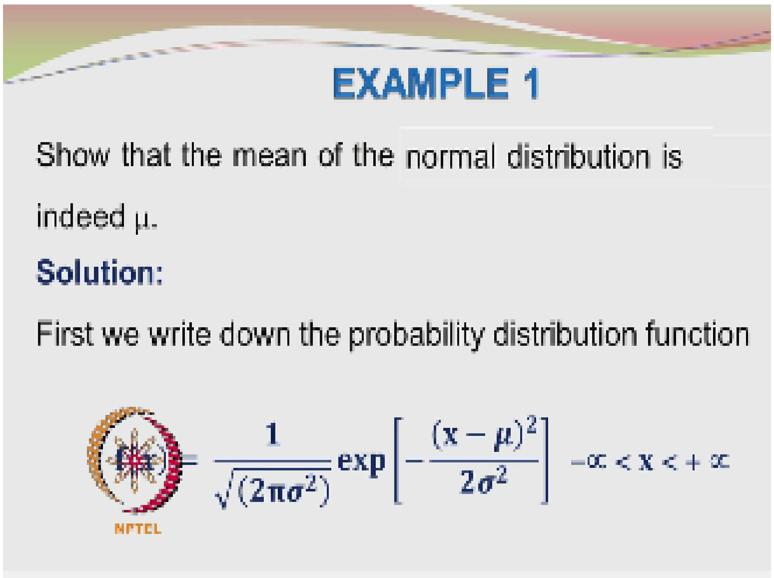
Statistics for Experimentalists
Prof. Kannan. A
Department of Chemical Engineering
Indian Institute of Technology – Madras

Lecture – 08
Example Set II

Hello, welcome back, we have already completed one example set in today's lecture, we will be looking at the second example set. As I had suggested, please read the problem statement and try to work out the problem on your own and compare the answers, after the answers are discussed; compare your answers with the answers that are discussed in the power point.

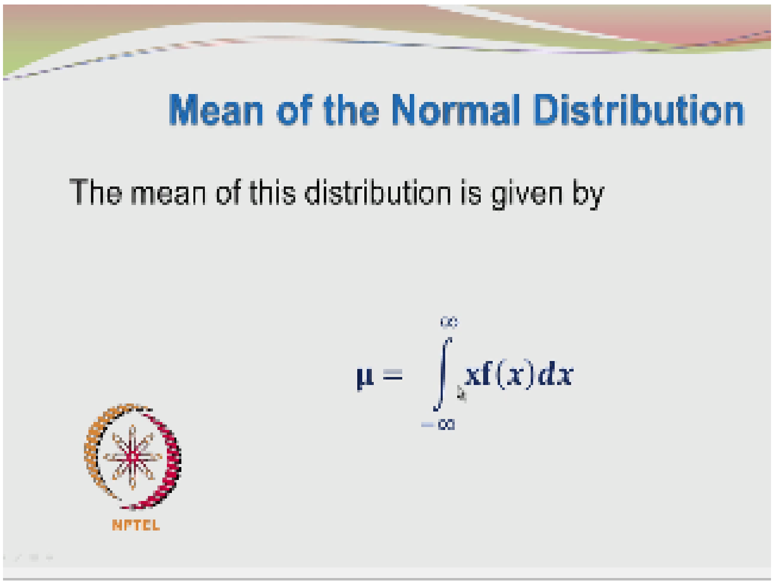
(Refer Slide Time: 00:52)



The first example, we are going to take the normal distribution. We know that the parameters of the normal distribution are Mu and sigma; Mu directly standing for the mean of the distribution and sigma standing for the standard deviation of the distribution, this is a unique case as in many other cases, the parameters of the distribution do not directly correspond to Mu and Sigma.

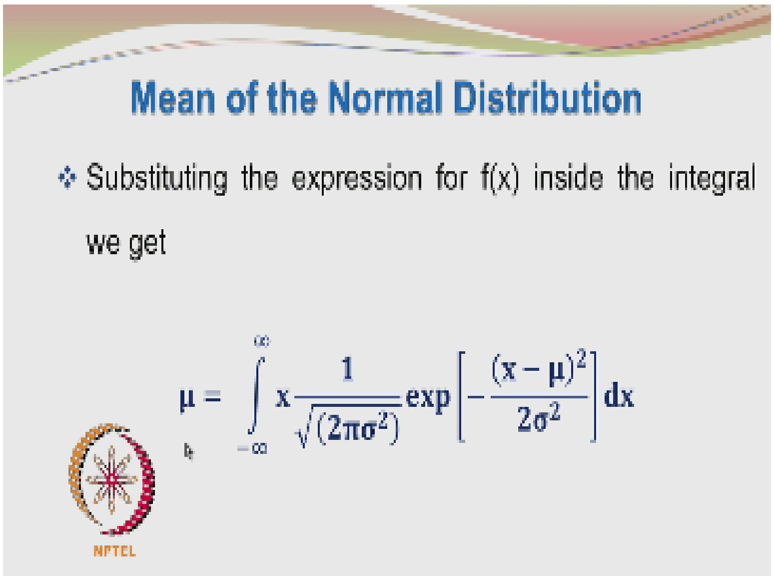
In the log normal distribution, you have seen that the parameters were quite different from the mean and variance of the log normal distribution. So, let us take the normal distribution, the probability density function is given by f of x is = 1/ root 2pi sigma square exponential -x - Mu whole square/ 2 sigma square. The two parameters are Mu and sigma, the range of this distribution is from - infinity to + infinity. The question is mean of the normal distribution is indeed Mu.

(Refer Slide Time: 02:04)



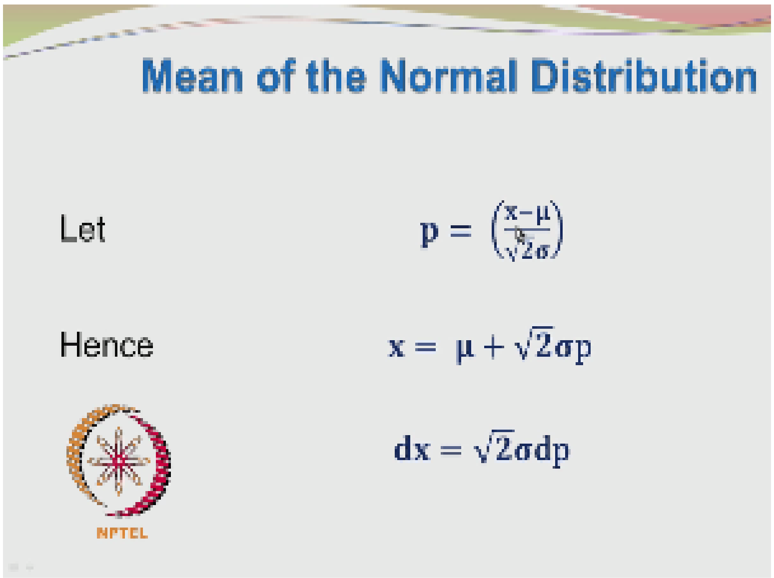
We will go to the definition of the mean for continuous probability density functions. This - infinity to + infinity x f of x dx, we plug in f of x here and then do the integration and see whether the answer on the hand side matches with the answer on the left hand side.

(Refer Slide Time: 02:27)



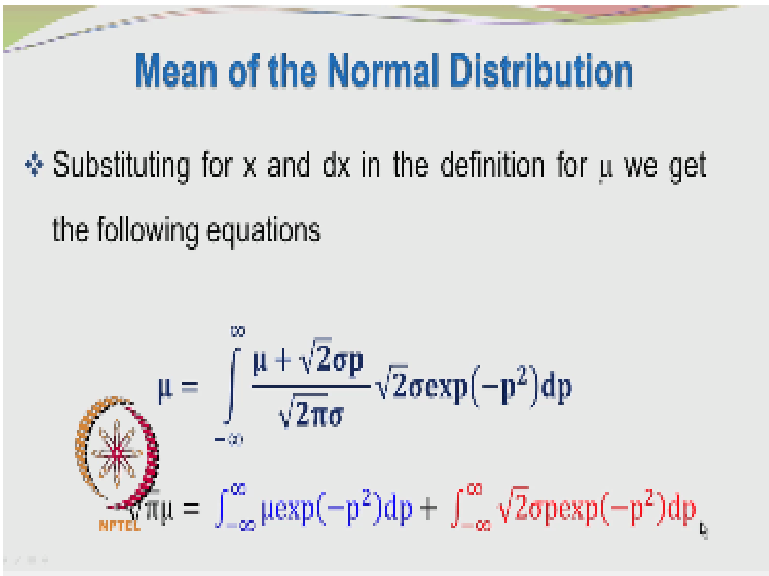
So, that is what we do here; we plug in 1/ root 2 pi sigma square exponential - x - Mu whole square/2 sigma squared here and this is the expression here. So, how do we go about solving this integral? I hope all of you have exposure to integral calculus, how to do integration and how to handle the limits, how to do integration by parts. What we do here is; we make use of the substitution, so that the integration gets simplified.

(Refer Slide Time: 03:14)



We take p is = x – Mu/ root 2 sigma, so when you simplify this, it becomes x is = Mu + root 2 sigma p. So, I differentiate x with respect to p, I get dx/dp equals root 2 sigma, elementary calculus. The important thing is to identify the correct form of the substitution for x in terms of p, so that the integral will get considerably simplified. We said p is = x – Mu/ root 2 sigma, so when you square it this expression inside the argument of the exponential term becomes - p squared.

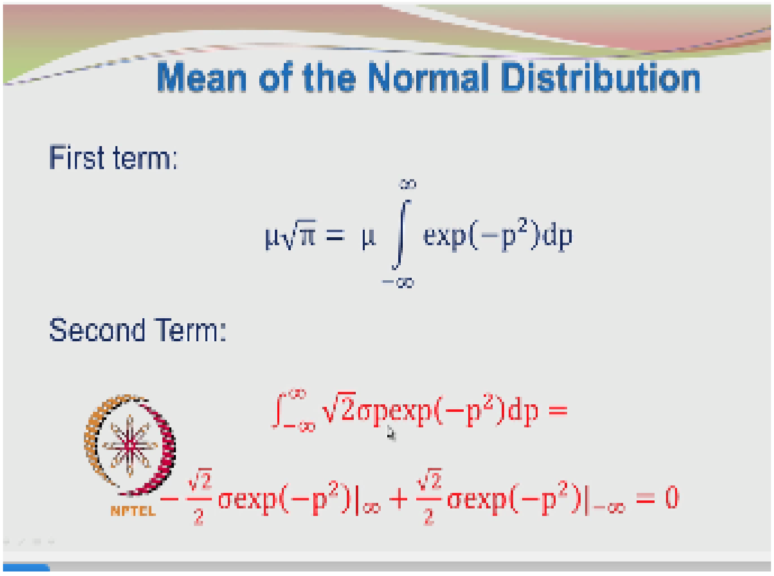
(Refer Slide Time: 04:07)



And x was replaced by Mu + root 2 sigma p, so you are having x; replacement for x here and this root 2 pi sigma is there in the original functional form and we substitute dx in terms of root 2 sigma dp. So, this can be integrated, I am just splitting the integral into 2 parts. So, you have the root 2 sigma cancelling out with this root 2 sigma, I am taking this pi; root of pi to the other side.

That is why you have root pi * Mu - infinity to + infinity Mu * exponential - p square dp + - infinity to + infinity root 2 sigma p exponential - p square dp. Please work out these steps on a paper, so that you can make sure that you are doing the manipulations correctly.

