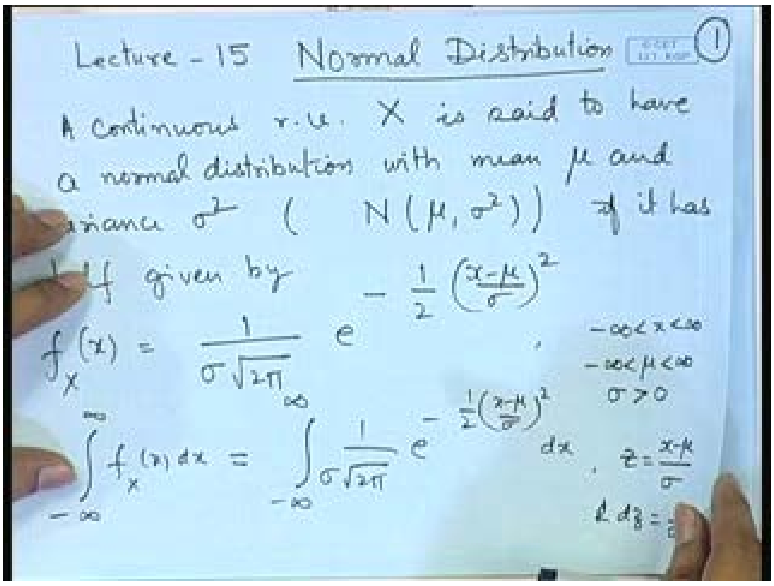
Probability and Statistics

Prof. Dr. Somesh Kumar
Department of Mathematics
Indian Institute of Technology, Kharagpur

Module No. #01
Lecture No. #15
Special Distributions-VI

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.

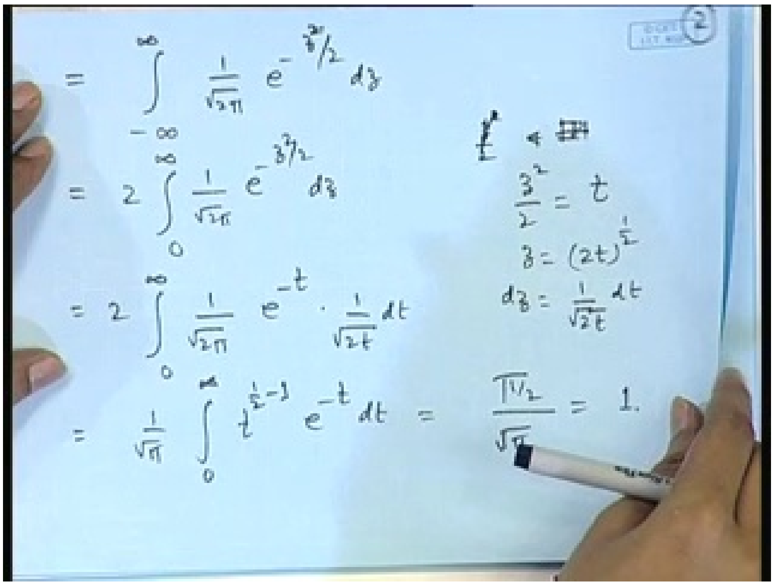
(Refer Slide Time: 01:14)



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 fx dx 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, dz will become equal to 1 by sigma dx - this is a one to one transformation over the range of the variable X.

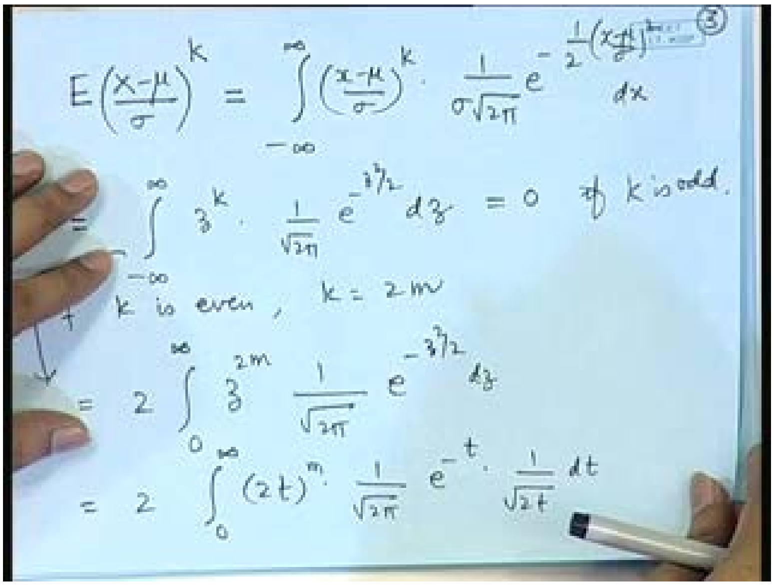
(Refer Slide Time: 03:36)



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 dz, 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 write this z square by 2 as less than modulus z, and here we can consider two 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 dz, over the range 0 to infinity we can substitute z square by 2 is equal to say, t that is, z is equal to 2t to the power half and dz is equal to 1 by root 2t dt, so, this becomes 0 to infinity 1 by root 2 pi e to the power minus t 1 by root 2t dt, that is equal to 1 by root pi d 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 suggests 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.

