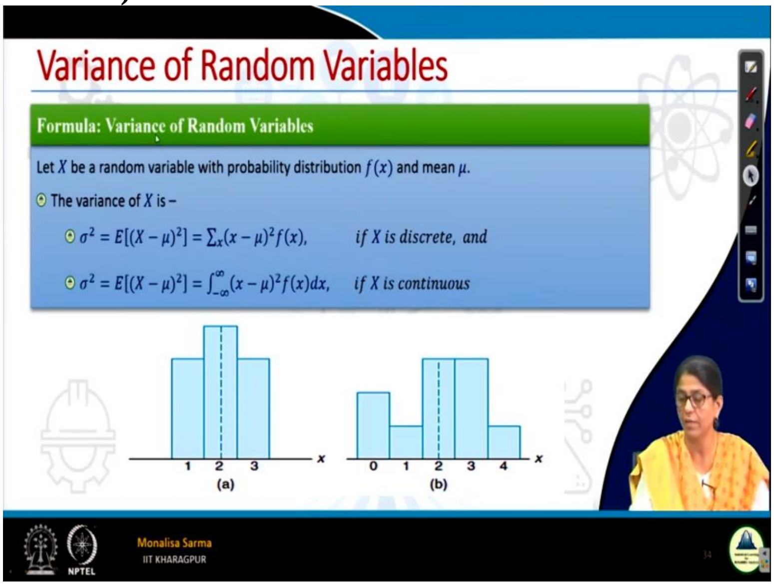
Statistical Learning for Reliability Analysis
Prof. Monalisa Sarma
Subir Chowdhury School of Quality and Reliability
Indian Institute of Technology, Kharagpur

Lecture - 08
Introduction to Probability Distributions (Contd.,)

Hello everyone warm welcome. So, now, in today's class we will continue from the last class like we were discussing probably distribution so, and then we have discussed the mean of an expected variable not we have discussed the mean of a random variable that we also call it as the expected value. So, and then today, we will start after that and that is we will proceed from there.

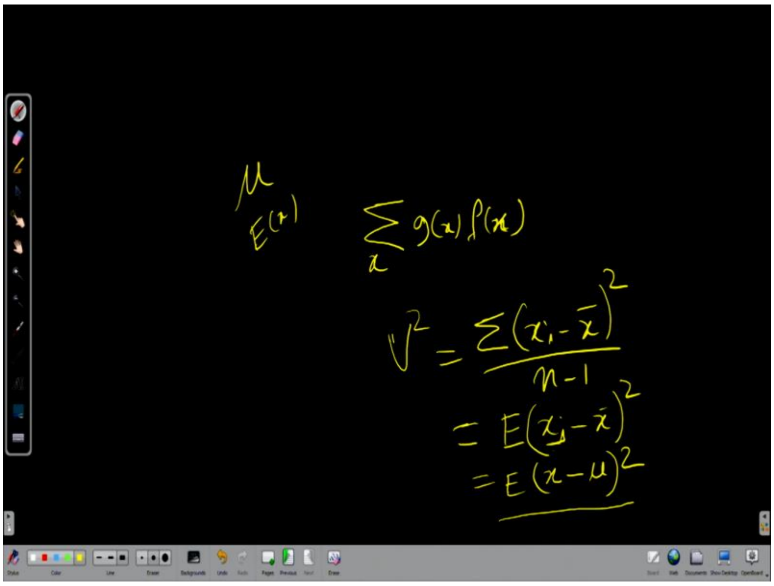
That is we will discuss now variants of random variables, how to calculate the variance of random variable. We have also seen while calculating the mean of a random variable, we have also seen that how to calculate the mean of g(x), what is g(x)?

(Refer Slide Time: 01:09)



Remember last class we have discussed that what is g x? g(x) is a function of x. So, how we calculate the mean of g x, mean of g(x) was same instead of x we have just written it g(x). So, let me just take the board and then explain it.