(Refer Slide Time: 05:22)

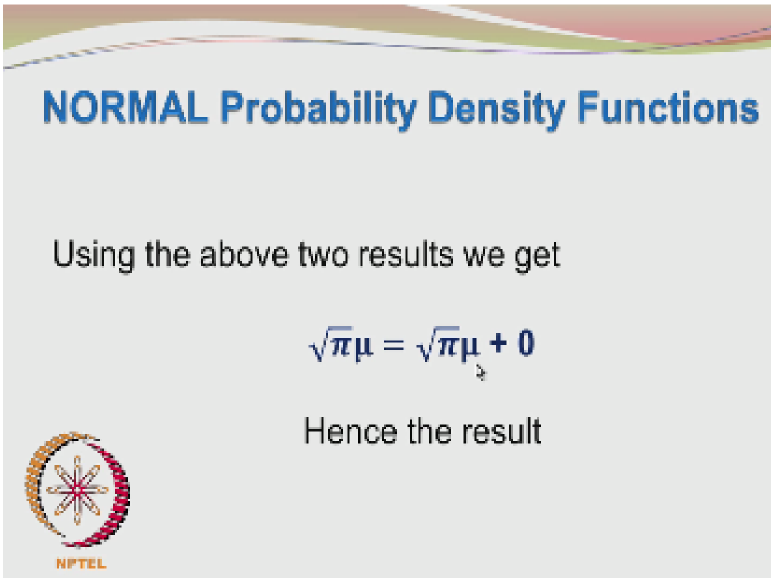


So, the first term is Mu * - infinity to + infinity e power - p square dp, this - infinity to + infinity e power - p square dp is a standard result, you will get root pi, so you have Mu root pi here and that is the same as the left hand side, which means that the second term here; since the first term here became root pi mu, the second term here should become 0 that can be shown rather easily and so you have here root 2 sigma p exponential - p square dp.

And this is in a very convenient form already, you have a e power - p squared and you have -1/2 of 2p here okay and * -1, the reason why I am saying -1/2 of minus 2p is when you differentiate this particular term - p squared, you get - 2p and so the integration becomes very simple and you will eventually get -1/2 * root 2 sigma e power - p squared + root 2/ 2 sigma e power - p square, you apply the upper limit here, you apply the lower limit here.

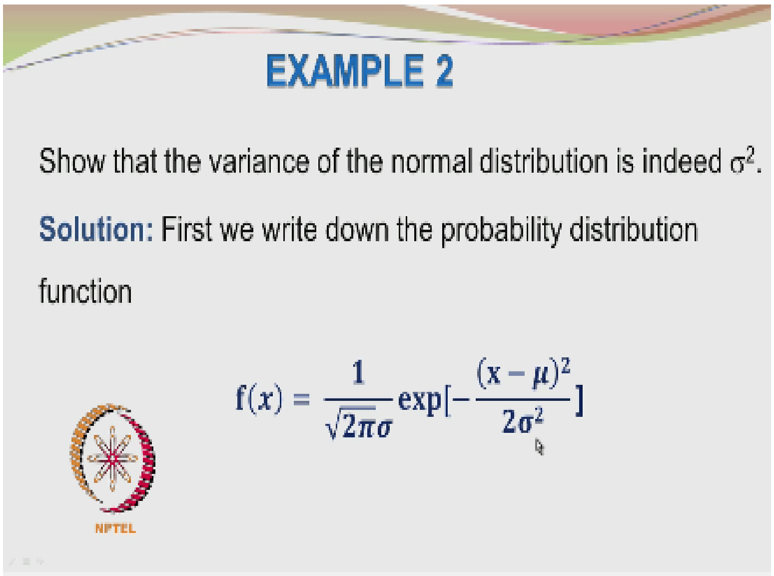
And you can see that the term vanishing, hence we have proved that the mean of the normal distribution is indeed Mu.

(Refer Slide Time: 07:10)



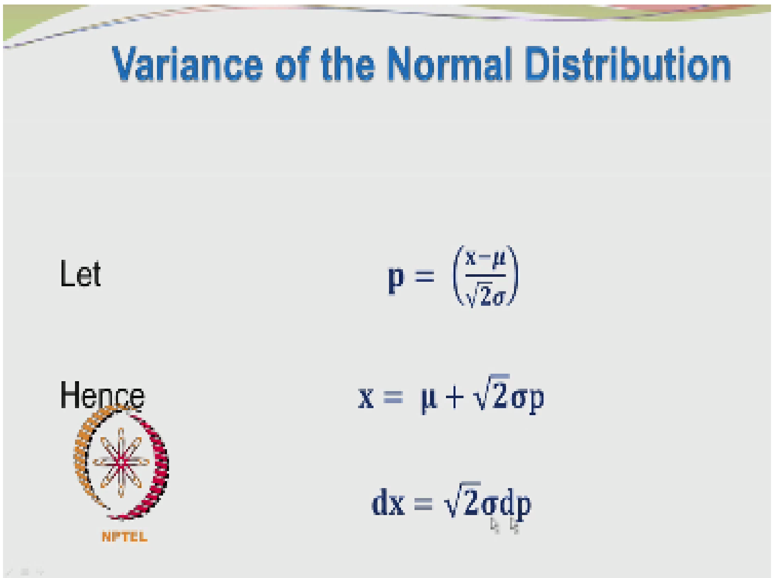
That is what is given in the slide. The second term on the right hand side vanished and you have root pi Mu is = root pi Mu.

(Refer Slide Time: 07:20)



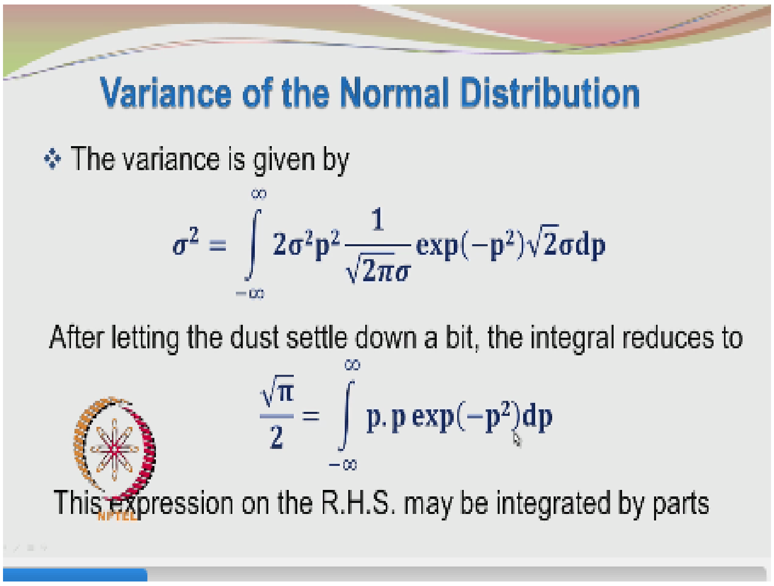
Show that the variance of the normal distribution is indeed sigma squared again, we write down the form f of x, which by now you should know almost by heart 1/2 root pi sigma e power - x - Mu whole square/ 2 sigma square.

(Refer Slide Time: 07:41)



And again, we do the same transformations p is = x – Mu/ root 2 sigma, x becomes Mu + root 2 sigma p dx is = root 2 sigma dp, very straightforward.

(Refer Slide Time: 07:54)

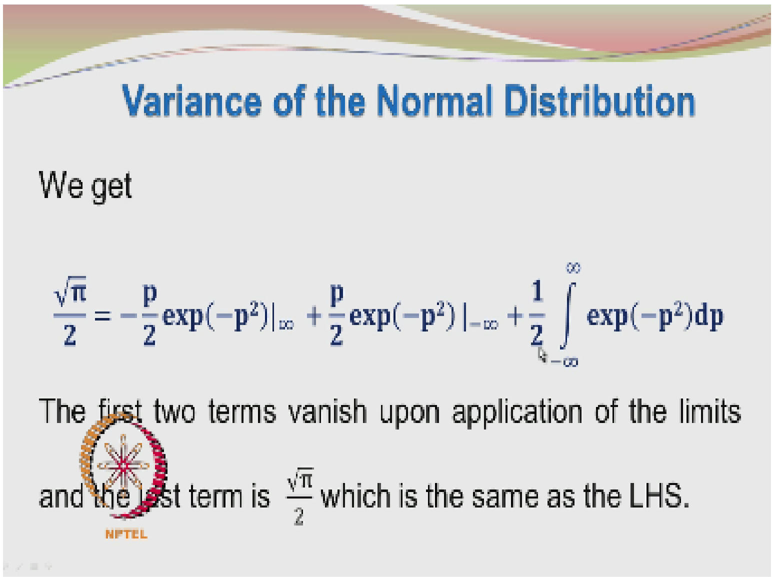


You know the variance is given by integral of x - Mu whole square * f of x dx, so when that happens and you substitute for x – Mu; x - Mu whole squared will become p squared * 2 Sigma squared and that is what is being written here, you have 2 sigma square p squared 1/ root 2 pi sigma exponential - p squared * root 2 sigma dp. Some steps have been omitted here, I would like you to do these steps and see whether you get this particular form of the integral all right.

Now, this integral is in fact, very easy to evaluate, I am taking root pi/2 to the left hand side and this sigma square will cancel out, so essentially we have to show that root pi /2 is = - infinity to + infinity p squared e power - p square dp. You may ask that since the x has been transformed into p, how come the limits have not changed. So, when x goes to + infinity, p also goes to + infinity and the next goes to – infinity, p also goes to – infinity.

So, the limits do not change even after the transformation from x to p, so you have root pi/ 2 integral - infinity to + infinity, I am writing p square as p * p e power - p square dp. The integration is now possible; we had already discussed how to integrate p e power - p square in the previous example.

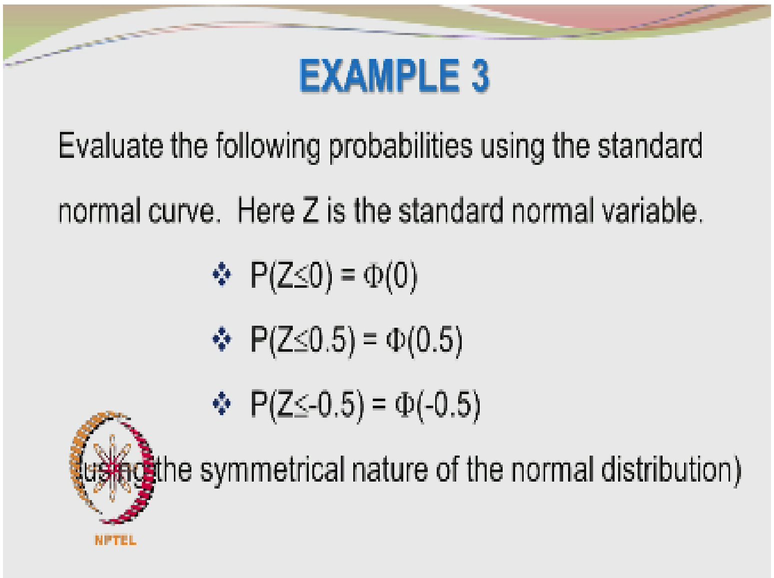
(Refer Slide Time: 09:58)



We get root pi/2 equals – p/2 e power - p square at infinity +p/2 e power - p squared at - infinity and we know that both these terms will vanish on the application of the limit and you are only left with 1/2 * - infinity to + infinity e power - p square dp. This result came after applying the integration by parts method and when you do this, you also by now know that - infinity to + infinity e power - p square dp is = root pi.

So, you have root pi/2 equals root pi/2 and hence equating the integral x - Mu whole square f of x dx, where f of x is the normal distribution, we can equate that integral; the second moment about the mean to the variance sigma square and the result is proved because we get the same answer on both sides after the manipulations have been carried out.