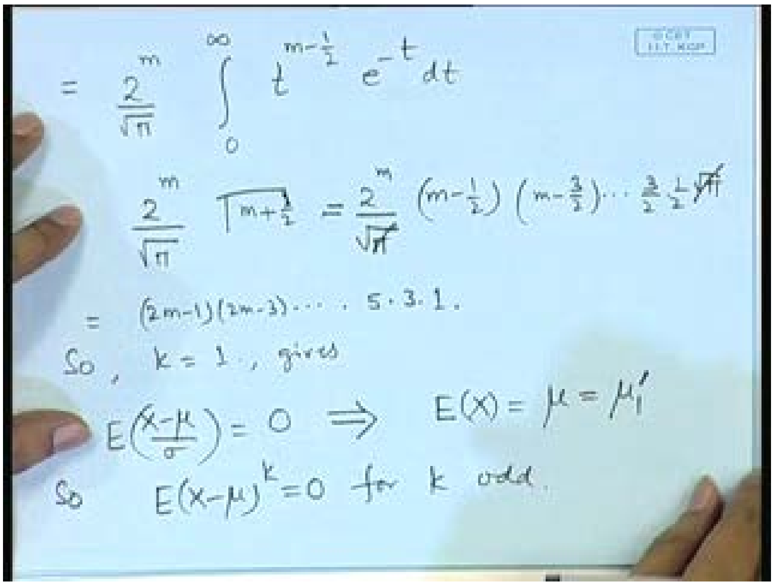
(Refer Slide Time: 06:09)



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 dx. 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 z to the power k 1 by root 2 pi e to the power minus z square by 2 dz. 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 2m, then this integral will reduce to 2 times 0 to infinity z to the power 2m 1 by root 2 pi e to the power minus z square by 2 dz.

At this stage let us consider the second transformation that we made that 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 dz is equal to 1 by root 2 dt. So, we can simplify this here, there are two square root 2 is in the denominator, so, that will cancel with this.

(Refer Slide Time: 08:34)

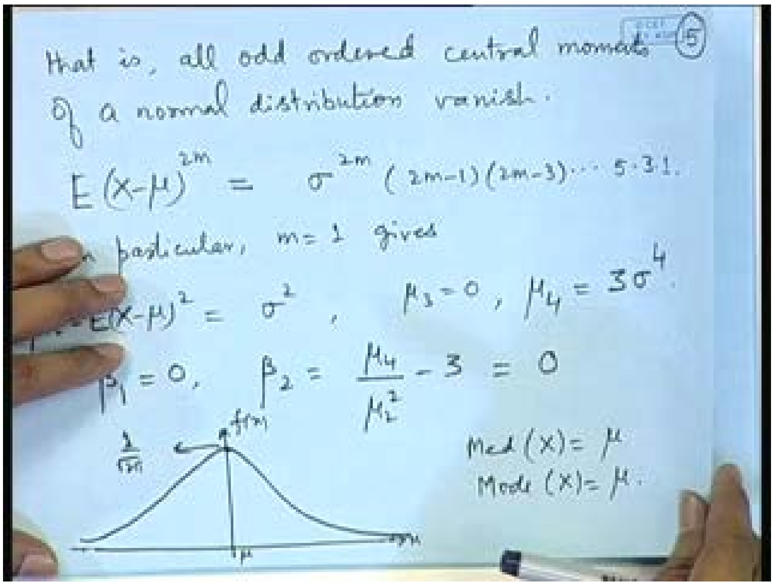


So, we are getting 2 to the power m by root pi 0 to infinity integral t to the power m minus 1 by 2 e to the power minus dt, which is nothing but a gamma function. So, this is equal to 2 the power m by root pi gamma m plus 1 by 2. So, if m is any integer here, m is equal to 1, 2 and so on that is, k is an even integer, then expectation of X minus mu by sigma to the power 2m is given by 2 to the power m by root pi gamma m plus half. Of course, we can further simply 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 is equal to 1, then this is zero. So, k is equal 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 that gives us a central moments of the normal distribution. Now, we have already shown that if k is odd, this is zero 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)

