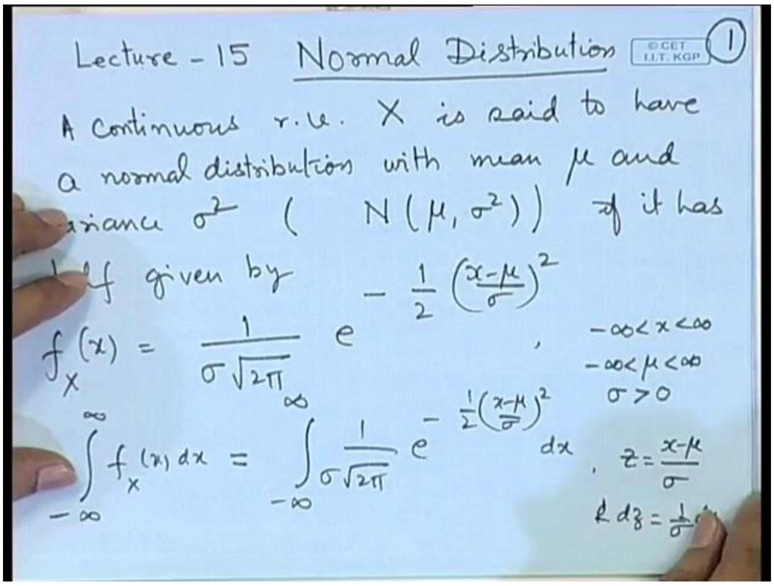
Probability and Statistics
Prof. Somesh Kumar
Department of Mathematics
Indian Institute of Technology, Kharagpur

Lecture – 29
Normal Distribution

Today I am going to introduce one of the most important distributions in the theory of probability and statistics; it is called the Normal Distribution. The normal distribution has become prominent because of one basic theorem in distribution theory, which is called the central limit theorem.

It tells that if we are having a sequence of independent and identically distributed random variables, then the distribution of the sample mean or the sample sum under certain conditions is approximately normal distribution or as N becomes large the distribution of the sample mean or the distribution of the sample sum is a normal distribution with certain mean and variance.

(Refer Slide Time: 01:15)

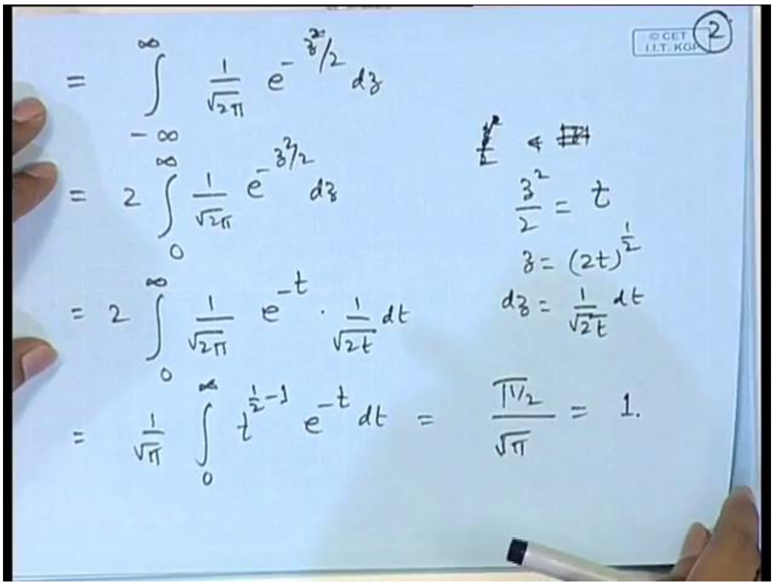


We will talk about the central limit theorem a little later; firstly, let me introduce the normal distribution. So, a continuous random variable X is said to have a normal distribution with mean mu and variance sigma square. So, we will denote it by N mu sigma square, if it has the probability density function given by 1 by sigma root 2 pi, e to the power minus 1 by 2 x minus mu by sigma whole square.

