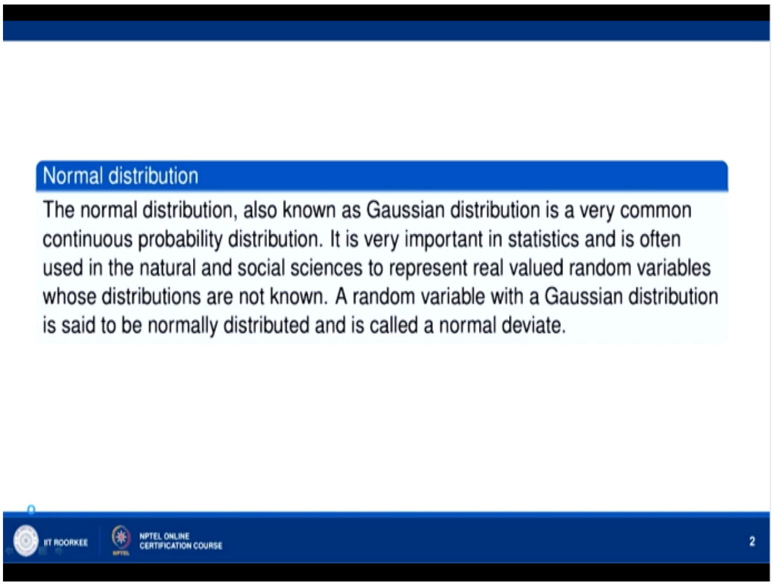
Advanced Engineering Mathematics
Prof. P. N. Agrawal
Department of Mathematics
Indian Institute of Technology – Roorkee

Lecture - 51
Normal Distribution

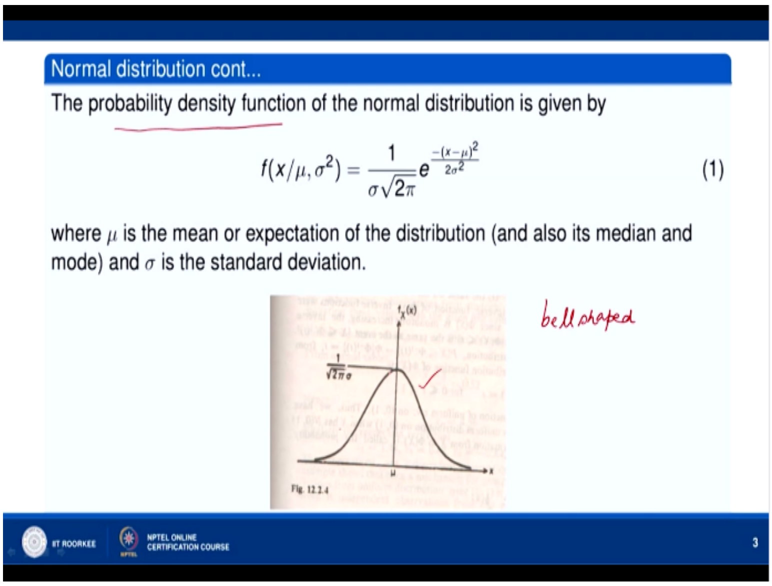
Hello friends. Welcome to my lecture on normal distribution. The normal distribution is also known as Gaussian distribution and it is very common continuous probability distribution.

(Refer Slide Time: 00:42)



It is very important in the statistics and is often used in the natural and social sciences to represent real valued random variables whose distributions are not known. A random variable with a Gaussian distribution is said to be normally distributed and is called a normal variate.

(Refer Slide Time: 00:55)

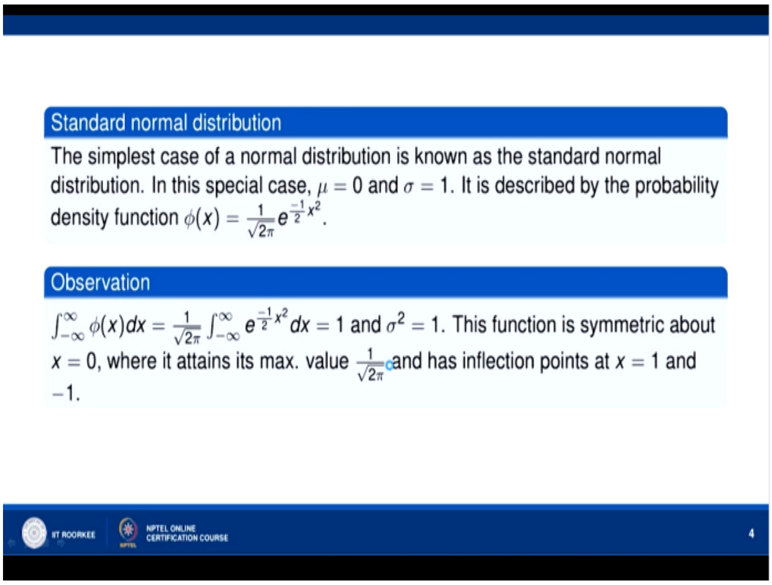


The probability density function of the normal distribution is given by f x mu, sigma square which is given by 1/sigma root 2 pi e to the power –x-mu whole square/2 sigma square where mu is the mean or expectation of the distribution and also its median and mode okay because at the mean, the mean, median and mode okay you can see in this figure, they coincide and sigma is the standard deviation.

This is the graph of the probability density function of the normal distribution okay. So this is the graph of the probability density function of the normal distribution. You can see that at x=mu the function fx, x assumes its maximum value which is given by 1/sigma root 2 pi and it is a bell-shaped curve okay. The shape of the curve is bell-shaped, so it is a bell-shaped curve okay.

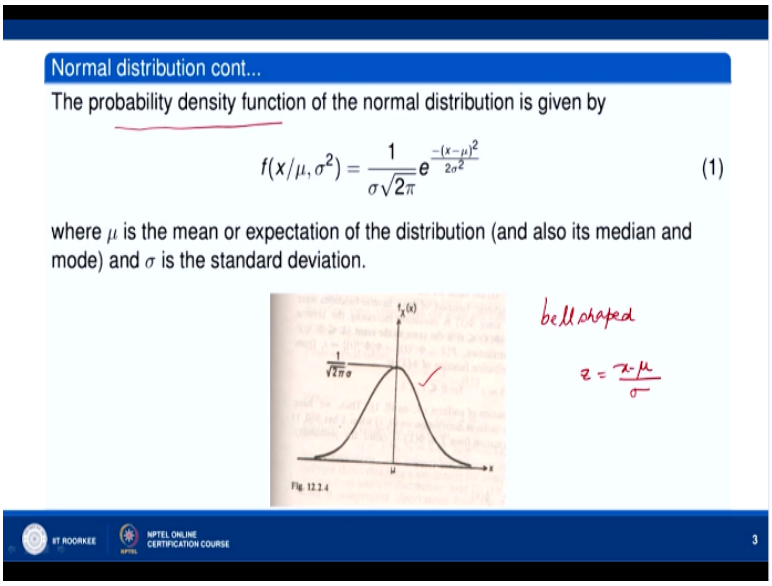
Mean, median and mode coincide for this distribution okay, mu is the mean and sigma is the standard deviation. The maximum value occurs at x=mu which is given by 1/sigma root 2 pi.

