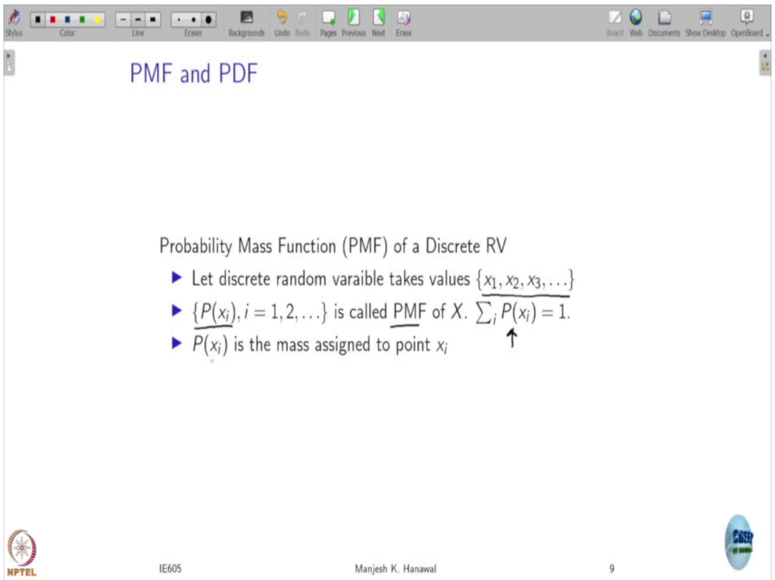
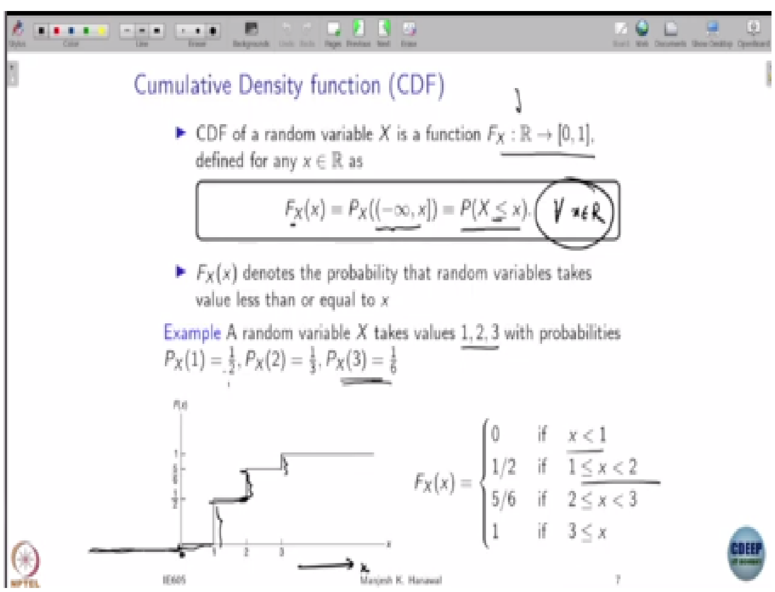
Engineering Statistics
Professor. Manjesh Hanawal
Department of Industrial Engineering and Operation Research
Indian Institute of Technology, Bombay
Lecture 07
Discrete and Continuous Random Variables - I

(Refer Slide Time: 00:14)

