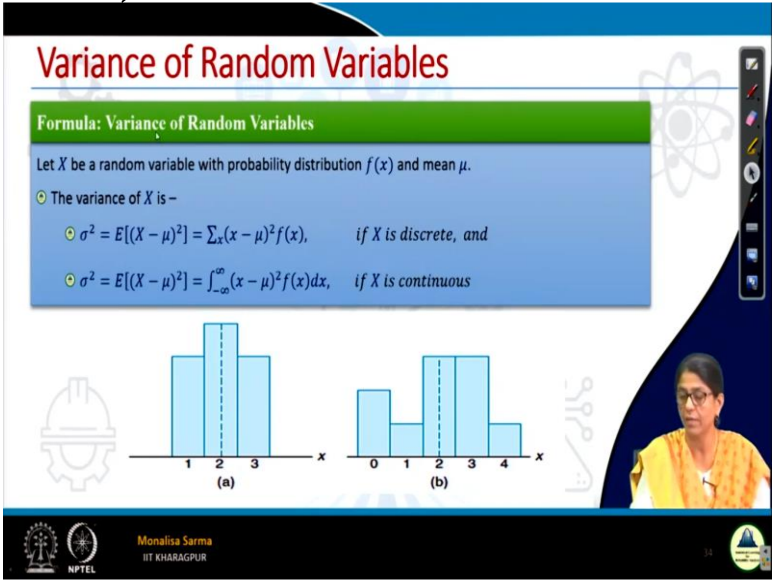
Statistical Learning for Reliability Analysis
Prof. Monalisa Sarma
Subir Chowdhury School of Quality and Reliability
Indian Institute of Technology, Kharagpur

Lecture - 15
Tutorial on Continuous Probability Distribution Functions (Part 2)

Welcome once again. So, last class, we have covered that tutorial on continuous probability distribution but we could not solve problems on all the distributions. We have only done some problems on normal distribution and lognormal distribution. In today’s class we will be seeing exponential distribution, problems on exponential distribution, problems on Gaussian distribution and also again some more problems on normal distribution.

Because normal distribution is a very important distribution, not only from the reliability perspective from and in other application perspective as well.

(Refer Slide Time: 00:59)



So, directly coming to the problem say the first question the length of time for one individual to be served at a cafeteria is a random variable having an exponential distribution with a mean of 4 minutes, see interesting the length of time for one individual to be served in a cafeteria it is at this length of time that is the time that is not a random variable and it is an exponential distribution with a mean of 4 minutes.

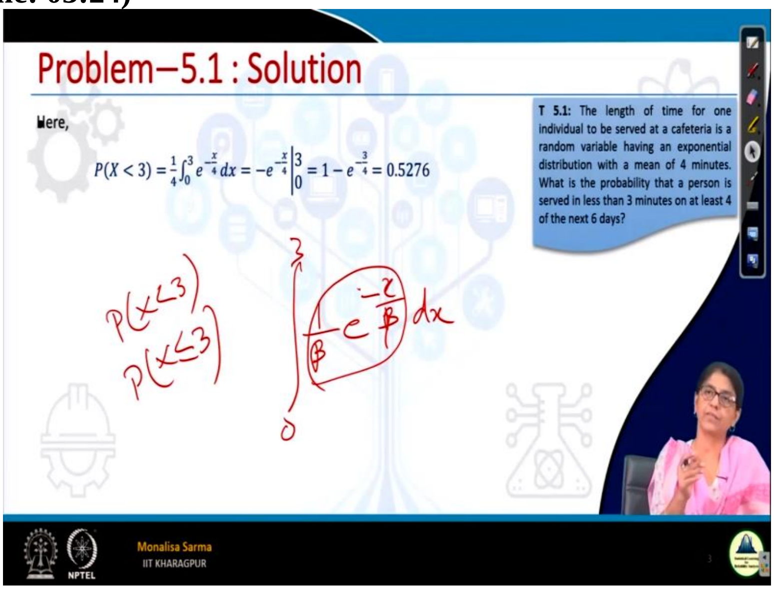
Mean of the time it takes for a person to be served basically, how much minute it takes for the person to be that means we can say that is the rate of the event. So that the rate of the event it is given specific that it is 4 minutes. What is the probability that a person has served in less than 3 minutes or at least for the next 6 days. See the second part, what is the probability that a person is served in less than 3 minutes.

That means we have to find out probability that if you are taking t as a random variable, t is less than equal to 3, probability of t is less than equal to 3 when we need to find out the person is serving less than 3 minutes, on at least 4 of the next 6 days. 4 of the next 6 days, this directly implies a binomial distribution as if there are 6 trials, out of 6 trials I need 4 success, 4 success and what is the probability of success my probability of success is this a person has served in less than 3 minutes.

This is my probability of success. Trials 6 trials out of 6 trials 4 success in binomial distribution, what are the parameters first is number of trials, number of success, probability of success these are the 3 parameters to define a binomial distribution. So, this is number 6 is the number of trials, 4 is the number of success and that the person is serving less than 3 minutes that is the probability of success.

Now, how do I get the problem to find out a probability of success second part I understood it is a binomial distribution, how do I get the probability of success, that is, what is that is exponential distribution, is not it? The time it takes for the first event to happen, time will take for the bus to arrive, time it takes for the failure to happen, or the time it takes for the person to be served in a team in a cafeteria. So these are the occurrence of an event. So that is the rate of an event. So this is and this directly, it is an exponential distribution.

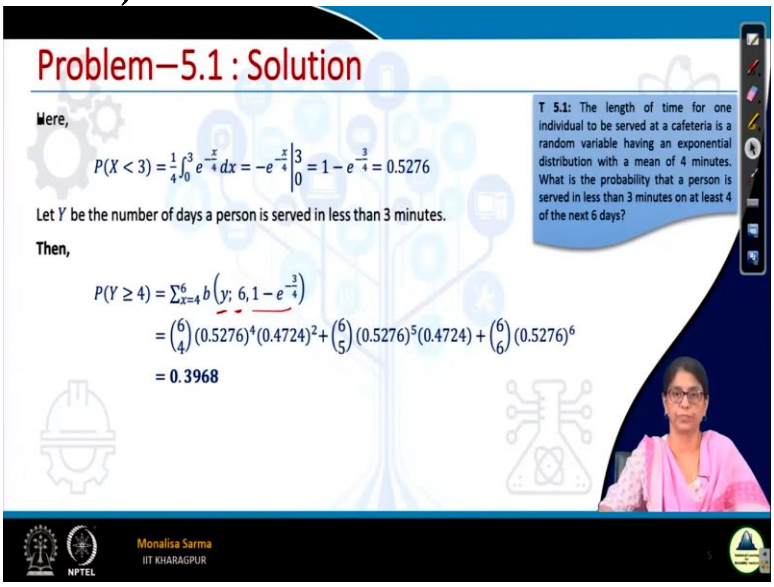
(Refer Slide Time: 03:24)



So, we know the formula for the exponential distribution, I mean, when I call it formula, I mean probability density function, I will not mention it again and again when I am telling it is a formula it means the probability density function. So, I know the expression for the probability density function, what is the exponential for exponential density function? It is 1 / β, 1 / β e $^{-x/\beta}$ this is my density function.