The range of the variable is the entire real line, the parameter mu is a real number and sigma is a positive real number. Now we will firstly, show that this is a proper probability density function and we will consider the characteristics of this. To prove that it is a proper probability density function, we should see that it is a non negative function which it is obviously, because here it is an exponential function and sigma is a positive number then we look at the integral of f x d x over the full range, here we make a transformation. So, 1 by sigma root 2 pi e to the power minus 1 by 2 x minus mu by sigma whole square, minus infinity to infinity.

Let us make the transformation here say Z is equal to x minus mu by sigma 1 by. So, d z will become equal to 1 by sigma d x, this is a 1 2 1 transformation over the range of the variable x.

(Refer Slide Time: 03:39)

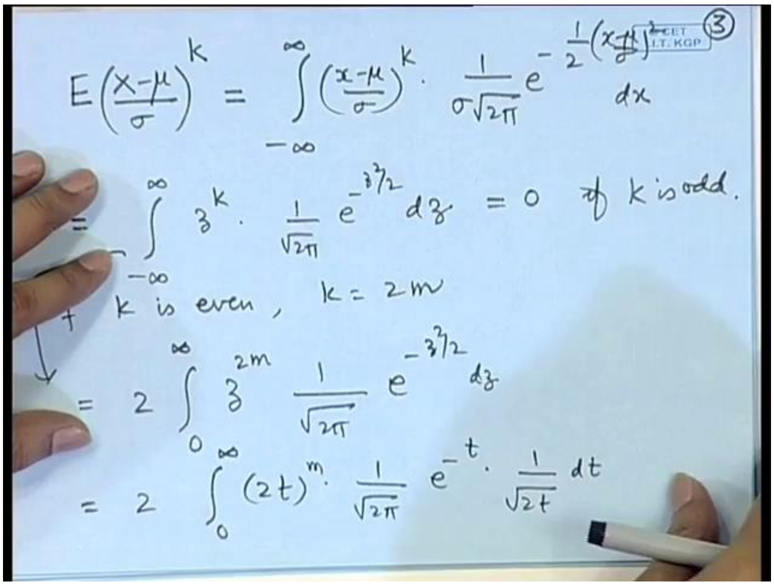


Therefore, this integral is reducing to integral from minus infinity to infinity, 1 by root 2 pi E to the power minus z square by 2 d z. This 1 by root 2 pi E to the power minus z square by 2 is also known as error function. So, we observe here that first of all it is a convergent integral, because we can right this z square by 2 as less than modulus z, and here we can consider 2 reasons: one is z is less than root 2 and z is greater than 2.

So, basically this entire quantity E to the power minus z square by 2 can be considered to be bounded and therefore, this is equal to 2 times integral 0 to infinity, 1 by root 2 pi E to the power minus z square by 2 d z. Over the range 0 to infinity we can substitute z square by 2 is equal to say t, that is z is equal to 2 t to the power half, and d z is equal to 1 by root 2 t dt. So, this becomes 0 to infinity, 1 by root 2 pi E to the power minus t, 1 by root 2 t dt that is equal to 1 by root pi t to the power half minus 1, e to the power minus t dt which is nothing, but gamma half by root pi.

Now, gamma half is root pi; so this is equal to 1, so this is a proper probability density function. We look at the moments of this distribution, now if we consider the transformation that we have made here that is z is equal to x minus mu by sigma, this suggest that it will be easier to calculate moments of x minus mu or moments of x minus mu by sigma so we will do that.

(Refer Slide Time: 06:11)