(Refer Slide Time: 11:15)



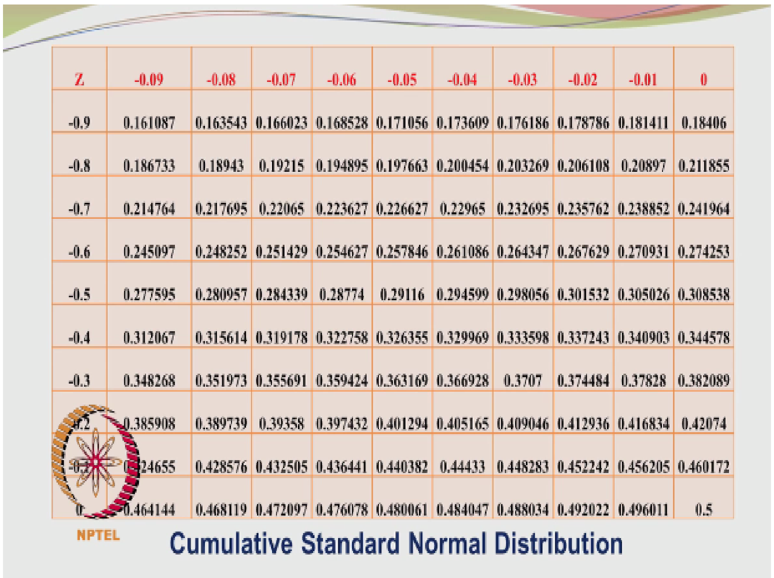
Now, we are going to look at the use of the normal probability plot. The normal probability plot runs to 2 pages as far as the normal distribution is concerned. The first page involves negative values of z and that corresponds to the left portion of the standard normal curve, you know that the standard normal curve has mean 0 and variance 1. So, when you have the standard normal curve, you have the negative part and you have the positive part.

Since the distribution is symmetrical, you need only one chart you do not need both the charts. If you have only the negative portion of the chart, you can even use it to find the probabilities for the positive portion of the curve. Anyway, I have provided data corresponding to both the negative as well as positive values of z, we will be referring to those charts to pick up the probabilities for different values of z.

Reading the chart is pretty easy, sometimes you may not get exactly the number you want. However, you can do interpolation between 2 values, if your z value falls between 2 tabulated values. Sometimes, you may also have to do the reverse, you are given the probability and then you are asked to find; what is the value of z that is going to give the desired probability, so this is the inverse problem.

We will first do a few examples, what is the probability of z, the standard normal random variable <= 0. This corresponds to the cumulative distribution function of 0, you do not even need a chart because the mean of the distribution is 0 and the curve is symmetrical, the area under the curve below 0 will be = the area under the curve beyond the 0 or above 0 and so the probability will be 0.5. To verify this, let us go to the actual probability chart.

(Refer Slide Time: 13:43)



Z	-0.09	-0.08	-0.07	-0.06	-0.05	-0.04	-0.03	-0.02	-0.01	0
-0.9	0.161087	0.163543	0.166023	0.168528	0.171056	0.173609	0.176186	0.178786	0.181411	0.18406
-0.8	0.186733	0.18943	0.19215	0.194895	0.197663	0.200454	0.203269	0.206108	0.20897	0.211855
-0.7	0.214764	0.217695	0.22065	0.223627	0.226627	0.22965	0.232695	0.235762	0.238852	0.241964
-0.6	0.245097	0.248252	0.251429	0.254627	0.257846	0.261086	0.264347	0.267629	0.270931	0.274253
-0.5	0.277595	0.280957	0.284339	0.28774	0.29116	0.294599	0.298056	0.301532	0.305026	0.308538
-0.4	0.312067	0.315614	0.319178	0.322758	0.326355	0.329969	0.333598	0.337243	0.340903	0.344578
-0.3	0.348268	0.351973	0.355691	0.359424	0.363169	0.366928	0.3707	0.374484	0.37828	0.382089
-0.2	0.385908	0.389739	0.39358	0.397432	0.401294	0.405165	0.409046	0.412936	0.416834	0.42074
-0.1	0.424655	0.428576	0.432505	0.436441	0.440382	0.44433	0.448283	0.452242	0.456205	0.460172
0	0.464144	0.468119	0.472097	0.476078	0.480061	0.484047	0.488034	0.492022	0.496011	0.5

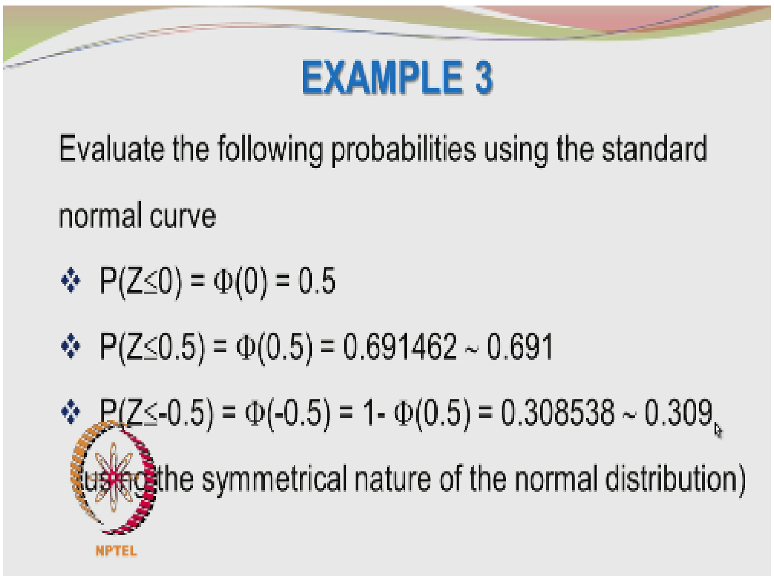
NPTEL

Cumulative Standard Normal Distribution

So, here we are the z value is 0 and we know that the distribution is expressed in such a way. The cumulative distribution function is expressed in such a way that it refers to the area under the curve below the listed z value, so you go to z corresponding to 0, this is -0.9 and so it comes to 0 here and you can see the probability to be 0.5. Now, let us look at the second part, what is the probability of z <= 0.5 okay.

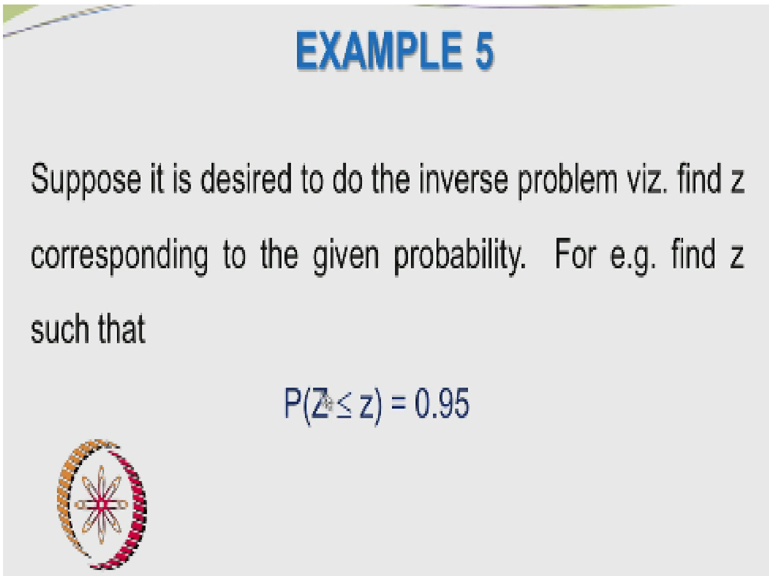
We have to find the cumulative distribution value of 0.5 and how do we do that? So, the z value is 0.5 and we go to 0.50 and this is 0.51, 0.52 up to 0.59 for the normal distribution at z equals 0.5, the area under the curve or the probability is 0.6915, so this is the value here.

(Refer Slide Time: 15:12)



And you can see here that is the answer we have also reported. Similarly, you can show that probability of z <= -0.5 will be 0.309 approximately and that happens to be 1 – 0.691. So, you can use the symmetry property of the normal distribution and find probability of z <= -0.5, the probability of z <= -0.5 would have been the same as the probability of z > or = 0.5, we just now found the probability of z <= 0.5 and that came to around 0.691, hence what is the probability that z > or = 0.5 will be? That will be 1 – 0.691 and we get 0.309.

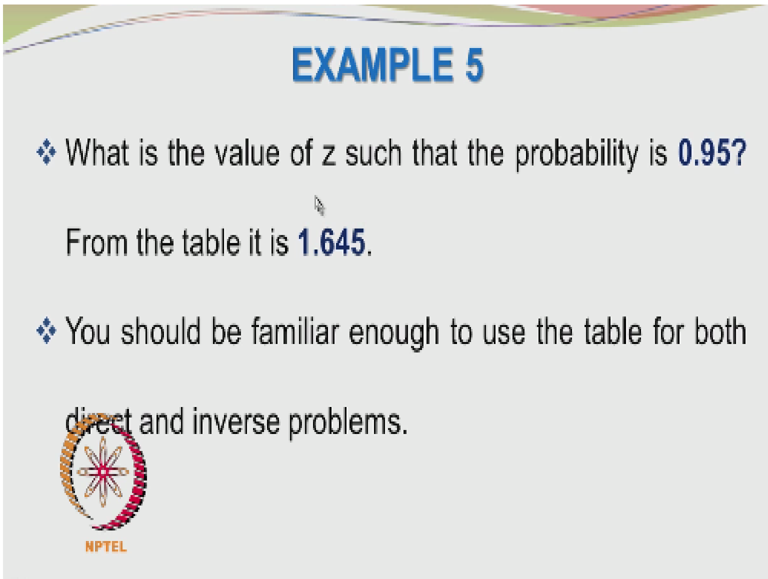
(Refer Slide Time: 16:26)



We have to do the inverse problem now. Suppose, it is desired to do the inverse problem to find z corresponding to the given probability, for example find z such that probability of z <= small z, this is a random variable and this is the value is = 0.95. We have to see what is the value attained by the random variable z such that the probability becomes 0.95. So, we go to the appropriate table and we search for the probability of 0.95 in this table.

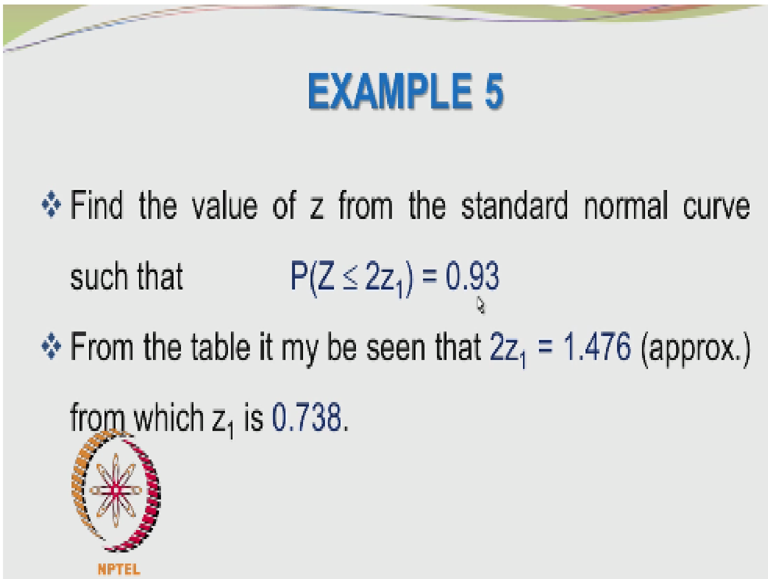
And we see that the probability of 0.95 falls between 1.64 and 1.65, you come horizontally along this direction here, you have 1.6, when you come to 1.64, the probability is 0.94, 95 and when you come to 1.65 the probability is 0.9505. So, the appropriate value of z would be somewhere between 1.64 and 1.65, so we can take it as 1.645, would be the close enough answer.