(Refer Slide Time: 02:15)



Now the simplest case of a normal distribution is known as the standard normal distribution.

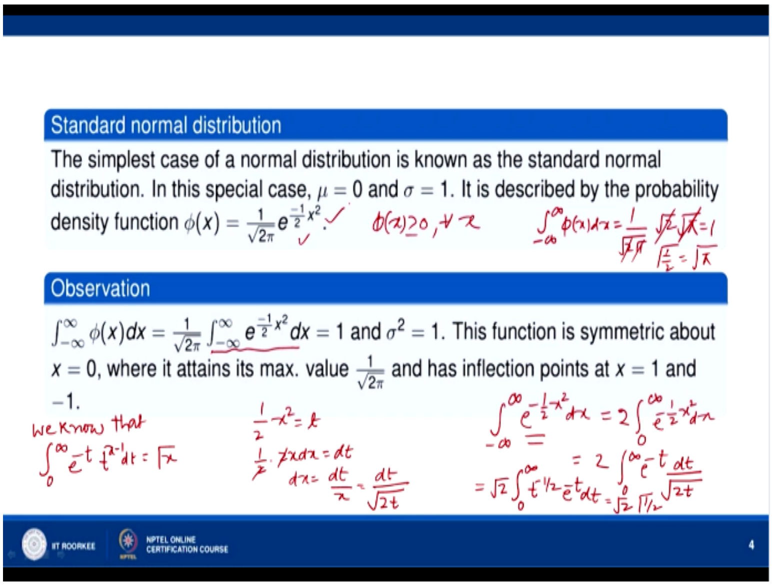
(Refer Slide Time: 02:24)



In this case, what we do is we put here z=x-mu/sigma okay. So when you put z=x-mu/sigma then what happens is the mean of the distribution, mean of the distribution for the standard normal distribution is=0 okay, mu becomes 0 and sigma becomes=1. It is described by the probability density function 1/root 2 pi e to the power -1/2 x square. So here if you take instead of x-mu/sigma z then this becomes e to the power -1/2 z square.

And the probability density function becomes 1/root 2 pi e to the power –z square/2 or we can also write it as 1/root 2 pi e to the power –x square/2. Now this is the probability density function for the standard normal distribution can be seen from here.

(Refer Slide Time: 03:28)



This you can see that phi x is>=0 for all x okay, phi x is>=0 for all x which is the first condition for the probability density function. Second is that integral/-infinity to infinity phi x dx must be=1. So that we can see here 1/root 2 pi integral/-infinity to infinity e to the power -1/2 x square dx. So let us integrate this. Integral over –infinity to infinity e to the power -1/2 x square dx.

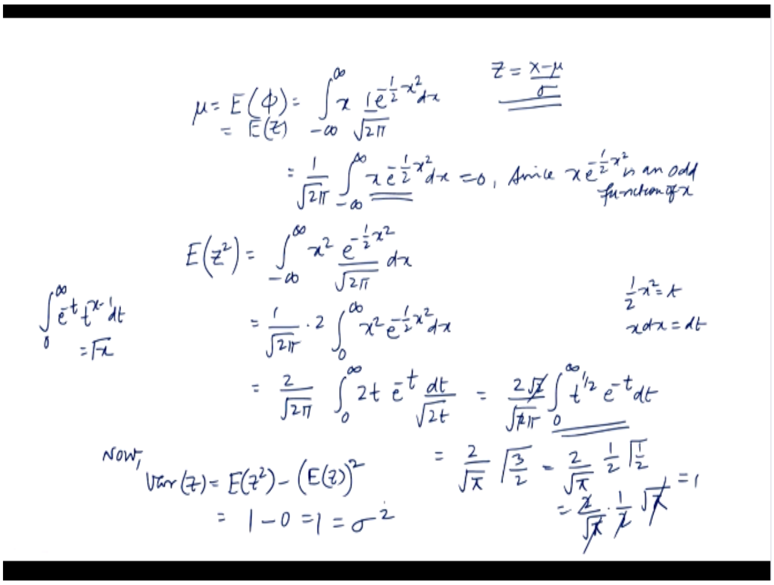
Since e to the power -1/2 x square is an even function, we can write it as 2 times 0 to infinity e to the power -1/2 square dx 1/2 x square dx. Now let us take 1/2 x square to be=t okay. So then what we get 1/2*2x dx=dt okay. So we get dx as dt/x, x is=root 2t okay. So dt upon root 2t okay, so let us put it here. Now when x is=0, t is 0 and when x is infinity, t is infinity, so the limits remain the same 0 to infinity.

We get e to the power –t and for dx we put dt/root 2t okay. So what we get then? This root 2 we can cancel with here 2 and we get root 2 and integral 0 to infinity t to the power -1/2 e to the power –t dt okay. Now we can write the value of this integral by using gamma function. We know that integral 0 to infinity e to the power –t t to the power x-1 dt is=gamma x okay. So here we have t to the power -1/2, we can write it as t to the power 1/2-1.

So this is gamma 1/2 okay. So this is root 2*gamma 1/2 okay. So what we get then okay thus we have found the value of this integral. Let us evaluate the integral over –infinity to infinity phi x dx now. So this will be=1/root 2 pi okay and here we get root 2*gamma 1/2 is root pi okay, so root 2*root pi okay. So this cancels with this and this cancels with this and we get 1 okay. Remember gamma 1/2 is=root pi okay.

So this integral over –infinity to infinity phi x dx=1 and therefore we can say that this phi x is a probability density function when we replace x-mu over sigma/z which we have taken as x here okay. Now we can also see that for this probability density function okay, the standard normal distribution mu is=0, sigma=1.

(Refer Slide Time: 07:12)



So let us find mu okay. So mu is the expectation of the standard normal distribution phi okay. So we have integral over –infinity to infinity okay, we have taken okay so x times e to the power -1/2 x square dx and we also have 1/root 2 pi okay, 1/root 2 pi is the multiple of, so let me put it like this x times this okay. So x times the probability density function which is 1/root 2 pi e to the power -1/2 x square dx.

Now this is=1/root 2 pi okay. We have integral over –infinity to infinity x e to the power -1/2 x square dx okay. Now you can see that this is an odd function of x okay. This is an odd function of x, so integral over –infinity to infinity x e to the power -1/2 x square dx is=0 okay. Since x e to the power -1/2 x square is an odd function of x okay. Now let us also find the variance of this.

So we need to find Ex square that is Ez square, this is actually phi we are writing here the z=x-mu/sigma okay, z is the standard normal distribution, so we are writing Ez square, this is actually Ez, you can write Ez here, expectation of z okay. So this expectation of z square let us find. This will be integral over –infinity to infinity, so x square we write x square okay e to the power -1/2 x square/root 2 pi okay.