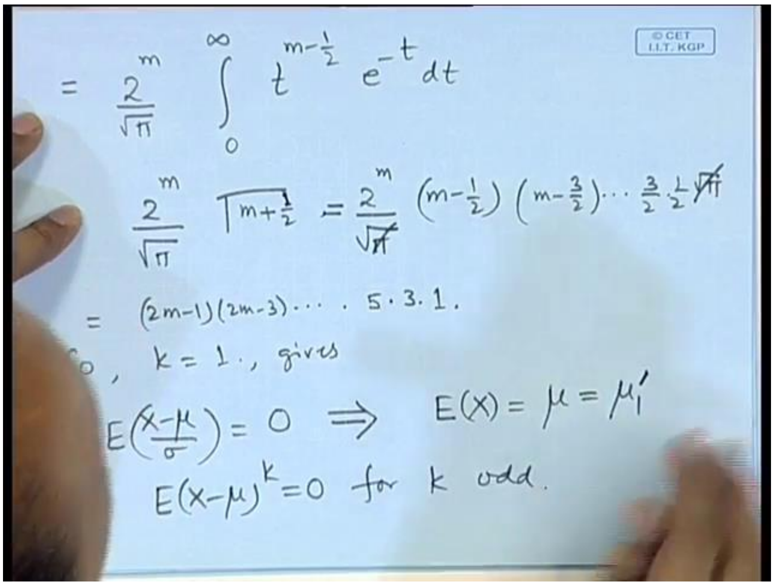


Let us consider expectation of X minus mu by sigma to the power k. So, this is equal to integral minus infinity to infinity x minus mu by sigma to the power k, 1 by sigma root 2 pi, e to the power minus 1 by 2, x minus mu by sigma whole square d x.

So, consider the transformation x minus mu by sigma is equal to Z. So, this will this particular integral will reduce to minus infinity to infinity, with z to the power k, 1 by root 2 pi, e to the power minus z square by 2 d z. If we look at this function, the function is an odd function if k is odd and therefore, this will vanish. So, this will vanish if k is odd and it is equal to if k is of the form say 2 m, then this integral will reduce to 2 times 0 to infinity, z to the power 2 m, 1 by root 2 pi, e to the power minus z square by 2 d z.

At this stage let us consider the second transformation that we made, there is z square by 2 is equal to t. So, if we make this transformation, then this quantity reduces to 2 times 0 to infinity. Now z square is equal to 2 t, so this becomes 2 t to the power m, 1 by root 2 pi e to the power minus t, 1 by root 2 t dt by considering d z is equal to 1 by root 2 t dt. So, we can simplify this here there are 2 square root twos in the denominator, so that will cancel with this.

(Refer Slide Time: 08:38)



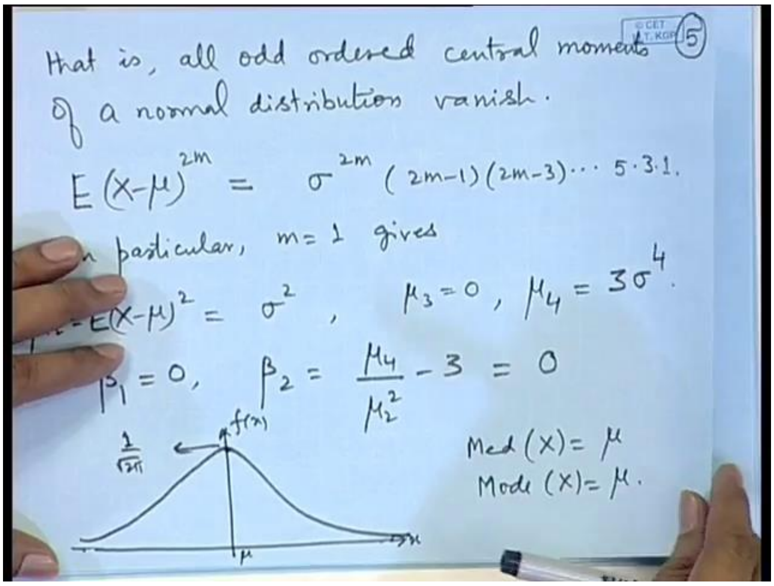
So, we are getting 2 to the power m by root pi, 0 to infinity integral, t to the power m minus 1 by 2 e to the power minus dt, which is nothing, but a gamma function.

So, this is equal to 2 to the power m by root pi, gamma m plus 1 by 2. So, if m is any integer here m is equal to 1 2 and so on that is k is an even integer, then expectation of X minus mu by sigma to the power 2 m is given by 2 to the power m by root pi gamma m plus half; of course, we can further simplify this to write in a slightly convenient looking form, we can write it as m minus half, m minus 3 by 2 and so on up to 3 by 2, 1 by 2 and gamma half that is canceling out, so it is equal to 2 m minus 1, 2 m minus 3 and so on up to 5. 3. 1.