(Refer Slide Time: 01:24)



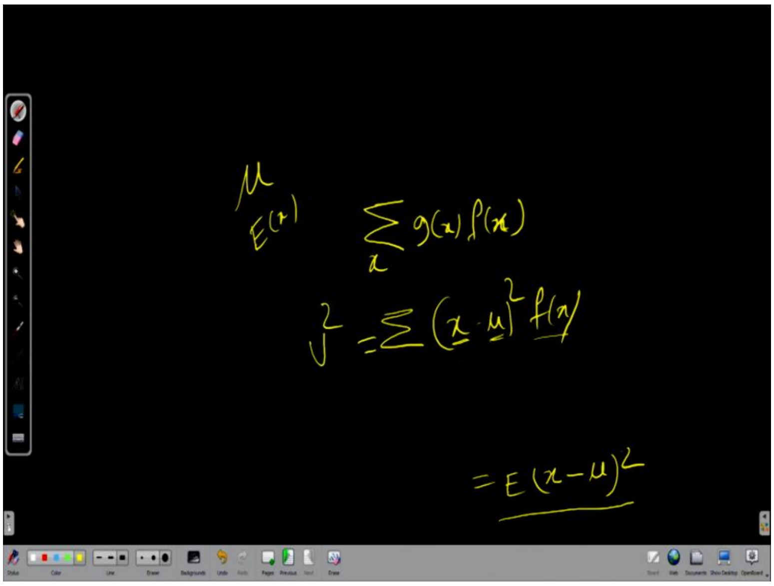
So, how we have calculated the mean of g(x)? Mean of g(x) is summation of g(x) f(x) and summation over all the values of x. So, what is g(x) is a function of x. So, now, when we are interested in finding the variance, variance of a random variable, what is variance basically? Basically when we calculate variance how do we calculate variance we calculate variance but it is v square we calculate variance as summation of $(x_i - {\sim}x)^2 / n - 1$.

This is the formula for variance now, this cannot I write it this way this is nothing but expect this is mean is not it when I am debating with n - 1 I am just finding the mean of all the values, all the values what I have get I just divided because total this much value that I will get and I divided it by that so, this I can just write it as mean of $(x_i - {\sim}x)^2$. This now, what is x_i? x_i means all the different variables that are random variability.

So that means, I can write it instead of x_i I can write it as x that has the different values that are random variability. And instead of x bar, so, we have already discussed when we talk of mean over random variable, either we use the term μ or we use the term μ E of x. So, let me call it as use mu. So, $(x_i - \mu)^2$ so, that means, when I am interested in finding the variance of a random variable.

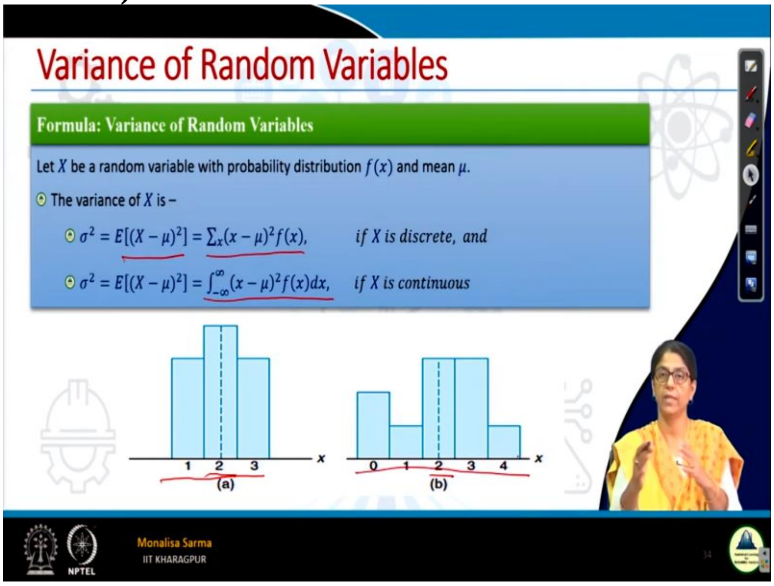
So that essentially I need to find expected value of $(x_i - \mu)^2$ now, what is x - μ also it is nothing but a function of x. So, what is how do I find the expected value of x - μ whole square just putting it in this formula.