(Refer Slide Time: 17:55)



That is the answer, which is given in this particular slide. What is the value of z such that the probability is 0.95 from the table; you find the value of z to be 1.645, so you should be able to use the table for both direct calculation of the probability and the inverse calculation of z for a given probability.

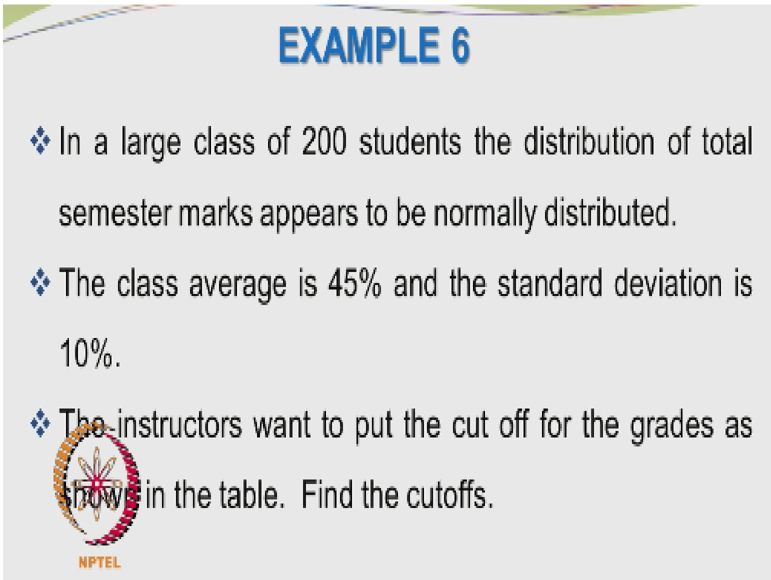
(Refer Slide Time: 18:21)



Now, let us go to the next example; example number 5, here we have the question what is the probability of the random variable <= to 2z1? So, that the probability is 0.93, we have to identify z1 for this situation, first we will identify what is 2z1, so from the normal probability chart, we can find that if 2z1 is 1.476, the probability of z <= 1.476 is 0.93. Please look at the standard probability chart and verify whether this is indeed so.

And then you can find out what is the value of z1 and that is coming to 0.738. Sometimes, what may happen is you may think that for whatever reason, probability of z <= 2z1 may be taken as 1 * probability of z <= z1 okay, then you may find the value of z1 but that value would be completely wrong, you cannot take the constant to or any constant here outside the; outside, if you had taken it like that; if you are equated probability of z <= 2z1, as 2 times probability of z <= z1, the z1 value would have been -0.088.

(Refer Slide Time: 20:30)



And this is completely different from the correct answer of 0.738. In my earlier lecture on normal distributions, I was telling that the distribution of student's marks is assumed to be normal or Gaussian and you can make the appropriate calculations for cut off. For example, mu + 2 sigma and above may be taken to be S grade and so on. We will take a simple example and see how the cut offs are fixed in this particular case.

So, the problem statement is; you are having a large class of 200 students and the distribution of total semester marks is assumed or it appears to be normally distributed. The class average is 45% and the standard deviation is 10%, the instructors want to decide or identify the cut off grades according to the following table. So, our aim is to find the cut offs.

(Refer Slide Time: 21:37)

EXAMPLE 6

Sl. No.	Grade	% of students	Cumulative %
1	S	5	5
2	A	15	20
3	B	30	50
4	C	25	75
5	D	15	90
6	E	5	95

So, you have the S grade, A grade, B grade so on to E grade. The instructors want to put the cut offs such that 5% of the class gets S grade, 15% of the class gets A grade, 30% of the class gets B grade, you expect the majority of the students in the class to get a grade somewhere in the vicinity of B and C okay. So, the majority of the class will have grades B and C in this case, 30% of the students should be having B grade.

And 25% of the students in the class should be having C grade and the D grade is 15% of the class and E graders is 5%. Here, you also have the column of cumulative percentages, what I do here is simply add the total percentage of students here, we have 5, 20, 50, 75, 90, 95 and if you include the students, who are getting U grade, then it comes to another 5% and that comes to 100. We are not going to look at the U grade okay, so we will be looking at the other grades.

(Refer Slide Time: 23:10)

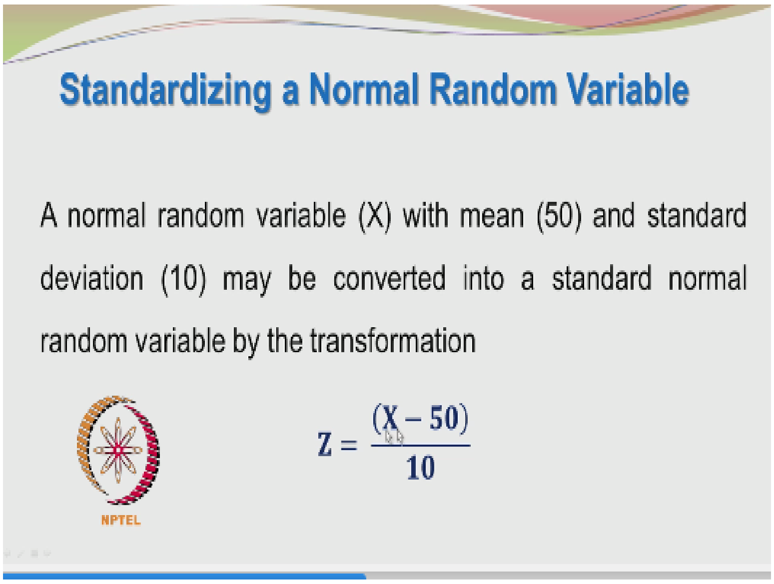
EXAMPLE 6

Solution:

We take the marks distribution to be normal with mean 50 and standard deviation 10. To use the normal probability tables in the standard form we need to convert the given distribution into the standard form. This is shown in the next slide.

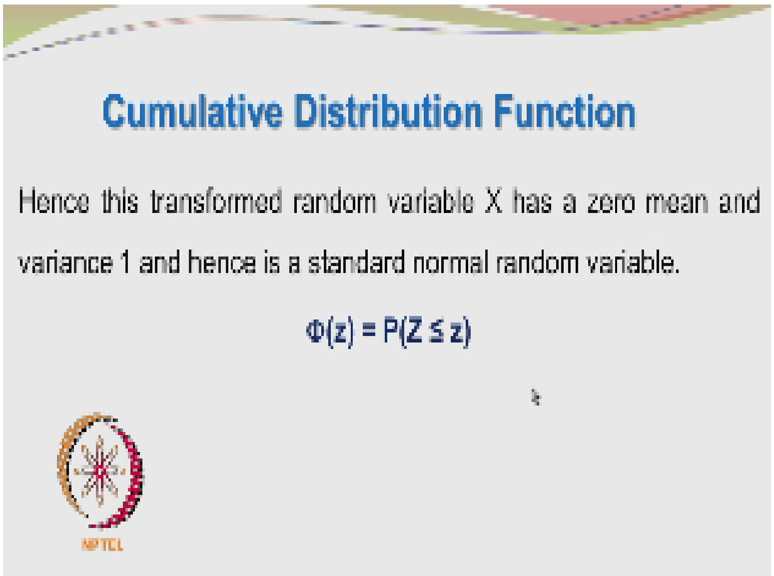
So, the mark distribution is assumed to be normal with mean of 50 and standard deviation of 10, okay. So, we have to first convert them into the standard form, so that the probabilities can be obtained.

(Refer Slide Time: 23:29)



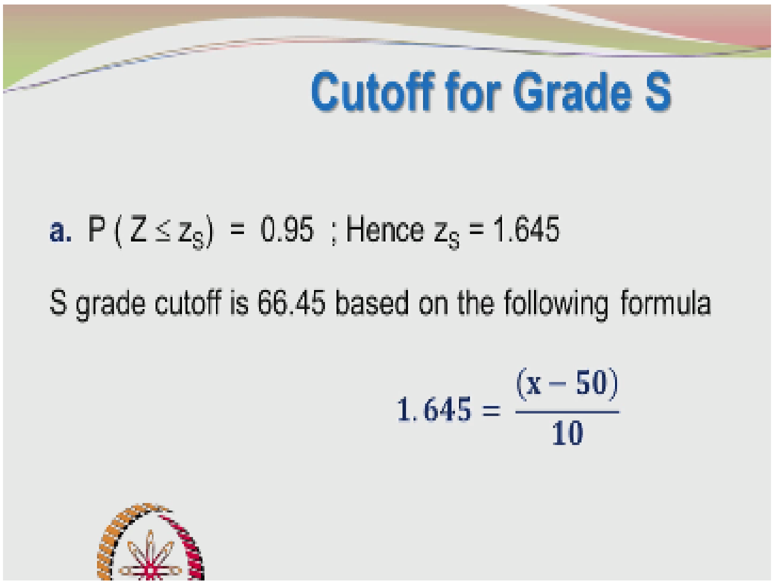
We have to convert the random variable X into Z, the random variable X is having its own mean Mu and the variance Sigma square, so that normal distribution had a mean of 50 and standard deviation of 10 or variance of 100. If you want to convert it into the standard normal, then it should be converted into a normal distribution of mean 0 and variance 1. For doing that, we take Z = X - Mu / Sigma, Mu is 50, Sigma is 10, so we convert X into Z by applying the transformation Z is = X – 50/10.

(Refer Slide Time: 24:23)



Now, we have to find the cumulative distribution functions for different values of Z.

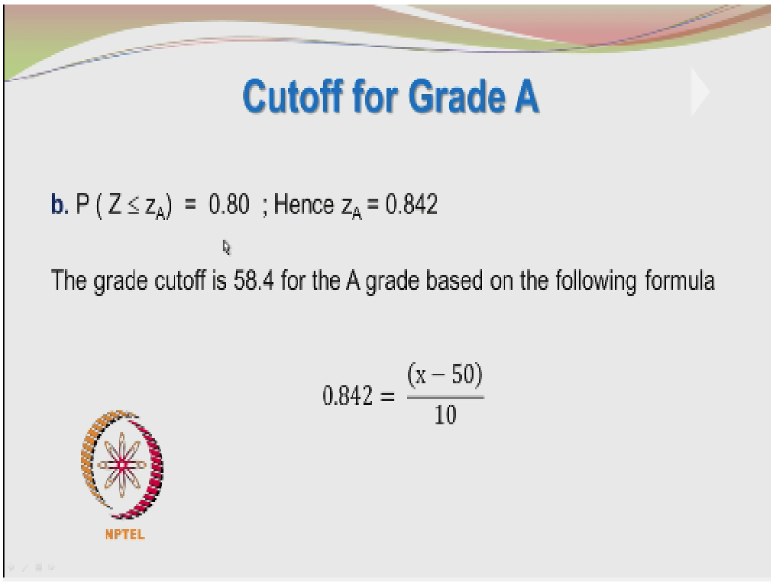
(Refer Slide Time: 24:29)



Let us look at the cut off for grade S. So, the grade S is defined such that the probability of the random variable Z, the standard normal random variable capital Z is <= z subscript s, this is the Z value corresponding to the S grade and that is = 0.95. We have seen from one of the earlier examples that the Zs will correspond to 1.645 and you substitute 1.645 here and that is = x – 50/ 10.

And when you do the calculations, 16.45 + 50 is 66.45, so the S grade cut off is 66.45. Any student getting value higher than 66.45 or keep it at 66.5 will be awarded the S grade.