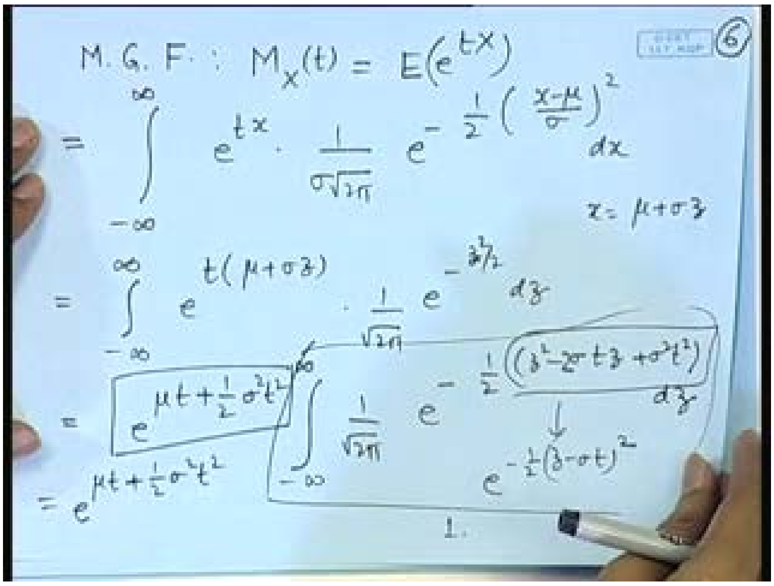


That is we can write that all odd ordered central moments of a normal distribution vanish. Now, this is quite important here because we are considering any parameters mu and sigma square, and for any parameters mu and sigma square all the central moments are vanishing provided they are of odd order. Now, let us consider even order. So, if you consider even order, we are getting the formula sigma to the power 2m in to 2m minus 1, 2m 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, 2m minus 1 is 1, so, 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 mu4 here, the fourth 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 fourth central moment is 3 sigma to the power 4; obviously, the measure of skewness is 0, measure of kurtosis that is, mu4 by mu2 squares minus 3 is also zero. That means, the peak of the normal distribution is a normal peak.

So, when we introduced 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 the mode of the distribution is also mu, that is a value at which the highest density value is taken. Let us consider the moment generating function of a normal distribution.

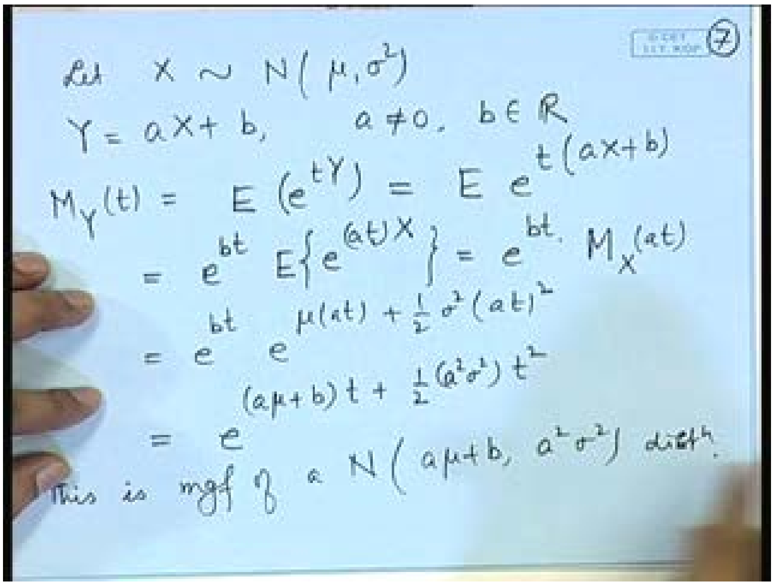
(Refer Slide Time: 14:56)



So, MX t that is, expectation of e to the power tx, this is equal to integral e to the power tx 1 by sigma root 2 pi e to the power minus 1 by 2 x minus mu by sigma whole square dx. 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 dz. So, since it is a quadratic in z we will again convert it into e to the power some term, which will involve a squaring z. So, we can write it as 1 by root 2 pi e to the power minus 1 by 2 z square minus sigma tz with a 2 here. This suggests that we should add sigma square t square and subtract it, if we 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 integrant denotes a probability density function of a normal random variable with mean sigma t and variance 1 therefore, this integral should be reducing to 1 and therefore, e to the power mut 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.

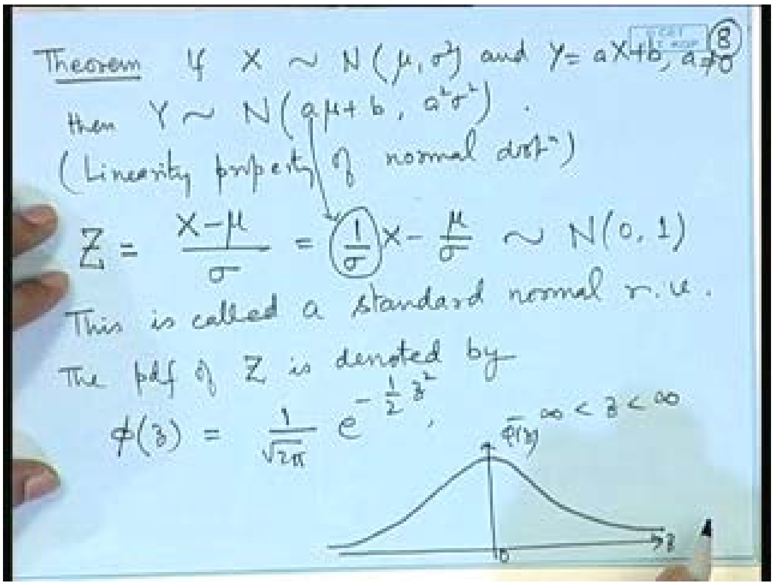
(Refer Slide Time: 18:08)



Consider say, so, let X follow normal mu sigma square; let us consider Y is equal to say aX plus b where a is any non-zero real and b is any real; consider the moment generating function of Y, that is equal to expectation of e to the power tY, this is equal to e to the power bt expectation of e to the power at X, this can be considered as the moment generating function of the random variable X at the point at. Now, the distribution of X is normal and moment generating function of X that is, MXt is given by e to the power mu t plus half sigma square t square, so we can substitute at in place of t in the expression of MXt. So, we will get here e to the power bt e to the power mu at plus half sigma squares at whole square. So, we can adjust the terms a mu plus bt 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.