(Refer Slide Time: 03:17)



I will just put it in this formula. So, if I put it in this formula what I get what is my g(x)? g(x) is $(x_i - \mu)^2$ f(x) so, what information I have? I have all the values of f(x) what the values are random variability I have all the values of f(x) that is probability of f(x) then of course, I can find out mu from the formula what we have learned and then with this I can find out my variance now my variance is this so, that is what.

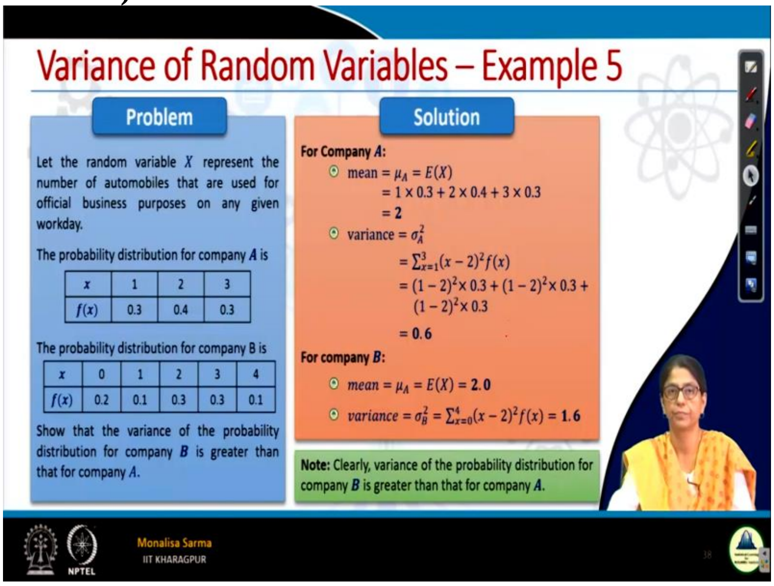
(Refer Slide Time: 03:52)



Here we see, you see this is the variance, variance is expected value of x - mu whole square if it is discrete we use summation, if it is continuous we use integration, so, what is see here this here in the example I have given 2 distribution so, distribution here mean is same for both the distribution mean is here also means is 2, here also mean is 2, here it is mean is 2 here a mean is 2, but the variation is different for both the cases.

This is here a variation is very less, here a variation is quite more it is more spread out variance means how spread out it is? It is more spread out.

(Refer Slide Time: 04:40)



So, here is an example. So, what it is given let the random variable X represents the number of automobiles that are used for official business purpose on any given workday, number of automobiles that if used in any given workday, in any given workday company may use 1 automobiles 2 or 3. What is the probability of using 1 is 0.3 probability of using 2 is 0.4 and probably using 3 is 0.3.

And again there is another company another company which may use in any particular day it can it may use no automobiles at all that is 0 or 1 2 3 or 4 and the probability of all this values are given for 0 is 0.2 and likewise it is given in this table this table. Now, what we have to see is so, that the variance or the probability distribution for company B is greater than for company A.

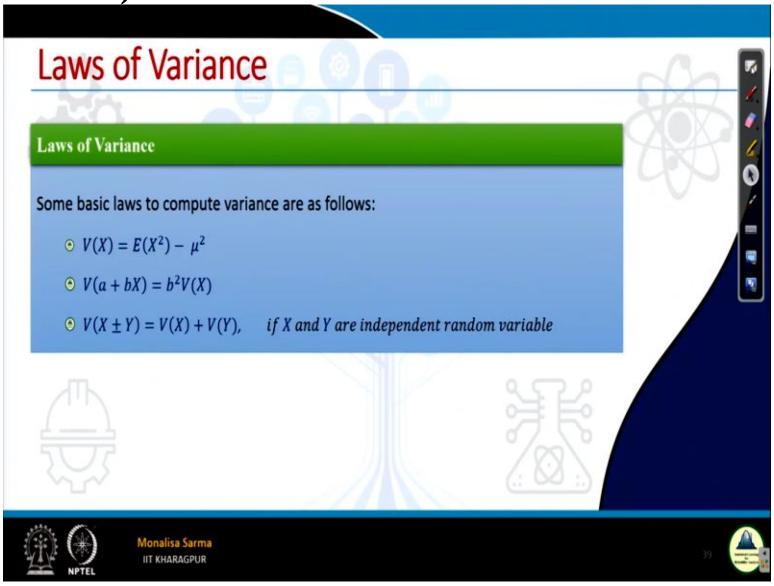
We have to show that the variance for company B is probability distribution of company B is better than company A that means, what we have to find out? We have to find out a variance of company A then we have to find out the variance of company B and from that we will see that the variance of company B is greater how we will do that now, how to find out the variance to find out a variance? First we need to find out the mean.