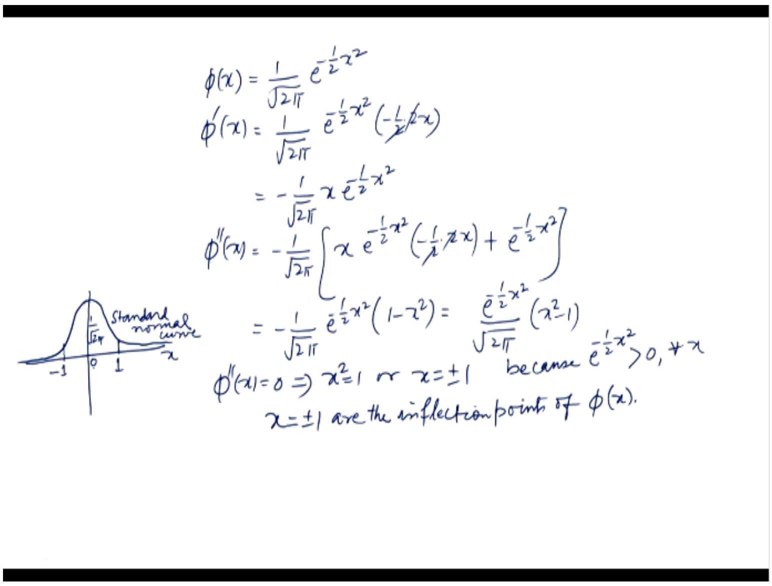
So now this is 1/root 2 pi x square e to power -1/2 square is an even function, so we can write it 2 times 0 to infinity x square e to the power -1/2 x square dx okay. Now let us put 1/2 x square=t as we have done earlier. So then this will be x dx=dt, so we shall get and the limits of integration will remain 0 infinity. So this is 2/root 2 pi 0 to infinity x square is=2t okay, x square is 2t e to the power –t and dx is=dt/x that is square root 2t okay.

So this is=2/root 2 pi, root 2t will cancel with 2t and will give you root 2t, so root 2 here integral 0 to infinity t to the power ½ okay, t to power 1/2 means t to power 3/2-1. So this is like this we get okay. So this root 2 will cancel with 2 here, root 2 here and we get 2/root pi and this is gamma 3/2 okay. This can be evaluated from here 0 to infinity e to the power –t t to the power x-1 dt=gamma x okay.

So this is t to power 3/2-1, so gamma 3/2 which is=2/root pi 1/2 gamma 1/2 okay. So this is 2/root pi 1/2 square root pi and this cancels with this, this cancels with this. So Ez square is=1 okay. Now variance of z is=Ez square-Ez whole square okay. This is Ez, so Ez whole square, so this is 1-Ez is 0 we have seen okay. So we get variance as 1. Variance is nothing but sigma square okay.

So sigma square is=1 okay. So here we can see that mean is=0 and sigma square=1 or we can say sigma=1. Now this function you can see 1/root 2 pi e to power -1/2 x square is symmetric about x=0 okay. We replace x by –x it does not change, so it is symmetric about x=0 where it attains its maximum value. It attains its maximum value at x=0 okay and it is 1/root 2 pi. It has inflection points at x=1 and x=-1, so this we can see.

(Refer Slide Time: 12:29)



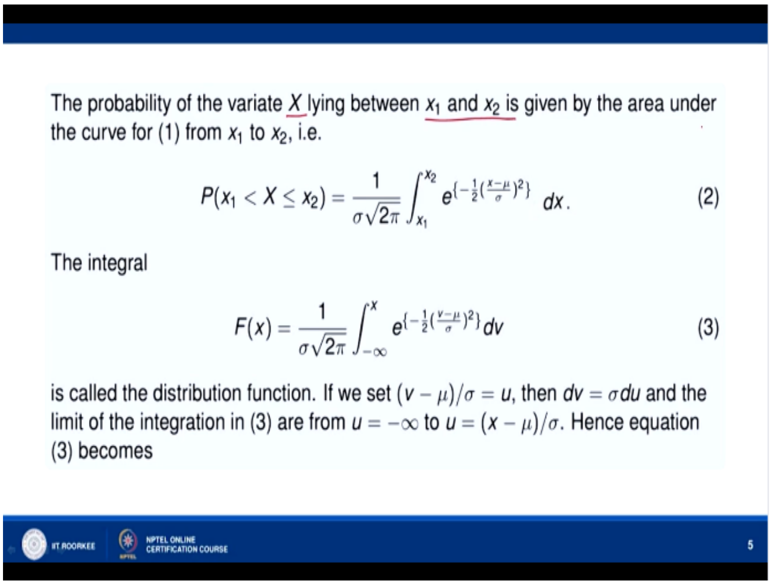
We have phi x=1/root 2 pi e to the power -1/2 x square okay. So let us find its first derivative phi dash x. So 1/root 2 pi, we have e to the power -1/2 x square*-1/2*2x. So we get -1/root 2 pi x e to the power -1/2 x square okay. Now let us find phi double dash x. So -1/root 2 pi and we have x*e to the power -1/2 x square*-1/2*2x okay+derivative of x that is 1. So e to the power -1/2 x square okay.

Now what do we get here? So this is -1/root 2 pi e to the power -1/2 x square we can write outside and we have 1-x square here okay or we can say e to the power -1/2 x square/root 2pi*x square-1 okay. So phi double dash x=0 implies that x square=1 okay or x=+-1 because e to the power -1/2 x square is never 0 okay, it is never 0, it is in fact >0 okay for all x okay. So if we have a differential function phi x and it turns out that phi double dash x=0, then the points where phi double dash x=0 gives us the inflection points.

What are the inflection points? Inflection points are those points on the curve where the concavity of the curve changes okay. So you can see how it changes. So x=+-1 are the inflection points okay of phi x. So if you draw the graph of phi x, it will look like something this okay. This is your x-axis, this is 0, x=0 at x=0 it assumes maximum value 1/root 2 pi and then it drops down okay and at x=+-1.

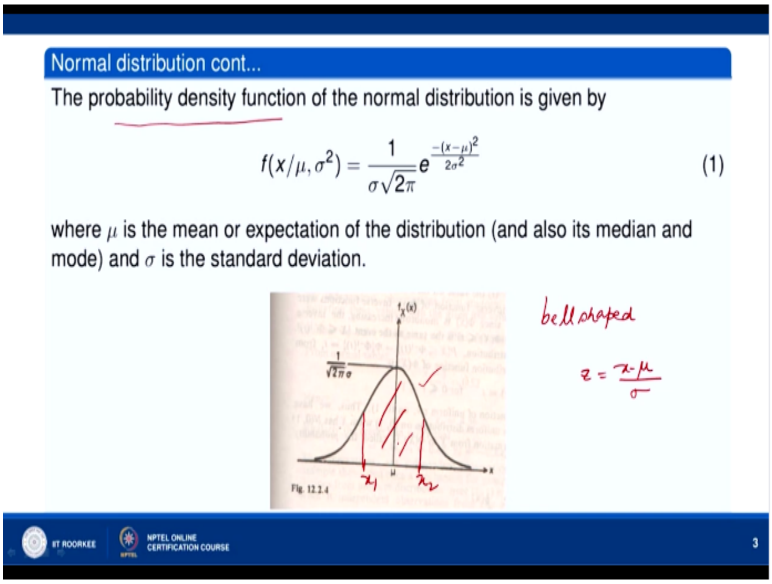
So this is the point x=1 okay where the concavity of the curve changes. It was downward earlier, now it becomes upward okay. So let me write again. So okay so here okay the concavity of the curve changes. This is -1 and this is 1, so concavity of the curve changes and this is 1/root 2 pi. This is your standard normal curve we call it, standard normal curve okay. So you can see it is symmetric about the y-axis. Now let us go further.