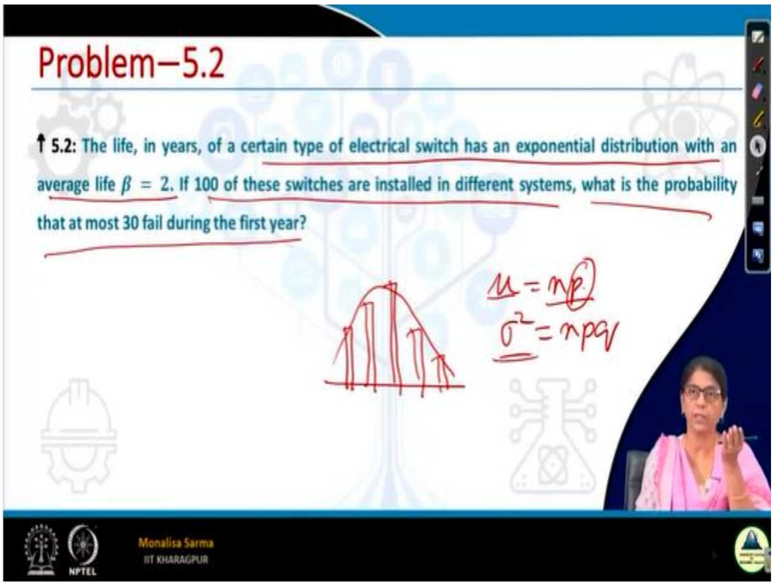
So, I need to find out what it is given certainly less than 3 minutes that has means I have to integrate it from 0 to 3 dx, this is my fx the feature to integrate it from 0 to 3 then I will get the value for probability of x less than equal to 3 whether I tell it is less than equal to 3 or in very important thing in continuous probability distribution. If I tell p of x 3 and p of x less than equal to 3, 2 are different things are same thing, 2 are same thing because that equal to 3 probability is 0. Similarly, if I tell greater 3 or greater equal to 3 both are same thing.

(Refer Slide Time: 04:41)



So Y be the number of days a person is serving less than 3 minutes, that is the Y be the number of successes what I told you. So I found the probability is 0.5276 probability of your success now simple binomial distribution, simple binomial distribution, specifying the value so, this is the probability, this is the number of successes, this is the number of trials simple put in probability distribution function or binomial and we will get the value.

(Refer Slide Time: 05:14)



Next question the life in years of a certain type of electrical switch has an exponential distribution when an average life $\beta = 2$, the life in years of a certain type of electrical switch has an exponential distribution with an average life of β is also equals to 2. Here β is specified as it is equals to 2. Life in years means the time it fails. So, rate of failure is that is $\beta = 2$ if 100 of these switches are installed in different system, say if 100 of these switches installed in different system what is the probability that at most 30 will fail during the first year.

So, I need to find out what is the probability that a switch will fail in one year, this in the first year, so, from 0 to what is the probability that a switch will fail in first year that we will find it by exponential distribution, then see the next portion 100 of a switches installed in different system, what is the probability atmost 30 will fail at most 30 when 0 can fail, 1 can fail, 2 can fail up to 30 it is again a binomial distribution.

Where n is the total number of 100 is the total number of trials and 30 is the number of successes at most 30. So, essentially it is a cumulative binomial distribution from 30 to 100. Now, there is one thing which you need to know is that when the n is quite large, like here, when n is quite large, then sometimes binomial distribution can a discrete distribution also can be estimated by approximated by in normal distribution.

Mainly in case of binomial distribution if the binomial distribution if it takes a bell shaped, binomial distribution, definitely we will our graph will not be a curve it will be either a bar chart or it will be a histogram whatever it is. So, if it is in the form of a, this sort of thing, so, it forms what? Bell shaped it is not it? So, and then the binomial distribution, if the curve is, the resulting curves, resulting diagram I should not occur.

The resulting diagram is in the form of a bell, if it gives a bell shape, then we can approximate binomial distribution by normal distribution, why we will do that is that now see for this case, you would have to find out what 100 is n, n is 100, our x is 30. And you did have to find the cumulative value and for cumulative value of 30. And n = 100. If we have the corresponding table for that for the corresponding probability, of course, the probability is also important.

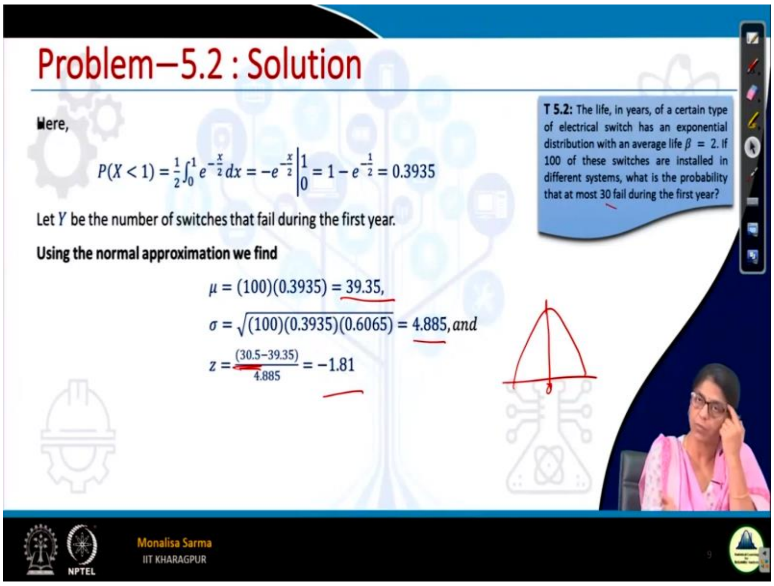
For the probability, if the corresponding probability if we can look it in a table it is fine, but if a probability in the table we do not have the table for all possible values of probability on all possible values are n and x, so, you do a table is not there to calculate such a big number it is a very, it is not a very easy task. So, in that case, if we can approximate it by binomial distribution things become very easy.