Mean is nothing but what is the formula for mean? Mean is summation of xf(x). So, what is the x value? x is 1 into 0.3 2 into 0.4 3 into 0.3. Now, what is variance? Similarly, variance is

$(x - 2)^2$ that is variances $(x - \mu)^2$ f(x). So, my μ is 2 $(x - 2)^2$ f(x), so, what is this x minus what value w will take? x will take 1, 2 and 3. So, 1 - 2 into probability of 3, so, this is 2.

So, 2 - 2 into probability of 2 there is a slight mistake in the slide I am sorry for that and then next is that it will take a value of 3. So, this value is 3 so $(3 - 2)^2$ into that probability that is 0.3. So, what we got this will cut out to be 0 2 - 2 is 0 and this from this and this will be getting 0.6. Similarly, we will find out the variance for company B as well. So, once we find out a company variance for company B same way we found out 1.6. So, the variance A is 0.6 variance B is 1.6. So, definitely company B's variances higher.

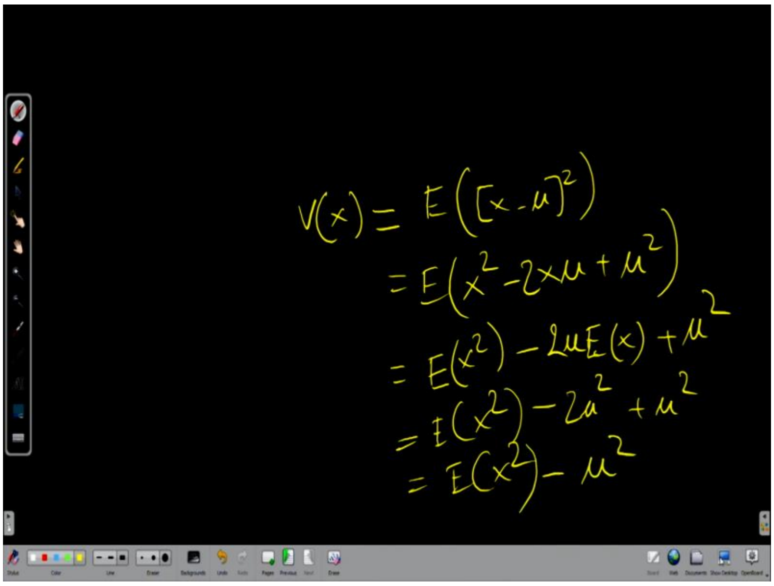
(Refer Slide Time: 07:22)



So, now, some basic laws to compute variance, variance that is we can tell it as VX, VX = E of x^2- μ^2, V of a + bX = b ^{2} V(X). V (X) + V(- Y) means 2 random variable X - Y which has 2 different distributions. So, whether these 2 random variables, we have added or sum it the variance what we get is V(X) + V(Y). This is a very important conclusion with we may need it many other problems when we are trying to find out the variance of 2 different populations.

And we are trying to variance of 2 different population if even it your earlier it was V (X + Y)or X - Y always it will be always the variance or added variants are never subtracted. So, now, how to we can also see how we can deduce this, I will be showing you 1 or 2 rest you can do it by yourself this meant a similar manner. So, first we will see the first one what is the first one?