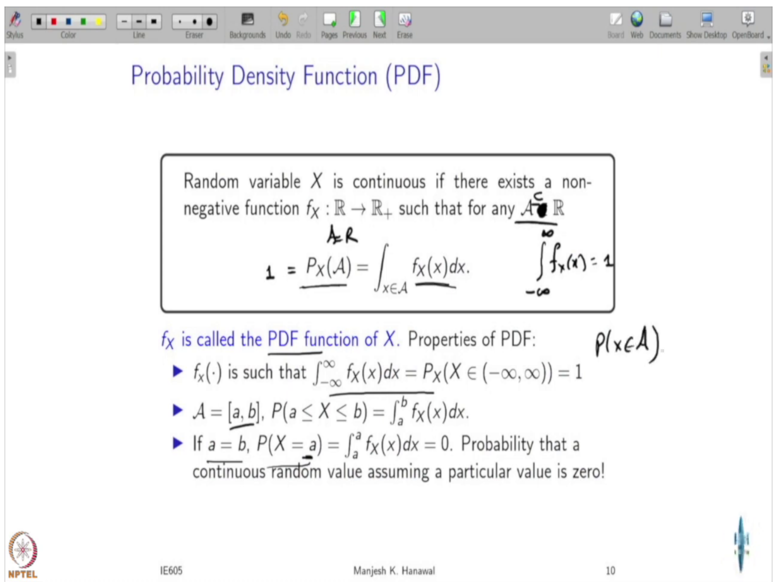




Now, we will talk about probability mass functions and probability density functions. Probability mass function is for discrete random variable and probability density function, we going to talk about continuous random variables. So, let us say my random variable x is discrete and it is take some discrete values x1, x2 like x3 like that. And the probabilities at these discrete points are going to be called a probability mass function. And naturally, the probabilities of all these values should add up to 1.

And what we are going to call is p(xi) we are going to call it as the mass assigned at point xi. And if you know recollect, masses corresponds to the jumps amount of the jumps at on this CDF curve.

(Refer Slide Time: 01:10)



Now, probability density function, so probability density function, we are going to talk about for the continuous case. And in this case, the way we will define probability density function is, if my probability of x on subset A, I can represent it as an integration of some function f(x). And this is true for any subset, then I am going to call my random variables to be continuous. And this function f(x) I am going to call it as a probability density function.

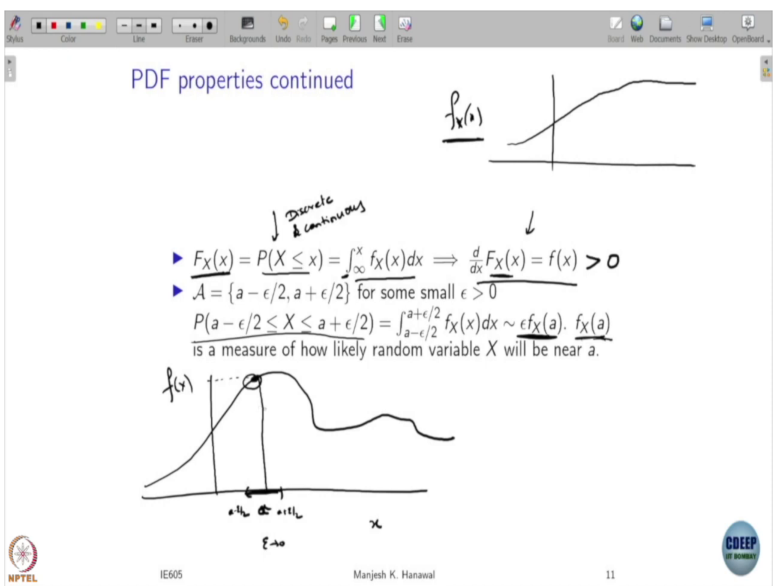
So, now let us look into some aspects. Now suppose this should have been A is a in this case A is a subset, A is a set. So, I should have A can we should have written this, $A \subseteq R$ here, $A \subseteq R$. Now, and this could be equal also. Now, let me see what happens if I take A to be R itself. I am basically saying probability that A on the real line, what is this value in that case, if I take A to be R real line enter real line. It is going to be 1, I am basically letting random variable to take any possible value.

So this left hand side is 1 that is the property of my probability. So now, this is saying in that case, what I am doing? I am integrating this function over the entire real line. So, this is like, basically, I am integrating it in between minus infinity and infinity. And then it is saying that this has to be 1. So, what it is basically saying is, if something is a PDF, the first natural condition, it has to satisfy if you integrate it, it has to equal to 1. So, what does this mean? The area under my PDF is going to be 1.

And now if I have to set a to b between some finite interval a and b, all I need to do is integrate that function between a and b. Now, if I set a = b now what I am basically saying is, basically I am saying what is the probability that X equals to a and by definition, if I integrate this function between a and a what I am going to get 0.

So, whenever we have a continuous random variable we will not ask the question what is the probability that it is going to take a particular value X equals to a because that value is 0. Whenever we have a, but this is fine, if it is a discrete random variable, there it corresponds to its mass. If it has a mass at that point, it will be that mass but in the continuous case, the, what you are saying is mass at each point is going to be 0. But the mass if I look into some region, that may be positive, but at a particular point, it is going to be 0.

