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

Lecture - 07
Introduction to Probability Distributions

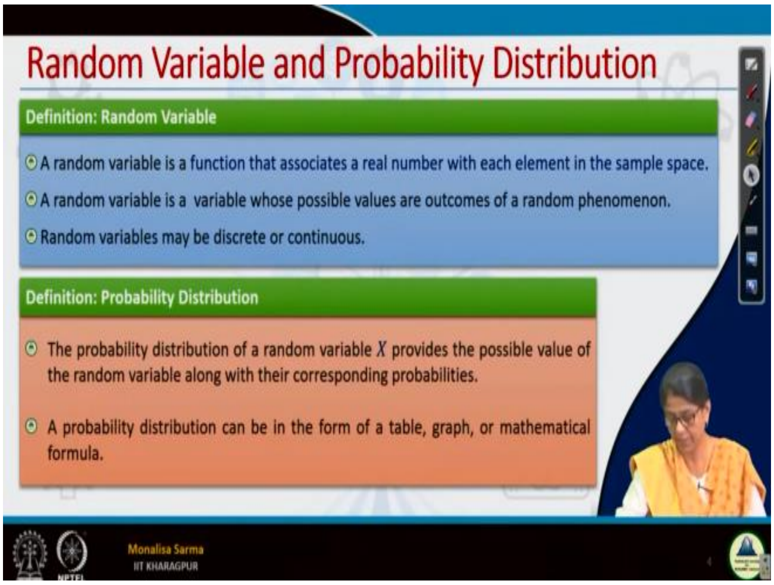
Hello once again so now, today we will be discussing probability distribution we have discussed probability but in probability theory we have in probability theory, we have seen different types of events then we have seen conditional probability total probability and Bayes theorem and now going a bit further. Now we will be seeing what is probability distributions.

(Refer Slide Time: 00:49)

So, for to in this probability distribution first I will be explaining what is a random variable then I will go on to explain what is a probability distribution? We will discuss discrete probability distribution what is a continuous probability distribution what is a discrete probability distribution we will also see what are the mean? We have seen what is the mean and variance given a set of numbers.

Now, we will see mean and variance with respect to a random variable mean and variance of a probability distribution essentially, then we will discuss what is a Chebyshev's theorem with that we will end this class.

(Refer Slide Time: 01:21)



Now, first is what is a random variable? So, random variable if you see the definition random variable is a function that associates a real number with each element in the sample space. Suppose in samples suppose of course, I have toss 2 coins, and when I toss 2 coins we are doing it let me take the pen, when I toss 2 coin, what is my sample space my sample space will be say, my sample space will be say head head, head tail.

Then again tail head, tail, tail this is my sample space. So what happened if I am interested in finding out from number of heads when I toss 2 coins so what the number of heads I get this is number of heads is 2 number of head is 1 here number of head is 1 here number of head is 0. So this is the random variable essentially. So it is a function that associated real number with each element, in the sample space.

A random variable is a variable whose possible values are outcomes of random phenomena. That is also we can define the variable so random variable now, this random variable may be discrete maybe continuous. So, now, with the knowledge of random variable, now, we can define what is a probability distribution when we do an X when any random experiment we will be getting different sort of outcomes. What are these outcomes?

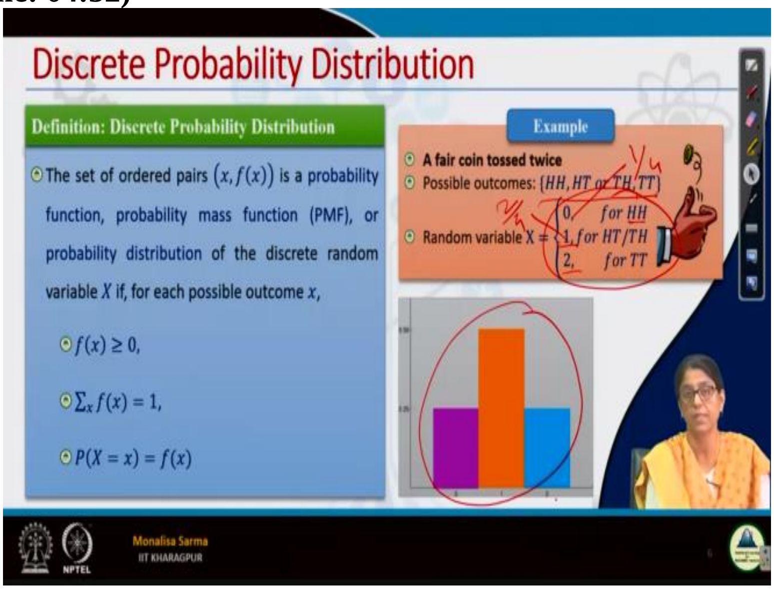
These outcomes are the values of the random variable. Now, this random variable does values they have a certain probability of occurrence like, if you toss a coin, what is the 2 outcomes head and tail we have a probability of occurrence of head and tail if I let me take it in this way. What is the probability of getting head? So probability of getting either we will get 1 head or we will get 0 head if it is a tail that is a 0 head and what is the probability of getting 1 head is 0.5 and probably of getting 0 head is 0.5.