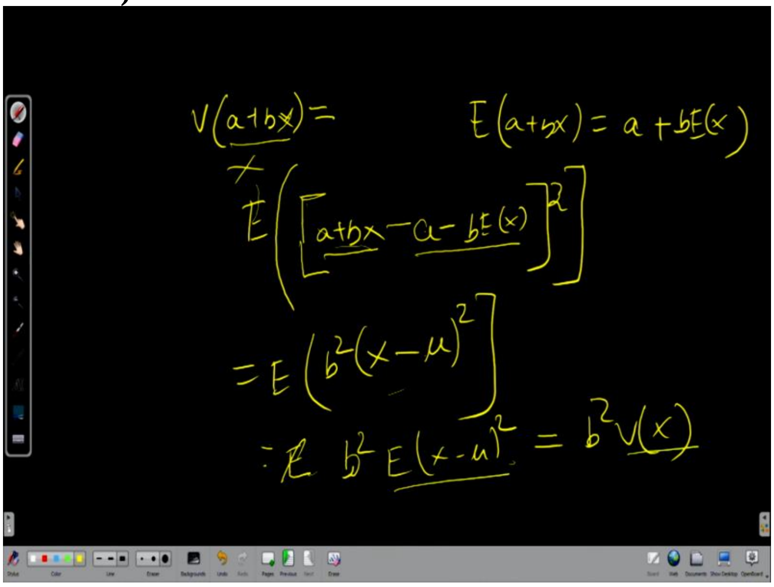
(Refer Slide Time: 08:32)



First one is your V of X.V of X how I can write V of X I can write as E of X – μ whole square. So, if I break it, how does X square minus twice X μ + μ^2, so, I am using the rule for expected value we have seen. So, this is E of X square this is 2 of X μ. So, this is 2 of μ into E of X because μ is constant. So, its expected value will be just μ only and then μ^2 what is mu square E of X? μ square E of X is E of μ^2 will be simple μ^2 that is a constant value.

So, what we got $E(X^2)$ - E (X) is what E of X is μ. So, we got 2 μ^2 + μ^2. So, we got E of X^2 - μ^2. So, this is the first one V(X) is V (X^2 - μ^2). Similarly, second one also I will show you but third one you please do it yourself. So, V a + bX = b^2 VX. How will you find this?

(Refer Slide Time: 09:49)



So, for this V of we need to find out V (a + bX), so what first we will see here, we already know what is E (a + bX)? E (a + bX) we already know we have seen in our last lecture that it will be a + b of EX. So, now, when we are interested in finding out V (a + bX) what is this

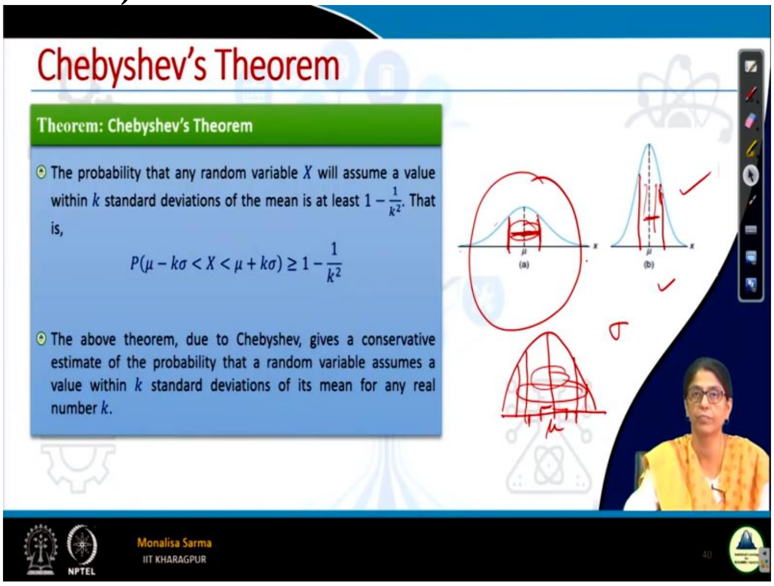
this will be equals to E of what X - μ ^{2} square this equal to E of this is if I consider this as X so, a + bX this is X – μ μ of means mu of this what is mu of this I have done it here that is - a - b of EX whole square.

So, what we got this a a get cancel b square remaining, what is there, what remains? So, it will be X - E of X is what E of x I can write as μ(X - μ)^2. So, what is this E of b square I can bring out let me cut this. So, this is b square b square E of (X - μ)^{2}what is (X - μ)^2? (X - μ)^2 is nothing but V. So, this I can write as b square V of X.