(Refer Slide Time: 25:36)



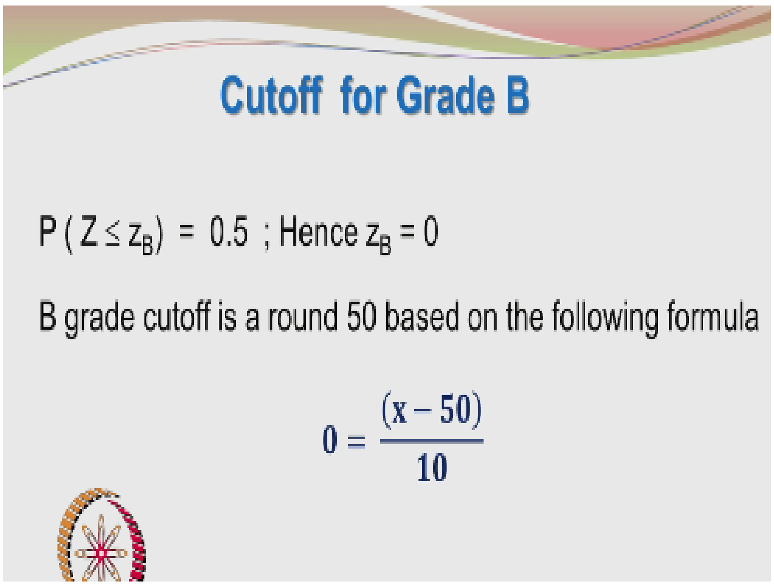
Now, you want to see what is the probability of Z <= zA, you have to identify zA such that the probability of Z <= zA equals 0.8. How did you get 0.8, we want to put the cut offs in such a manner that the percentage of students having the cut off mark or above should constitute the 20% of the class. The A cut off should be so fixed that 20% of the class should have A grade or S great okay, so where is the cut off for A.

Even though the cut off for A is 15%, we have to add but the percentage of students, we have to add the 15% to the percentage of student having the S grade also. What would be a nice idea here is to draw the normal distribution and on the normal distributions right hand side put some numbers as cut off marks and you can put the S grade and A grade there, okay. So, you have to add the 5 to 15.

So, that the area beyond the cut off for A will be 0.2, so to find that; since the probability chart gives the area below the z, the area below z will be 0.8. If the area above zA is 0.2, so the cut off is decided based on the probability below the za being 0.8 and in such a case, we can see that zA is 0.842, I show this. So, when you come to this particular table, you locate z value of 0.84 but the probability is only 0.788.

So, you have to increase the value of z some more and so you see here corresponding to 0.84, the probability is 0.7996 and corresponding to probability of 0.802, you have the z value of 0.85. So, the required z value would be somewhere between 0.84 and 0.85 for all practical purposes, you may take the z value to be 0.84, I have used a software, so I have a more accurate value here and I plug in 0.84 or 0.842 here, x – Mu/ sigma and then this becomes 8.42 and so 58.42 would be the cut off for A. So, the cut off for A is 58.4.

(Refer Slide Time: 29:24)

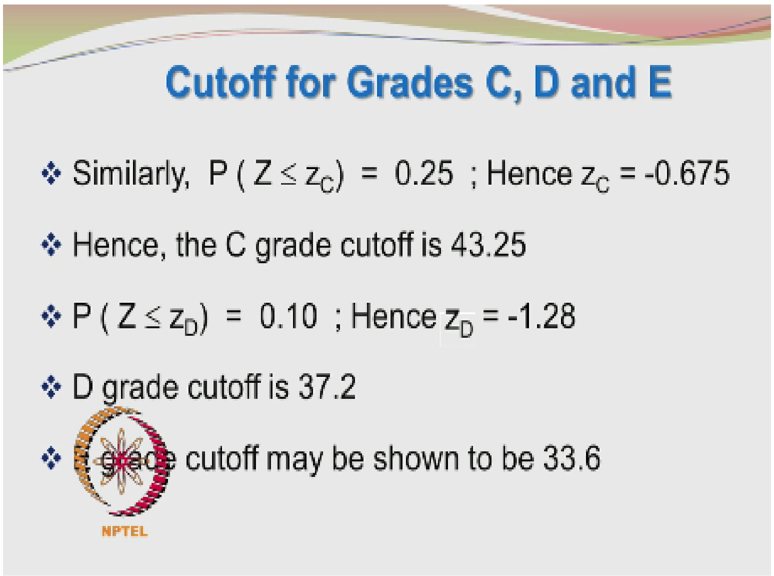


Now we have an interesting case, the cut off for grade B should be placed such that 50% of the students should have a grade B or above. And 50% of the students should have the grade lower than B, you can see that even though, the cut off for B should be such that it encompasses 30% of the class, we have to look at the cumulative distribution and see that 50% of the class will have either grade B or higher.

But when there are students, who are crossing a certain cut off like 58.4 in the present case, they will automatically get the A grade that would constitute about 15% of the class. And if they cross a certain value corresponding to the cut off for S, they will automatically get the S grade and that will encompass about 5% of the class. So, you can do the calculations and you will find that since the cumulative percentage is 50.

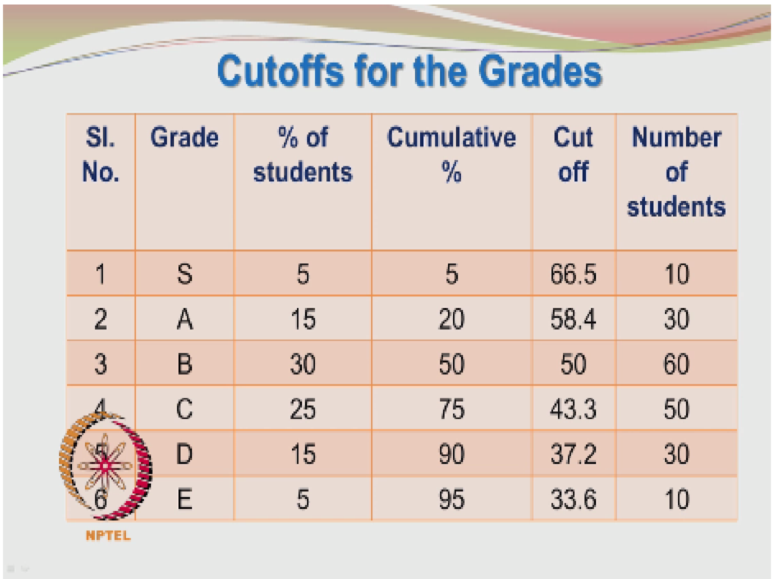
Or the cumulative probability is 0.5 that will definitely correspond to z value of 0 and the standard normal random variable taking a value of z equals 0, means the cut off for the B grade is 50. Since zB is = 0, we put 0 here and so the cut off for the B grade is around 50.

(Refer Slide Time: 31:29)



Similarly, you can find out the cut offs for the C grade and D grade and E grade, the same exercise is followed; the zC values have been identified using the same procedure as I outlined earlier. Now, you can find out the cut off for grade C, cut off for grade D and the cut off for grade E.

(Refer Slide Time: 31:57)



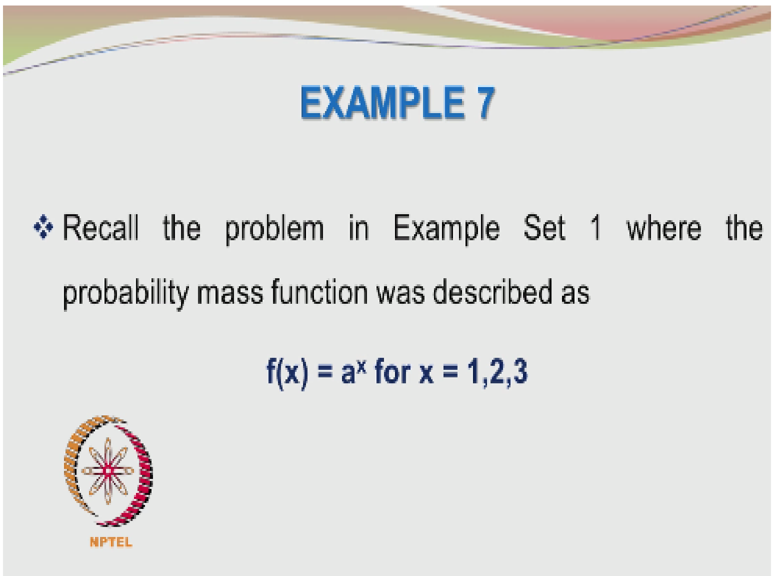
Cutoffs for the Grades

Sl. No.	Grade	% of students	Cumulative %	Cut off	Number of students
1	S	5	5	66.5	10
2	A	15	20	58.4	30
3	B	30	50	50	60
4	C	25	75	43.3	50
5	D	15	90	37.2	30
6	E	5	95	33.6	10

You can summarize the results, you can see that the students who are getting 66.5 or above get a S grade and that encompasses about 5% of the class, so in a class of 200 students, 10 students get S grade. So, the number of students who get the A grade is 30 and the cut off for the a grade A is 58.4, so any student getting grades or marks rather between 58.4 to 66.5 will be awarded A grade and 30 such students meet this criteria.

The cut off for the grade B as we discussed earlier is 50, for grade C is 43.3, grade D is 37.2, grade E is 33.6, okay and any student getting marks below 33.6, would be automatically assigned a value of U, he may have to write the supplementary or the repetition of the course. If you count the total number of students, you can see that the number comes to 10 + 30, 40, 100, 150, 180, 190, so 10 students have failed the course unfortunately.

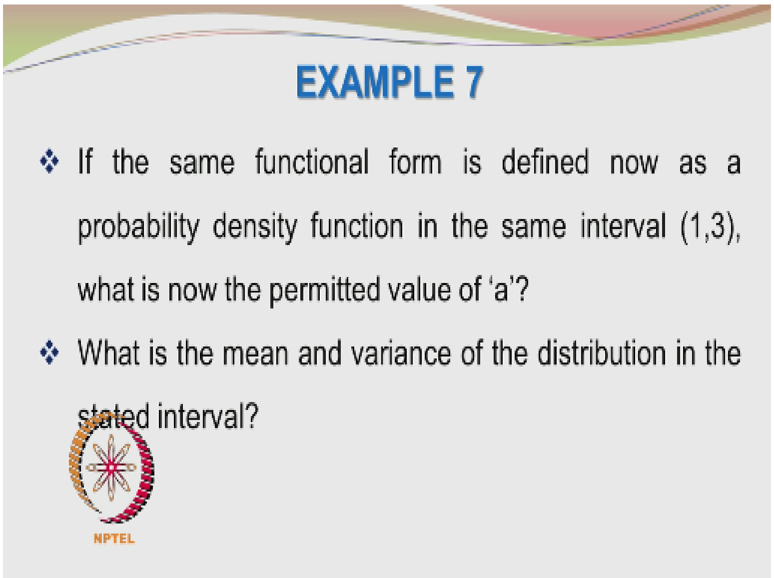
(Refer Slide Time: 33:22)



Let us now go to another problem so far, we have been discussing about the normal distribution, now we have a case where the distribution is described by a funny kind of expression f of x is = a power x, x can take values 1, 2 and 3. We had discussed this problem earlier in the case of discrete probability distribution, x was taking values 1, 2 and 3 but now, when you discuss the same problem for the continuous distribution case, then x will be ranging from 1 to 3.

