Introduction to Stochastic Processes
Professor Manjesh Hanawal
Industrial Engineering and Operations Research
Indian Institute of Technology Bombay
Lecture 07
Discrete and Continuous random variables

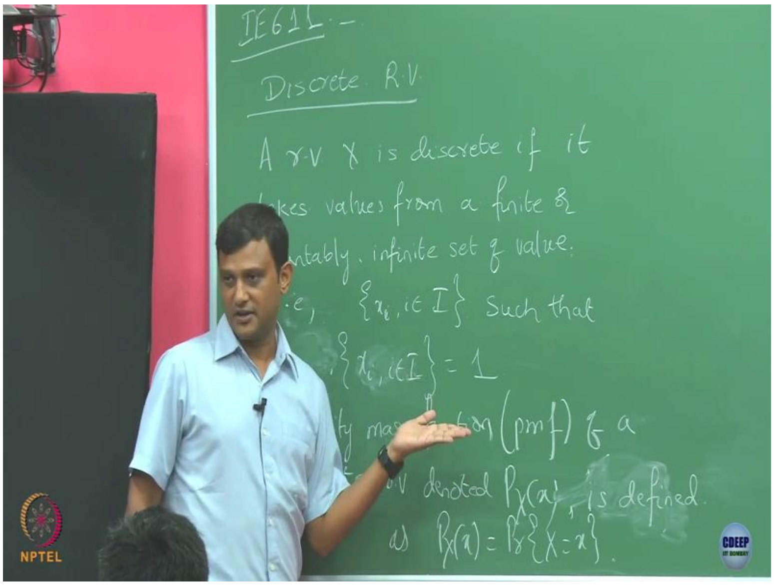
Let us start. What we did in the last class is we said what is the CDF of a random variable and then we defined what are the properties the CDF satisfies. So we gave a statement which was if and only if, that is if there is a CDF for a random variable, it has to satisfy certain properties, and if you come up with a function which has this properties there will be an associated random variable for which this function is going to be CDF.

What we are going to do today is we will build on whatever the random variables we try to characterize some more properties of a random variable using this CDF's another PDF that I am going to introduce in the class. So before that, so far what all the examples I gave and what all the way I defined my CDF, it was for a random variable that took finite number of outcomes.

For example in the coin toss it was just heads or tails which we mapped to 1 or 2 and then in the case of dice it was 1, 2, 3, 4, 5, 6. So like that, like the outcomes was only finite.So it is not necessary that the random variable x has to take finite outcomes. It could be taking continuous of outcomes, right?

For example height of the population in India; it could be like any value, you say some number; the possible values are something between 2 feet to 8 feet any value is possible. So we try to make this continuous and discrete random variable notion bit more formal today.

(Refer Slide Time: 02:17)



So we are going to say a random variable x is discrete if either it takes finite set of values or countably infinite set of values. So for example let us have index set I, I. This could be simply; this index set could be simply 1, 2, 3, 4, 5, 6.

That is, then it is going to take 6 different values of random variable. Or this i could be just consist 1 or 2, in this way, in this case my random variable takes only two possible values, x1 and x2. Or it may happen that this i is countably infinite.

For example, this I could be all possible integers, positive integers, 1, 2, 3, 4, 5, 6 like that and I have that many possible outcomes that my random variable can take. For example, if you want to model something, let us say number of atoms that you are going to see in some experiments, I mean they could be uncountably many. Like there is no bound on how many atoms you could...

So you could be like indexing them, 1 atom, 2 atom like that, that many. Butyou are still able to index them. And you can enumerate them. So that is why we are going to… so you understand the meaning of countably infinite, that means basically you are able to enumerate all the possible outcomes.

So whenever it is the case that I have finite values, that finite outcomes my random variable takes, or it can take countably infinite set of values then we are going to call it as discrete random variables

So in this case I know that there are only certain values that are going to take, that are going to be realized when I conduct my random variable. And then the probability will be only associated with these points, right, all the points this outcome can take. So any other values there will be no associated probability because that values are never going to arise.

So when we have such a discrete random variable, the more interesting, instead of looking at directly the CDF, may be you just want to look at something called probability mass function.