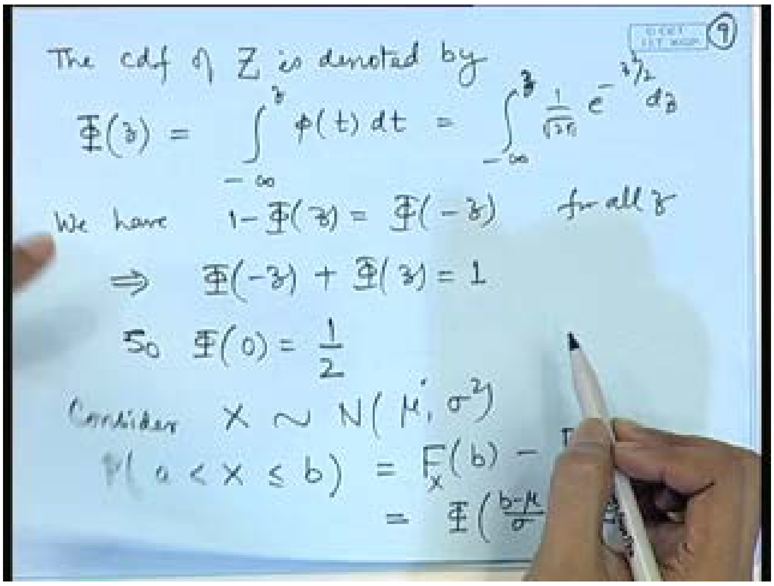
(Refer Slide Time: 20:34)



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 aX plus b where a is not 0, then Y is also normally distributed with parameters amu 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 type, 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 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, 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: 23:22)

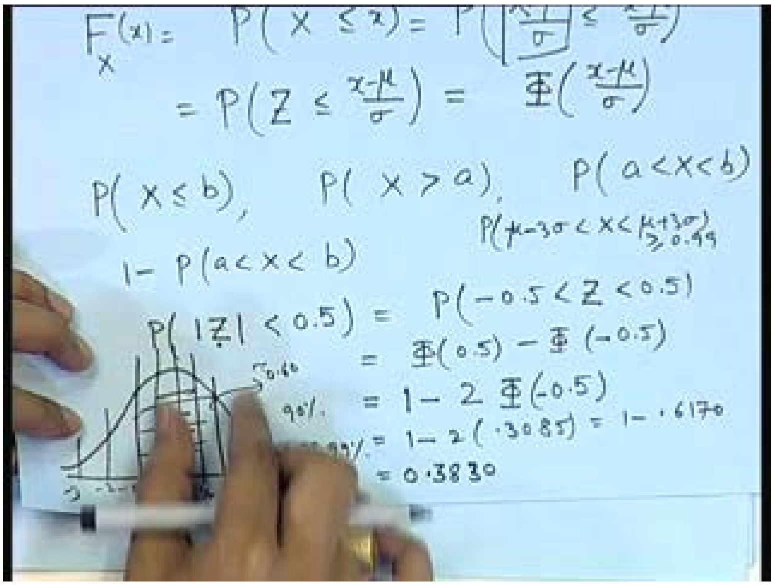


The cumulative distribution function of standard normal random variable is denoted by capital Phi of z that is, integral from minus infinity to z say, phi t dt where a small phi t is the probability density function of a standard normal random variable. Now, before going to the problems, let us look 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 as 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, small phi of minus z is equal to small phi of z for all z, that is because of the symmetric 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 help 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.

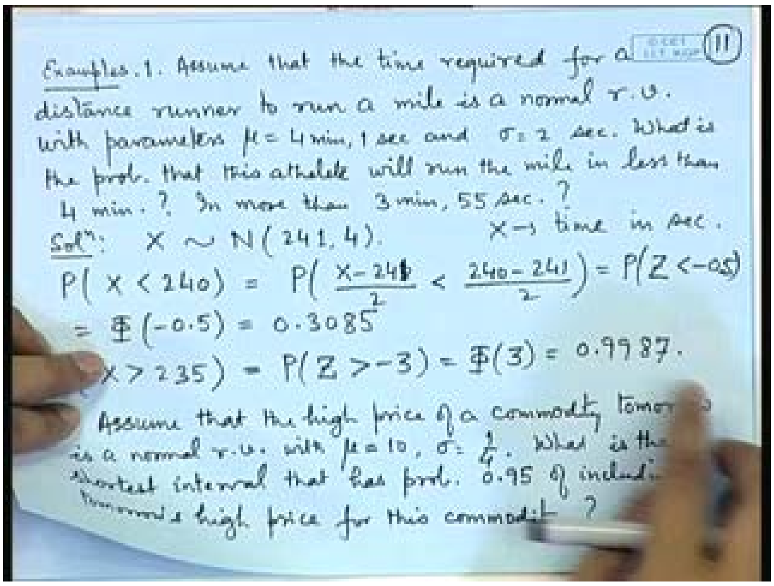
(Refer Slide Time: 26:34)



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 FX that 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 probabilities related to normal distribution can be calculated in terms of probabilities related to standard normal distribution. Now, how would 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 dz. If we made 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 probability related to any normal distribution, we will firstly, convert it to a probability related to standard normal distribution and then utilize the tables or 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, or 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. One more point here is that since the values of the cdf can be evaluated using Phi of minus z plus Phi of z is equal to 1 therefore, many times the tables are tabulated only for either positive arguments of z or negative arguments of z.