So, we are able to obtain a general moment of X minus mu by sigma. So, if we utilize this suppose I put k equal to 1 then this is 0. So, k is equal to 1 gives expectation of X minus mu by sigma is equal to 0, which means that expectation of X is equal to mu; that means, the parameter mu of the normal distribution is actually the mean of it at first non central moment therefore, the terms expectation of X minus mu to the power k they give us the central moments of the normal distribution.

Now, we have already shown that if k is odd this is 0; that means, all odd ordered central moments of the normal distribution are 0. So, expectation of X minus mu to the power k is 0 for k odd.

(Refer Slide Time: 11:06)



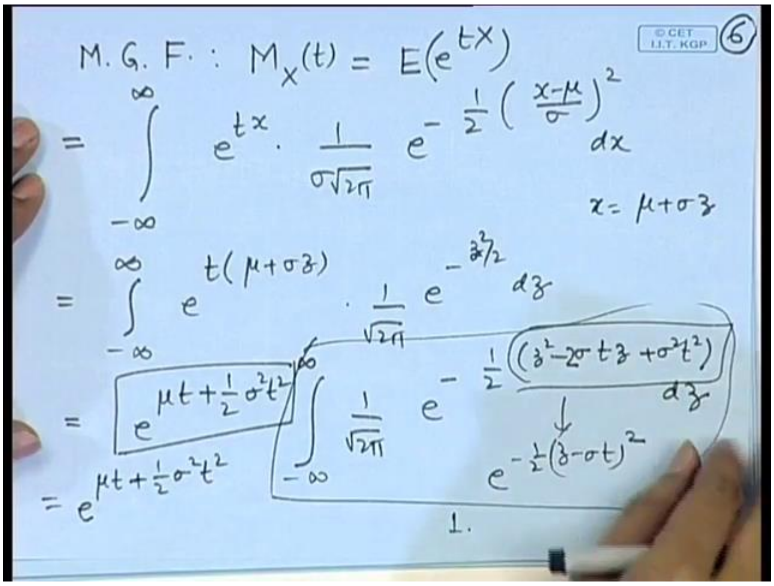
That is we can write that all odd ordered central moments of a normal distribution vanish; now this is quite important here, because we are considering any parameters mu and sigma square and for any parameters mu and sigma square, all the central moments are vanishing provided they are off odd order.

Now, let us consider even order. So, if we consider even order we are getting the formula sigma to the power 2 m into 2 m minus 1, 2 m minus 3 up to 5. 3. 1. In particular suppose I put m is equal to 1 here then I get by putting m is equal to 1, 2 m minus 1 that is sigma square. So, in particular, if I put m is equal to 1 this gives expectation of X minus mu square that is equal to sigma square that is mu 2, the second central moment of the normal distribution that is the variance is sigma square.

As we have already seen that generally mu and sigma square are used to denote the mean and variance of a distribution. So, the nomenclature comes from the normal distribution where the parameters mu and sigma square are actually corresponding to the mean and variance of the random variable. If we look at so obviously, mu 3 is 0, if we look at mu 4 here the forth moment that is if I put m is equal to 2 then here I will get 3 this is 1 so this will become 3 sigma to the power 4. So, the forth central moment is 3 sigma to the power 4; obviously, the measure of is skewness is 0, measure of kurtosis that is mu 4 by mu 2 square minus 3 is also 0; that means, the peak of the normal distribution is a normal peak. So, when we introduce the measure of kurtosis or the concept of peakedness, we said that it has to be compared with the peak of normal distribution or a normal peak.

So, basically the peak of the normal distribution is considered as a control or a standard. So, if we look at the shape of this distribution, the normal distribution it is perfectly symmetric around mu, and the peak of it is normal distribution. The value at x equal to mu is 1 by root 2 pi that is the mode of the distribution the maximum value, since it is symmetric about mu it is clear that the median of the distribution is also mu and mode of the distribution is also mu that is the value at which the highest density value is taken.