Student: (())(6:58)

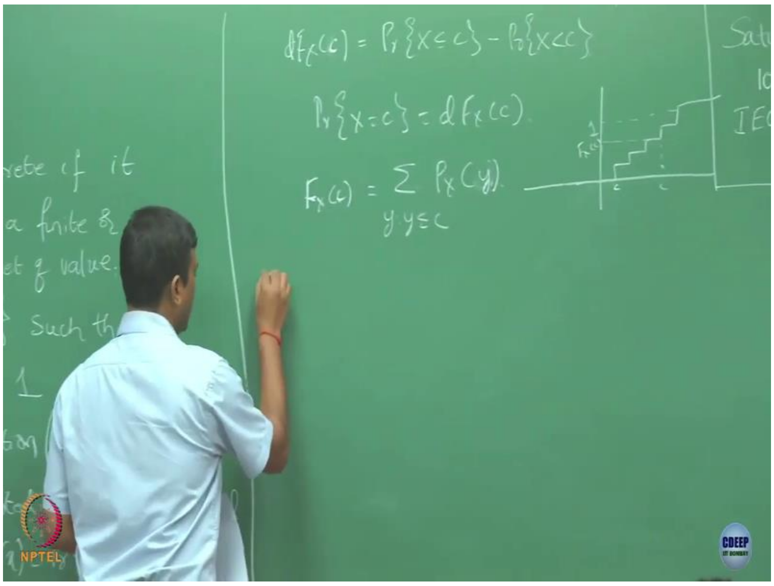
Professor: Yes? Right, not clearly because see I already said this is for all I coming from this set I, So if I don't have this and if I write like this… then write, I need to include this, but the way earlier I have written is it is just like my x i, I am taking i from this.

Now this means this is already including all possible values that my random variable is taking. That is why this probability is already 1. So that is simply taking probability of the points at which, which are being taken by this random variable x.

So what we are basically doing is I am going to call something probability mass function which is going to just give me probability of the points that my random variable is taking, right. So here maybe I should write, I am just, so let us say I am going to take my random, my x takes only 5 outcomes.

So if I give probability for each of the possible 5 values then I am going to take it as probability mass function. And it could be like countably infinite also but at each of the points if I have the corresponding probability that is going to define the probability mass function of that random variable. So is there a relation between this probability mass function and CDF then.

(Refer Slide Time: 09:28)



So we have defined something called d of f of x, f of x, right? We defined in the last class. What did we say? So this was like, I mean we just said that this is nothing but probability that x is less than or equal to c minus x less than c.

We also defined what we mean by probability that x strictly less than c,. Thendo you see any relation between probability mass function and this d of f of x, see? They are going to be same, this is just like the, probability that …

So I have now expressed this. I have expressed my probability mass function in terms of my CDF function. So do you think I can express my CDF in terms of the probability mass function? How is that? So f of x of c, how should be d?

Student: (())(11:06).

Professor: It is correct? So look for all the points which are less than or equals to c and thus add them so then this will result in the cumulative probability till the point c, so this is basically cumulation, so I recall that for discrete... yes we had come up with one example where my CDF look like, right? Like this value being 1, we said that this jump corresponds to, let us say whatever this point; let me call this as c. This jump here corresponds to that d f of x.

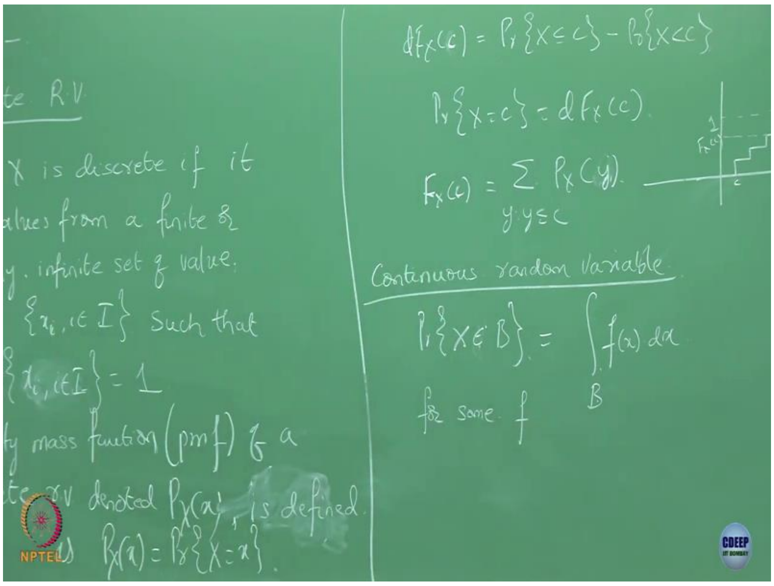
Now we are exactly calling that jump for all the points that wherever it is happening as the probability mass function. And then you know that this function is nothing but the accumulation of all these points till any, suppose if I take this as c, so then I get f of x c by accumulate all these jumps and that is what this denotes.