(Refer Slide Time: 16:20)



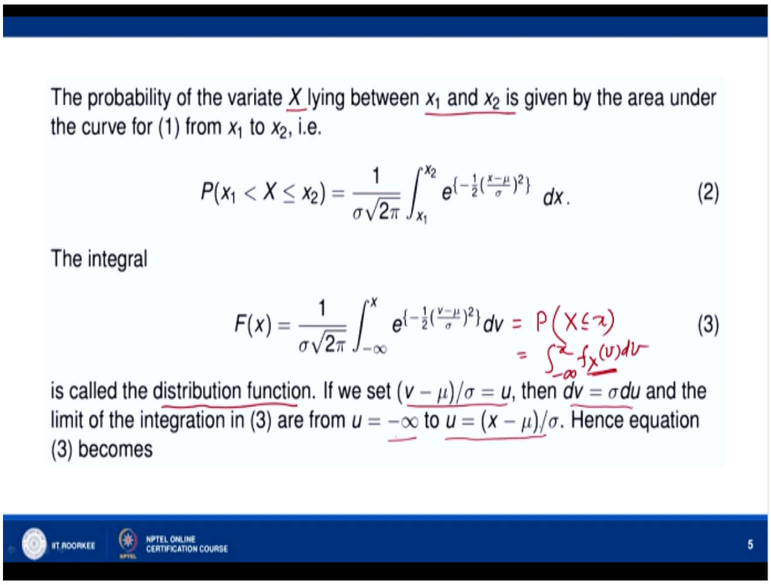
The probability of the variate X the normal variate okay, the probability of the normal variate X lying between x1 and x2 is given by the area under the curve for 1 under this curve okay under the curve for 1.

(Refer Slide Time: 16:35)



Say x1 is here, x2 is here okay, then the area under this curve from x1 to x2 okay, this area okay. So this area will give us the probability of the normal variate from x1 to x2. So that probability is given by integral over x1 to x2 1/sigma root 2 pi integral/x1 to x2 e to the power-1/2 x-mu whole square, x-mu/sigma whole square dx okay. The integral Fx okay Fx=1/sigma root 2 pi integral over –infinity to x e to the power -1/2 v-mu/sigma whole square dv then gives us the cumulative distribution function of x.

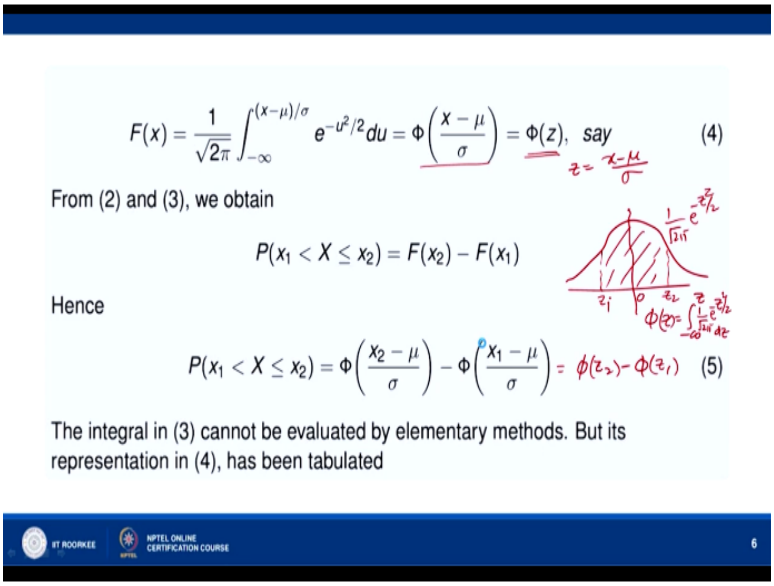
(Refer Slide Time: 17:25)



This is nothing but you can see probability that X is<=x okay, probability that x is<=to x is what? It is given by integral over –infinity to x probability density function of x okay. This we know. So its probability density function here is 1/sigma root 2 pi okay. Instead of x we are using the variable v, so we can write variable v here okay. So integral over –infinity to infinity Fx v dv and this is 1/sigma root 2 pi integral over –infinity to x e to the power -1/2 v-mu/sigma whole square dv okay.

This is called as the distribution function or cumulative distribution function for the normal variate okay. Now if we put v-mu/sigma=u here okay, v-mu/sigma=u if you put then you can see dv/sigma=du, so dv=sigma du and the limits of the integration change from –infinity to x to –infinity to u=x-mu/sigma okay.

(Refer Slide Time: 18:40)



Hence, equation 3 becomes Fx=1/root 2 pi –infinity to x-mu/sigma e to the power –u square/2 du and this is a function of x-mu/sigma, so we can write it as phi x-mu/sigma and this phi x-mu/sigma then becomes phi z where z is the standard normal variable okay, z=x-mu/sigma as we have discussed earlier, z=x-mu/sigma. So x-mu/sigma is replaced by z. Now from 2 and 3 okay from this and this okay let us see what happens.

From 2 and 3, we obtain that probability that x1<X<=x2 is then Fx2-Fx1 because Fx is the cumulative distribution function okay. It gives the probability that x is<=X, so Fx2-Fx1 will give us the probability that x1 is<X<=x2. So if you put the value of x2 here, then Fx2 becomes phi x2-mu/sigma, so probability that x1 is<X<=x2 can be now written as phi x2-mu/sigma-phi x1-mu/sigma.