Again I am repeating how I have done so, I am interested in finding out V of a + b what is V of a + bX? V of variance of X is (X - μ)^2. So, this is X so, (X - μ), μ means E of X so, this is X and this portion is E of X E of this I have seen it is this. So, I have subtracted this a a get cancel what remains is b into X - E of X, E of X is μ (X - μ) whole square, so b square and bring it outside.

So, I got b ^{2} E of (X - μ)whole square and now, what is this this is nothing but equals to V of x. So, similarly, you can find out try and do the third one.

(Refer Slide Time: 12:29)



Now, there is an important theorem, important theorem in the sense like see now, you understood what is a variance? Variance means how spread out my data is, now this you see in this figure and this figure both the figure may have the same μ but say this figure is more spread out the second figure, this figure is more spread out than compared to this figure, this figure.

This figure is more spread out, the spread out means its variance is more and suppose I am interested in finding out what is the in any random experiment for any particular event, I am interested in finding out what is the probability that my value will lie in this interval some distance from μ similarly, I have considered the same distance from μ what is the probability of a value line in this interval.

So, in this case the probability of lying in this in this case my probability of lying when this interval will be high compared to this my probability of lying in this interval, lying in this interval will be high compared to this probability distribution, why this probability is nothing but in case of continuous probability it is nothing but the area if you consider the area this area is because I hold it this whole area corresponds to one.

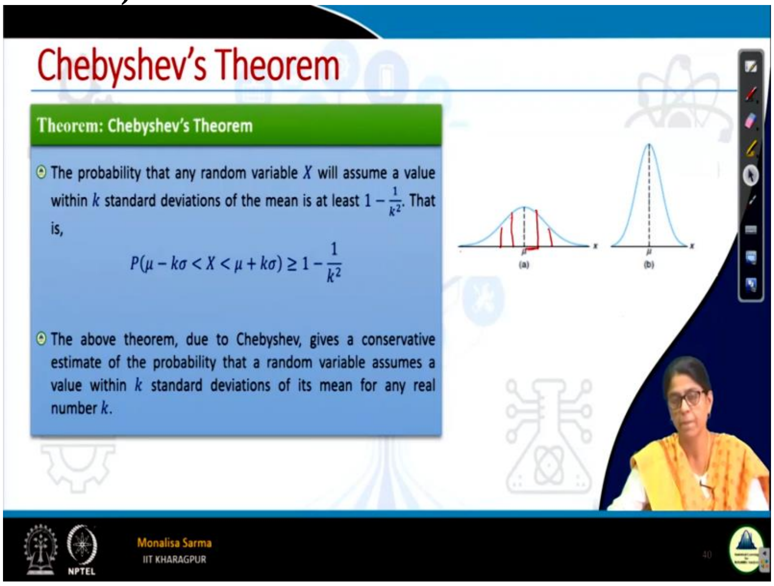
So, definitely this area will be less compared to this area. So, that means my probability will be higher lying in this range. So, Chebyshevs theorem basically gives a very conservative estimate of that. So if I, in any random exponent for a particular event, the probability of lying within a particular distance from mu what is the probability of lying getting a value within a particular distance from μ that Chebyshevs theorem gives that.

So, what is that the probability that any random variable X will assume a value within k standard deviation of the mean, k standard deviation means one standard deviation. So, suppose I have drawn a figure suppose this is my figure, suppose this is mean and suppose this is one standard deviation, this distance is standard deviation, standard deviation is variance route overall variance this my standard division This is my standard deviation.

So, this is one standard deviation then 2 standard deviation maybe this point what is the probability of lying this region within this region, 3 standard deviation maybe this point what is the probability of lying within this region. So, what the Chebyshevs theorem says the probability that any random variable X will assume a value within k standard deviation is at least $1 - 1 / k^2$.

But here k, this is applicable when $k > 1$ it is not applicable for $k = i$, so, above theorem due to Chebyshevs gives a conservative estimate of the probability that a random variable assumes a value within k standard deviation of its mean for any real number k, it is clear?