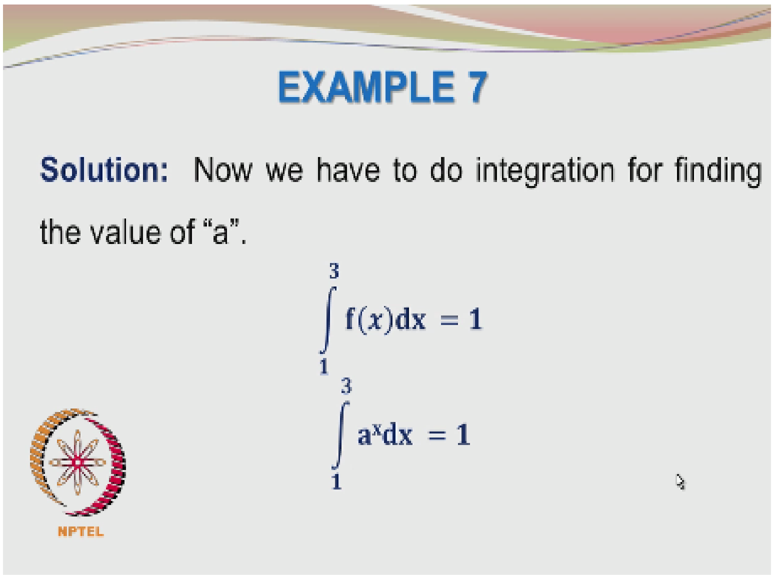
Or x can take any value between 1 to 3, here, x can take values 1 or 2 or 3, if the probability distribution function is discrete, we are talking about a random variable, which is discrete in nature in that case, x can take values only 1, 2 and 3, if the random variable x; capital X is continuous, then the range for X will be from 1, 2, 3 from 1 up to a value of 3, okay.

(Refer Slide Time: 35:00)



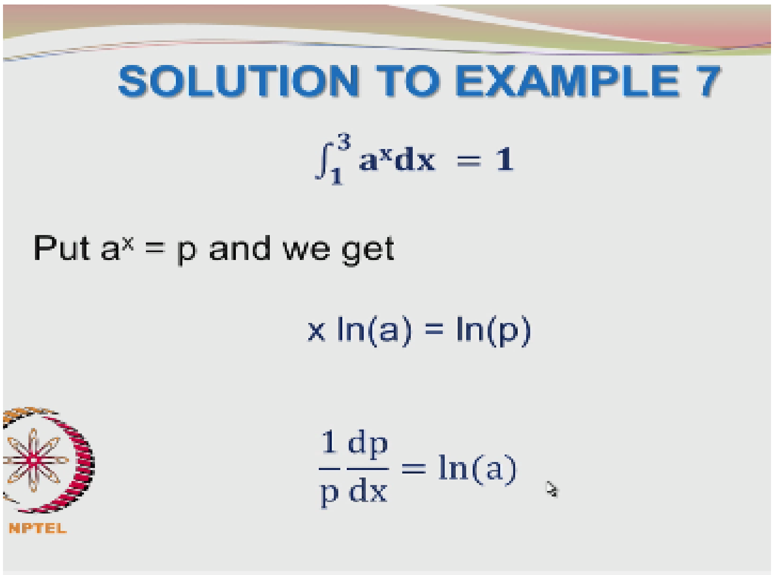
So, the interval is now 1, 3, what is now the permitted value of a, and what is the mean and variance of the distribution in this stated interval?

(Refer Slide Time: 35:14)



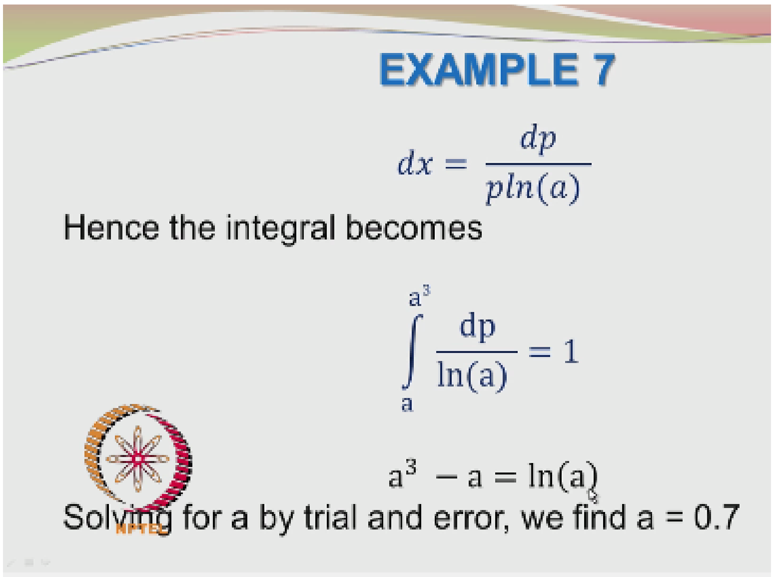
To find the value of a, we use the criterion that between the lower limit to the upper limit, the area under the curve should be = 1, so using this criterion, we integrate a power x between the permitted limits 1 and 3 and see for what value of a, we will get this integral to be unity.

(Refer Slide Time: 35:47)



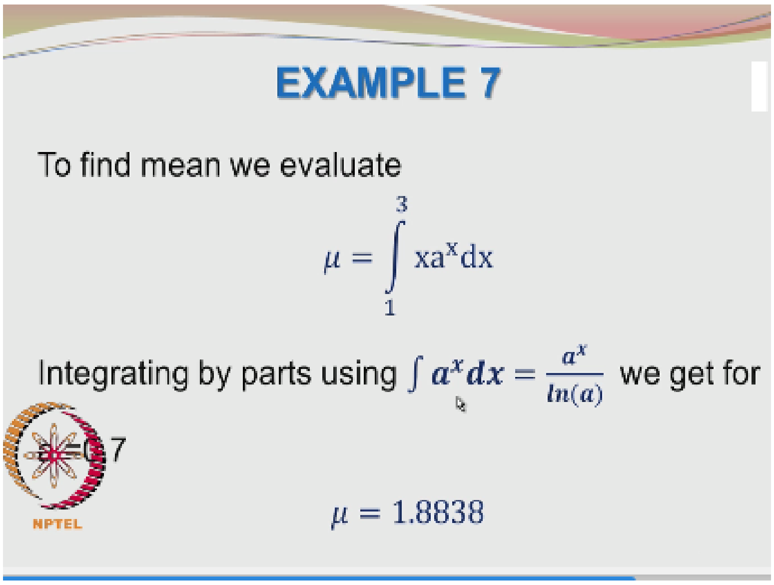
So, we carry out the integration, the integration is quite interesting, you put a power x is = p, then take natural log on both sides x log a is = log p and then you differentiate p with respect to x and you will get by 1/p dp/dx is = log a.

(Refer Slide Time: 36:07)



Hence, dx becomes dp/p log a and the integral attains eventually this following form, the lower limit becomes a and the upper limit becomes a cube dp/log a is = 1 and that is nothing but p / log a, so you get this expression a cube - a is = log a. By trial and error, which can be quite easily done with the help of a spread sheet, you can find that a takes the value of 0.7.

(Refer Slide Time: 36:51)

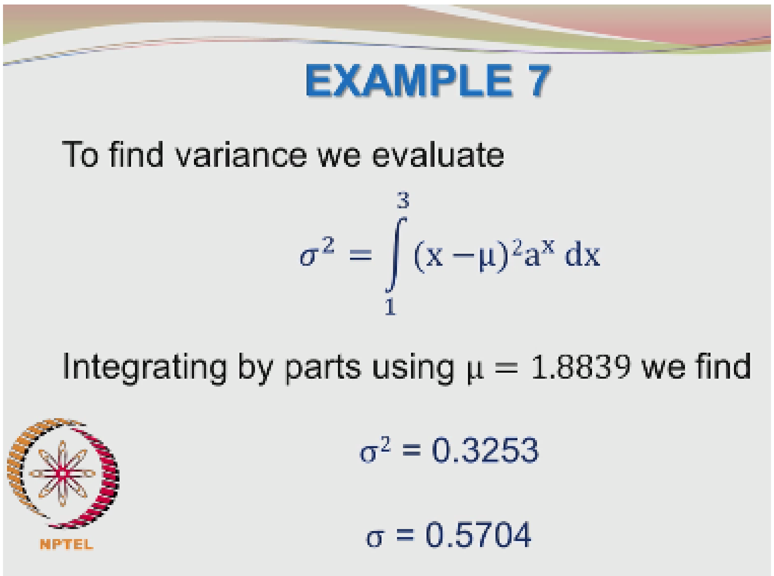


To find the mean value, we evaluate Mu is = 1, 2, 3 xa power x dx and this would require integration by parts, recall that Mu for a continuous probability distribution function is given by Mu is = lower limit upper limit integration x f of x dx, here the lower and upper limits are 1 and 3 respectively, we have xa power x dx, we can do integration by parts and integral of a power x dx is = a power x by log a.

We get the value of Mu to be 1.8838, this integration I am not showing all the steps, I am expecting you to do it on your own and see whether you get the answer. Sometimes, you will carry out the integration; sometimes the integration maybe lengthy and after doing it, you may want to check whether you have done it correctly. There are a few software, which can do the integration between limits and give you the numerical value.

One such software is the MATLAB okay, if you have access to MATLAB, you may want to use the integral option to see whether the answer is correct.

(Refer Slide Time: 38:31)

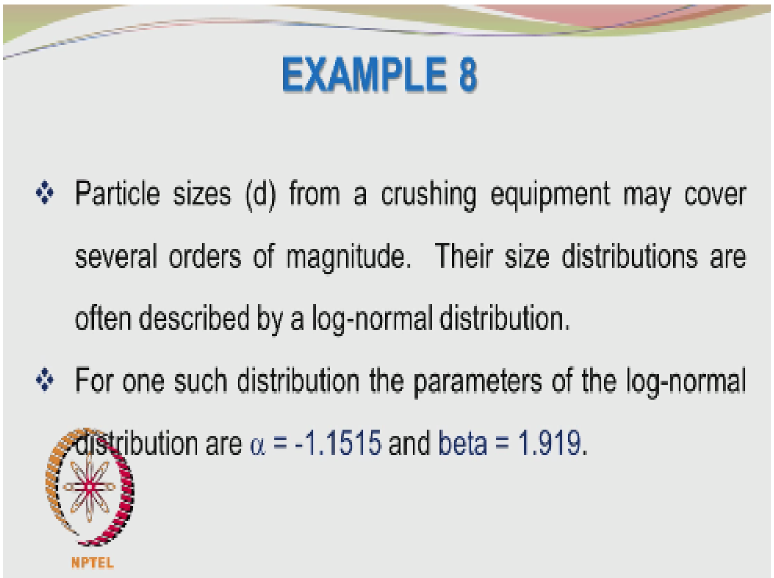


In the next case, we want to find the variance; variance by definition is x - Mu whole square f of x dx, so f of x is given by a power x. One important thing you have to remember is the value of x is varying between 1 to 3 but you have fixed the value of a to b 0.7 that was the first subdivision of the present exercise. What was the permitted value of a, so that the criterion for the probability distribution is satisfied.

The criterion for the probability distribution was integral of 1, 2, 3 a power x dx is = 1 and after carrying out the integrations and putting in the limits, you found that a, was 0.7. Now, you have to put the value of 0,7, wherever you see, a. What I have done is kept a, as it is, so that the integration is first done and the substitution is finally done. So, in this case, a actually should be 0.7 here but I left it as, a because to help me in my integration.

Finally, I will substitute the value of 0.7 into a, so once you do that, you can again do the integration by parts and find out that Sigma squared is 0.3253 or sigma is = 0.5704, I am taking the square root of this value, to get 0.5704.