(Refer Slide Time: 29:49)



So, suppose that the time required for a distance runner to run a mile is a normal random variable with parameters mu is equal to 4 minute 1 second and standard deviation 2 seconds, what is a probability that this athlete will run the mile in less than 4 minutes or in more than 3 minutes 55 seconds, if we consider X as the time required and we consider it in the time measured in seconds, then X will follow normal distribution with mean 241 seconds, here 4 minutes 1 second is 241 seconds and sigma square is 4 seconds. So, what is the probability of running in less than 4 minutes that means, X is less than 240, what is the probability of this event?

So, utilizing this relationship, we can write this as X minus 241 by 2 less than 240 minus 241 by 2, which is probability z less than minus 0.5, which is Phi of minus 0.5. So, this value will see from the tables of the standard normal cdf and this value turns out to be 0.3085- most of the tables are given up to four or five decimal points and we can also use a numerical integration rule such as Simpson's one-third rule, etcetera, to evaluate this. Similarly, if we want to calculate what is the probability that he will run in more than 3 minute 55 seconds, then probability X greater than 235 that is, probability z greater than minus 3, that is if we put 235 minus 241 divided by two, so it becomes z greater than minus three; so, if we look at the shape of the distribution minus 3 suppose here then 3 are here, so this is equivalent to Phi of 3, which is 0.9987, it is extremely high probability. So, here you can see that mean is 4 minute 1 second that means, almost surely he will complete the race within 3 minute within, in more than 3 minute 55 seconds.

This also brings out another important property of the normal distribution. The normal distribution is having high concentration of probability in the center, since you are having in the density function e to the power minus z square by 2 the terms go rapidly towards 0, the convergence towards 0 is very fast; therefore, in the range around the mean mu most of the probability is concentrated. (Refer Slide Time: 26:34)

So, if we consider say, probability of say, modulus z less than 0.5 that is, minus 0.5 less than z less than 0.5, here z denotes standard normal distribution; so, this is equal to Phi of 0.5 minus Phi of minus 0.5. So, we can write because the value of one of them needs to be seen, we need not see both of them, either we see Phi of 0.5 or Phi of minus 0.5 and the other one we can write in terms of 1 minus that; so, this we can write as 1 minus Phi of minus 0.5; so, this becomes 1 minus twice the value of Phi minus 0.5 as 0.30; so, this is equal to 1 minus 2 into 0.3085, that is equal to 1 minus 0.6170, that is equal to 0....

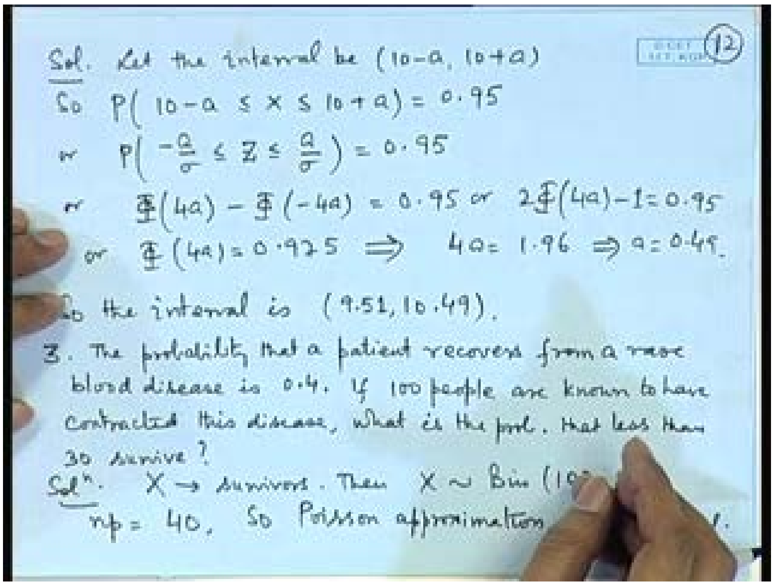
That means, within a very short distance that is, minus 0.5 to 5, itself almost forty percent of the probability is concentrated. If we consider say, minus 1 to 1, this consist of almost 60 percent of the probability; if we consider minus 2 to 2, this consist of more than 90 percent of the probability; if we consider minus 3 to 3, this consist of more than 99.99 percent of the probability.

So, in terms of mu sigma this means that mu minus 3 sigma less than X less than mu plus 3 sigma is greater than 0.99, these are known as 3 sigma limits then, minus 2 to 2 is called as 2 sigma limits. So, in the industrial applications where certain product requirement such as, the width of certain bolts produced, diameters of certain nuts, etcetera, or various kind of quality control features which are implied in industry if they follow the normal distribution, then the industrial standards specify that in order that the product we defined as a good item or proper item, the specification should be within 3 sigma limits. So, in industry industrial standards these things are quite useful. (Refer Slide Time: 29:49)

Let us consider another application. Assume that the high price of a commodity tomorrow is a normal random variable with mu is equal to 10 and sigma is equal to 1 by 4, what is the, what is the shortest interval that has probability 0.95 of including tomorrow's high price for this commodity?

