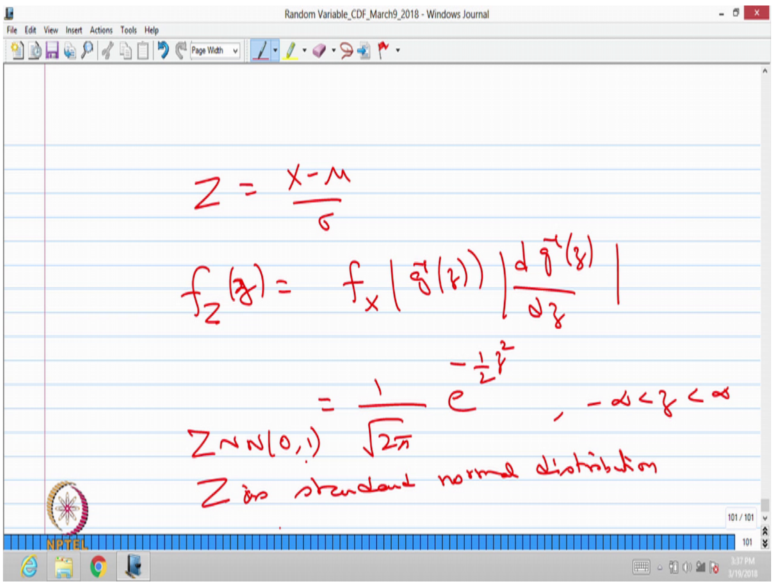
We can draw the probability density function of this suppose the mu takes a value here then you can draw the diagram of f of x. So, it is basically asymptotically 0 at minus infinity as well as plus infinity and it keep increases and it has the maximum value at mu.

So, this is going to be the probability density function of normal distribution, that means, a the area below this curve is going to be 1 the mu is called a location parameter and the sigma square is going to be called as a scale parameter. That means, whatever the values of mu that will fix where the peak will come that will fix the location. And the sigma square that will fix as the spread of the distribution or the variance. So, you can have a different diagram for the different sigma value.

So, let me give another diagram the probability density function for x is equal to mu supposed this is going to be the probability density function of. So, this is a peak, x is equal to mu and for example, suppose a this is going to be sigma is equal to 2, suppose you want to have a the same mu and sigma has to be the different value suppose you one can draw the sigma is equal to 4 with the same mu it is going to be. So, this is going to be sigma is equal to 4.

That means, the mu will be same for both the probability density function whereas, the sigma will give a spread more sigma we will give the more spreader and if the sigma is going to be lesser and lesser you will have a less spread and you will have a peak. One can go for making a standard normal distribution in the form of making a transformation of X to another random variable called is Z with the substation Z is equal to X minus mu by sigma

(Refer Slide Time: 18:05)



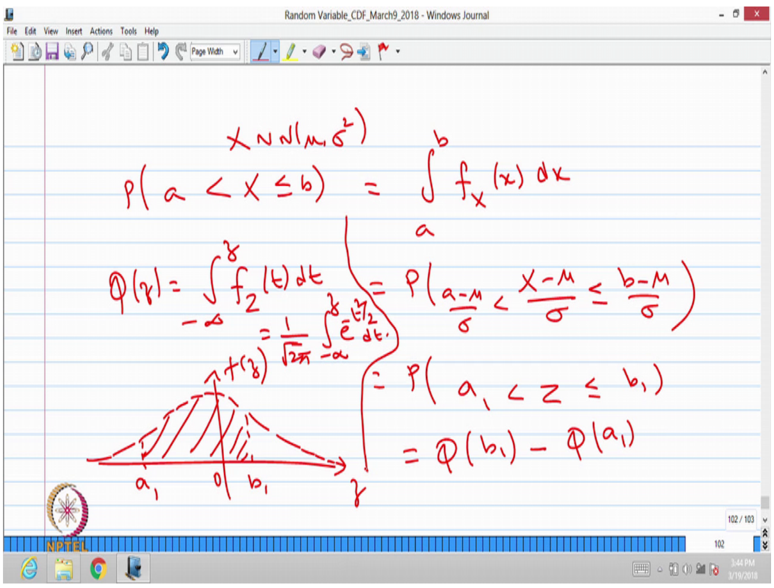
By using a function of a random variable you can get the probability density function of Z as a function of Z by using the definition of the probability density function of this g inverse of z then derivative of g inverse of z with respect to z So, you can use this, you can get the probability density function of X standard normal.

You can find the distribution of z that is going to be 1 divided by square root of 2 pi e power minus 1 by 2 z square, where z is lies between minus infinity to infinity the probability density function is 1 divided by square root of 2 pi e power minus 1 by to z square.