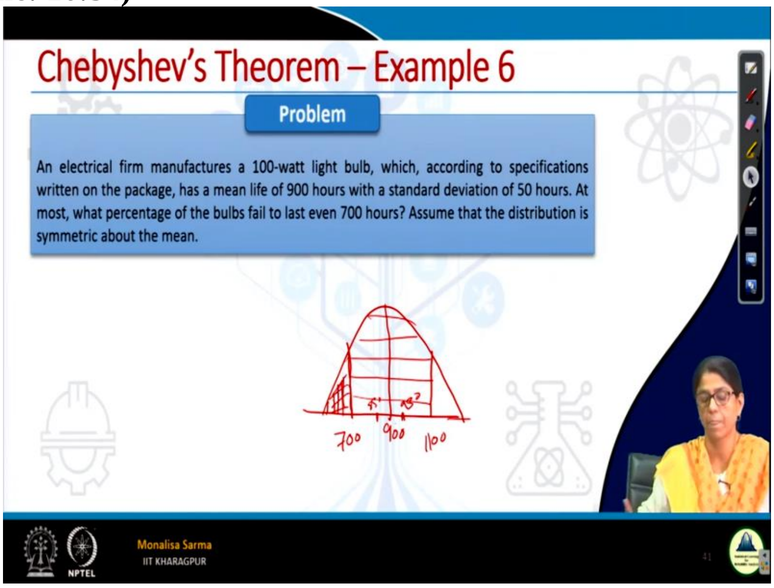
(Refer Slide Time: 16:00)



Essentially what the Chebyshevs theorem tells? Chebyshevs theorem tells what is my probability of getting a value in a particular range, this range is specified by the standard deviation what is the probability that I will run in a random experiment and I will be getting a value how much standard deviation from the mean, mean is the middle portion, from this mean how much standard deviations?

If it is one standard deviation means maybe this, one standard deviation maybe this mean, this may be one standard deviation, maybe 2 standard deviation maybe this, this is what Chebyshevs theorem says, this gives very conservative, conservative estimate of course. But still this gives us an idea in an experiment based on a variance we can find out what is the probability of getting a particular getting a value within that range.

(Refer Slide Time: 16:54)



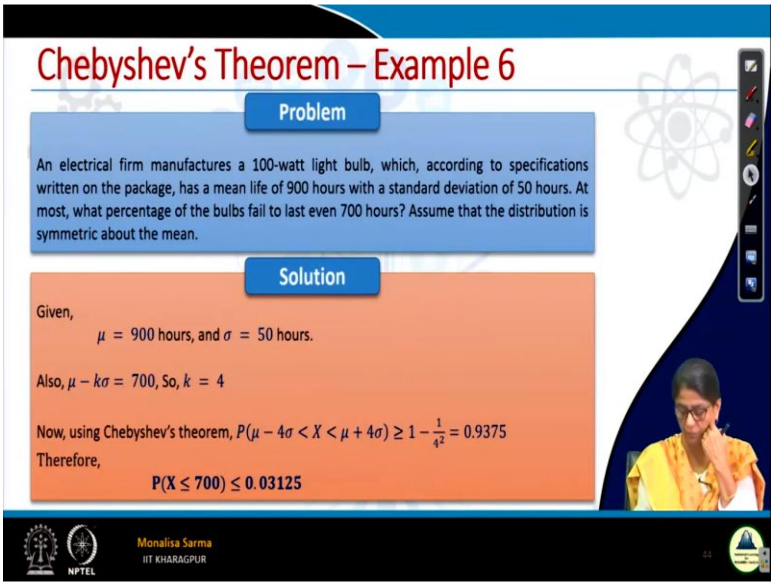
So, one example will make it more clear you see an example here an electrical firm manufactures 100 watt light bulb which according the specification written on the package has a mean life of 900 hours and a standard deviation of 50 hours. So, this electrical bulb, so it has suppose I am using a simple distribution suppose it is this type of distribution. Till now, I have not explained the type of continuous probability distribution of this problem.

So, let us assume that normal distribution so, this is what is given mean is 900 hours, so, this is maybe the mean, mean in standard hours this side it will be greater than 900, this side will be less than 900. And standard deviation of 50 maybe is 50 is this point that means, it is 950 and this is 850 here that is the standard deviation, standard deviation is 50. So, if I am considering one standard deviation, one standard deviation maybe 950 at this point.