Or life becomes easy if you can approximate that is the end when you see if we approximate it, our value will not be significantly change, their value will change slightly, which it can be considered. So, that is why in such cases we can approximate by normal distribution. Now here in case we will be approximated by normal distribution. So, when is the normal distribution again, we need 2 parameters, what are the 2 parameters the parameters are μ and σ.

So, for normal distribution, what is μ? μ is n x p and what is σ? Σ is npq, sorry variances npq. So, to find out for normal distribution, we need μ and σ square here μ and σ square nothing is given. But here we will be able to find out the p probability that it will fail during the first year using the exponential distribution we will be able to find out that will be my p we have n, n is 100.

So, once I find my pq is 1 - p is q. So, I got my μ, I got my σ square, then I will be able to find out my, I will be able to find out the required probability using approximating by normal distribution.

(Refer Slide Time: 09:47)



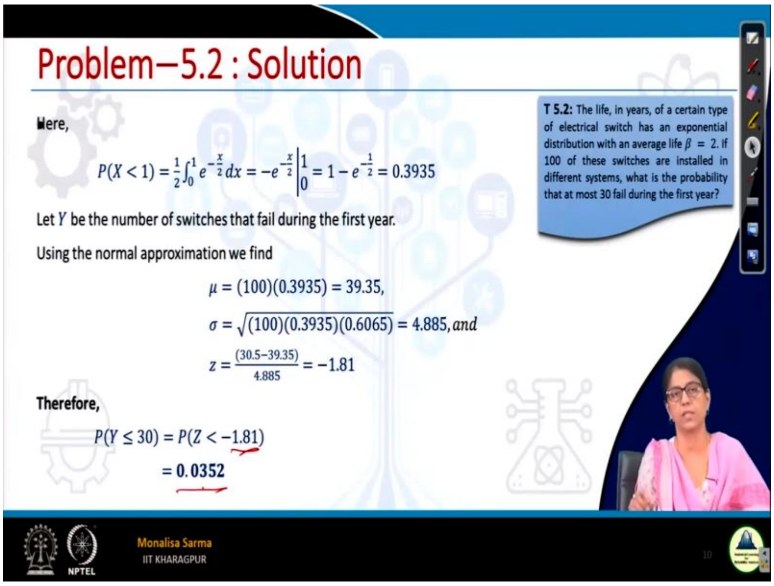
So you see how first we found out the probability using exponential distribution, my probability density function exponential distribution where β is 2 is given. So this is how this is the probability we got so, now we will be approximating using normal distribution. So, what we took μ is n x p we found is this then we found out the standard deviation we found this then we need to find out the z value corresponding to at most 30 fails.

So, it is this is 0 and something 30 means whatever, whatever value will come to for whatever value will come for 30, whatever z value now see when we are trying to approximate using a discrete number when you are trying to approximate by continuous variable, we will have to considered as what to say it is near about value like 30 because probability at exactly 30 will be 0.

So and when we are considering a 30 that means, it will be in a very small interval nearby 30. So, when it is given at most 30, it will fail at most 30. So, we cannot consider an exactly 30 here, see here, we have considered 30.5 instead of 30. Because when we approximate in for continuous, we will have to take a closer interval nearer to 30. So, I have taken is 30.5 here. So, if in other cases, if it is specified that it is less than 30.

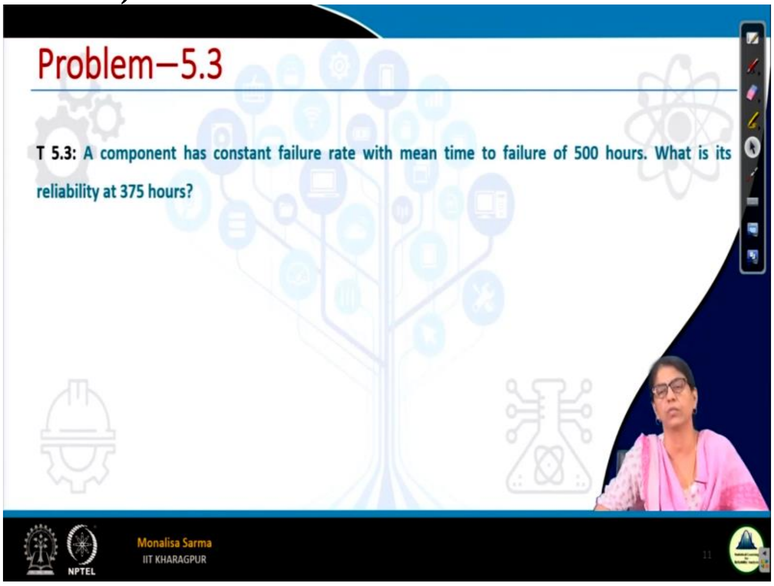
Then I would have considered, I would not have considered 29, I would have considered 29.5. So, I found the z value here -1.81. Once I found the z value, then getting the probability and just a matter of looking into the table that is all.

(Refer Slide Time: 11:30)



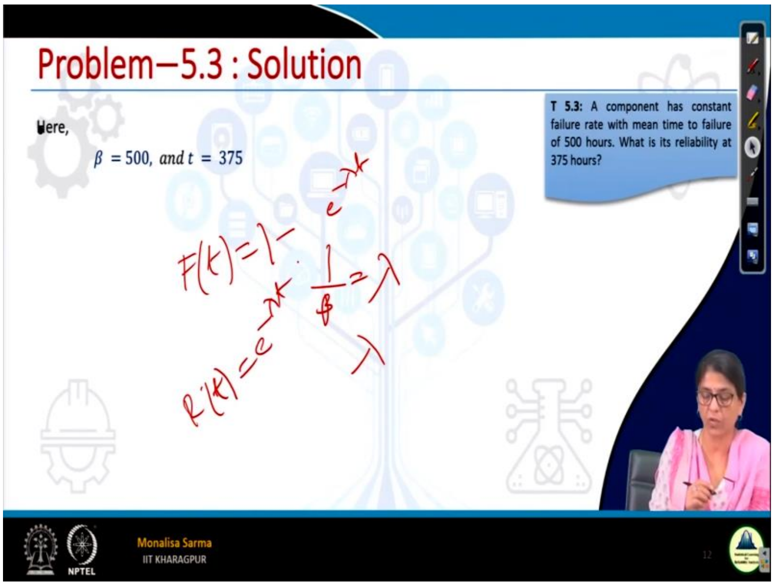
So that is the -1.81 there is a value from the table I got this. So this is the probability.