(Refer Slide Time: 04:44)



Now, what is the meaning of the PDF? What does PDF is saying? So, in the PMF, it was clear in the discrete case, when we talked about probability mass function, we know that PMF at a particular point is basically how much is the mass there. But, if I just say, P(x) is equals to this, what does this mean? Let us say I have a function x, and this is my let us say I have some function like this, what does now I have some particular value x. What does this mean? What is the value? What does the value indicate? Does it indicate the probability at this point?

Student: No.

Professor Manjesh Hanawal: No, we said probability is 0 at that point, but then what it is indicating

Students: (())(05:41)

Professor Manjesh Hanawal: It is kind of giving some rate of change of the mass in that neighborhood. Let us try to understand that. So, before that we know that CDF is simply probability $X \leq x$ and here X can be discrete or continuous. So, we can always define discrete or continuous. So, we can always define PDF for a continuous as well as discrete function and we said that PMF we defined for the continuous and PDF for that, PMF for the discrete and PDF for the continuous case.

Now, by definition if it is a continuous random variable, this should be minus infinity to X here and by our definition of probability, sorry, differentiation function the relation between PDF and CDF is this. Because F(x) is nothing but integration of my small f(x) between minus infinity to x. So, if you just differentiate it, you will get this differentiation. So, that is what like one way to interpret is F(x) at point x is the rate of change of my cumulative density function at that point x. And is does this indicate that it is a positive quantity F(x) at any point?

Students: Yes.

Professor Manjesh Hanawal: Why is that?

Student: Integration of positive values is positive.

Professor Manjesh Hanawal: Is not decreasing. And differentiation may not always make sense for the discrete random variables. As long as this function is differentiable, this is fine. If my function I cannot differentiate, I cannot define this at that point. So, that is why remember that I am talking about now this continuous functions, this is not what I am talking about is a discrete function this is for the continuous and naturally it is telling that if such a probability density function exist, then your cumulative density function for that continuous random variable is differentiable at every point. And then how, so does differentiability implies continuity?

Students: Yes.

Professor Manjesh Hanawal; Yes that means it is saying that if my continuous random variable is.

Student: (())(08:30)

Professor Manjesh Hanawal: If my if my continuous random variable has certain PDF, it is already saying that my PDF, CDF has to be something like this. It has to eventually go and one end saturate and it has to be increasing. But at every point there cannot be jumps. It has to be continuous and differentiable at every part. Now, let us go back and apply our definition.

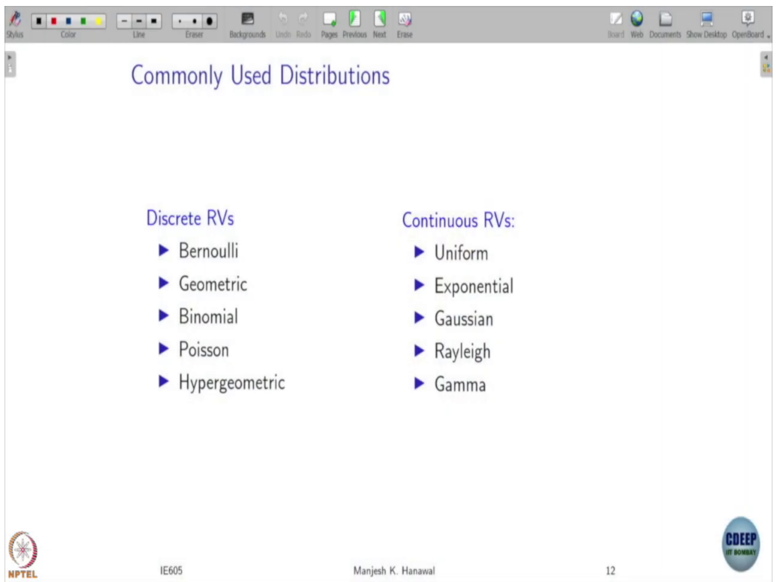
Let us take a small interval around the point a, I am writing it as a let us take this point a and take a small point around this, this is a plus epsilon by 2 and this is a minus epsilon by 2. I am now taking it in the small neighborhood of a and now I am asking the question, what is the probability that my x takes value in this range? By definition, this is nothing but integration of your f(x) between a minus epsilon by 2 and a plus epsilon by 2.