So, this that means the values of the random variable when we do an experiment the values that we get the random variable and the probability of occurring of those values these 2 things put together is probably distribution. So, what is probability distribution? Again I am repeating probability distribution is nothing but the values of the random variable as well as the probabilities of occurrence of these random values all the random variable in an experiment. So, how we define it?

The probability distribution of the random variable X provides the possible values of the random variable all possible values of the random variable along with their corresponding probabilities that is the probability distribution. So, we will be having since random variable may be discrete random variable may be continuous, discrete means it will take some infinite values 1, 2, 3, 4 some finite values may not be one to other but it will take some finite value and what it is continuous?

Continuous means it can take any value within an interval. So, if the random variable is discrete then we call it discrete probability distribution if the random variable is continuous, we call it a continuous probability distribution. So now a probability distribution we can describe a probability distribution in the form of a table in a form of a graph or in the form of a mathematical formula, we will see.

(Refer Slide Time: 04:52)



Now, first we will define we will see the condition of a probability distribution function. So what is the probability distribution? First when I tell that is random, suppose my random variable is X, then what is the probability of X? That I will define it as an f(x), what is the probability that X stays that I can let me define it as f(x) or you can tell P(X) anywhere and it is just a function of X basically the probability of X, f(x) is nothing but the quality of those the set of ordered pairs x f(x).

The set of ordered pairs f(x), f(x) is a probability function or we can call it a probability mass function or probability distribution we can call it a probability distribution function or probability mass function of the discrete random variable x. Now, here one thing please note it see I have mentioned here of the random variable x and here I have used capital X in here I have used small x, this is a general convention when we talk about the random variable, we always use capital letters.

And when we talk the value that the random variable takes we usually denote it by small letters, now, this x is the value that this capital X takes my random variable is capital X, and x is the value that a random variable takes. Now, when I am tossing a coin tossing 2 coins, I was getting 2 head 1 head, 0 head. So, this value is 2, 1, 0 that I will denote it with a small letter, but the random variable is denote it with a capital letter it is a general convention followed in all the books.

So, if for each possible outcome x so for a discrete x f(x)to a when we can say it is a probability mass function or a probability distribution function it satisfies this following conditions what is this first of f of x should be greater or equal to 0, what is f of x? f of x is nothing but f of x is the probability of x so it should be greater or equal to 0 in fact, it should be less or equal to 1 as well.

And what is summation of x f(x)? Or sorry sorry summation of f(x), x summation of f(x)for all values of x will be always 1 probabilities will be always 1 is not it so the probability of getting 2 head probability of getting 1 head probability of getting 0 head all probability put together summation should be always 1. And probability of x = x is nothing but we designated by f(x) so f(x)we can tell it is a probability density function.

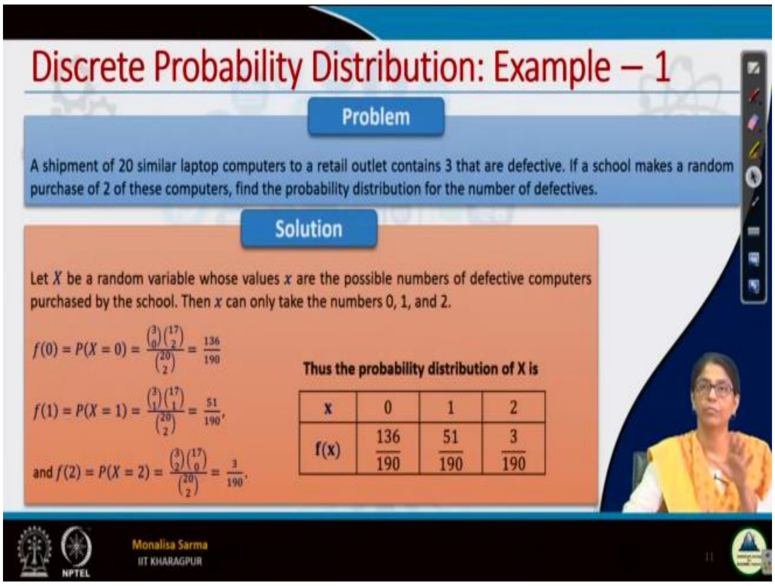
sorry probability distribution function when it satisfies these 2 condition basically f(x)will be greater or equal to 0 and f(x)summation of f(x)would be 1. So here is an example so a coin I have tossed a coin twice the example what just now that I have given, so possible, these are the possible outcomes. So HH, HT, TH and TT. So even if I am considering is the getting heads.

So, what is my getting 0 sorry getting maybe getting T so, getting 0 tail is when I get HH that means I got 0 tail both I got head then getting 1 tail is either HT, TH getting 2 is TT. Now, what is the probability of this says how do we find a probability? So, sample space size is 4. So, I got 0 head how much I got? What is the just only how much is favourable to this event of getting 0 head that is 1. So, what will be my probability? 1 / 4.

So, probability for this will be my 1 1 / 4, now, for 1 getting 1 head, my favourable outcomes are 2 HT and TH so, what it will know what will be my probability I am writing here probability will be 2 1 by 4 and getting 2 head sorry getting 2 tail is just 1 again it is 1 1 by 4. So, this if I can do when in the last slide I told this probability distribution I can define it in the form of a table in the form of a graph in the form of a function.