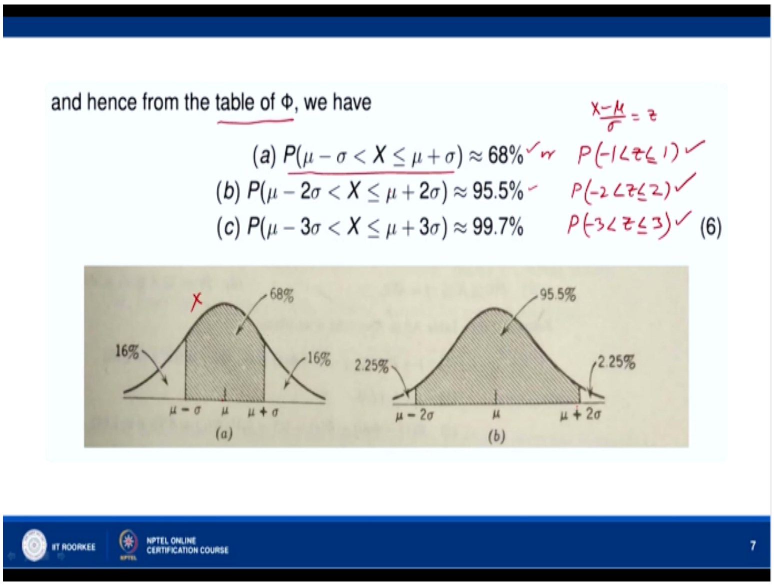
This can then be written as phi z2-phi z1 which will be nothing but the area under the standard normal curve okay, standard normal curve, this is standard normal curve phi z okay from z1 to z2 okay, phi z2-phi z1 okay. This is the area under the graph of phi z from z1 to z2. So the integral in this phi z represents the area, remember phi z represents the area under the graph of the density function okay.

Density function is this. Let me write density function, density function is 1/root 2 pi e to the power –z square/2 okay. So this is the graph of 1/root 2 pi e to the power –z square/2 okay and phi z represents the area phi z is the area from –infinity to z okay 1/root 2 pi e to the power –z square/2 dz okay. So phi z2-phi z1 gives this area okay under the standard normal curve okay.

The integral in 3, now integral here cannot be evaluated by elementary methods okay, this integral, you can see this integral cannot be evaluated by the elementary methods that we know in integral calculus. So it has been numerically evaluated and it is available in the form of a table. Now it is given in all standard engineering mathematics books at the end as an appendix. So you can refer to that table to determine the value of this integral okay.

Now this integral has been evaluated for the standard normal curve okay for this curve okay, 1/root 2 pi e to power –z square/2 that means this phi z has been evaluated okay. This area under the graph of the function 1/root 2 pi e to power –z square/2 has been evaluated for any given z. You can take any given z here when the area under the graph has been evaluated. So that we can use, we shall see just now when we do examples how to use that table. So we will evaluate the probability using that table okay for the standard normal curve.

(Refer Slide Time: 22:54)



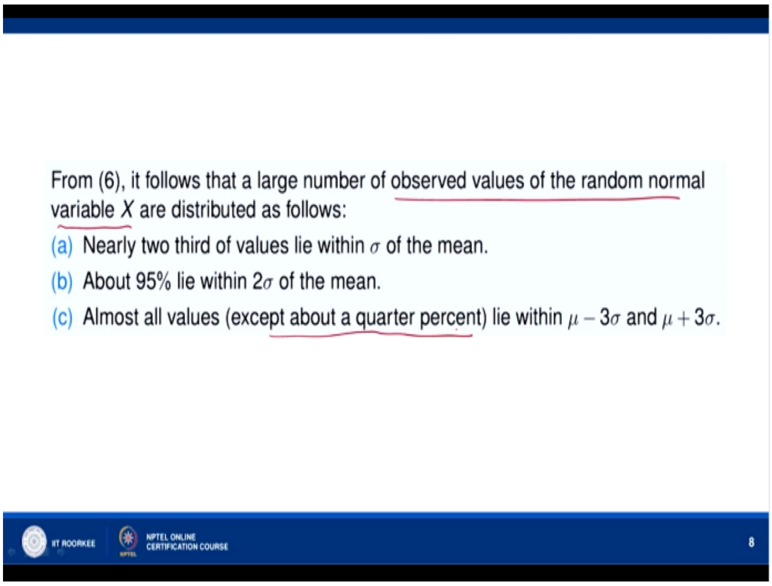
Now from the table of phi okay, from the table of phi it follows that the area or you can say the probability when x lies between mu-sigma to mu+sigma okay is 68%, this is say this is your mean okay, mean of the normal variate x okay, normal variate x, so on either side when you go to by a distance sigma, so then from mu-sigma to mu+sigma 68% of the area lies under the graph of the density function of x okay.

That means two-third of the values of x lie in the region between mu-sigma to mu+sigma approximately and then from mu-2 sigma to mu+2 sigma 95.5% values of x lie and when you take mu-3 sigma to mu+3 sigma 99.7% of the values lie that means about a quarter percent remain outside this region which lies between mu-3 sigma to mu+3 sigma. If you convert this to standard normal variable z then what it means?

You put x-mu/sigma is=z okay so this is same as probability that -1 is<z<1 okay and this means probability that -2 is<z<=2 okay and this means probability that -3 is<z<=3. So from the table of phi from the standard normal curve, we have calculated the area which lies between z=-1 to z=+1 that area is 68%, so that area is 68% for the normal variate x here. We have converted this z to this x here okay.

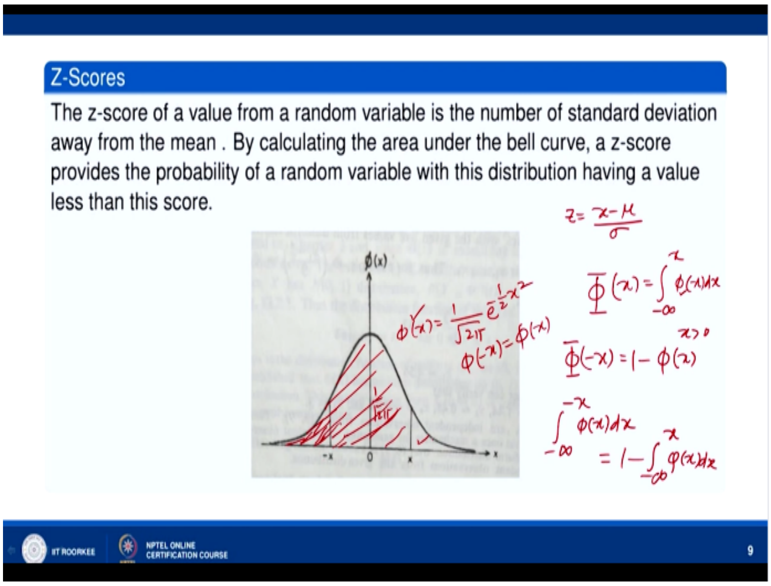
So from these areas we have concluded about the values of x that lie in those ranges mu-sigma to mu+sigma it is 68%, mu-2 sigma to mu+2 sigma it is 95.5% and mu-3 sigma to mu+3 sigma it is 99.7% okay.

(Refer Slide Time: 25:32)



Now from these inequalities, from these values it follows that a large number of observed values of the random normal variate X are distributed as follows. Nearly two third of values lie within sigma of the mean, about 95% lie within 2 sigma of the mean, almost all values except about a quarter percent okay 0.25%. If you leave 0.25%, all other values lie between mu-3 sigma to mu+3 sigma okay.