Now, if this epsilon is very close, very short, like almost epsilon is almost tending to 0. You are integrating in a very, very narrow region and in that region, you can approximately assume that your f(x) remains constant. In the small region when you are trying to integrate it, in that you can approximately assume that is a constant and then this probability is nothing but epsilon into f(x) into a.

So, it is basically saying that probability in that region is epsilon times that quantity that value given by this PDF in a way this is telling you that this is the weight or like it is the kind of rate with which the probability is changing in the neighborhood of your epsilon and that is why you can, even the rate when that epsilon is very small this probability we can simply take it as a product of epsilon and the value of your PDF at that point. And this is all argument true only when x is that interval you are looking at small. If it is interval is large, I mean this is does not make sense.

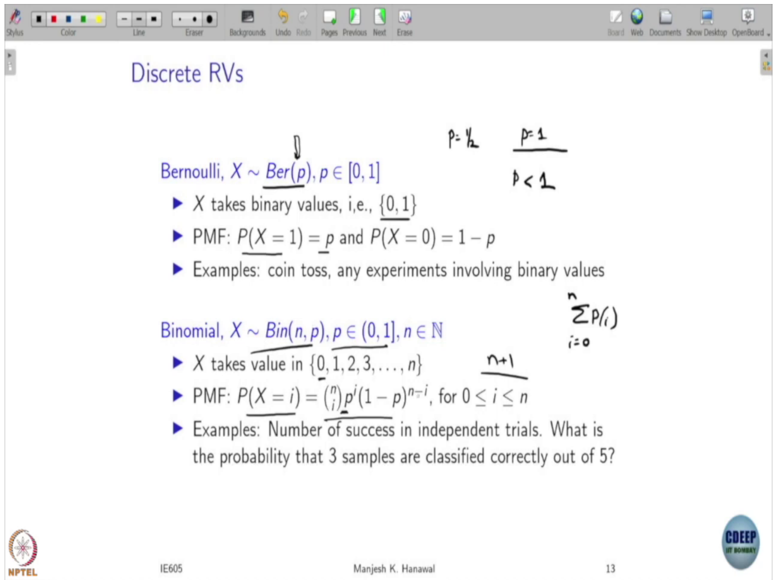
So, in a way, what we can interpret is therefore, f(x) at point a is a measure of how likely random variable x will be near point a, I mean, when you conduct an experiment, what is the likelihood that outcome is near my point a. It is not exactly saying the probability of getting a, it is saying that your outcome is going to be in the neighborhood of a, what probability the small neighborhood. With this understanding of what a CDF what is PDF, what is continuous random variable what is discrete random variable and those their properties, now we will study about some commonly used distributions which I have listed here.

(Refer Slide Time: 11:57)



So, in IE621 these are already covered. Now, let us today start talking about this continuous random variables. We are going to focus on some standard discrete random variables and we are going to focus on them because when we want to model something. It is better that you have some distributions which you know better and these are the some standard discrete random variables about which we know better. And, it is also like these random variables are comes pretty handy in studying various applications. Let us look into them.

(Refer Slide Time: 12:39)



Bernoulli random variable X. It is denoted by Bernoulli and it comes with a parameter p. So, whenever I talk about random variable, I need to talk about what are the possible outcomes. It outcomes are 0, 1. And now I have to give probabilities to them. I am going to say that probability that X is taking value 1 is p and X is taking value 0 is 1 minus p and that p is the success parameter p and this p can be anything between 0, 1. That is why this Bernoulli is a parameterized random variable with this parameter p.

Now, where does this random variable can be useful? It could be useful to model your coin toss, when head comes you call it tail, why you want to call it tail instead of that you simply call it is 1 and when tail comes you call it a 0. Now depending on your application, you may call other way also tail you call 1 and head as 0.

And now depending on the bias of the coin head is going to come with probability p and tail is going to probability 1 minus p and all the time we are talking about a fair coin for which p equals to half. But why p equals to has to half p can be anything. Like p can be 1 in which case your are in a Sholay movie or like a p can be strictly less than 1.