So, this is if I express this maybe in the form of a function this is what it is shown here. And it is this one this maybe it is in the form of a function, this is in the form of a graph. See and y axis I have given a probability x axis is the different values that random variable take random variables may take the 0 value random variable is with respect to an event always. So that what is my event? Getting a tail. So, the random variable maybe 0, 1 and 2 and this side it is the y axis is the probability and this is if I just explain it with the help of a graph.

(Refer Slide Time: 09:54)



A small example now to understand this concept, what it is given? First go to revision shipment of 20 similar laptop computers to a retail outlet. A shipment of 20 similar laptop computers to a retail outlet contains 3 that are defective out of 23, 3 defective if a school makes a random purchase of 2 of these computers find a probability distribution for the number of defectives.

Say we here it is asked that we have to find the probability distribution for the number of defective if the school has purchased 2 computers, what may be the number of defects in this case? That is purchased 2 computers. So, what may be the number of defectives? The number of defectives may be there may be 0 defective there may be 1 defective there may be 2 defective atmost 2 defective.

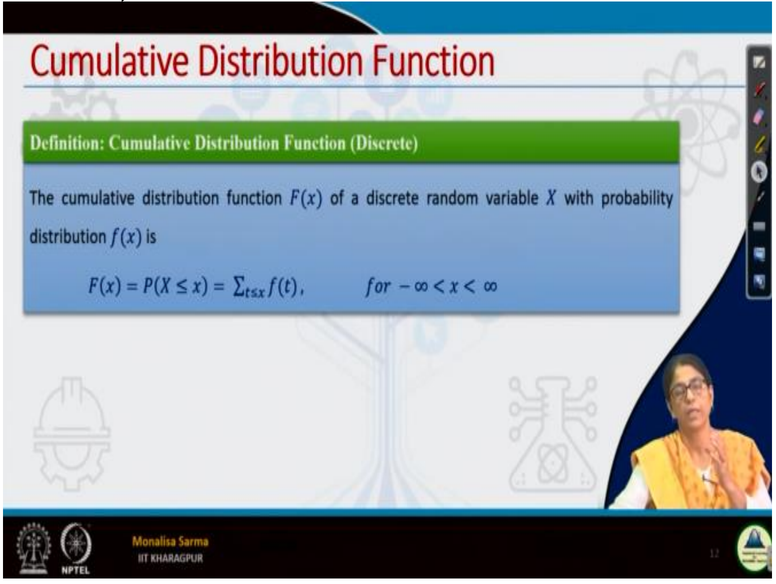
So, number of defective will be 0, 1, 2, that means, the random variable will take the value 0, 1 and 2 now we have to find out what is the probability of 0? What is the probability of 1? What is the probability of 2? This is the probability distribution is not it? So, that is what probability distribution is the value of the random variable all the possible values of the random variable along with this probability of occurrence.

So, how do we find out what is the probability that 0 defective? So, what is the total sample space? Total sample space is the total 20 computers out of 20 I am picking 2. So, it will be 20 C 2 we have seen remember what is where we will use permutation where we will use combination now, here 2, these are indistinguishable is not it? So, it will be 20 C 2 the whole sample space is 20 C 2.

So, and how much what we want to see? 0 defective that means, in 20 there are 3 defectives, how many non defectives are there? Non defects are 17 and we want to find out what is the probability of picking 0 defective that means, the 2 computers that we have purchased that means, we have we got it from the good ones good 17, 17s are the good ones. So, 3 C 0 x 17 C 2 see this one.

Similarly, for 1 same denominator is 20 C 2 and that means, we have 1 defective means from that 3 defective we got 1 and 2 another one we got it from the 17 similarly, 2 from 3 we got 2 and from the 17 we got 0. So, we write in tabular form this is the tabular form in the tabular form the probably distribution of x so, we have seen how we write in a tabular form we have seen how we write it in expression form how we write it in a figure graph.

(Refer Slide Time: 12:44)

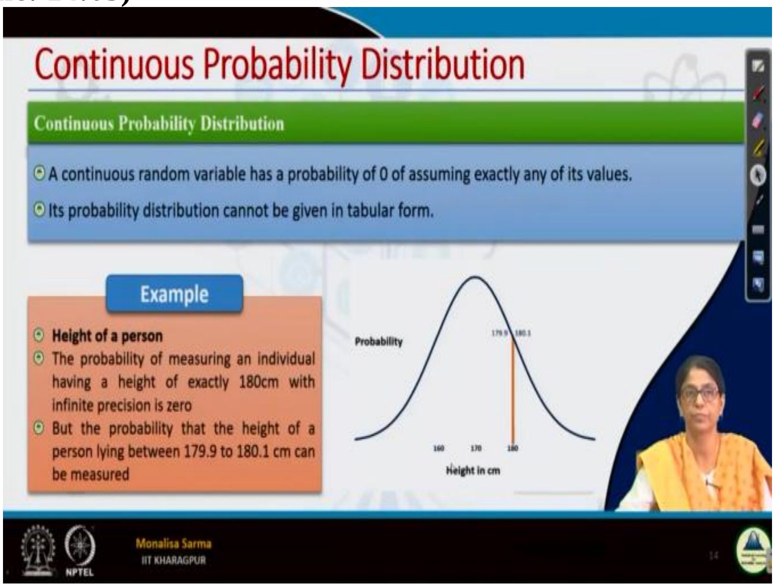


So, now, what is cumulative distribution function? What we have seen here more probability of getting 0 defective, 1 defective, 2 defective we are concerned of a particular number. Now, if I am asking this question of head and tail only 1 question of getting when I am tossing 2 coins what is the probability of getting either 0 tail, 1 tail or 2 tail maximum 2 tail I can get. So, I cannot obviously get 3 tails now, if my question is what is the probability of getting at most 1 tail?