(Refer Slide Time: 14:58)



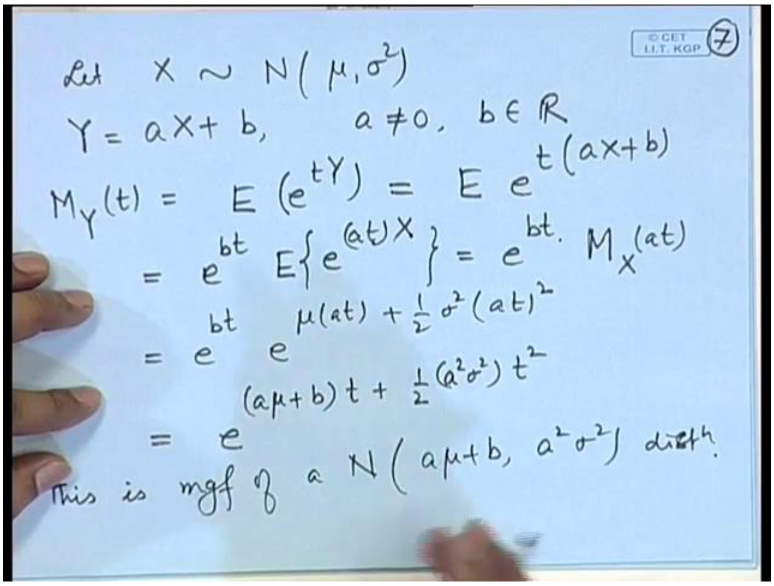
Let us consider the moment generating function of a normal distribution. So, M x t that is expectation of e to the power t x, this is equal to integral e to the power t x, 1 by sigma root 2 pi, e to the power minus 1 by 2, x minus mu by sigma whole square d x. So, we will still consider the same transformations which we introduced for the evaluation of any integral involving the normal density function, that is z is equal to x minus mu by sigma and z square by 2 is equal to something.

So, here if we write x minus mu by sigma is equal to z then we are having x is equal to mu plus sigma z. So, the integral becomes e to the power t, mu plus sigma z, 1 by root 2 pi e to the power minus z square by 2 d z. So, since it is a quadratic in z, we will again convert it into e to the power some term, which will involve a square in z. So, we can write it as 1 by root 2 pi e to the power minus 1 by 2 z square, minus sigma d z or with a 2 here.

This suggests that we should at sigma square t square and subtract it. If you subtract it then the term will be half sigma square t square. So, if you look at this particular term this is z minus sigma t whole square. So, the integrand denotes a probability density function of a normal random variable with mean sigma t and variance 1 therefore, this integral should be reduce in to 1, and therefore e to the power mu t plus half sigma square t square becomes a moment generating function of a normal distribution with parameters mu and sigma square.

Using the moment generating function of a normal distribution, we can prove an interesting feature consider say; so let X follow normal mu sigma square.

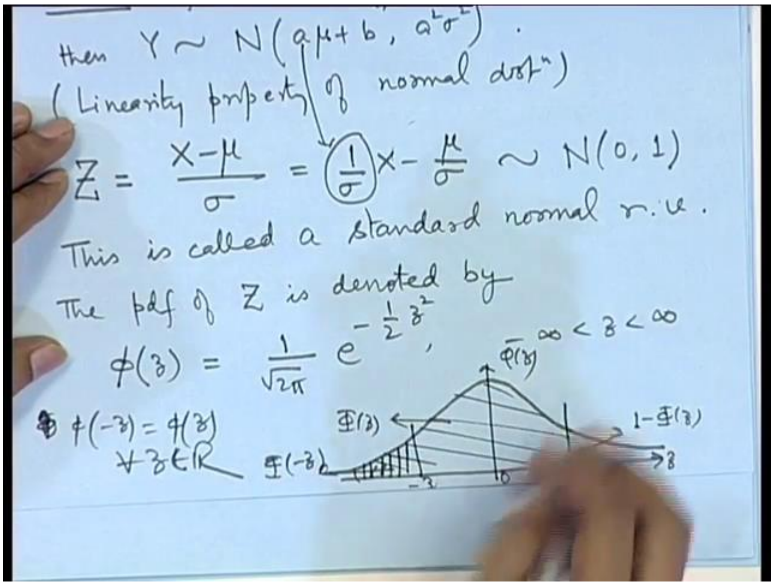
(Refer Slide Time: 18:10)