This is 850 one standard deviation, 2 standard deviation this side it is 1000, this side is 800 within 2 standard deviation. So, now, what is the question at most what percentage of the bulb fail to last even 700 hours, failed to last even so, this 900 maybe this much is 700. So, the question is what percentage of the bulb failed to last even 700 hours. So, that means, my bulb but the value that I got that means that my working life of the bulb is all in this range lesser than 700.

It is not in this range greater than 700 is in this range. So, what I will do is that I will try to find out the probability within range from this 700, this side 700 means, it is from 900 to 700 how many standard deviations it is? It is 50, it is 4. So, here so 1100, so, what is the probability of finding in this range, I will find out using Chebyshevs theorem and then I will subtract it from that 1 minus I will get this area is not it.

(Refer Slide Time: 19:17)



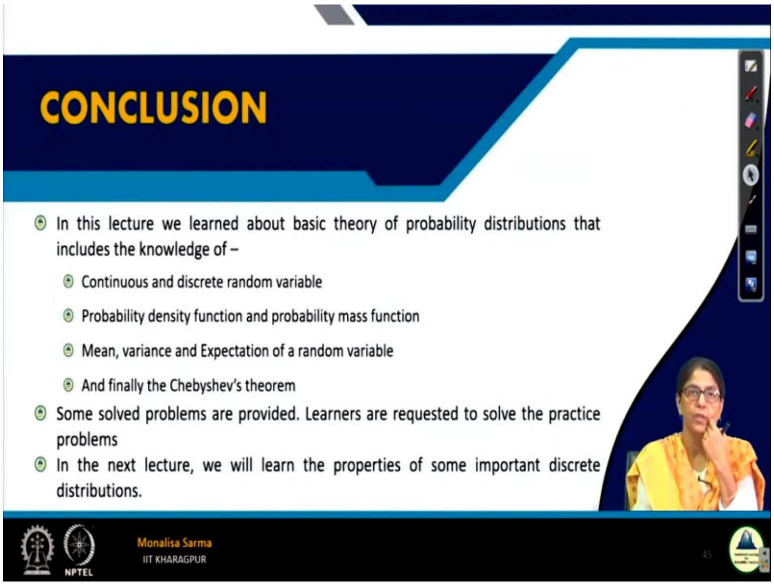
So, see here my 900 is, μ is 900 σ is 50. So, μ - k σ is 700. So, definitely k goes to 4 that all without solving also we can say because it is sigma is 50 it is very easier to tell. So, now, using Chebyshevs theorem, but we get probability of this is at least this value that means greater than equals to 1 minus this probability that it will lie in the range of 700 to 1100. It is this much 0.9375.

So, what is the probability that it will lie that it lesser than 700, this is 900 so, this is 700, this correspond to 700, this corresponds to 1100. So, the probability that it will lie in this range is just 0.9375 Now, I am interested in finding out this probability. So, 1 minus of this probability I will be getting this and this together. So, it will be lesser than this that will be divided by 2.

So, it is this much, we can just do it and one minus of this and then divided by 2 because it is this lesser than this as well as greater than 1100, it can greater to both the portion because we have just found out a property of the middle portion and we want to find out what is the property that has appeared in the lesser than 700. So, then what we got one minus of this it will greater to both the portion, lesser than 700 as well as greater than 1100.

1 minus so what that value 0.9375 will give that and we are interested in only lesser than 700. So, whatever value we will get probably that will be divided by 2, then we will get 0.3125.

(Refer Slide Time: 21:13)



So, I conclude here. So, in this lecture let I should not say in this lecture, in this lecture and a previous lecture that was lecture 7, but we learned about the basic theory of probability distributions earlier we have learned probability now in this last 2 lectures we have learned what is the probability distribution, what is a random variable then we have now learned what is continuous probability distribution, discrete probability distribution, we have learned how to find out the mean of a distribution, variance of a distribution.

And also we have discussed Chebyshevs theorem like last time again I am repeating solve as many problems as possible. Now, in my next lecture, we will learn different discrete probability distributions. With that I end this class thank you. Here the references of the different books and thank you guys.

(Refer Slide Time: 22:04)