That means, at most 1 tail means there may be 0 tail as a 0 and 1. So, I have to find the probability of occurrence 1, 0 probability of occurrence of 0 and probability of occurrence of 1. So, probability of occurrence of 0 plus probability of occurrence of 1 that will give me at most 1. So, what is that? That is the cumulative distribution function we have seen cumulative frequency distribution remember, so, does the same portion same concept.

So, cumulative distribution function is F of x this probability that x that is the random variable is less than or equal to x is this summation of f(t) where t is range less than equals to x. So, here at most 1 means f (0) + f (1) that is cumulative.

(Refer Slide Time: 14:05)



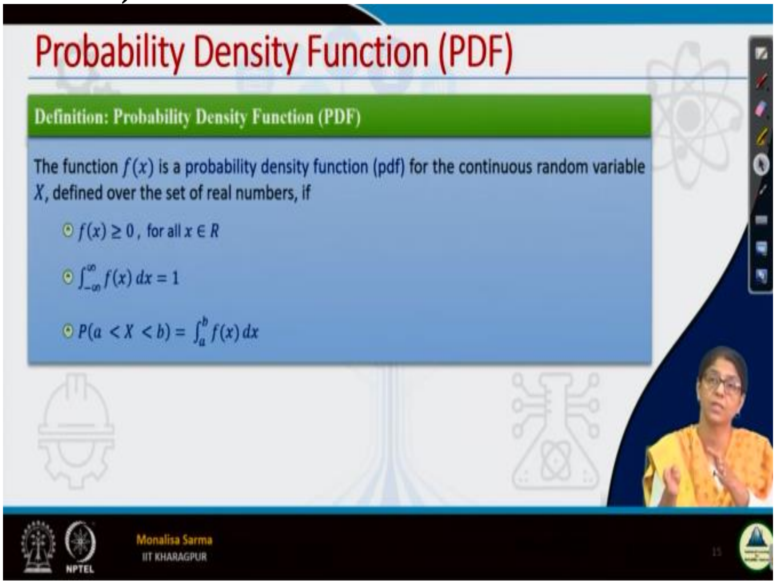
Now, continuous probability distribution, continuous probability distribution when the random variable is continuous and a random variable contains any value in the range than the distribution that we get is a continuous probability distribution that will not be a discrete probability distribution. Now, taking any value in the range and for the when we tell that so, it can be any infinite value any value within the range it can.

So, now like suppose if I asked a height of a person, you can tell this height of a person is a 5.8, but will it be exactly equal to 5.8 or it is equals to 5.800001 or 5.812345 something like that, you it is very difficult to tell exactly exact value it is almost the probability of getting an exact value is almost equal to 0, you cannot it is such a minute scale measurement is not there to measure.

So that is why what we do in case of continuous probability distribution, what we tell is that a continuous random variable has a probability of 0 assume exactly any of its value. And when we talk about continuous probability distribution, we have a probability of 0 assuming exactly of any of its value like in the discrete we are telling what is the probability of 0, 0 head, 1 head, 2 head exactly, but when you are talking about continuous probability distribution.

I have probability of taking exactly a particular value we cannot define so, that is basically 0 and since that is the situation so we cannot write a probability distribution table we cannot represent it in the form of a tabular form, in tabular form we have to give specific value. So, in continuous probability distribution, we can only represent it either using a function or using a graph. So, this example whatever example I told you this now.

(Refer Slide Time: 16:06)



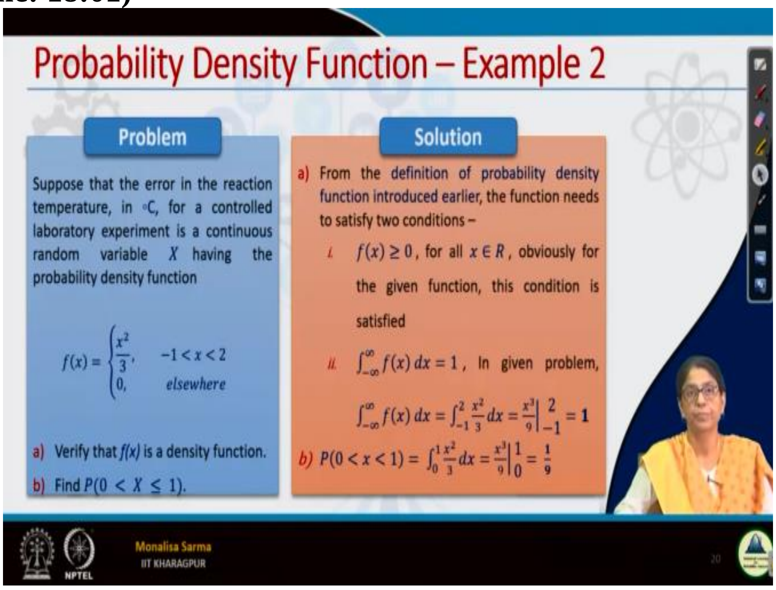
So, now, like probability distribution function what we have seen our probability mass function for discrete random variables similarly, for continuous random variable we call it a probability density function. So, for probability density function the condition it has to satisfy saying like the that is the f of x is greater or equal to 0 for x belongs to any real number and integration of f(x) dx is 1 here it is integration.