And can you think of any other example instead of just a coin toss? It could be anything in many any experiments it will be just interested in whether something happened or not. Whether I passed or failed or like I won or loss, you put money in a lottery and you will be interested in just knowing 1 and all there you can use this Bernoulli random variable wherever the outcome is just binary.

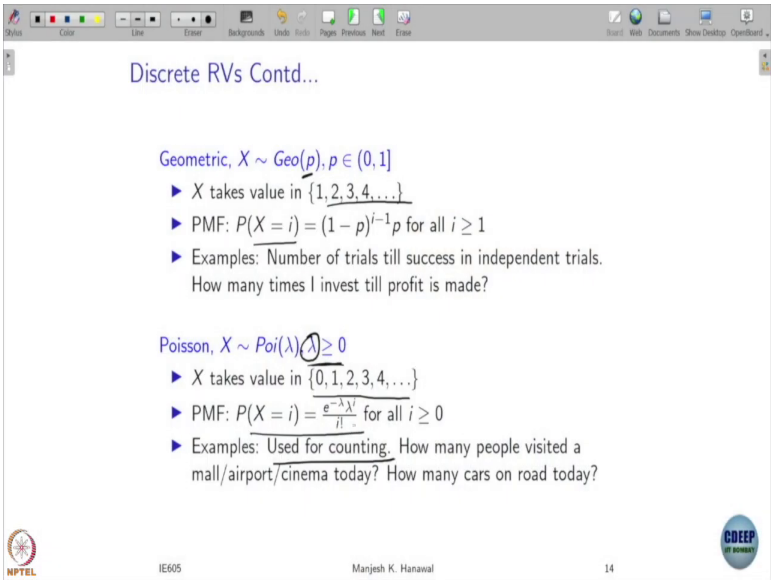
Like this if this is pretty much used in machine learning nowadays. You will be interested in a classification. You if you were given a photo you need to tell whether in that photo dog is there are not. Only 2 outcomes yes or no, there you can use these Bernoulli random variable. Next is binomial, binomial comes with 2 parts n and p, where n is p is again between 0 1, but n is an integer. And X takes here value between 0 1 2 all the way up to n. So, how many possibilities n plus 1 possibilities are there.

Now, I have to assign probabilities to them probability that X is going to i is n choose i p to the power i 1 minus p to the power n minus i, Why i did like this? We will see later, but right now take it, but once I say this, it should not be arbitrary PMF. Arbitrary values when I say PMF they should add up to 1. Do they add up to 1? They do, check. So, now, where this is going to be useful?

This is going to be useful when you are going to do like counting in successive iterations when you are repeating something and then you want to count. Suppose let us say you have coin you have thrown n times. Now, you want to ask how many times head happened? In all those n trials head may never happened in that case it is 0. It may happened 1 or it may happen n times. And you may want to ask the question, out of this e trails how many times 3 times head happened or 5 times head happened like that.

So, then you can use this and naturally if you little bit carefully look into this what you saying is if are interested i like number of heads is i. It is saying that the different ways 3 can come in n choose i, out of n trials, 3 heads can come anywhere. In how many different possible ways they can come that is n choose i and whenever they came p corresponds to success 1 minus p corresponds to failure. So, you are going to multiply and it will comes in that way.

(Refer Slide Time: 17:22)



Geometric distribution another one which comes with parameter p and this is like you can think of like till the first head appears and that can take value head can happen in one second or three all the way. So, this is takes all the values 1 to n and probability that it is going to take value i is 1 minus p to the power i minus 1. So, that is the example I gave, number of trial head happens. Geometric and another one is Poisson.

Poisson is also is over infinite sorry accountably all numbers including 0 and now the probability that x equals to i is given by e to the power minus lambda lambda to the power i by i factorial and here lambda is a quantity which is greater than or equal to 0. And again if you sum this, we are going to see it as 1. So, Poisson comes with a parameter lambda. And geometry comes with a parameter P and where is Poisson use?

Poisson use mostly used in counting. Like how many how many buses crossed IIT main gate today? Or how many cars passed? Or how many people entered IIT? There, like you may want to count and where the count value can be anywhere between 0 12 3 4 like this and there you may want to use such a Poisson model and this rate is going to kind of control in that rate is going to describe these probabilities. So, these are the main four we will talk about and we will stop here.