(Refer Slide Time: 26:00)



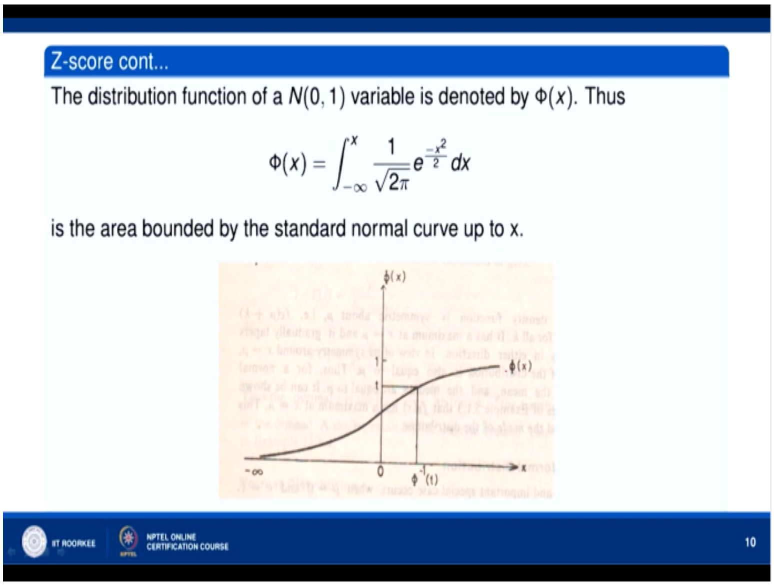
The z-score of a value okay z is x-mu/sigma we call it z square. So z square of a value from a random variable is the number of standard deviation away from the mean. You can see from the definition, z is nothing but the number of standard deviation away from the mean. By calculating the area under the bell curve, this is bell curve okay. This is the curve for z okay standard normal curve.

We have written here x for z okay. You can see at x=0, it assumes the maximum value 1/root 2 pi. So this phi x here is 1/root 2 pi e to the power -1/2 x square okay. So this is 1/root 2 pi. You can see that phi-x is=phi x okay, so it is symmetric okay. It is symmetric with respect to y-axis and the area under capital phi x okay gives the area under the graph of phi x okay –infinity to x phi x, note the difference between this phi and this phi okay.

This is capital phi okay. It is the area under the graph of this probability density function phi. This is probability density function phi x=this probability density function. So under the graph of phi x from –infinity to x, so this area okay. Suppose x is positive okay, if x is positive then you can see because of the symmetry of phi x about the y axis, the area under the graph of phi x from –infinity to –x okay.

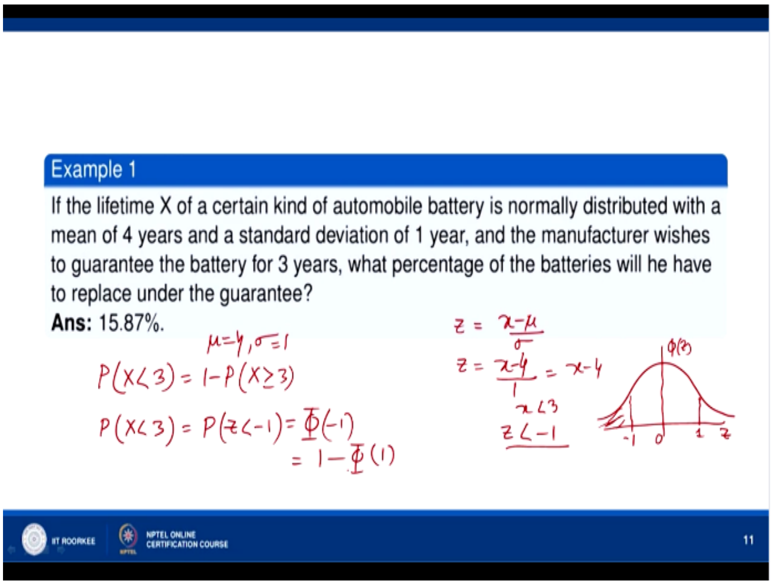
Phi-x is=1-phi x okay. If you want this area okay, to obtain this area okay, what we do, the total area under the curve is=1, from 1 we subtract this entire area okay. So then we will get this area. So this area integral over –infinity to –x okay, phi x dx okay area under the graph of phi x from –infinity to –x can be obtained by subtracting the area under the graph of phi x from –infinity to x okay because of symmetry okay.

(Refer Slide Time: 28:43)



So now the distribution function, this is capital phi x okay like which I have talked about here this capital phi x okay. So this is capital phi x, capital phi x is the cumulative distribution function of a small phi x okay. It is denoted by given by –infinity to x 1/root 2 pi e to power -1/2 x square dx and is the area bounded by the standard normal curve up to x okay. This is the graph of standard normal distribution okay, this phi x okay, this is the graph okay.

(Refer Slide Time: 29:18)



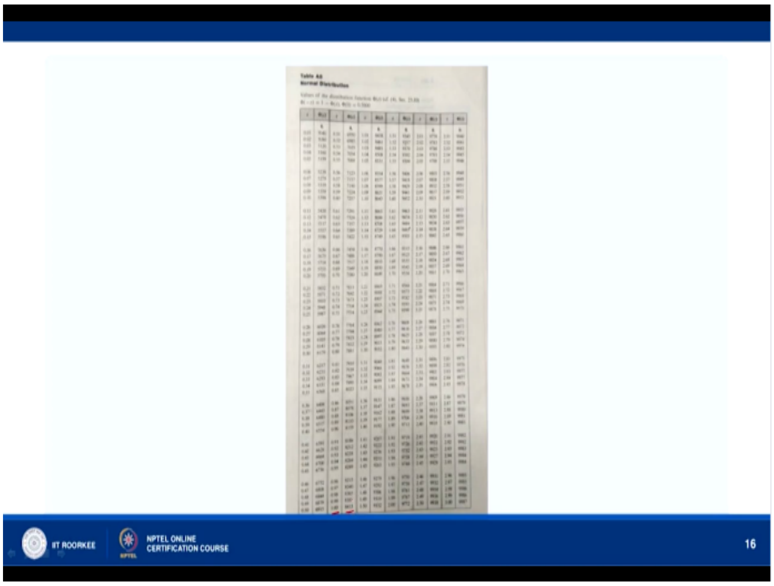
Now what we do? Let us go to first example. If the lifetime X of a certain kind of automobile battery is normally distributed with mean of 4 years. So mu is given to be 4 years okay, standard deviation sigma is given to be 1 year okay and the manufacturer wishes to guarantee the battery for 3 years. That means he says that the battery will at least work for 3 years okay. What percentage of the battery will he have to replace under the guarantee?

That means we want to find the percentage of those batteries which will not last for 3 years okay. So probability that X is<3 okay, this is what we want okay and probability X<3 is=1-probability that X is>=3 okay. Now we will change this probability X<3, we will change this variable X to z okay, so z=x-mu/sigma okay, so mu is=4, so z will be=x-4/1 okay, so z=x-4 okay. We want the probability that x is<3 okay.