Now, here one thing which is very important to remember to note that as in when I am talking of discrete probability distribution, then value of f(x)cannot be greater than 1 because in case of discrete probability distribution, what is f(x)? f(x)is the probability of the occurrence of x , f(x)is nothing but the f(x)is the probability of x so, probability of x cannot be greater than 1. Any probability means that it has to be within 0 and 1.

So, in case of discrete f(x)is the probability of x, but in case of continuous probability distribution f(x)is not the probability of x it is the probability density function its value can be variable it value can be variable greater than 1 as well. So, it is a density function it just in a curve when we draw the curve, it shows how much probability is concentrated at that point near that point.

Basically, f(x) basically will give us the shape of the curve and when we have to find out the area at a particular interval, it will always be the area in a particular interval, so, we will have to find the area so, that is why it is integration. So, remember for this probability density function my effects can be greater than 1. So do not get confused why is if this is probability, why it is greater than 1, it is not probability it is a probability density how much it is concentrated at that point, how much probabilities are concentrated at that point.

(Refer Slide Time: 18:01)



So, now, this is a small example, suppose that error in the reaction temperature is degrees centigrade for a control laboratory experiment is a continuous random variable X having the probability density function, probability density function is given. So, it is given f(x) = x square by 3 for a particular range what is the range? x varies from - 1 to 2 and in other places it is 0. So, we have to verify that f(x) is a density function.

See, when x takes a value very near to 2, then we will get f(x)greater than 1 you see, that is what I have told in the last slide, that is we have already discussed now, we have to verify that f(x) is a density function, how will you verify that f(x) is a density function? To verify that f(x) is a density function it has to satisfy those 2 conditions what we have seen what are the 2 condition?

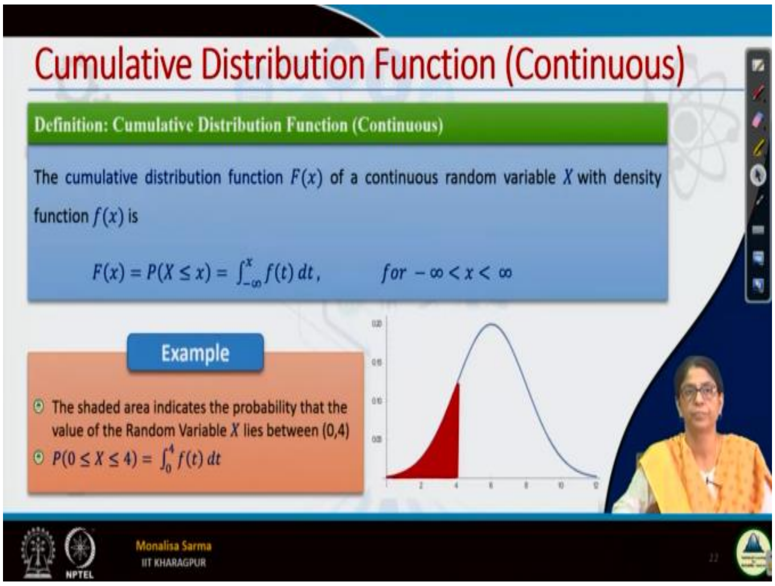
First is f(x) should be greater or equal to 0 and integration of f(x) what the range should be equals to 1. So, returning from this expression, what is given it is obvious that it is greater or equal to 0 from - 1 to 2 it is x^2 1 / 3 square definitely it is not less than 0 it cannot be less than 0 and elsewhere it is 0. So, it is greater than or equal to 0 first question condition is already satisfied second condition from infinity to $-\infty$ to ∞ it should be 1 that is the second condition.

And if we just do the integration, then it will just see the integration it is given here, then we will get 1 so, it satisfies both condition then we can say f(x) is a density function. Next question, we have to find out the probability between x of 0 to 1 nothing just f(x) integration from 0 to 1. So, what is f(x) ? f(x) is x^2 1 / 3 we have to just do the integration from 0 to 1. So, see, gradually we will be getting many such problems with your knowledge integration and differentiation will be necessary.

So, please you have already learned it in your school level and high secondary level may or may not be in school level of high secondary will definitely please brush up your knowledge on differentiation and integration for many problems like here we have done this a simple integration of course, all of you will remember that, but if any of your forgotten please brush up that so else it will be difficult for you to understand the problem.