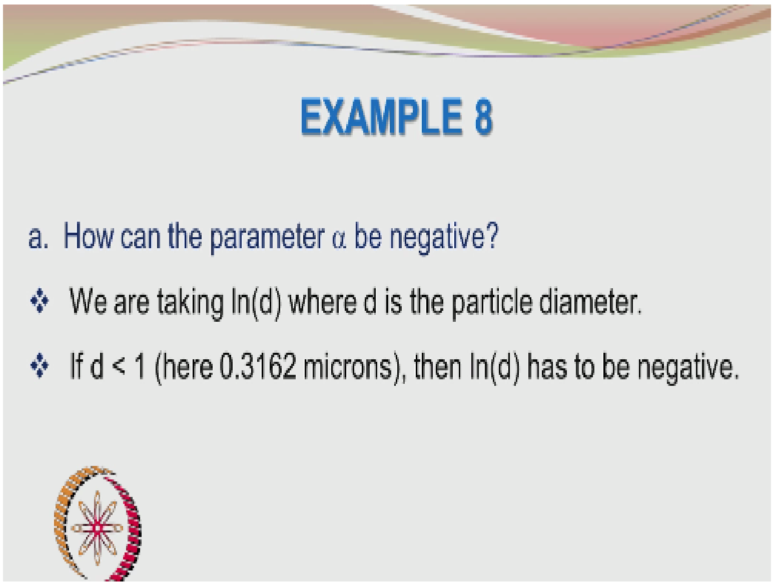
(Refer Slide Time: 40:41)



The next example is involving the log normal distribution. We were discussing the log normal distribution in one of the earlier lectures and we saw that when the random variable x was subject to a transformation okay and you took the natural logarithm of this random variable x, it became a new random variable and this new random variable started to behave or obey a normal distribution okay. So, this was a very useful result and what we can do is; do the transformation for our data.

And use the properties of the normal curve in the analysis of the data, so the problem statement is particle sizes of; represented by d from crushing equipment, may cover several orders of magnitude and their size distributions are often described by a log normal distribution. For one such distribution, the parameters of the log normal distribution are alpha is = - 1.1515 and beta is = 1.919.

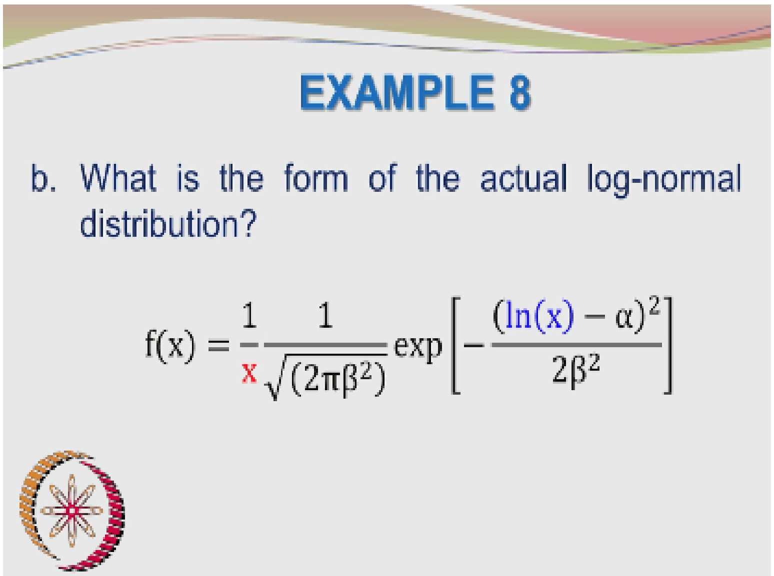
(Refer Slide Time: 42:02)



So, these are the values of the parameters. The log normal distribution parameters are -1.1515 and 1.919. So, the question is how can the parameter alpha be negative, is it making physical sense? When we are converting d to log d, okay, we are making the transformation from the original variable to the transformed variable in this particular case, we are taking the natural logarithm and so we have log d.

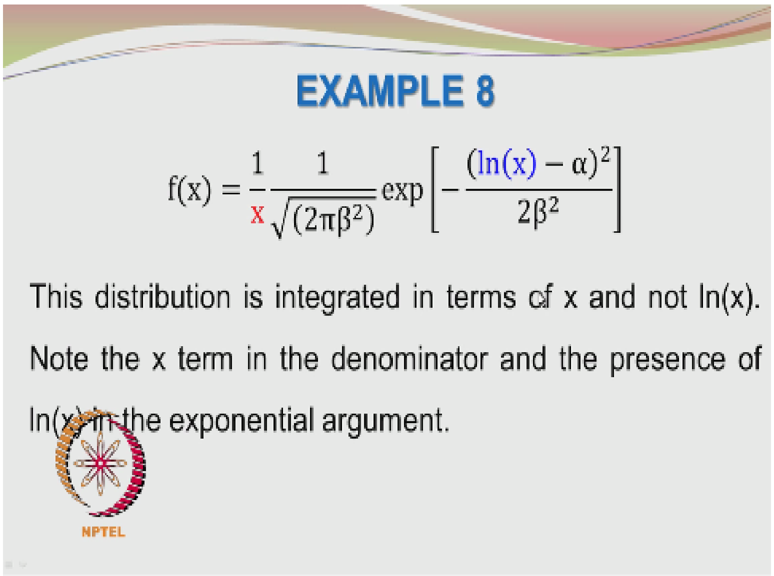
So, if the value of d is < 1, log d can become negative, so the d value here was 0.362, the d value here was 0.3162 microns, so log d became negative.

(Refer Slide Time: 42:59)



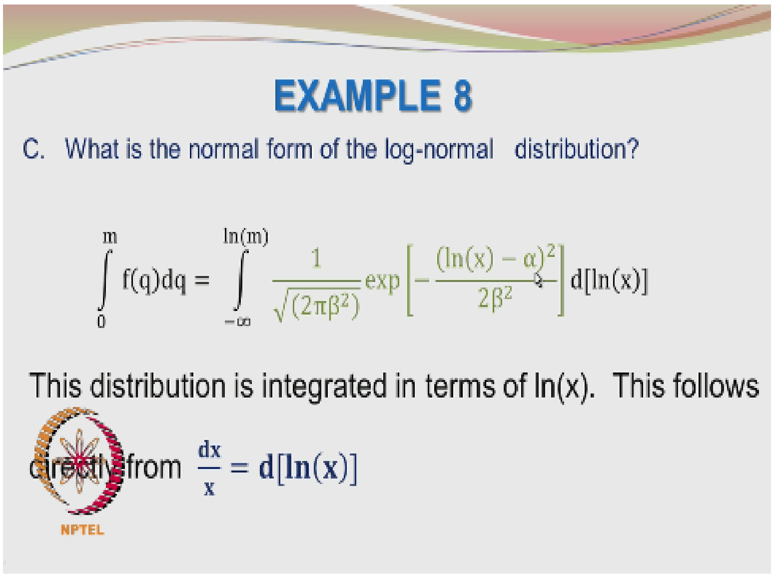
The next question is quite simple what is the form of the actual log normal distribution okay? The form of the actual log normal distribution was f of x is = 1/x * 1/ root 2 pi beta squared e power – log x - alpha whole square by 2 beta square. The difference from the normal distribution was you have log x here, instead of x and then you also have an additional 1/x term here.

(Refer Slide Time: 43:35)



So, whenever you want to find the cumulative distribution using this function, you have to be careful, you have to integrate the function in terms of Xx and not log x okay. When you want to find the probability of a < x < b, it should be integral of a to b f of x dx and so, this will be integral of a to b and then this particular function into dx, okay, when you are using this form directly.

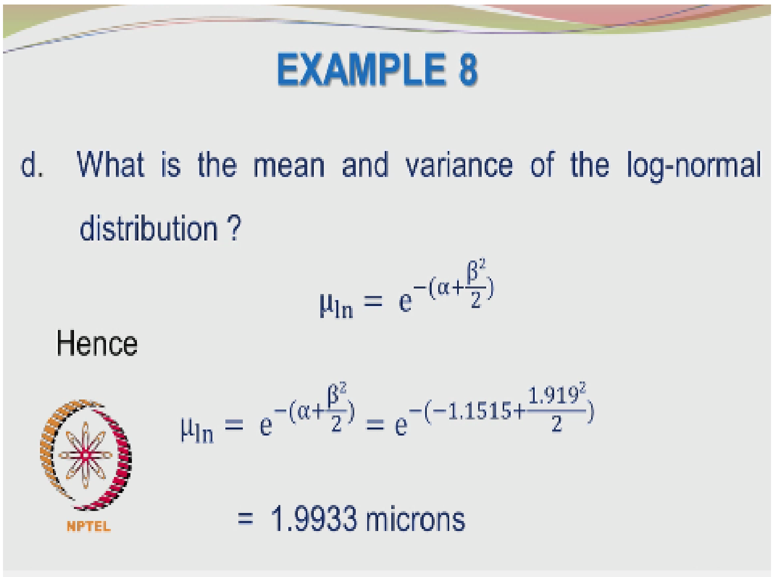
(Refer Slide Time: 44:20)



The question is what is the normal form of the log normal distribution okay, so you have changed the random variable x to log x, so you have to consider log x as random variable; a new random variable and think always in terms of log of x okay. So, to do that, you can put log x as some variable p and you can write dx/x as d of log x okay and so, dx/x is d of log x, so you are going to have dx/x replaced by d of log X as shown here.

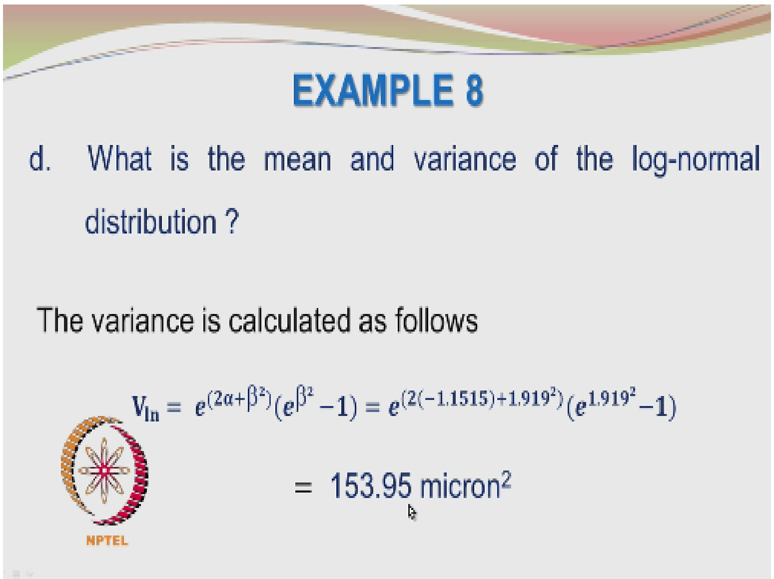
And then you are also having log x here and then 0 to m, in fact becomes - infinity to log of m, so you take log x as the random variable and if you start thinking in terms of log x, this is nothing but a normal distribution with mean alpha and standard deviation beta. I would suggest to you to carry out this derivation or this kind of transformation and integration on your own to convince yourself. And become familiar with the use of the log normal distribution, otherwise things can become slightly confusing.

(Refer Slide Time: 46:11)



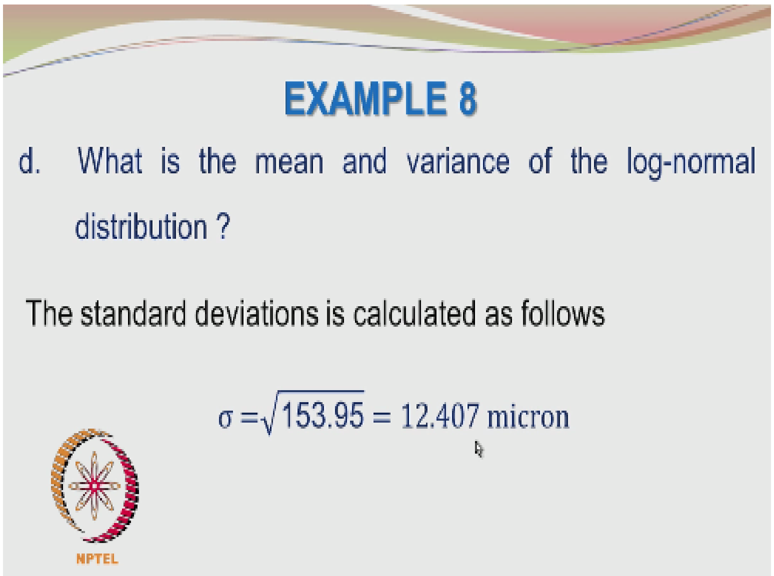
So, what is the mean and variance of the log normal distribution, they are not directly alpha and beta, you have to use the formula, Mu of the log normal distribution is e power - alpha + beta square/2 and in this case, it comes to 1.9933 microns.