Now, we call Z as standard normal distribution. Why it is called a standard normal distribution because this is also a normal distribution with the parameters 0 comma 1. If you compare the probability density function of the normal distribution with the probability density function of z you can come to the conclusion mu is 0 and sigma is 1. That means, you do not need to specify the values the value is always mu is equal to 0 and the sigma square is one therefore, we call Z as a standard normal distribution.

So, this is obtained by transforming a normal distribution with the parameters mu and the sigma by using X minus mu divided by sigma that gives a standard normal distribution. That means, whenever you have a normal distribution you can always do this transmission X minus mu divided by sigma gives a standard normal distribution.

(Refer Slide Time: 20:30)



The advantage of this is when you find out the probability of X lies between a to b when X follows normal distribution with the parameters mu and sigma square. Since it is a continuous type random variable you have to compute the integration from a to b the probability density function of X. This is very difficult to compute because the integration is e power minus x minus mu by sigma whole square form. So, it one cannot compute the close from expression for this integration, that means, you need a numerical integration to evaluate the probability of X lies between the a to b.

So, what we do is first we transfer from this whole probability computation into standard normal distribution form, then we give the table of a numerical integration to evaluate the integration therefore, you can get the probability. Let us let us do the same thing. So, this is same as this is same as the probability of X lies between a to b same as a X minus divided by sigma is lies between mu b minus mu divided by sigma and the left side it is a minus mu divided by sigma this is same as the probability of.

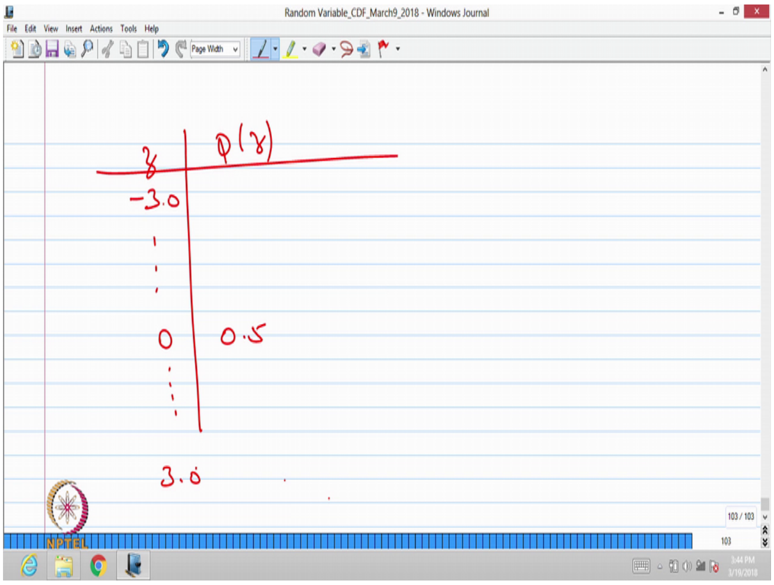
Let me write X minus mu divided by sigma is a Z, less than or equal to let me treat b minus mu by sigma of or known value of mu in sigma the whole thing is known therefore, just let me write as the b 1 point and a minus mu divided by sigma that I make it as a 1. That means, the probability of a X lies between a to b that is same as a probability of Z lies between a 1 to b 1. That is same as in a standard normal distribution if the probability density function is ok; suppose a 1 is somewhere here and suppose b 1 is somewhere here. So, Z lies between a 1 to b 1 that is nothing but this shaded area.

The way I have taken a 1 is in the left hand b 1 is in the right of 0 and you know that it is symmetric about z is equal to 0 therefore, the left side area is 0.5 and the right side area is 0.5. Therefore, this is same as for that I am going to do one more notation psi of z is nothing but minus infinity to Z probability density function of the standard normal distribution psi of z is a CDF at the point z for the standard normal distribution.

Therefore this becomes psi of a b 1 minus psi of a 1. The probability of a standard normal distributed lies between a 1 to b 1 that is same as a psi of b 1 minus psi of a 1, since it is a continuous type random variable there is no mass at any point. Therefore, this probability competition is valid whether you make it both are closed interval or both are open interval or one is open and 1 is closed and so on. For all the 4 types the probability lies X lies between a to b that is same as probability Z lies between a 1 to b 1 that is same as psi of b 1 minus psi of a 1.

Now, psi of b 1 you can compute numerically, similarly psi of a 1 you can compute the difference is going to be probability lies between a 1 to b 1.