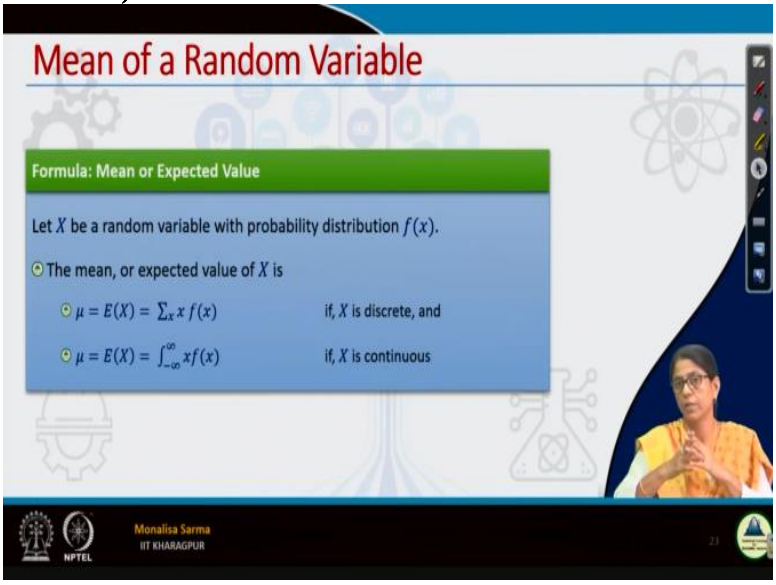
And as I have promised in my first introductory lecture that I will be taking lots of examples see for all each concept I will explain and accordingly I will give 1 example.

(Refer Slide Time: 20:38)



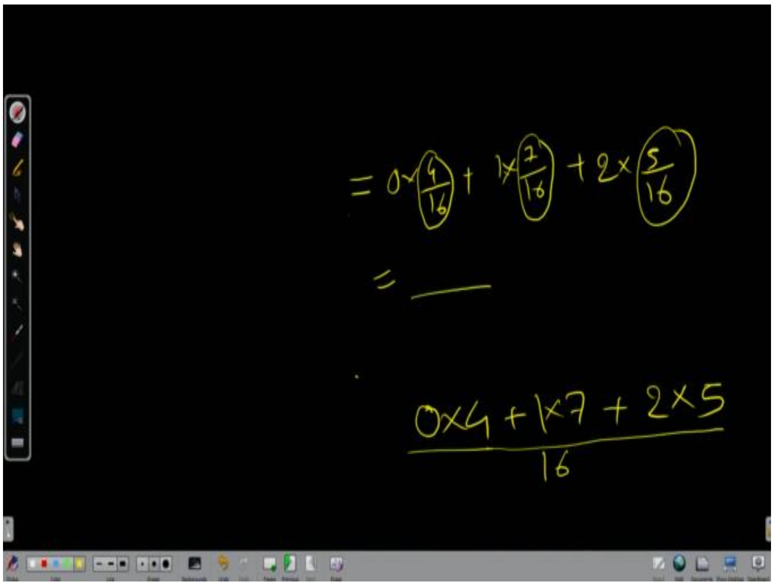
Now, what we have discussed for cumulative distribution function for discrete case same is the case cumulative distribution function for continuous case cumulative means from till x whatever x we specify from minus infinity to till x that is the cumulative distribution function same the concept of discrete.

(Refer Slide Time: 21:06)



Now, how do we find the mean of a random variable we have seen how to find out the mean of a set of numbers. Now, instead of taking let me give you another example suppose, I have 2 coins and suppose I have toss the let me take the graph paper I have toss 2 coins.

(Refer Slide Time: 21:32)



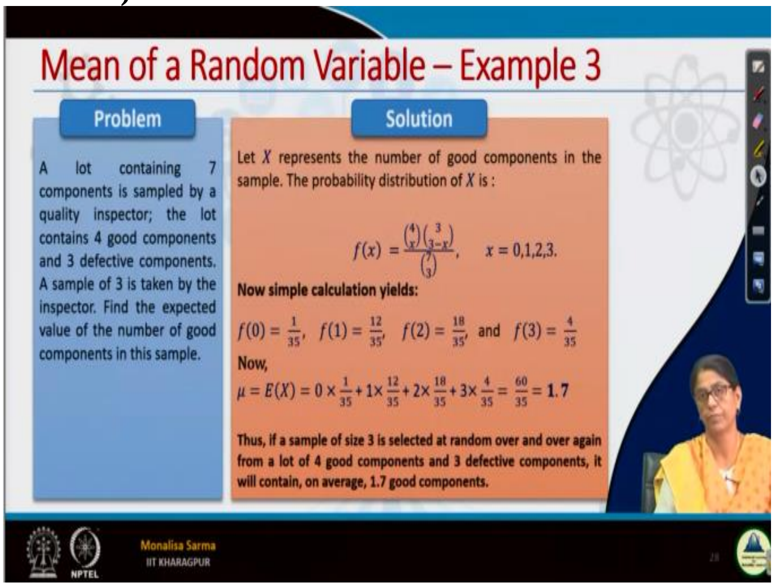
Suppose I have toss 2 coins 16 times and 4 times when they say 4 times I got no head only tail, tail, tail, tail, tail, tail. And suppose 7 times I got 1 head and suppose 16 so it is it will be 5, 5 times I got 2 head. So now what is the average number of heads per toss mean is average, is not it? So what is the average number of head per toss? I am tossing 2 coins, so what is the average number of heads per toss? Average number of head how will you find out?