Now, from the properties of the normal distribution that we discussed just now, we should consider here the interval to be asymmetric interval around the mean, because we are requiring a shortest interval, so shortest interval will be symmetric around this, because if we take non symmetric interval then this will become slightly longer because of the tail is converging faster and there is more concentration of the probabilities towards the center.

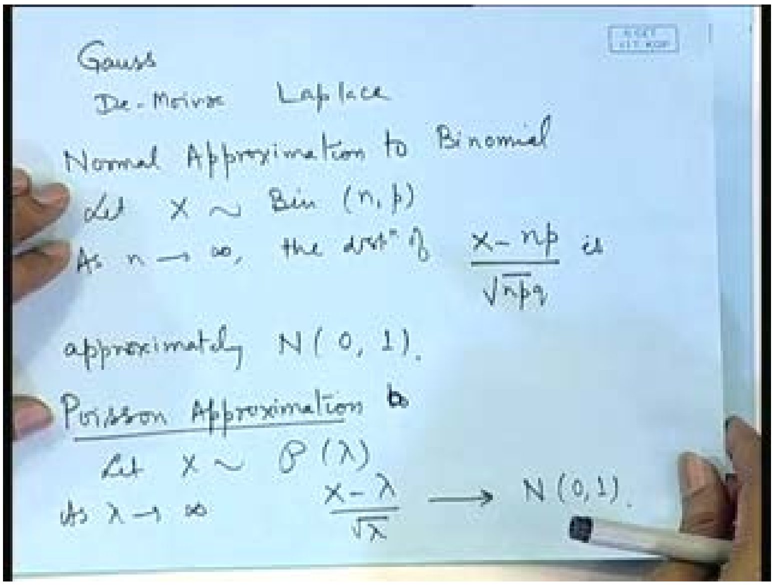
(Refer Slide Time: 37:25)



So, we can consider, since here the mean is 10, we can consider the interval of the form 10 minus a to 10 plus a. So, we want the value of that the probability of X lying in this interval 10 minus a top 10 plus a is 0.95. So, consider the transformation X minus mu by sigma. So, here mu is 4, mu is 10 here, so, X minus mu by sigma, so, it becomes minus a by sigma less than or equal to z less than or equal to a by sigma, and sigma is 1 by 4, so this is 4a; so, this is becoming Phi of 4a minus Phi of minus 4a is equal to 0.95; what is the value of a for which this is satisfying?

Once again we utilize the relation between capital Phi of X and capital Phi of minus X. So, this becomes twice phi of 4a minus 1 is equal to 0.95, and this gives us Phi of 4 a is equal to 0.975- and from the tables of the normal distribution we can see that the point up to which the probability 0.975 is 1.96- so, after evaluation a becomes 0.49; and therefore, the interval 10 minus a to 10 plus a reduces to 9.51 to 10.49. So, if the mean price is 10 and the standard deviation is 0.25, then the interval which will have the high price with probability 0.95 is 9.5 to 10.5 approximately, which is basically 2 sigma, because here sigma is 0.25, so 2 sigma becomes 0.5, so around 10 the interval of 0.25 is, so, in 2 sigma limits we have more than 95 percent of the probability here.

(Refer Slide Time: 39:56)

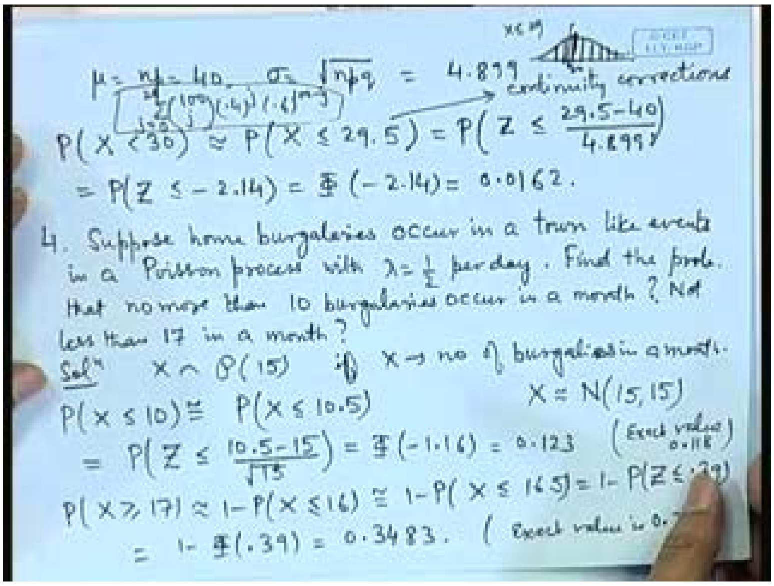


Another important point which I mentioned was that the origin of the normal distribution, the normal distribution was derived as a distribution of the errors by Gauss. So, he was considering astronomical measurements, now, he did not consider one measurement, he considered several measurements and considered the average of those measurements to consider that as the estimate of the actual measurement. So, since in each measurement some error will be concentrated and therefore, if we look at the distribution of the errors, gauss observe that it is normal distribution, that is why it was called, normal distribution is also called the law of errors or the error distribution, etcetera. And it turns out that the sample mean or the sample sum is normally distributed. However, even apart from the gauss, mathematician such as De-moivre and Laplace, etcetera, they obtained the normal distribution as an approximation to binomial distribution, or distribution of Poisson approximated to normal distribution.