(Refer Slide Time: 46:32)



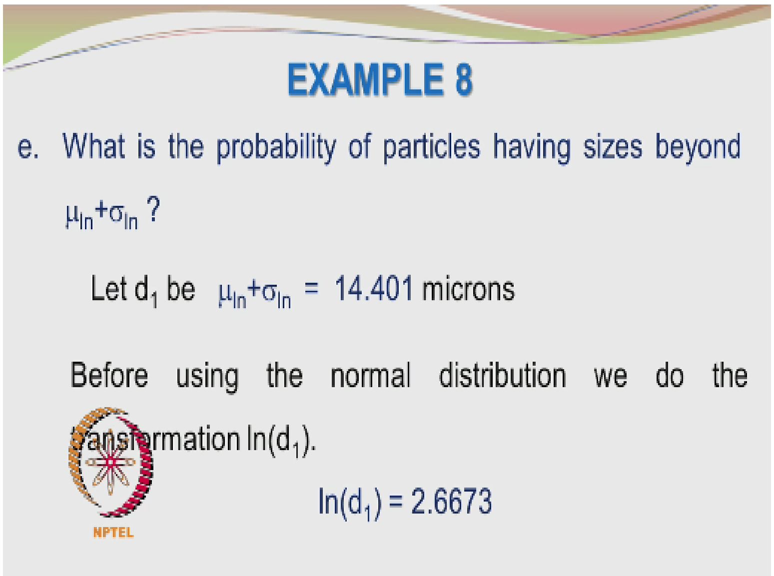
The variance is e power 2 alpha + beta square * e power beta squared - 1 and that comes as 153.95 micron square.

(Refer Slide Time: 46:43)



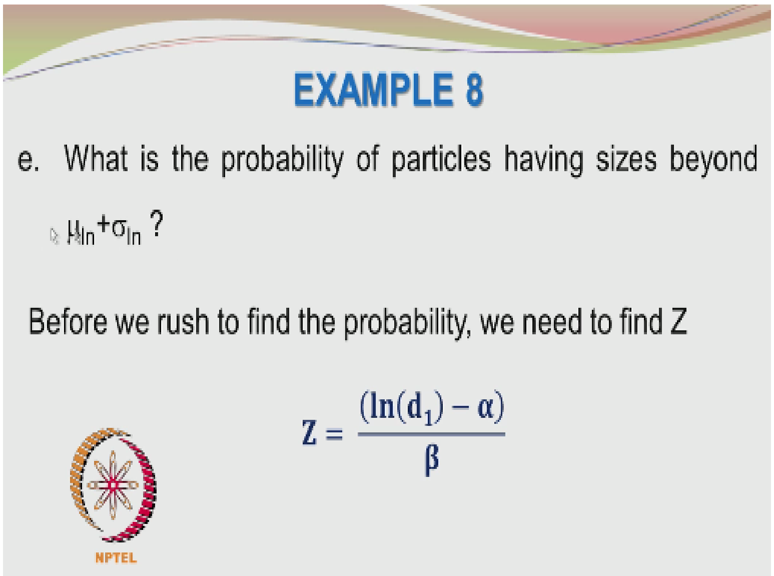
And the standard deviation comes to 12.407 micron.

(Refer Slide Time: 46:49)



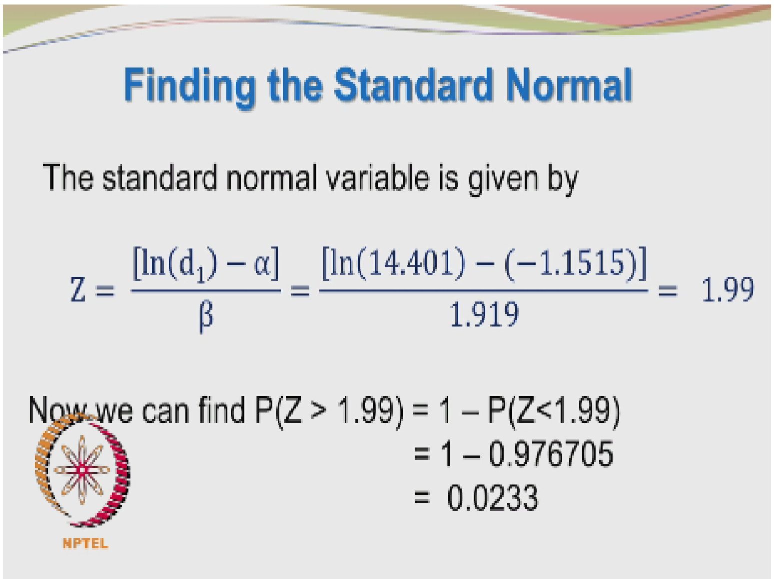
So, there is another subdivision or a couple of subdivisions to this example 8, which we will see shortly.

(Refer Slide Time: 47:00)



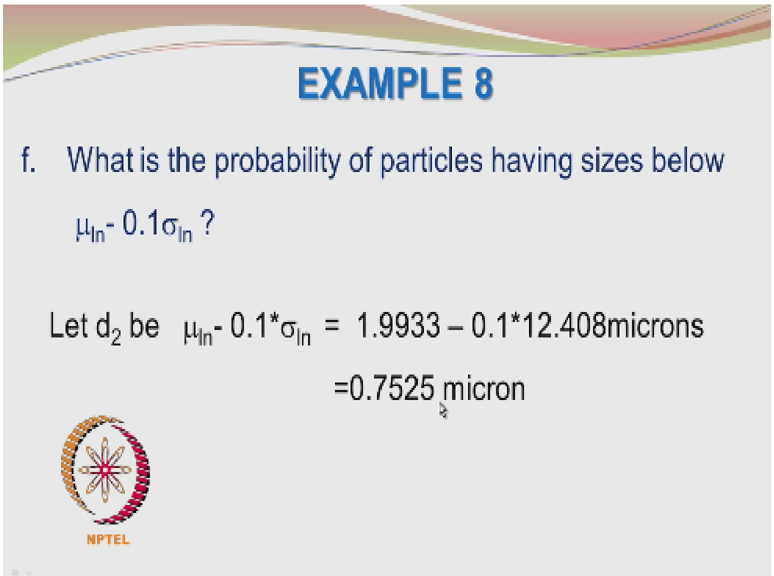
Coming to party of this example, we have really come to the party, what is the probability of particles having sizes beyond Mu of the log normal distribution + sigma of the log normal distribution. Remember that both the mean and standard deviation have the same units, so in this case, it is microns, what we have to do is; to find the probability, we have to convert to *Z*, the standard normal form by the transformation. Log of d – alpha/ beta, so this d will be based on what is the expression given here.

(Refer Slide Time: 47:49)



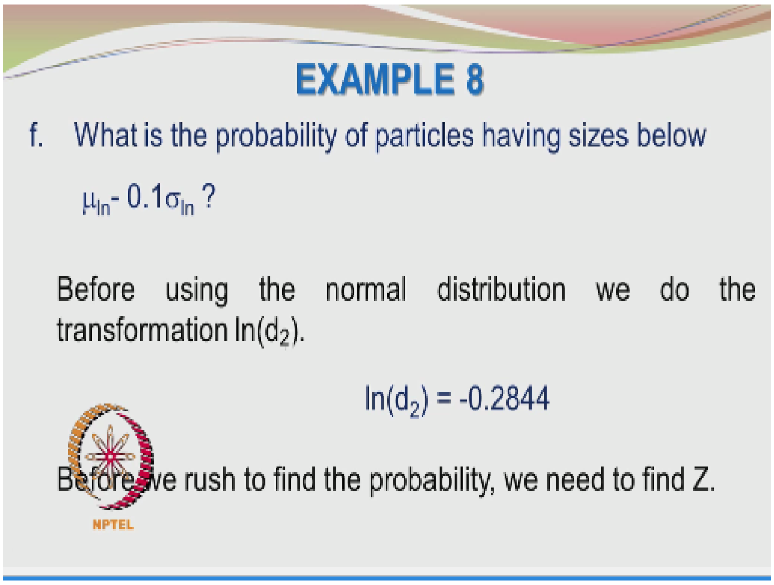
So, log of d1 - alpha would be log of Mu + sigma comes to 14.401, you may want to verify that; - alpha is -1.1515 and beta is 1.919 and you get z value of 1.99 and what is the probability? Z > 1.99 is 1 - probability of Z < 1.99 and we get the answer as 0.0233, so the probability of finding particles beyond Mu + sigma of the log normal distribution is pretty small at 0.233.

(Refer Slide Time: 48:36)



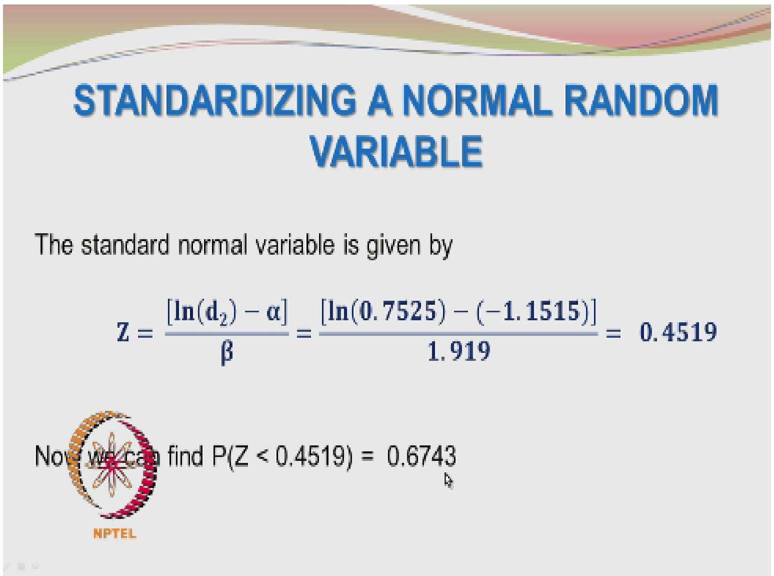
Part f; what is the probability of particles having sizes below Mu log normal – 0.1 Sigma log normal and let d2 be Mu log normal – 0.1 * Sigma log normal that comes to 0.7525 microns.

(Refer Slide Time: 48:59)



And this we have to convert to log; so log of 0.7525 is -0.2844 and then we have to convert this into Z.

(Refer Slide Time: 49:14)



Log of d2 – alpha/beta, which is log of 0.7525 – of -1.1515/1.919 and Z value comes to 0.4519 and the probability of Z < 0.4519 is = 0.6743. So, I have given a few typical examples for the log normal distribution as well as the normal distribution, there are many variants and versions of the examples, you can do. I would suggest to you to take up any standard statistics and probability textbook.

And work through some of the problems to see not only whether you are using the probability charts correctly but whether you are able to understand the problem statement correctly and then do the necessary calculations. Sometimes, you may do the calculations correctly but the answer is not what the problem statement was looking for, so it is important to understand the question properly and then do the mathematical calculations also correctly. Thank you.