Per toss is so 4, is not it? So this 4 I got 0 head 0 x 4 + 1 x 7 + 2 x 5 divided by how many? 16. This, of course, let me rub it, this thing I am writing in top this can I not write it, 0 x 4 1 / 16 + 1 x 7 1 / 16 + 2 x 5 1 / 16. And out of 16 I got 4 times 0 out of 16 I got 7 times 0, 7 times 1 head out of 16 I got 5 times 2 head what is this? This is relative frequency distribution relative frequency distribution is nothing but probability we have discussed it remember in our second lecture, so this same expression, I can write it this way. So what if I calculate it, I will be getting something whatever you find, calculate it and find it out. So this is the average of that so that is how we find a mean. Similarly, so, what how we found out the mean of what that means, the problem what are the different values that a random variable takes as well as the probability of that value 0 into what was that 0 x 4 / 16, 4 times we got 0 head.

So 0 x 4 / 16, 0 is a value of the random variable and 4 x 1 / 16 is a probability that is the relative frequency that I got 0 here. So what is 4 x 1 / 16 is nothing but a probability. So, similarly, so when I am interested in the mean or a mean is also called the expected value, it is also termed as E mean of a random variable it is designated as E is described as μ sometimes we call it as expected value, it has written it as mean it is also written as E.

So what is μ or E of X? It is nothing but x summation of x f(x) what we have just now we have seen, but now you have seen remember this one, this is 0, 1, 2 these are what? x And what is this? This is f(x)summation of x f(x). This is if x is discrete summation of x f(x). If x is continuous, then just integration of x f(x) is the same thing. So, that is how we find the mean of an expected value.

(Refer Slide Time: 25:02)



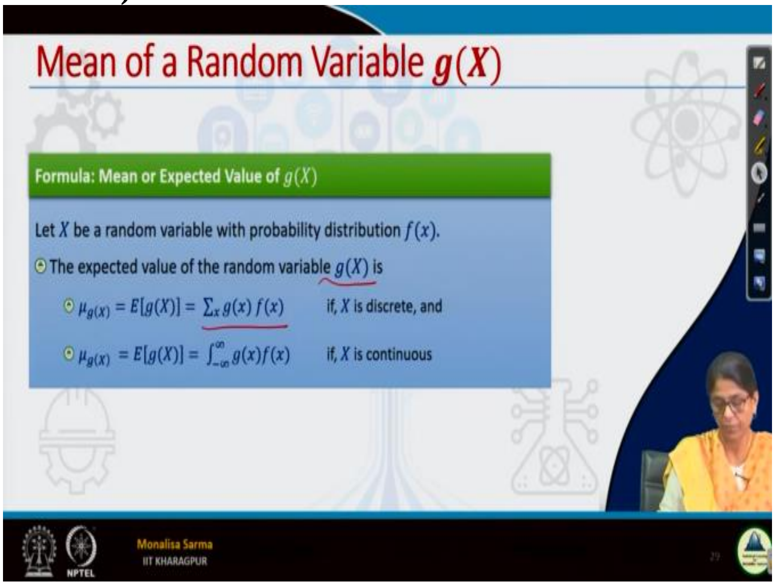
So it is a simple example is there a lot containing sample 7 components is sampled by a quality inspector and the lot contains 4 good components and 3 defective components they have total 7 components out of them 4 is good and 3 is defective a sample of 3 is taken by the inspector find the expected value of the number of good component in the sample see, X represents the number of good components assembled.

So, we will have to find out the probability distribution of X how many good components or what the value what value that a random variable X can take then a variable where we are interested in number of good components. So, how many good components are there? There are 4 good components, but we have taken 3 so, the random variable can take either 0 good component 0, 1, 2, 3 is not it?

So, basically, x can take the values you do 1, 2, 3 see here x can take. So, what is f of x? f of x is the 7 is this one, 7 C 3 is the total sample space and out of 7 how many good components? Good component is 4 we are interested in finding a good component. So, we are taking all the x components from 4 and remaining how many are remaining? 7 - 4 is 3. And from this 3 how many were picking total we are taking 3 components out of 3.

This x is the number of good components and from this 3 we are picking the other component that is 3 - x. So this is the function called the distribution function. So we can calculate this so we found out for f0, f1, f2, f3, we can calculate it. So now we need to find out the expected value expected value is summation of x f(x). So, this X is 0 into 1 1 / 35 + 1 x 12 1 / 35, 2 x 18 1 / 35 + 3 x 4 1 / 35, just put it next question and get developed you can check it again. So, on an average, it will contain 1.7 good components.