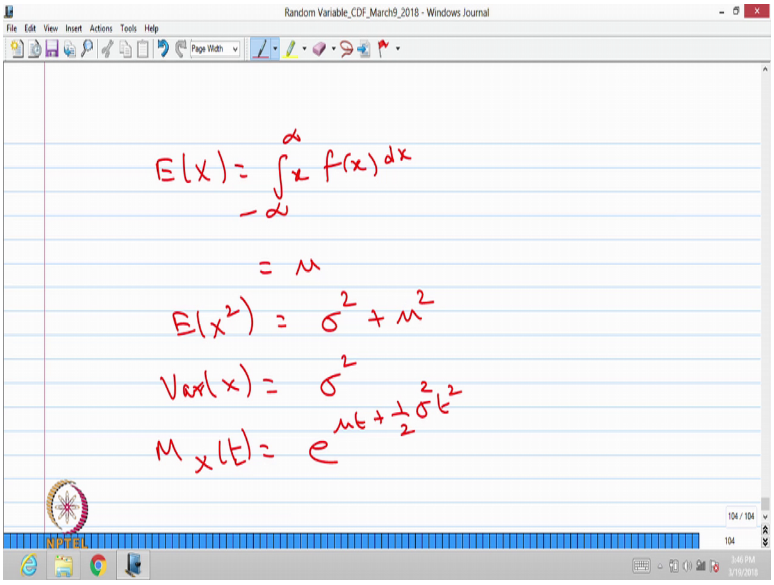
(Refer Slide Time: 24:49)



So, there is a table for different values of z you can have a psi of z. It start from, usually it start from minus 3 and for 0 psi of 0 is nothing but integration from minus infinity to 0 and since it is symmetric therefore, this value is 0.5 and it will be keep going till 3.0.

So, as far as the exam is concerned we will supply the psi of z values with the notation psi of z means integration from minus infinity to z, the probability density function of a standard normal distribution, the dt. I can write one more line also that is 1 divided by square root of 2 pi minus infinity to z e power minus t square by 2 dt both are on the same.

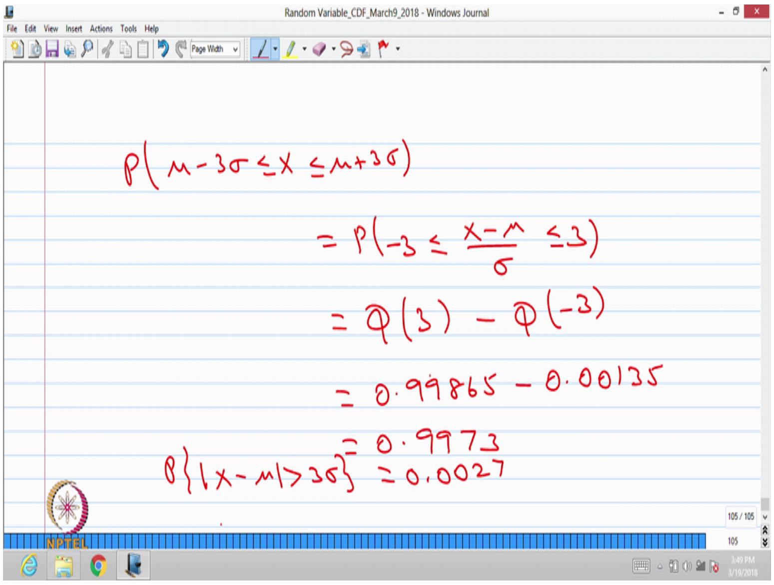
(Refer Slide Time: 25:54)



Now, you can go for computing mean and variance of a normal distribution mean is a minus infinity to infinity x times probability density function of this you can simplify by substituting the f of x value and so on one can get the value mu. And similarly if you compute E of X square, E of X square you will get a sigma square plus mu square. Therefore, the variance of X is going to be E of X square minus E of X whole square therefore, you will get sigma square..

One can get the mgf also, mgf of normal distribution, whenever I use the word X that is a normal distribution, whenever I use the word z; that means, is a standard normal distribution; So, mgf of normal distribution that is going to be e power mu t plus 1 by 2 sigma square t square. One can evaluate this is the mgf of a normal distribution with the parameters mu at the sigma square.