Probability that x<3 means okay, this x is<3 means z is<-1 okay because x is<3, so this z is<3-4 so z is<-1. So this is=probability that z is<-1 okay. Now let us look at the graph okay, graph of the standard normal curve. This is standard normal curve okay 0 and z=this is zx okay, this is phi z okay the probability density function phi z. So this is z=-1 here okay. As I said the area under the graph of phi z okay from –infinity to -1 can be found by subtracting the area from –infinity to 1 okay, -infinity to 1 from 1.

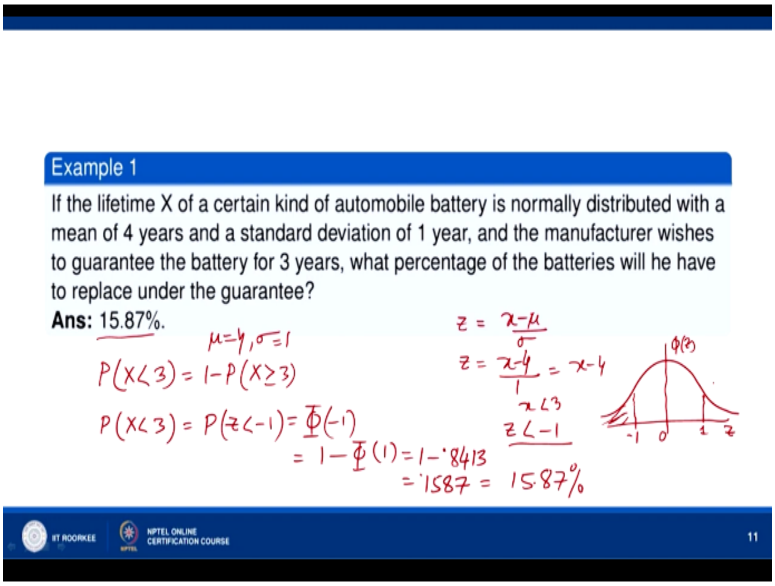
So this is=phi-1. Let me write like that in terms of the capital phi function, so phi-1 okay. This is=1-phi 1 okay. Now let us go to the table. See the value of phi 1 from there.

(Refer Slide Time: 31:59)



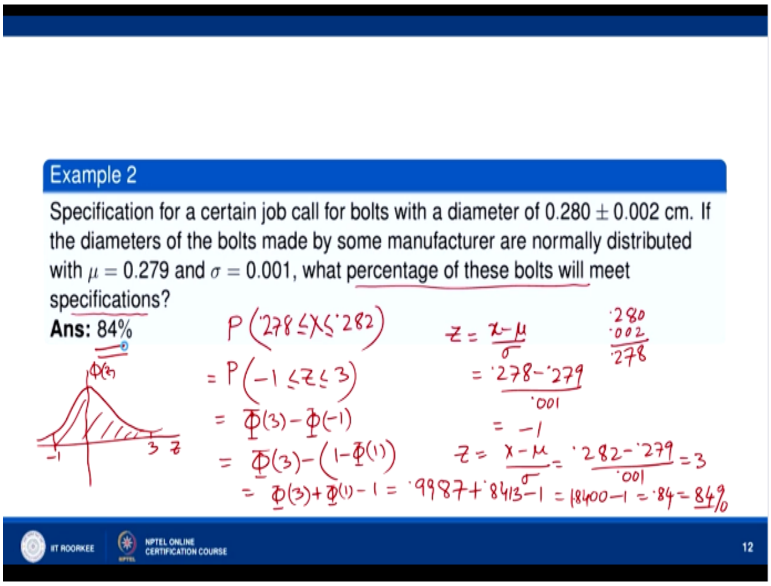
This is the table, you can see in this table when z is=1, phi is=8413 okay, this is the value of z and this is the value of capital phi z okay. So when z is=1, this is 8413 okay, so let us put that value there.

(Refer Slide Time: 32:21)



So this is 1-0.8413 okay, so this is 1587 okay. So this means that it is=15.87% okay. So he will have to replace 15.87% of the batteries because they will not last for 3 years or more okay. So this is how we will use the area under the graph of the standard normal curve to solve the problems.

(Refer Slide Time: 33:00)



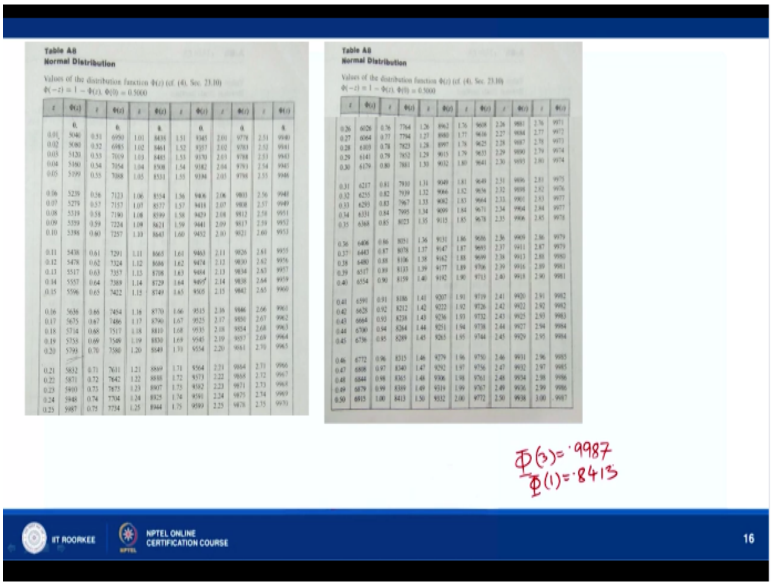
Now let us go to a specification for a certain job call for bolts with a diameter of 0.280+-0.002 centimeter. If the diameters of the bolts made by some manufacturer are normally distributed with mu=0.279 and sigma=0.001, what percentage of these bolts will meet specifications? So specifications are that the diameter of the bolt should lie between this range okay.

So let us see X should lie between 0.280-0.002. Let us find the range of values, so we subtract first 0.002 and get here 278 okay. So 0.278 okay and then when we add 0.002 to 0.280, we get 0.282. So we want the percentage of bolts which will meet the specifications that means we want the probability that the X denotes the diameter of the bolts. So diameter if it lies in the range 0.278 to 0.282, then it is acceptable okay, it needs the specification.

So now let us calculate the corresponding values of z. So z=x-mu/sigma okay. So we get z=0.278-mu=0.279/0.001. So this is=-1 okay and z=x-mu/sigma when x=0.282, so we get 0.279/0.001. So this is=0.003/0.001, so this is 3 okay. So this is=probability that z lies between -1 to 3 okay. Now let us look at the graph okay again. So suppose this is graph, so -1 is here okay, this is your z-axis okay, this is your phi z function okay, probability density function phi z.