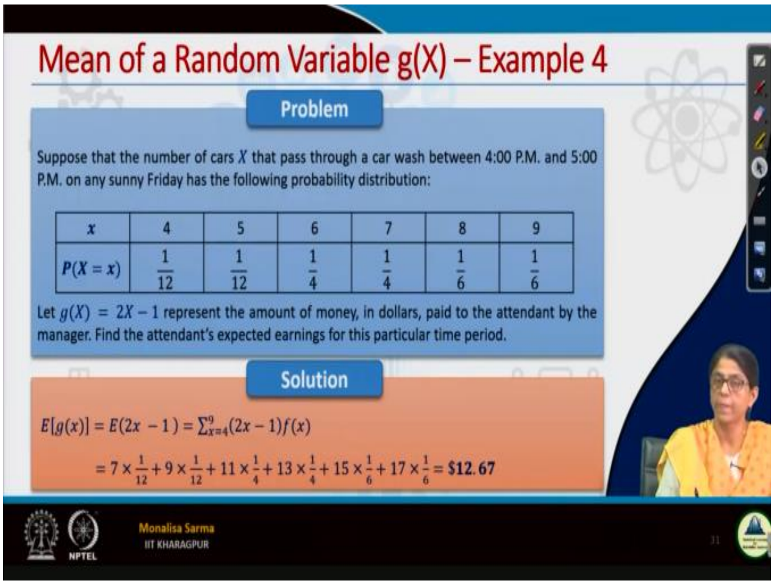
(Refer Slide Time: 27:16)



Now, mean our expected value of g(x) if we are interested in finding the expected value of a function of x, we have seen how to find out when X is a random variable, we have seen how to find the expected value of expected value of x is nothing but summation of x f(x)if it is discrete. Now, we are interested in finding out the expected value of a function of x, say x, maybe 3x x maybe x^2, x^3, or 3x + 2 anything.

So if you are interested in finding out the expected value of a function of x with a function of x, I am terming it as a g(x) here. So it is same, there is just summation of g(x) instead of x. I am writing g(x) that is, that is the difference.

(Refer Slide Time: 28:02)

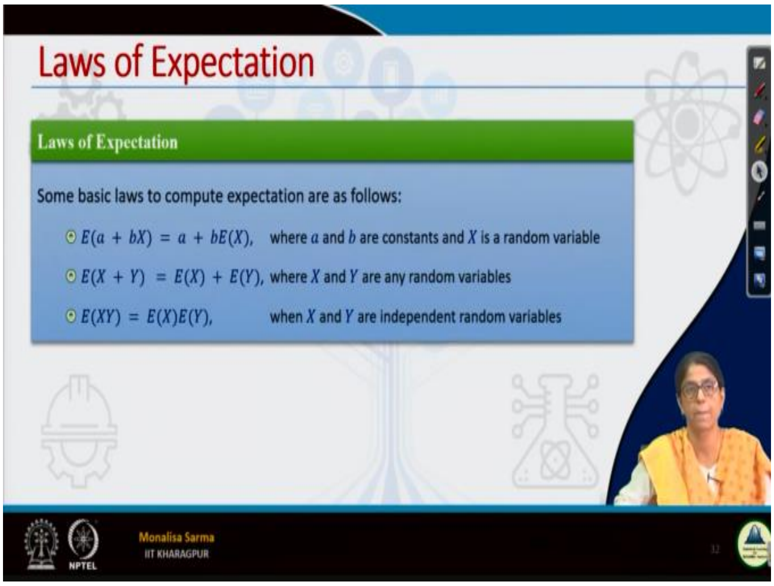


So, using this, this one simple example is there, it is very simple, you will be able to do it yourself only, just the weekly I will go through it, suppose that the n number of cars just pass through a car between 4pm and 5pm on a sunny Friday as the following probability distribution value added a random variable can take is 4, 5, 6, 7, 8, 9 and this is the second row is the probability of this particular thing.

So, now, we have to find out the attendant's expected earning g(x) = twice x - 1 represent the amount of money in dollars paid to the attendant by the manager. So, this represent amount of money in dollars paid to the attendant by the manager. So, we have to find out the attendant's expected earning. So that means the average earning average earning how will I find so this is g(x) f(x), so g(x) 2x - 1 is g(x). So, what x will take what value?

x will take 4 first actual day 5x will take 6, 7, 8 and 9 and this into and into f(x) whenever x is taking value for we will multiply it by 1 / 12 and by x is taking value 5 we will take multiply it by 1 / 12, x is taking 6 will multiply by 1 / 4 likewise see, just we have substitute different values for X and accordingly whichever values we have taken for x, that probability will be considering that is all.

(Refer Slide Time: 29:35)



So now there are simple laws of expectation some basic laws we have. So what is the expected value of a + bX? Always remember the expected value of a constant is the value itself the constant itself. So here expected value of a + bX expected value of a is a itself. Expected value of 2 is 2, expected value of 10 is 10. So, there is no error we cannot find an expected value that is average value of a constant number.

So, here it is E of a + bX so, obviously, it will be a + b into E of X. So clearly expected value of X + Y if 2 random variable X +Y what is the very interested in finding out the expected value so, it will be expected value of X + expected value of Y, expected value of E of XY = E of X x E of Y. However, this is applicable when X and Y are independent random variable. So, now I think I will stop this lecture here.

And though initially I have to told that I will be covering variance also and Chebyshev’s theorem also in this lecture but time may not permit. So, I am stopping this lecture here, in the next lecture I will start from the variance of the random variable. Thank you.