(Refer Slide Time: 27:17)

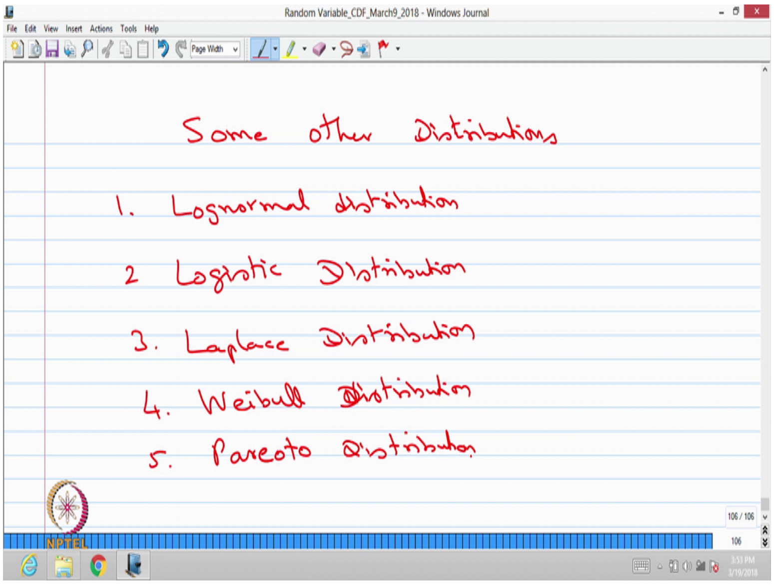


One can get easily the probability of X lies between mu plus 3 sigma mu minus 3 sigma to mu plus 3 sigma that is same as probability of X minus mu divided by sigma lies between minus 3 to 3 that is same as psi of 3 minus psi of minus 3.

If you compute psi of 3 from the table one can get a 0.99865, and the psi of minus 3 is 0.00135 therefore, you will get a 0.9973. That means, when X lies X is a normal distribution and the probability of a mu minus 3 sigma to mu plus 3 sigma 99.73 percentage accumulator in that range that is equivalent of same the probability of the absolute of X minus mu which is greater than 3 sigma that probability is 0.0027. So, away from mu minus 3 sigma to mu plus 3 sigma that probability is 0.002 or within the interval X lies between mu minus 3 sigma to mu plus 3 sigma that is probabilities 0.9973 is a important observation.

Like that one can compute what is the probability of X lies between mu minus 2 sigma to mu plus 2 sigma. Similarly one can compute probability of X lies between mu minus sigma to mu plus sigma.. These are all the standard results for the normal distribution. The main use of this normal distribution is a central limit theorem that I will explain after I introduce a multidimensional random variable and so on, so that I will discuss at the end of the probability part.

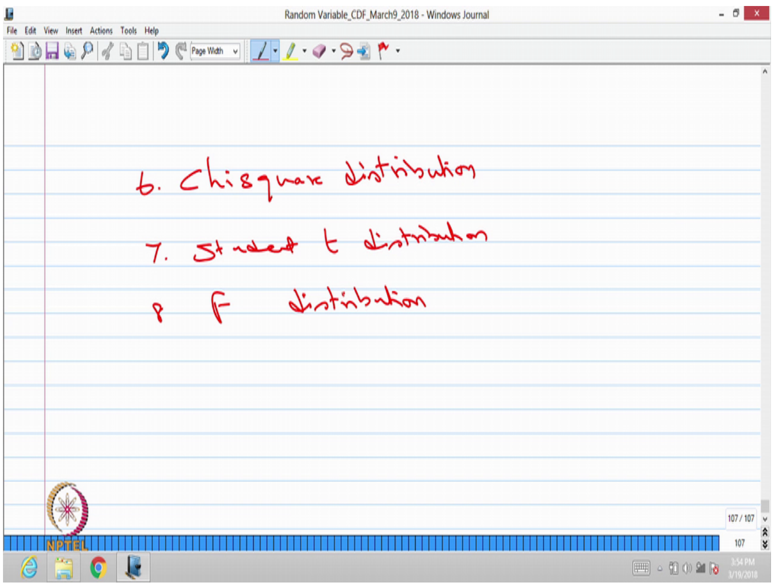
(Refer Slide Time: 29:44)



Some other distributions the first one which I am not going to discuss in detail I am just going to give the name of the distribution that is Lognormal distribution. This distribution is the function of a normal distribution in which the range is between 0 to infinity therefore, it is of the important.

The second one that is logistic distribution, this is also a important distribution which is of in the form of exponential function; Third one Laplace distribution, and the fourth one Weibull distribution, and fifth Pareoto distribution.

(Refer Slide Time: 31:11)



Then we have a chi square distribution, then we have a student t distribution, then we have F distribution. This 3 distributions that is a chi square distribution, student t distribution and F distribution, this 3 distributions are very important distribution for statistical inference. So, whenever you do the course on statistics then we need this distribution to discuss the statistical inference. So, we are not going to discuss in detailed about this distributions as for as this course is concern.

Like that there are many more standard distributions which is or not interest towards therefore, now I am stopping here.