So the way we interpret, so even though I have written as the summation, summation means always we are adding finitely many terms, but here I am writing it as like y less than or equals to c. That means there are uncountably many ys. The way to interpret is all those points where, all those ys where this guy has positive value, we are only summing over them.

Fine so most of the time when we have a discrete random variable, may be either, you give me CDF or probability mass function they have kind of same amount of information. And most of the time when we are dealing with probability mass function we will just deal with the, so whenever we are dealing with a discrete random variable we just deal with the probability mass function.

That means it is basically saying that, you have these many discrete points that are possible outcomes, probability mass function just tells you what is that probability that each of these points can be taken by my random variable. That is it, now let us move on to continuous random variable.

(Refer Slide Time: 13:47)

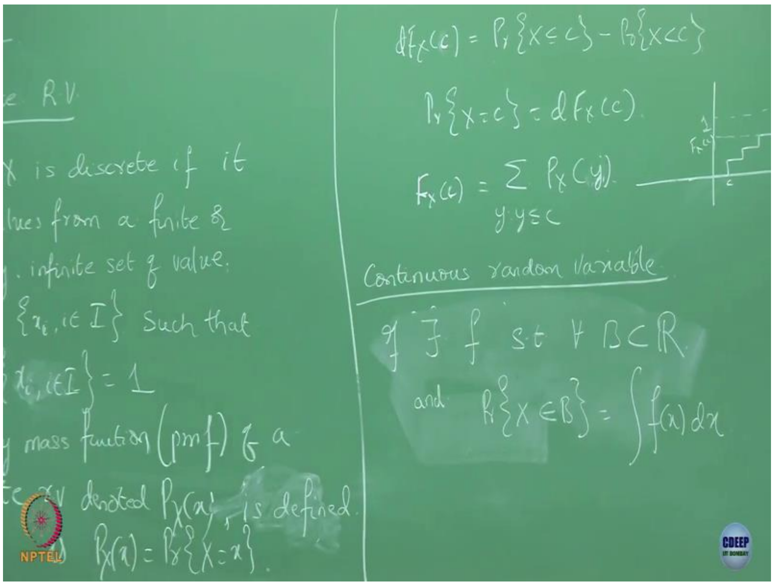


So to make this notion of CDF or to make the notion that what is that the probability of a continuous random variable that is falling in some interval? So let us say I am interested in finding... so what this means?

Suppose, so B is some set. B is some subset of my real numbers. Now if x is the continuous random variable I expect it could be taking any positive number, right. Like it is not like it can only take some finite number or countably finite number.

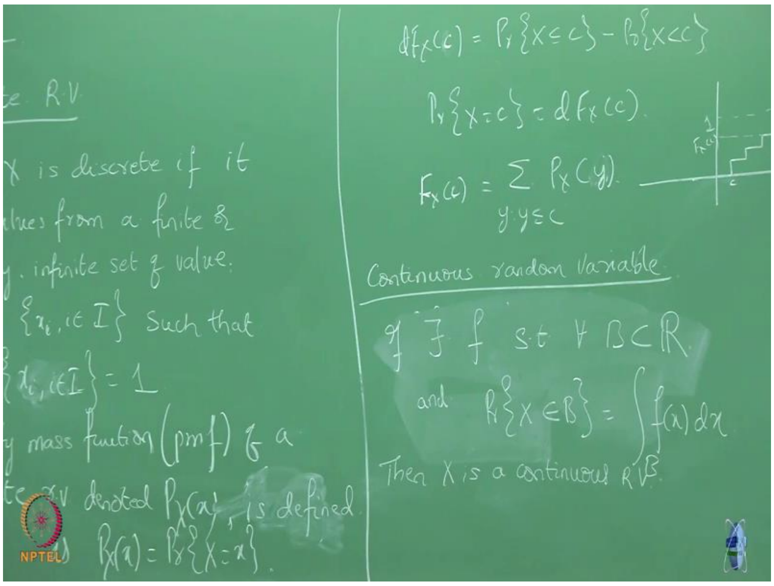
But now let us say this B is some interval or whatever set. Now I want to ask this question that what is the probability that x belongs to that set, B? Now we are going to say that. May be let me write it in this fashion. Yes...(())(15:44) B which is like subset of R .