(Refer Slide Time: 11:40)



So, a come next component has a constant failure rate with mean time to failure of 500 hours what is its reliability at 375 hours similar type of problem.

(Refer Slide Time: 11:54)

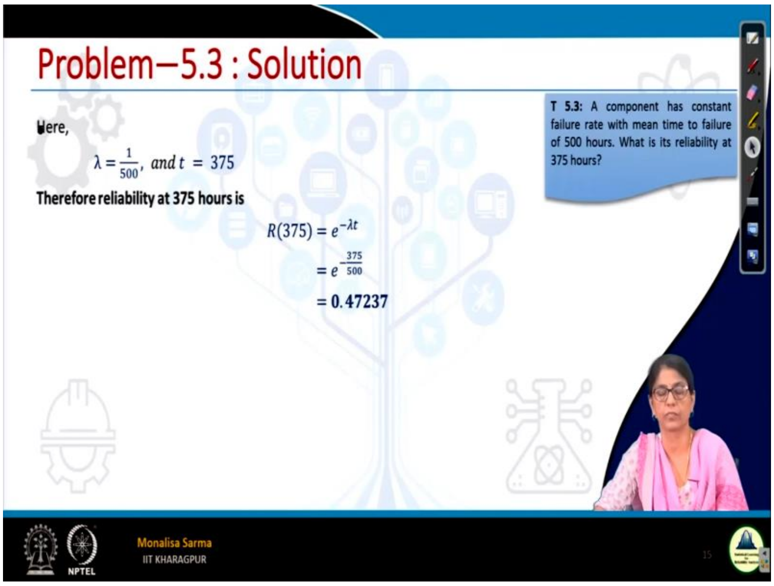


So, here β is given it is 500 hours, mean time to failure that is 500 hours. So, then we have to find out the reliability this is again an exponential distribution exponential distribution we know it is $1 / \beta\ e^{- x \text{ square } \beta}$. Again one more thing the $1 / \beta$ is also equals to λ we have seen that is not it? $1 / \beta$ is nothing but it was λ we have already discussed.

So, instead of writing that, we can also write we have also instead of writing $1 / \beta$ we can also write λ. So, when we find try to find out the reliability of this first we will have to find out F of t. So F of t will be nothing but integration of the probability density function what we have said once we integrate that what we will get? We will get $1 - \lambda\ e - 1 - e^{- \lambda t}$ not λ it is not there $e^{- \lambda t}$.

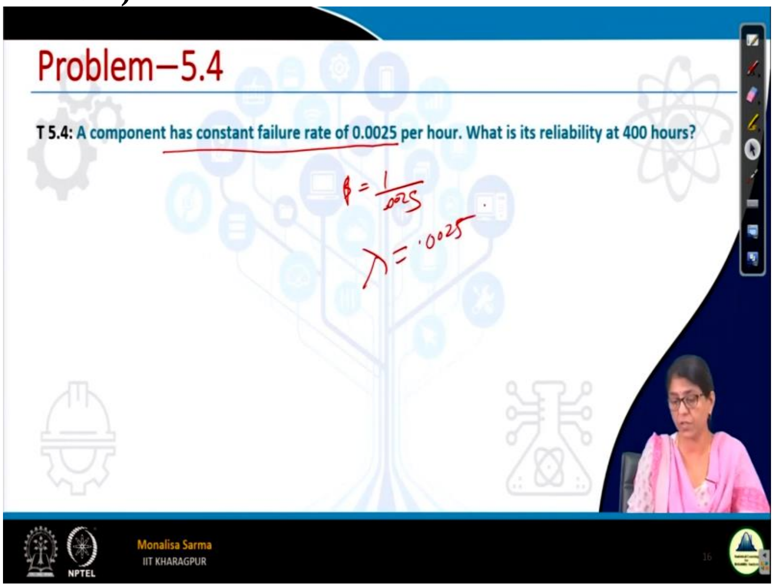
I can you can write it in terms of β or you can write it in terms of λ whatever it is. So, if this is Ft and what is my Rt? Rt will be just $e^{- \lambda t}$ or $e^{- 1 / \beta t}$ whatever you may write in terms of λ, you may write in terms of β.

(Refer Slide Time: 13:12)



So, that is all here it is given I have written in terms of λ.

(Refer Slide Time: 13:20)

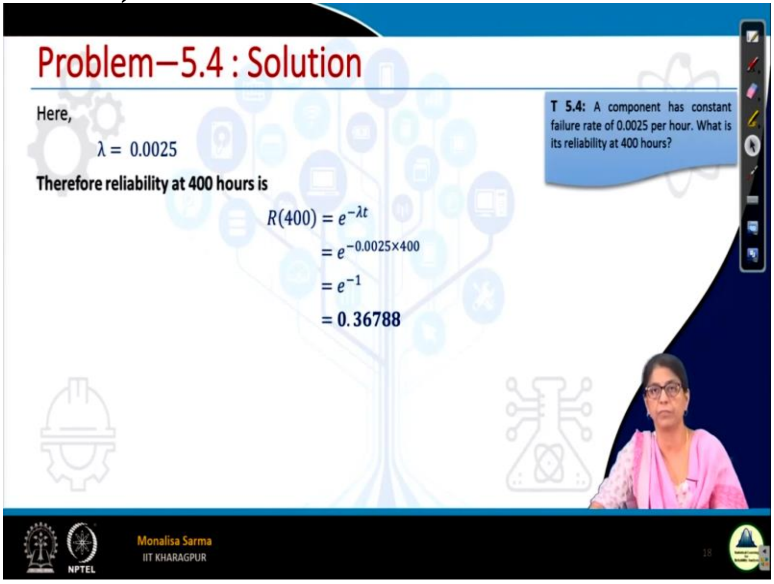


So, a component has a constant failure rate of 0.0025 per hour, failure rate of this much this much failure 0.0025 failure per hour see here, please notice the difference between this question and the previous question, it is given that component has a constant failure rate of 0.0025 per hour that means, 0.0025 failures per hour, if this is 0.0025 failures per hour that means, what will be my then that means, this is my λ value this is not β value.