So, let us look at it, normal approximation to binomial. So, let X follow binomial n, p. As n tends to infinity, the distribution of X minus np divided by root npq is approximately normal 0. Similarly, Poisson approximation too. So, let X follow poison lambda distribution. As lambda tends to infinity X minus lambda by root lambda, this converges to normal 0. Basically, these were the original central limit theorems, the modern versions are for any random variable X. So, let us look at applications of this year. (Refer Slide Time: 37:25)

The probability that the patient recovers from rare blood disease is 0.4, if 100 people are known to have contracted this disease, what is the probability that less than 30 will survive? So, if we consider here X as the number of survivors, then X follows binomial 100, 0.4, and we are interested to calculate the probability that X is less than 30.

(Refer Slide Time: 43:25)



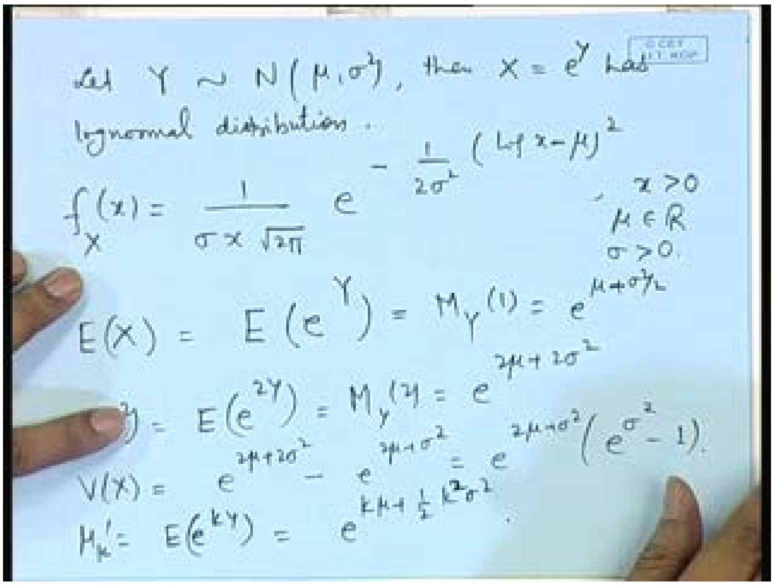
Now, if we look at the actual calculation of this, using the binomial distribution, then this is reducing to sigma 100cj 0.4 to the power j 0. 6 to the power 100 minus j, j is equal to 0 to 29. You can look at the difficulty of the evaluations, of the complexity of the terms here, we have factorials involving 100c0, 100c10, 100c20, 100c25, etcetera, and then the powers of numbers which are smaller than 1, so, the large powers of these numbers will yield lot of computational errors and also the terms will be complex. However, here we can see that N is large, so if we can consider, if we try to look at np is equal to 40 and we want to use Poisson approximation, then that will also be very complicated because that will involve the same summation e to the power minus 40 40 to the power j by j factorial, which again involves large terms here. So, in place of that we will use the normal approximation, so np is 40 and npq is 24, so square root of that is 4.89. So, probability of X less than 30, now, here what we do, we apply so called continuity corrections, this continuity correction is required to approximate a discrete distribution with a continuous distribution.

Consider like this, in the binomial distribution the curve is like this; now, if we are approximating it by a normal distribution and suppose this is 30, so, it could be also less than 29, X less than or equal 29, so in that case if we had approximated it by normal, we should have attained X less than or equal to 29 whereas, here if we write straight away, we will write X less than or equal to 30 because in continuous distribution the probability of a point is negligible, so, a better thing would be to take a middle value between 29 and 30 as 29.5- so, this is called continuity corrections. Now, we make use of the fact that this X minus mu by sigma is approximately normal, so, 29.5 minus 40 divided by 4.899, this is z less than or equal to minus 2.14 that is, the cdf of the standard normal distribution at the point minus 2.14. So, from the tables of the normal distribution we can observe it is 0.0162. So, the probability is quite small that less than 30 will survive.

Let us look at another example where the Poisson distribution is approximated by normal distribution. So, consider a large town where the home burglaries occur like events in a Poisson process with lambda is equal half per day that is, a rate, find the probability that no more than 10 burglaries occur in a month, or not less than 17 occur in a month. So, if we consider this, then X follows Poisson 15 if I denote by X the number of burglaries occurring in a month's time. So, since lambda is half per day in a month we assume 30 days, so the parameter lambda will become15, so, probability X less than or equal to 10. Again we make use of the continuity correction, since it is less than or equal to 10 it is also same as probability X less than11. So, in normal distribution, it could have been calculated as X less than or equal to 10 or X less than or equal to11. So, as a continuity correction we take the midpoint X less than or equal to 10.5. So, if we make use of the normal approximation here, then X minus lambda by root lambda is approximately normal as lambda becomes large, so, here, it is 15 here, so, 10.5 minus 15 by root 15, which are after simplification minus 1.16, so, the value of the normal distributions cumulative distribution functions at this point is 0.123.