(Refer Slide Time: 15:39)



All of you understand what I mean by this? If I am able to express whatever this probability in this fashion then I am going to call my random variable x as continuous.

(Refer Slide Time: 16:44)



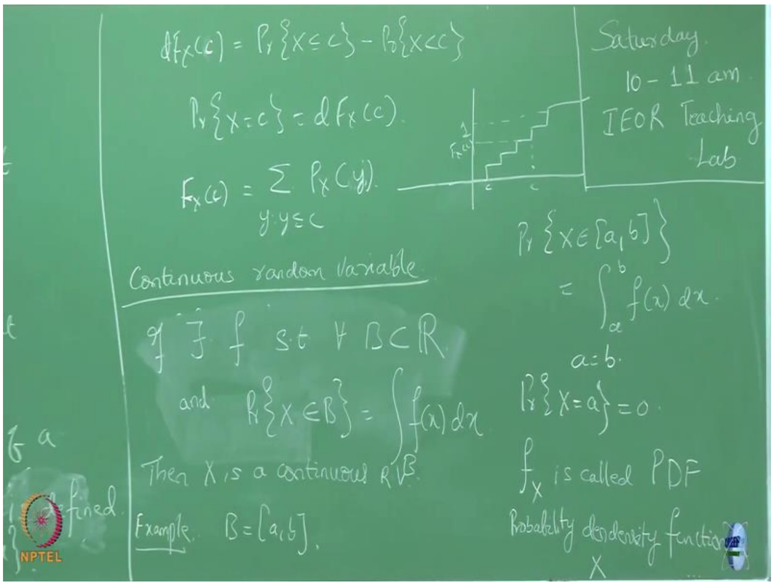
So, let us ponder on what we are trying to say here. Earlier we know that like if I want to find… so in the discrete case when we have a, like if I want to find what is the probability that my x takes some value, you just try to add the probabilities of all the possible values that your x can potentially take.

But now there you are able to add because x only take at most like finite number of values or countably many. But now you do not have that. What it is, now you have to deal with x taking continuous set of values or uncountably many. So now you have to define they are through integration.

So that is where this integration is coming. And now you are saying that if there exists such a function, if there exists such a f such that this holds for any B, you are going to ask the question, okay whether my x belongs to B or whatever B you like. It is not like you are going to ask it for any specific B. If this hold for any B that is a subset of R then you are going to say that your x is continuous.

So do not worry of this, this is just like formal definition. But what we mean by continuous random variable is something that takes all continuous set of values then.

(Refer Slide Time: 18:38)



So for example, your B could be simply an interval. Your random variable is continuous. So now you are going to ask the question what is the probability that x takes value in the interval a b? Then what is this going to look like? This is simply going to look like integration over a b f of x d x, right. This is just like area of this curve in the interval a b.

So in a way this function f is acting as a corresponding probability, but not exactly probability, in some sense it is trying to give us that notion here. So if I had a continuous random variable, so I am going to say if my random variable x is continuous, if at all I can come up with such an f such that this is satisfied. So now let us further look at this.

Now suppose, I have taken B to be interval here. Suppose let us say a equals to b. Then what is this? So in this case it is going to be 0, right? Then what we are saying? Probability that x is equals to a, let us say x is equals to 0 because this integral value is 0.

So what we were saying is, if x is a continuous random variable, the probability that it taking particular value is 0. It is possible that it can take value positive probability in an interval but it taking a particular value is 0.

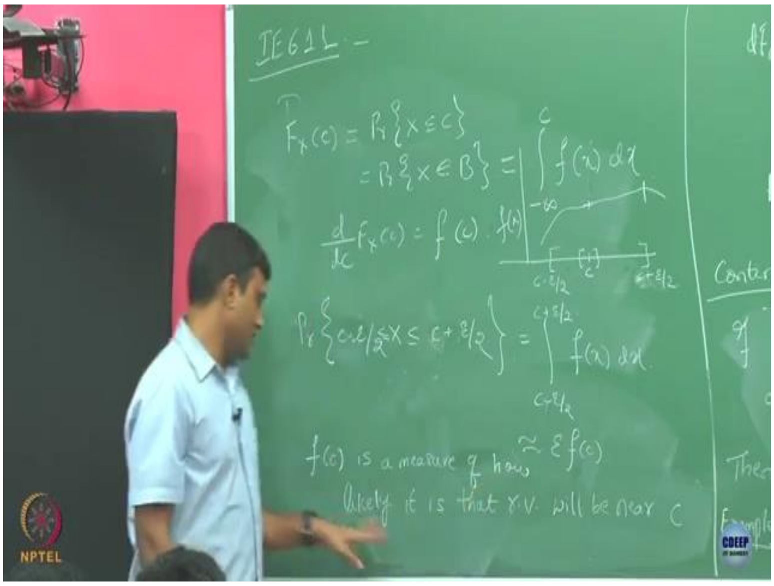
So that is what, if you are going to have, looking at the height of the population in India, like population of India is in billions. If I am going to ask the question that what is the probability that a given, a selected person is going to be exactly of height 5 feet?