That means, if I consider β, β will be equals to 1 / 0.0025 because this is considering 0.0025 per hour 0.0025 failures per hour when I consider β is? B is 1 failure per unit time. So, it will be β = 0.0025 whether I can say here it is specified α is specified sorry λ is specified. So, here λ is also 0.0025. See the difference between the previous question what does the previous question as constant failure with mean time to failure of 500 hours.

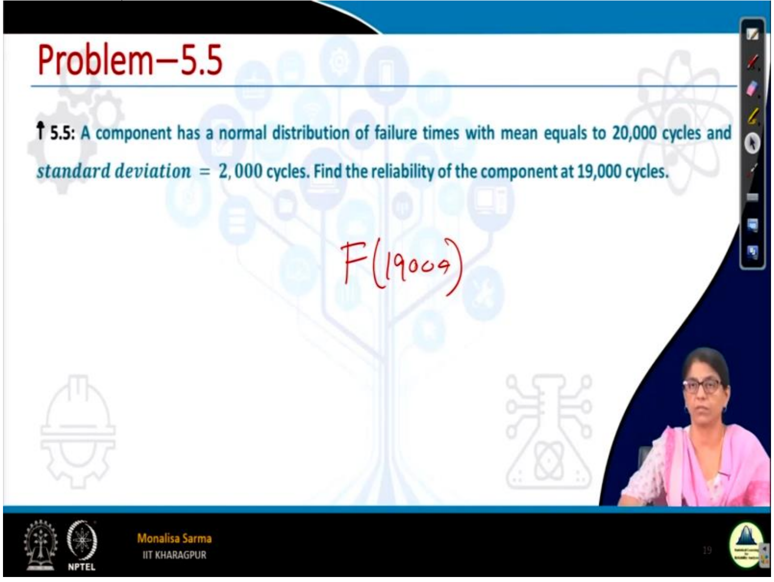
Mean time to failure means that is the time it has time for the first failure or the whatever the one failure is maybe has occurred the time it has taken for the next failure to happen. So, it is within 500 that is the first failure that has happened. So, that is β see the difference between the α and β here.

(Refer Slide Time: 14:51)



So, here sorry not α λ, λ and β here. So, this is λ is given so directly we can find out a reliability $e^{\lambda t}$.

(Refer Slide Time: 15:01)



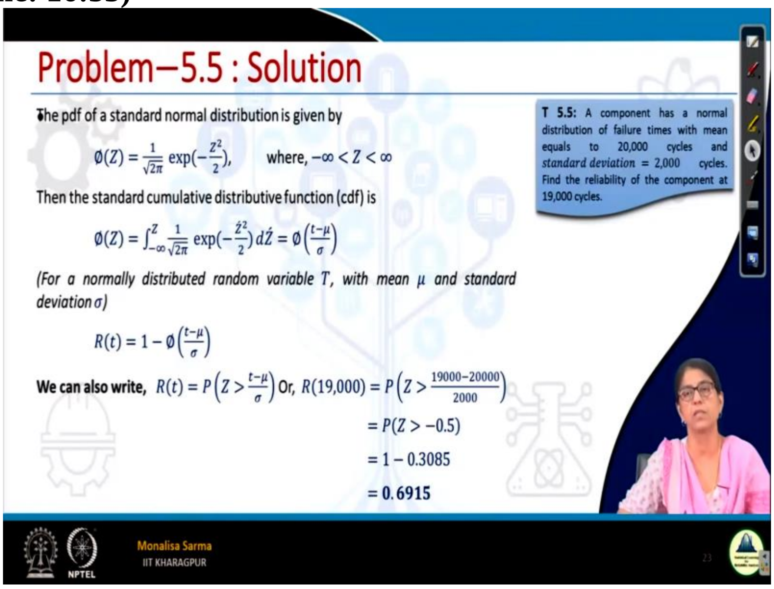
Again, a component has a normal distribution of failure times with mean equals to 20000 cycles, standard deviations because the 2 cycles find the reliability of the component at 19,000 cycles. I have given lots of question on reliability, why mainly, we will have an idea on reliability as well plus we will be practising more and more problems on different distribution.

So, when you talk about reliability, it is nothing just you find out a first you find a cumulative, if I consider the variable when we talk of reliability, then the random variable we consider easily in terms of time it is time to failure. So then if t is my time to failure than small f of t is my failure density function, then I get F of t, but this F of t implies F of t is the cumulative failure till time t.

That means, from - ∞ to t that is the cumulative failure. So, once I talk of cumulative failure, that is F of t from time ∞ to t that is the cumulative failure, then when I what will be my reliability? Reliability will be simply 1 - F of t. So that is just the difference between the F of t and R of t 1 - F of t is R of t, that is why I am telling that is the reliability of the component.

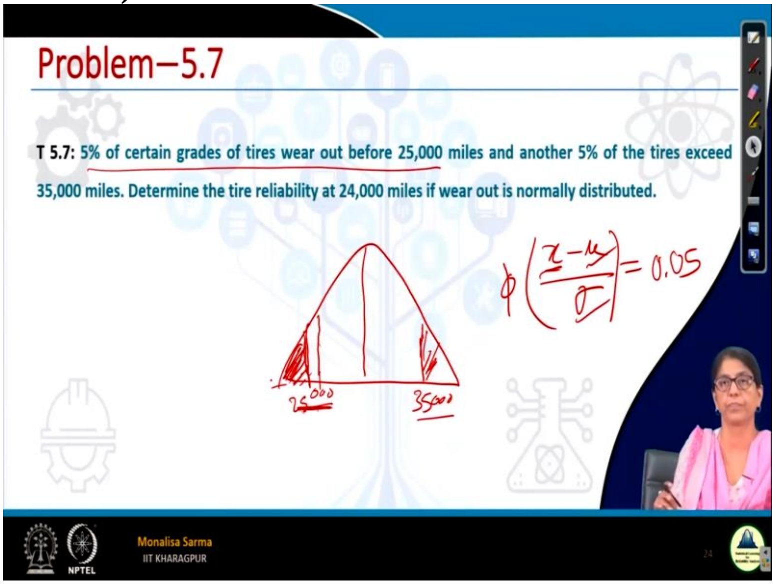
Now, when I am asking reliability of the component at 19,000 cycles. So, that means, first I have to find out what is the failure probability till 19,000 cycles that is F of 19,000, so that means, till 19,000 F of 19,000 means from - ∞ to 19,000 so, it is a cumulative so, we can look it from the table we can get the value and whatever my 1 - of that will give the value.