Let us consider Y is equal to say a X plus b, where a is any non 0 real and b is any real; consider the moment generating function of Y, that is equal to expectation of E to the power t y. This is equal to E to the power b t, expectation of E to the power a t x; this can be considered as the moment generating function of the random variable x at the point a t, now the distribution of x is normal and moment generating function of x that is M x t is given by e to the power mu t plus half sigma square t square.

So, we can substitute a t in place of t in the expression of M x t. So, we will get here e to the power b t, e to the power mu a t, plus half sigma square, a t whole square. So, we can adjust the terms a mu plus b t plus half, a square, sigma square, t square. If we compare this term with the moment generating function of a normal distribution with parameters mu and sigma square, then we observe here that mu is replaced by a mu plus b and sigma square is replaced by a square sigma square. So, we can say that this is mgf of a normal a mu plus b, a square sigma square distribution.

(Refer Slide Time: 20:38)

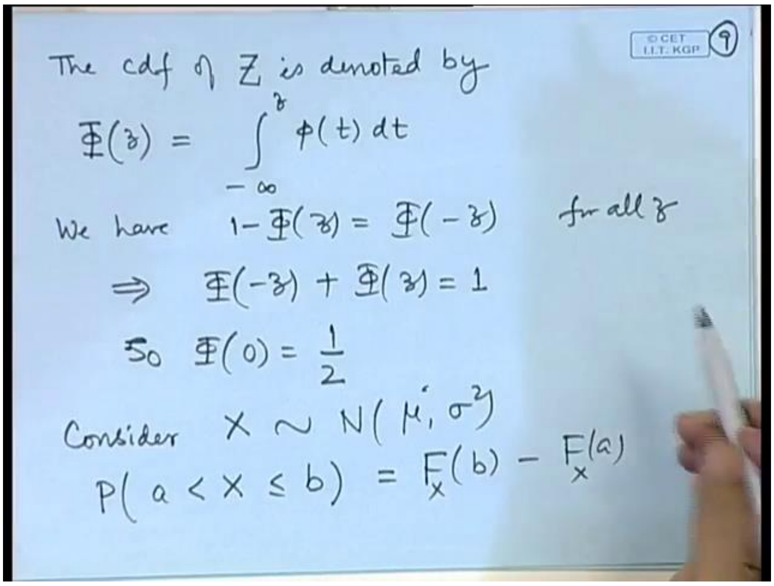


So, by the uniqueness property of the moment generating function, we have proved that if X follows normal mu sigma square and Y is equal to a X plus b, where a is not 0 then Y is also normally distributed with parameters a mu plus b and a square sigma square, this is called the linearity property of normal distribution; that means, any linear function of a normal random variable is again normally distributed.

Using this let us consider a random variable Z defined as X minus mu by sigma. So, if x follows normal mu sigma square and we make a linear transformation of this types. So, it is 1 by sigma x minus mu by sigma. That means, if we compare here then a is 1 by sigma and b is minus mu by sigma. So, if we substitute here, we will get mu by sigma minus mu by sigma that is 0 and this will become 1 by sigma square into sigma square that is 1. So, this will follow normal 0, 1.

A random variable Z which has a normal distribution with mean 0 and variance 1 is called a standard normal random variable. Let us look at the density function the pdf of Z is denoted by. So, there is a standard notation it is phi of z is small phi of z, it is 1 by root 2 pi e to the power minus 1 by 2 z square. We can see the shape of it, this is symmetric around z is equal to 0.