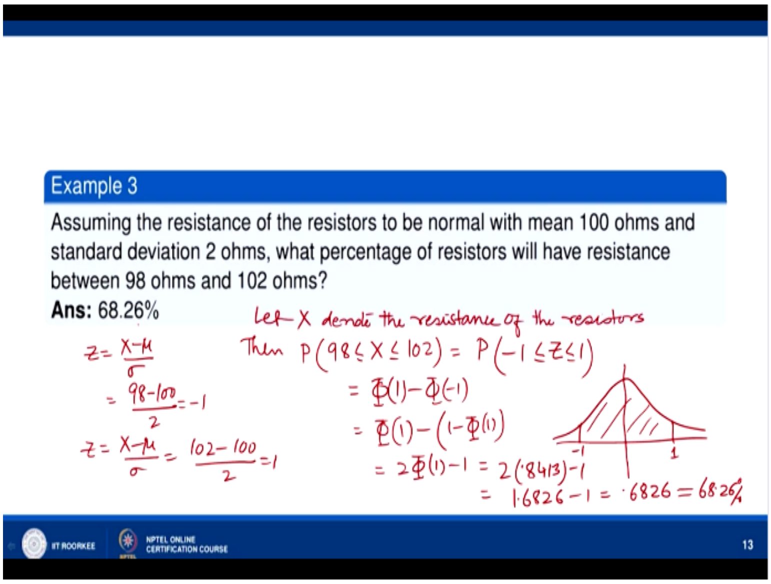
Now this probability is we have to find, so -1 to 3 we want okay this is 3 here, this area we want -1 to 3 okay. This means what? We find the area under the graph of the curve from –infinity to 3 and subtract the area from –infinity to -1. So this is phi 3-phi-1 okay and phi-1 can be found by subtracting the value of phi at 1 from 1 okay. So phi 3-1-phi 1 okay, so what we get phi 3+phi 1-1 okay. Now let us see the value of phi 3 and phi 1 and then we will add them and subtract 1.

(Refer Slide Time: 36:49)



So we can see here the value of phi 3 okay, phi 3 is 0.9987 okay, this is 0.9987, from here you can see, phi 3=0.9987 okay and phi 1 we had already seen, phi 1=0.8413 okay. So let us use these values. So 0.9987 +0.8413-1 okay, so how much is that, 7+3=10, 1 carry out we have 9 1 0 and then we have 184 okay, so 18400-1 okay, so 0.84 that means 84% of the bolts meet specifications okay, so 84% of the bolts meet the specifications.

(Refer Slide Time: 38:05)

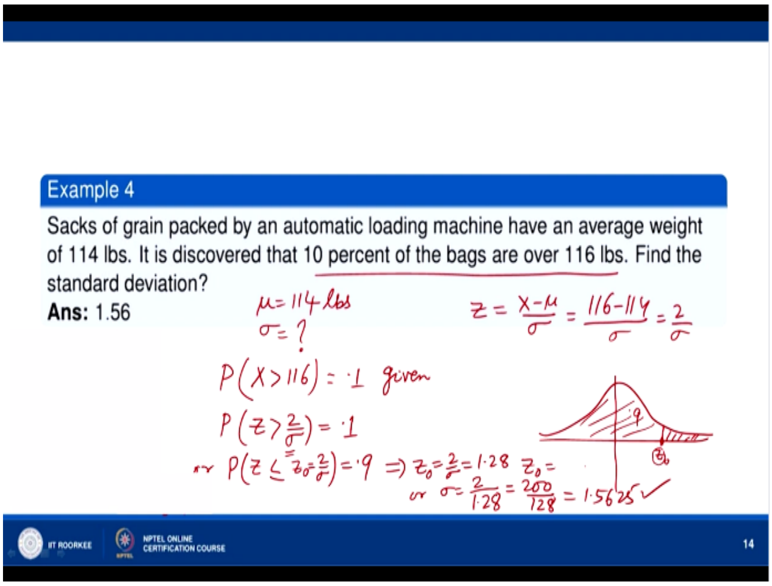


Now assuming the resistance of the resistors to be normal with mean 100 ohms and standard deviation 2 ohms okay. What percentage of resistors will have resistance between 98 and 102 ohms? So we have let us say X denotes the resistance of the resistors okay. Let X denote resistors okay. Then, we want the probability that X lies between 98 and 102 okay. So this is=probability that z lies between the z=x-mu/sigma, so z will have the values 98-100/2, so that is=-1 okay.

And when X is 102, it is +1, 102-100/2, so +1, so z lies between -1 and +1 okay. So now what we will do, let us see okay. So this is your -1 okay. This is +1 okay. We want the area under the standard normal curve from -1 to +1 okay. So what we will do? We will find the area under the standard normal curve from –infinity to 1 and from that we subtract the area under the standard normal curve from –infinity to -1.

So this is phi 1-phi-1 okay and phi-1 we know is 1-phi 1. So phi 1-1-phi 1 okay. So this 2 times phi 1-1 okay and we have seen phi 1=0.8413, so we can use this value here. So 2 times 0.8413-1 okay, so this is 1.6826-1 okay. So this is 0.6826 meaning that it is 68.26% okay. So 68.26% resistors will have resistance between 98 ohms and 102 ohms okay.

(Refer Slide Time: 41:12)



Now we go to this last question. Sacks of grain packed by an automatic loading machine have an average weight of 114 pounds. So here we are given mu, mu=114 pounds. It is discovered that 10% of the bags are over 116 pounds. Find the standard deviation. We have to find sigma okay; we have to find sigma here. We know that z is=x-mu/sigma okay. We are given that probability that x is more than 116 pounds okay is=0.1 okay, this is given to us.

The 10% of the bags are over 116 pounds, x denotes the weight of the bag, so probability that x is>116 is=0.1 which is given to us. So z will be=let us see the corresponding value of z, so z will be=116-114/sigma. So 2/sigma okay, so probability that z is>2/sigma is=0.1 okay. Now we have to see from the table that for what value of z we get the probability to be 0.1, you see let us look at this okay.

Suppose this 0.1 is the area okay, so suppose this value of z okay and z is>2/sigma. So this area is given to us okay. We want to find the value of z for which when z is>2/sigma okay z is>z0 suppose this is z0 okay. So when z is>z0, the area becomes 0.1, so that means that the area to the left of z0 okay. Suppose 2/sigma is z0 okay, so we want this area, this area is 0.9 okay. So let us find the under the graph of standard normal curve for what value of z0 okay we have the area 0.9 okay.

So let us go this one. So here you can see. This is I think 1.26, 1.26 gives you the area close to 0.9 okay. So we have let us take, so z0 is 1.26 okay. So then this is or you can say probability that z is<=z0 okay, z0 is=2/sigma okay is=0.9. So then we obtain z0=1.2 sigma, so z0=2/sigma implies that 2/sigma=1.26 or sigma=2/1.26, so how much is that? This is 200/126 okay.

So what we have? And then we have 1.28, so 2/1.28 okay, so we get okay so 1.5625 okay. Let us put the exact value okay. So what we get is 1.5625. So the standard deviation is 1.5625 okay. So with this I would like to end my lecture. Thank you very much for your attention.