(Refer Slide Time: 16:53)



So it is, these are the expression but just for a quick recap I have given you this is the failure density function for normal distribution, if we find out the cumulative this is the value and this is the value will look up at the table that is the z value from the z value z is close to - 0.5. So, corresponding to -0.5 we will get the value from the table and what we get is 69.15

(Refer Slide Time: 17:30)



Similar type of question 5% of certain grades of tires were out before 25% 25,000 miles another 5% of tires exceed 35,000 miles determined the tire reliability at 24,000 miles the wear out is normally distributed, we given the wear out is normally distributed. So, first let me draw the normal distribution what it is given? 5% of the certain tire whereas before 25,000 5%, so, 5% means always this 5% means the probability that is in this area.

So, this is corresponds to 25,000 miles. So, 5% were out before 25,000 miles another 5% of tires exceed 35,000 miles exceed 5000 and exceed 35000 definitely it will be this side, definitely it will be this side another exceed 35000. So, this side is exceed, this side is less so, this is 35,000. So, what it is given what is the probability corresponding to this value it is given, probability corresponding to this value it is given.

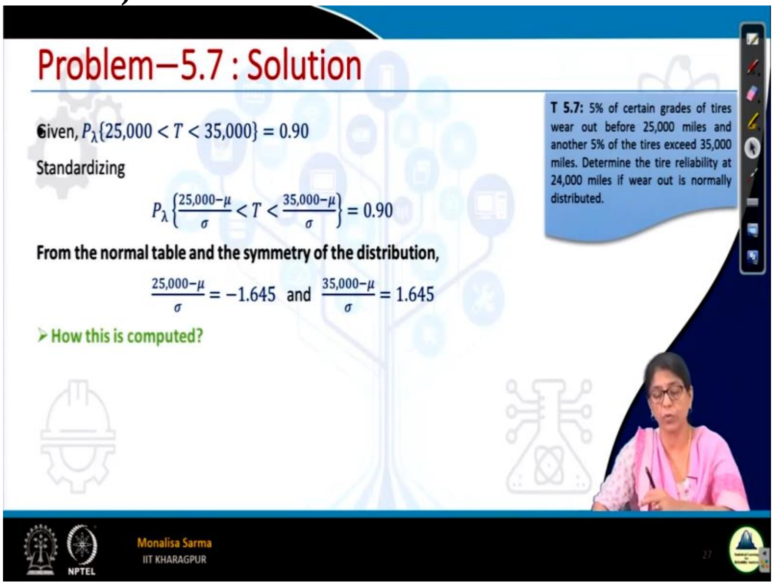
In earlier cases till now, what we have solved what we have we did not know the probability corresponding to a particular z value, particular z value we could find out how if the random variable was x - μ / σ it gives us the z value from the z value we found out the corresponding probability from the table. Now, here the probability is given and my z value is given that means, my say x - μ / σ is equals to this phi of this I get the probability value.

Now, this probability value is given it is 0. 5% and 0.05% 0.05. This is given and what I know? I know this x, x is 25,000. But I do not know μ I do not know σ. Similarly, this is also given, this is also same x - μ / σ this is also 5% 0.05 I do not know x, I do not know μ, I do not know σ. So, 2 questions are there, 2 equation 2 unknown variables. So, I can find out the values of μ and σ that is by solving these 2 problems this 2 equation.

Now determine the tire reliability at 24 tires and 24,000 miles if wear out is normally distributed. So, I have to find out tire reliability at 24,000 miles. 24,000 miles means it will be somewhere here because let me rub this and it will be easier for me to show this was this, this was 25,000 so, I need 24,000 miles 24 it means this is 25,000 so 24,000 will be left of 25,000 so 24,000 will be somewhere here this value.

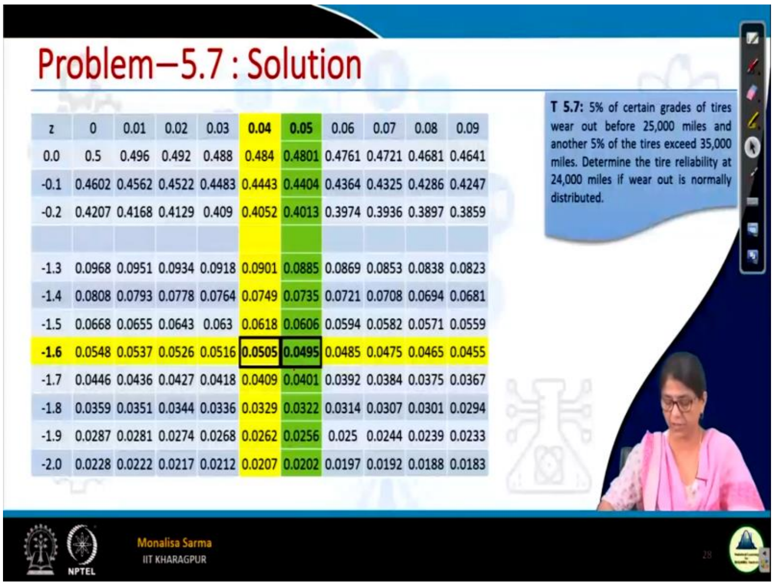
So, I need this reliability at 24,000 miles, to reliability of 24,000 miles this side which will give me F of t cumulative failure after that. So, reliability will be this whole thing is 1, 1 minus of this will give me reliability at 24,000 miles what is the reliability that what is the probability that it will not fail probability failure if so, it will be 1 minus of this same thing does the same concept everywhere.

(Refer Slide Time: 21:03)



So, here this is given so, it both side is 0.5, in this figure what we have seen this is 0.05. This is 0.05 total put together is 10% so, if this is 0.05, this is 0.05 what is the remaining this? Remaining this will be 0.90. So, this value is equal to 0.90, now we know what is the problem the corresponding to this, we know what is the probability corresponding to this then we found out the value of μ and σ, from the normal table similarly the same way now, this one is we got the value 1.645.