(Refer Slide Time: 23:23)



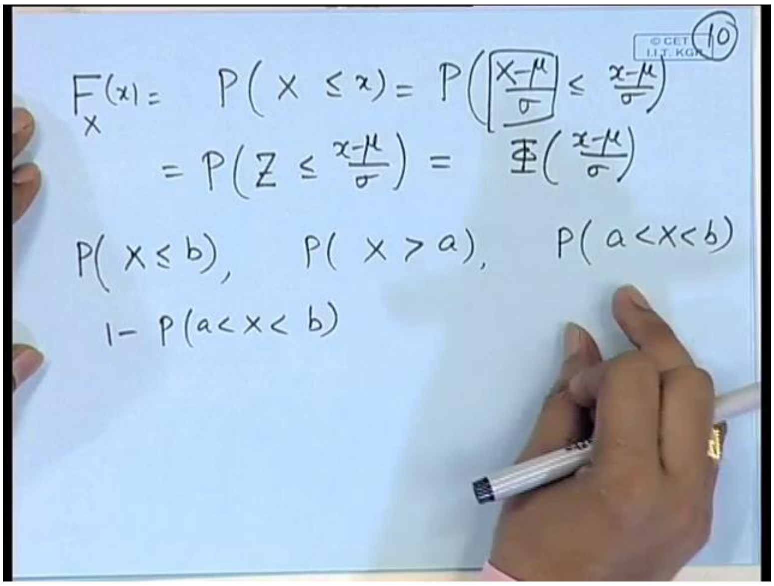
The cumulative distribution function of a standard normal random variable is denoted by capital phi of z, that is integral from minus infinity to z say phi t dt, where small phi t is the probability density function of a standard normal random variable.

Now, before going to the problems, let us look at the properties of this distribution. The standard normal distribution is symmetric about z is equal to 0. So, if we are considering say if this is the point z then phi z is actually this area. So, this area will become equal to 1 minus phi of z, if we call this area is capital phi z then this is 1 minus phi of z. By symmetry of distribution if we consider the corresponding point say minus z here, then the area here is phi of minus z, which shows that 1 minus phi of z is equal to phi of minus z.

So, we have 1 minus phi of z is equal to phi of minus z, this is true for all z; that means, we can write phi of minus z plus phi of z is equal to 1. And another thing of course, we could have observed here is that phi of a small phi of minus z is equal to a small phi of z for all, that is because of the symmetry property of the distribution.

In particular we can put z is equal to 0, then this will give phi of 0 is equal to half which is true, because the median of the standard normal distribution will be 0. Now this will helps us in evaluation of the probabilities related to any normal distribution. So, if we are having a general normal distribution, that is normal mu sigma square and we are interested to calculate say probability of a less than x less than or equal to b, then it is equal to F of b minus F of a.

(Refer Slide Time: 26:37)



However, if we consider the result here that X minus mu by sigma will have a standard normal distribution then this can be shifted to F x it is probability of X less than or equal to x, this we can write as probability X minus mu by sigma less than or equal to small x minus mu by sigma.

Now, this is Z. So, this is equal to phi of x minus mu by sigma; that means, the probability is related to normal distribution, can be calculated in terms of probability is related to a standard normal distribution. Now how do you evaluate this? Capital phi of z is equal to integral of minus infinity to z; e to the power minus z square by 2 d z. If we make the transformation z square by 2 is equal to t after suitably altering the ranges, so that it is a 1 to 1 transformation, it is reducing to an incomplete gamma function.

So, the incomplete gamma function can be evaluated using numerical integration say Simpson's one-third rule etcetera and tables of the standard normal distribution are available in all the statistical books. So, if we want to evaluate the probabilities related to any normal distribution, we will firstly convert it to a probabilities related to standard normal distribution and then utilize the tables are numerical integration here. In particular if we consider say probability any particular probability say X less than or equal to b, say probability X greater than a probability a less than X less than b are 1 minus probability of a less than X less than b.

So, these are some of the usual probabilities that are required in normal calculations. So, all of this can be evaluated using the properties of the standard normal cumulative distribution function. So, we stop here.