Similarly, not more, less than 17 in a month that is, probability X greater than or equal to 17, which is 1 minus probability X less than or equal to 16, or 1 minus probability X less than or equal to 16 by 0.5- that is the continuity correction- and after shifting by 15 and dividing by root 15, it becomes 0.39, which we can see from the tables of the normal distribution, and 0.3483 is the value if we compare with the exact value which we could have calculated from the e to the power minus 15 15 to the power j by j factorial; in the first one we would have j is equal to 0 to 10, then this value is actually 0.118, so, we can see that there is a very small margin of error even for lambda is equal to 15; in the second one the value is 0.3483 by approximation and the exact value is 0.336, so, the error margin is extremely small.

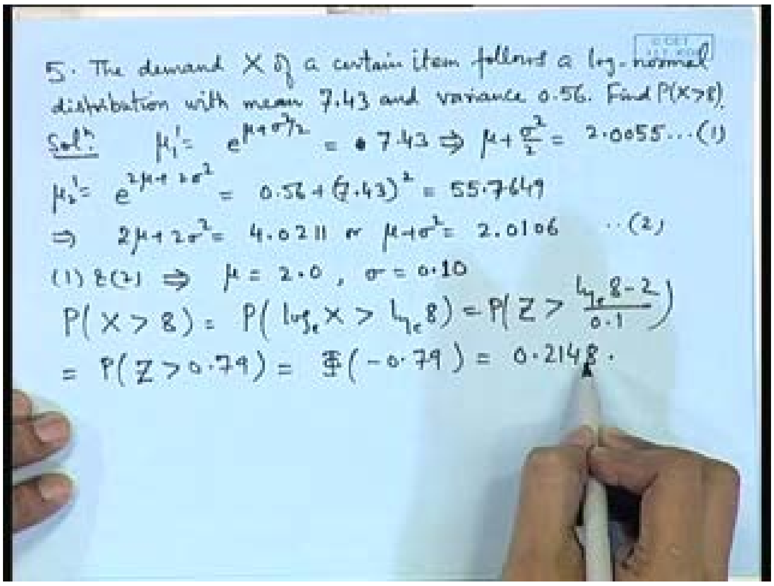
(Refer Slide Time: 50:03)



A related distribution to normal distribution is called lognormal distribution. So, if we say Y follows normal mu, sigma square, then X is equal to e to the power Y has lognormal distribution. So, the density function of log normal distribution will be given by 1 by sigma x root 2 pi e to the power minus 1 by 2 sigma square log x minus mu square, here x is positive and mu and sigma as usual. The necessity of this kind of distribution can be understood like this, that many times X observations may be very large and log x observations may be useful on the other hand, if Y observations are very small, then e to the power Y observations may be of the reasonable size, so if we take e to the power Y, then that will have a lognormal distribution.

The moments of the lognormal distribution are obviously, in the form of the moment generating function of, so, this is nothing but the moment generating function of Y at 1 that is, e to the power mu plus sigma square by 2. If we consider the second moment, then it is equal to expectation of e to the power 2y that is, My 2 that means, variance of X is equal to e to the power 2 mu plus 2 sigma square. In general, mu k prime is equal to expectation of e to the power ky that is equal to e to the power kmu plus half k square sigma square. So, moments of all orders of the log normal distribution exists and they can be expressed in terms of the moments of the moment generating function of normal distribution.

(Refer Slide Time: 52:45)



Lets us look at one application here. Suppose the demand of a certain item follows a lognormal distribution with mean 7.43 and variance 0.56, what is the probability that X is greater than 8? So, if we utilize this formula here, the mean of a lognormal distribution is e to the power mu plus sigma square by 2, the second moment is e to the power 2 mu plus 2 sigma square, so, we have mu1prime is equal to 7.43, which yields the equation mu plus sigma square by 2 is approximately 2, and mu2prime is variance plus mu1prime square, so, 0.56 plus 7.43 square and we substitute here the value for this, so, after simplification this gives the equation mu plus sigma square is equal to 2.0106. So, if we solve these two equations we get mu is approximately 2 and sigma is approximately 0.1.

So, now, the probability of a lognormal random variable can be calculated using normal distribution. So, probability of X greater than 8 reduces to probability logx greater than log8, which is probability z greater than log8 minus 2 divided by 0.1, which is probability z greater than 0.79, and from the tables of the standard normal distribution we can see it is 0.2148. So, this distribution is quite useful in various applications and since it is directly related to the normal distribution the calculations related to this are quite conveniently handled using the properties of the normal distribution. Another thing that you can observe here is that this distribution will be skewed distribution, X is having symmetric distribution, but logx is having skewed distribution here. We can actually the third moment, the fourth moment, etcetera, to consider the measures of skewness and kurtosis, etcetera.

We have considered almost all the important continuous distributions which arise in practice of course, one can say that we are looking at any given phenomena, then what will be the distribution corresponding to that that can be considered by making a frequency polygamma ,(()) histogram, and looking at the data we can see that what kind of distribution will be best suited to describe that data. The distribution that we have discussed so far are the ones which are more important in the sense that they arise in lot of practical applications and also historically they are important as they were considered as certain phenomena like, some physical phenomena, or some genetic phenomena etcetera where they arise. In the next lecture will be considering various applications of these distributions. So, we stop here.