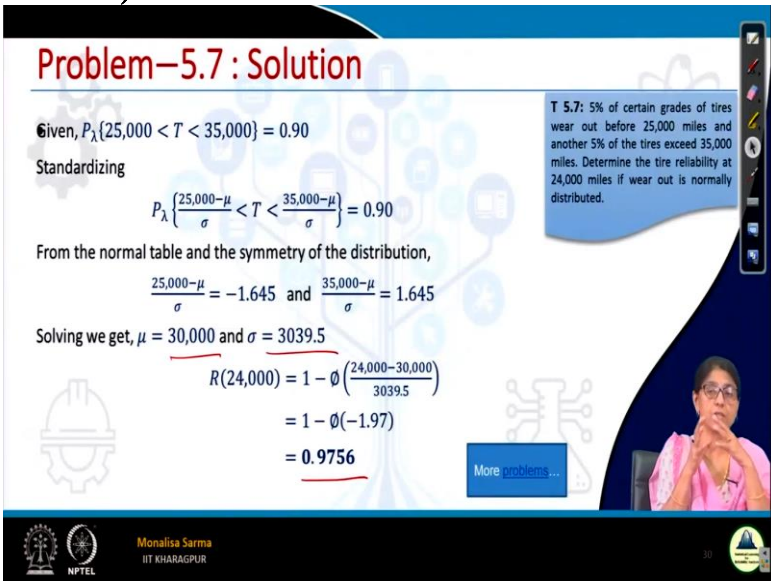
(Refer Slide Time: 21:55)



z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5	0.496	0.492	0.488	0.484	0.4801	0.4761	0.4721	0.4681	0.4641
-0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
-0.2	0.4207	0.4168	0.4129	0.409	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
-1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
-1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
-1.5	0.0668	0.0655	0.0643	0.063	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
-1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
-1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
-1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
-1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.025	0.0244	0.0239	0.0233
-2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183

Again as I told you in the 1.645 it is not in the 1.64 is there 1.65 is there how we can get the value for 1.61 just we will add these 2 values and divide by 2 in simple interpolation.

(Refer Slide Time: 22:07)



So, that is how we got this value and by solving this we could solve this, we could get μ is equal to this, σ is equal to this now simply putting the values for that for finding F of t 1 - phi of F of t I will get reliable data 24,000 simple. So, there are some more problems which I have given a link here there is more problems if you see which are these are problems on Weibull distribution.

As I already mentioned in my 13th lecture that Weibull distribution I will not be discussing in this class because it is this distribution it is a bit difficult to understand for people with reliability background they find it more easier to understand. So, industry people if an industry people or even a professionals who are dealing with reliability or any quality or reliability if they are interested they can see the reliability distribution slides also already I have kept there, already mentioned in the class and if they want to practice any problems.

They can also see the problems are here the link is here. They have around 2 3 problems solve problems here. So now but not only that any students who are interested and M.Tech students if they are interested even B.Tech students who are interested because Weibull distribution is widely used distribution, it is widely used in reliability and quality studies. So, if they are interested, they can go through the problems as well as the concept which I have given in my 13th lecture.

And any doubt of course, I will take up in the doubt clearing session. So it is but I am repeating it is not included in the course that means questions, when I will be setting questions for exams, question will not come from this portion you can be assured of that. So, so now, so from here we have almost come to end of these tutorial classes. Now, by solving so many problems by now, I think you could understand why we need probability distribution.

So, I thought I should ask this question at the end of this class, we have done many problems and problems in probability distribution. So by now, I am sure all of you could answer, all of you could get it, why we need a probability distribution? So, from this from the type of problems we can solve, do not you see that if given if a probability distribution is given then from the probability distribution, we will be able to find out probability of occurrence of certain values.

Suppose, the weight of students or weight of child in a particular locality, suppose weight of child in a particular locality weight of child of a certain age in a particular locality. Suppose in the age of say 15 to 20, 15 to 20 in a particular locality, the weight of the children is suppose it is mean weight is say 40 with a standard deviation of say some x whatever may be the standard deviation.

So, now if I am interested if that means, if it is given the probability distribution here the probability distribution is given what it is given, we know that it is normally distributed, we know the mean, we know the standard variation standard deviation. Now, if we are interested what is the probability that of getting a child of particular weight probability that of getting with 30 or 20 or maybe 50 or 60 given the distribution we will be able to tell that.

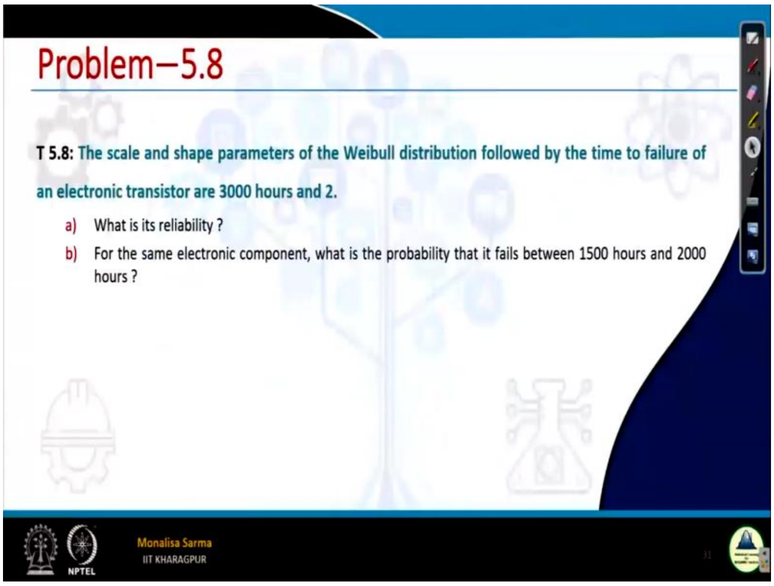
Now, the question is that how do we know given a setup data first initially when we are talking about this child weight in a particular area, who told me that distribution how do I know what distribution it is I just I will just have the data how will I how I will have the data there are some I have taken the weight of all the children in that all I cannot take I have taken the weight of the few students suppose in that areas there are around 1000 children out of one suppose, I have taken the weight of around 1000 children I have taken the weight.

So, from the weight how do I know this weight is normally distributed. So, for that there are again different techniques which I will not be discussing. So, given the set of data, so, from the set of data we can find out this set of data falls in which distribution and there are different techniques for that one of the techniques is probability plotting the many such techniques and power from, from the set of data.