That may be like negligible fraction, compared to the entire population. So in that way this continuous random variable is capturing the notion that this random variable taking a particular value is going to be 0.

But if you are going to ask the question, okay what is the probability that the height is going to lie in the interval, let us say 4 to 5 feet? Then maybe there is a non-negligible amount of mass which will, has, going to have that, right? So then that case you expect this probability to be strictly positive. And you expecting like particular height value, you expect that height value to be very, very small or may be like negligible compared to the size of the population.

Suppose now further I want to express, now I want to connect this, by the way we are going to call this f as CDF of, sorry PDF, what is PDF? Probability density function and to denote that this is a PDF of an associated x, we subscript this f with x?

(Refer Slide Time: 22:53)



So now I know that my CDF of a random variable is nothing but, by definition, this is nothing but x by c. So let me express this in terms of this, in terms of a set. If I want to express this like this what should be this B?

Student: (())(23:18).

Professor: Minus infinity to c, so then applying our definition here, this is minus infinity to c... suppose my random variable x is continuous already.

So if my random variable is continuous, I know this definition already holds, So my CDF is integral of my PDF. So we are just going to say that my d by dc of (())(24:01) is f of fine.

We just get this by differentiating both sides. Now to get bit more intuition about what does this function f mean? Okay, fine mathematically we have to defined there exists such a function which need to satisfy this property and we also said that this function f seems to be corresponding proxy for probability in my continuous domain, but what does that actually mean?

Now suppose you want to take the probability. So what is this probability sign? I want x to be between c epsilon by 2 and c minus epsilon by 2. So I am asking my x to be in this interval where this is c plus epsilon by 2 and this is c minus epsilon by 2.

And now let us say I also want to now look at my, in this region I want to look at, like let us say how...so suppose I want to plot my f of x in this region, whatever like, let us say f of s looks something like let us say, for some reason, let us say f of x looks like this.

Now if you are going to apply your formula here and what you are saying, this is between c minus epsilon by 2 and c plus epsilon by 2. And now suppose I let epsilon shrink, or epsilon becomes smaller. So if I let epsilon become smaller and smaller what means, this interval is shrinking. It become shrinking and shrinking, and if I choose epsilon small enough, can I approximate this integral?

Student: (())(26:58)

Professor: How?

Student: F of c (())(27:05).

Professor: Just f of c.

Student: (())(27:08).

Professor: f c into epsilon, So I can approximate this as epsilon. So first thing we note is, suppose if you let epsilon becomes small, that means you are basically asking the question x is exactly equals to c.

That is already going to be 0, if you let epsilon go to 0 that we have already observed. But now what it is saying is if you look at a small interval neighborhood of this c, what this function f is telling is basically it is giving you, if you... the area of that, the neighborhood of this f, that is actually giving you the probability that my x takes the value in that interval.

So this f is itself not a direct proxy for f, but in a way this f is capturing the probability of this event through this. So assume that this is in the small interval, this is like almost constant; just taking the value of f of c in that small interval. Then this will come out. Then you will be left with epsilon by 2 plus epsilon by 2. So that will give you an epsilon, right in that range.

So in that case we can interpret f of c is measure of how likely it is that (())(29:16) it will be near C so in a way if you are going to look at this curve like this, so if you are going to get a curve like this. May be if this value is large, in a sense we can say that my random variable is may be, likely possibly can be taking value here.

But that is only when we interpret it in this fashion, when you are going to take this interval very small and then look what is the probability and that probability is simply turning out to be epsilon times f of c. Okay so if this quantity is large this probability is also largethis f of c is large. In that way f of c is a measure of how likely it is that the random variable is taking value close to c, okay.