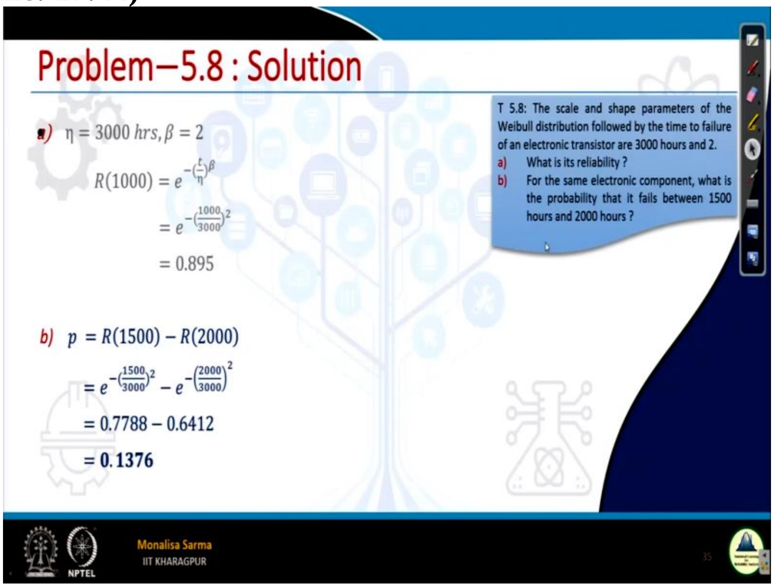
So, first step is that we have the data, from the data we come up with the probability distribution again from the once we have the probability distribution again from the set of data itself by parameter distribution, we can estimate the parameters that means, what are the parameters, if it is normal my parameters is μ and σ, if it is gamma my parameters is α and β, if it is exponential my parameters is only β, β or I can say λ or whatever I can say.

So, once we know that distributions, then the parameters are estimated of the distribution. So, we have the distribution, we have the parameters of the distribution, we have all the information. So, any probability for that we can find out from this distribution that is basically useful application of probability distribution. So, we are learning probability distribution mainly from this application perspective or not only in reliability studies and any studies and any analytics in any business decision and I think we will have quite a use quite application of this probability distribution.

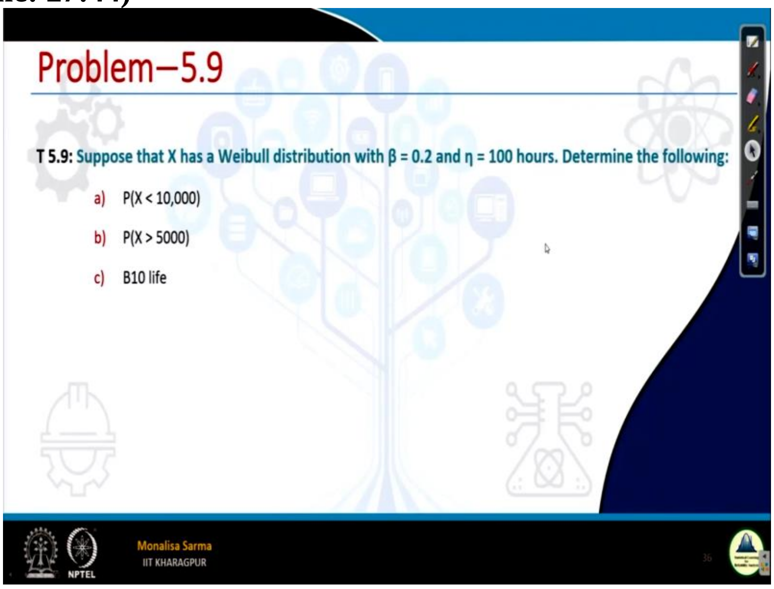
(Refer Slide Time: 27:41)



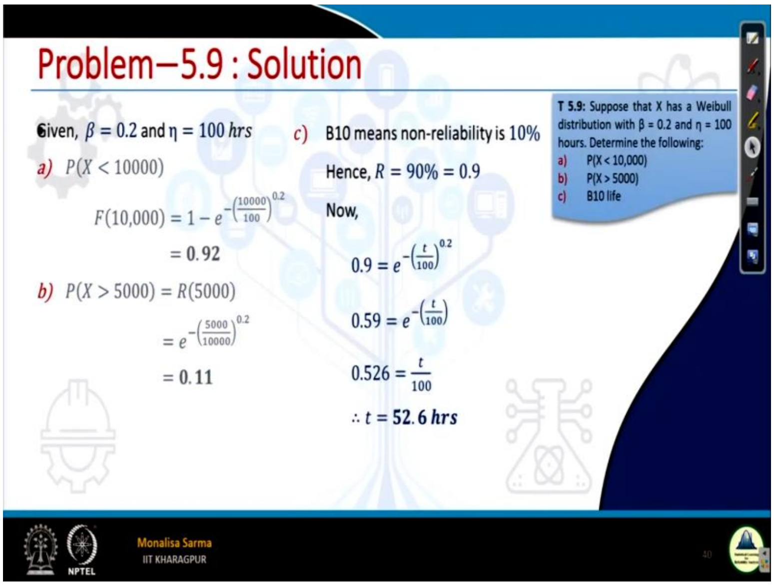
(Refer Slide Time: 27:44)



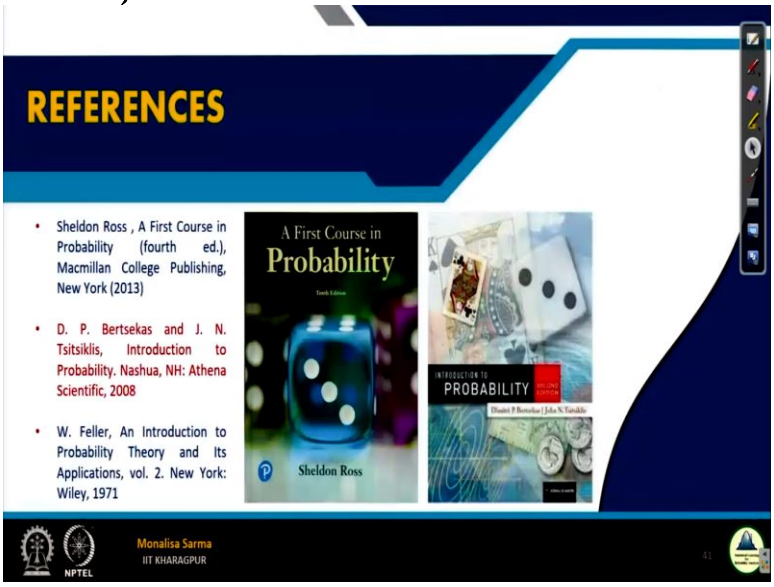
(Refer Slide Time: 27:44)



(Refer Slide Time: 27:46)



(Refer Slide Time: 27:47)



So, these are the questions for which are not be solving for this reliability problems. I mean for Weibull distribution and then I will conclude this tutorial class with the references and thank you guys. So, next class will start a new topic, a very interesting topic that is sampling distribution. Thank you.