Introduction to Stochastic Processes
Professor Manjesh Hanawal
Industrial Engineering and Operations Research
Indian Institute of Technology, Bombay
Lecture No. 08
Expectation of random variables and its properties

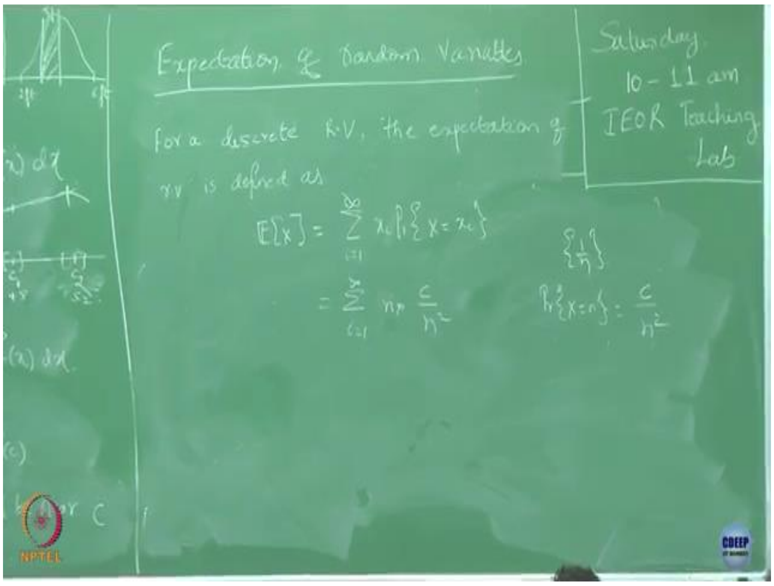
Okay, so, now, I have defined a random variable with whatever we discussed so far, we try to distinguish what we mean by a discrete random variable and a continuous random variable. So, when the things are discrete, things are a lot easier to visualize and understand that your random variable is just taking some finite number of values I, if I describe what is the probability of each of these possible value? I kind have almost all the information.

Now, to move to the continuous case, we try to make it more formal like we know that like when I say random variable is continuous that means it is just like it can take possibly more than unaccountably many possible values. But to make that formal, we introduced this notion of probability density function associated with our continuous random variable and then just try to correct what is the interpretation of that probability density function.

So, often we just may not be interested in what is the value taken by the random variable like if you do, if you have if you go for a gambling or something, how many matches you won? How many matches, in which match you won which match you lost, you do not care, what you care is at the end how much money I made. An aggregate behavior if you are going to go to some like let us say if this happens in inter IIT sports.

That I do not care on which event I lost, which event I won, whether I got the largest number of medals across all the competition so that we are the champions. So, often instead of looking at the specific value outcomes, you want to look at the average behavior. For example, what is the average rainfall in Mumbai? You do not care like how much rain today, how much it rained yesterday, all the water that comes, maybe gets stored in our lakes there. So, at the end what you care about is how much of the water got accumulated. That is what interest you throughout the year like not like on which day I have got how much of the rains.

(Refer Slide Time: 2:53)



So, instead of looking at the micro information, maybe I want to get kind of macro information that is like a average. So, we would be next looking at what we call as expectation of random variables. So, if I say, expectation like, I want to get a global information about a random variable, what you could, what could be a natural way to define the expectation? Let us say if you conduct experiment, I do not want to look at the specific outcome, but I want to see if you are going to repeat that experiment again and again, what is the outcome that you are going to see on an average?

So that is fine, some of the outcomes divided by number of outcome, but that you have actually conducted the experiment. But a priori I tell you, this is the description of your experiment. If you do this experiment, your experiment is going to take these possible values and it has this associated P, M, F from that, can I say something before conducting given the experiment? How? So, what you, one possible way you want to do is you want to multiply each outcome with its associated weight, what could be associated weight?

Student: Probability.

Professor: The probability itself, right? Because the each outcome is going to happen to certain out probability. So, maybe you just want to multiply these two and sum over all possibilities. So, for the discrete random variable, this is straightforward. Or it could be an infinity. So, let us say I have a discrete random variable as I said it is going to take either finitely many possible outcomes or it could be accountably infinite.

But whatever it is, I can find the sum. I know what all the possible outcomes, so Xi's are the possible outcomes. And this is going to give you the associated probability. So, this P Xi is basically the PMF associated with that random variable. So this value, which is nothing but the weighted sum of my outcomes and the weights here are probability, I am going to call this as the expected outcome, I am going to see if I do a this is the expected outcome of my random variable x.

So, if it is or if it only takes finitely many terms, then this summation will not be over infinity it is going to be only over the possible number of outcomes. Right now we are just saying this, this quantity here could be finite or infinity. It is not necessarily that even though my random variable takes always finite values, but the expectation could be infinity. So, can you think of an example where this expectation can be infinity for a discrete random variable?

Student: If a series does not converge.

Professor: If a series does not converge. Okay, your series does not converge but and what about the probabilities? So, let us take your series, you want to maybe just take this says your series itself like 1, 2, 3, 4 is your possible outcomes. Can you give me a, can you put a probability mass on this so that if I look at the expectation it is.

Student: 1 (())(7:51)

Professor: 1 by?

Student: n square.

Student: 1 by n square.

Professor: So, will it makeup probability mass function?

Student: (())(8:06).

Professor: No, but this ratio as summit over and it has to add up to 1.

Student: (())(8:19).

Professor: So, you, so we can appropriately scale it, that scaling happens to be pie by 6.

Student: pie Square by 6.

Professor: Pie square 6. So, I will just write that as a constant, we can come up with a constant, so that this will make it a probability mass function. So, now if you apply this, what is this is going to be? And will this converge? Now, this is going to diverge. So, and this could be infinity in this case. And so, this is infinity, but even when we have unaccountably many possibilities, the expectation can still be finite.

For example, I can come up with like why outcomes has to be and it could be n by n, 1 by n itself, my possible outcomes are 1, 1 by 2, 1 by 3, 1 by 4 like that. So, if you just then, then if I make it 1 by n, this guy is already definitely I know this guy is going to converge and this is done in that case it is going to be finite, okay fine. So, we have an expectation defined for a discrete random variable like this. Now, what about continuous random variable?

Student: We are getting 1, 2, 3, 4, 5, 6 so in a normal way if I ask what is your expectation? So, that values will belong to that say 1, 2, 3, 4, 5, 6?

Professor: What?

Student: What I am expected this by throwing a dice?

Professor: Yeah, what your?

Student: Expectations means what I am expecting should be there, so I will get 3.5 (())(10:19).

Professor: See like, let us say, let us say you defined your expectation such a way that it turned out to be one of the possible outcomes, like 4 or 4 or 5. But let us say that happened to be 4. Now, if you throw a dice, do you think your expectation is met? Because the dye can be taking one of the six values and you may not 4 may not come when you throw it. So, your expectation is not met. So, by expectation here, yeah fine. So, by expectation here what we are just saying to mean is and the actual interpretation we are going to do is, if you are going to do this experiment again and again and again and then you do average all possible values. What is the value you are going to say? That is the expectation.

Student: (())(11:21).

Professor: What is that exactly?

Student: Sir even if I role infinitely then that.

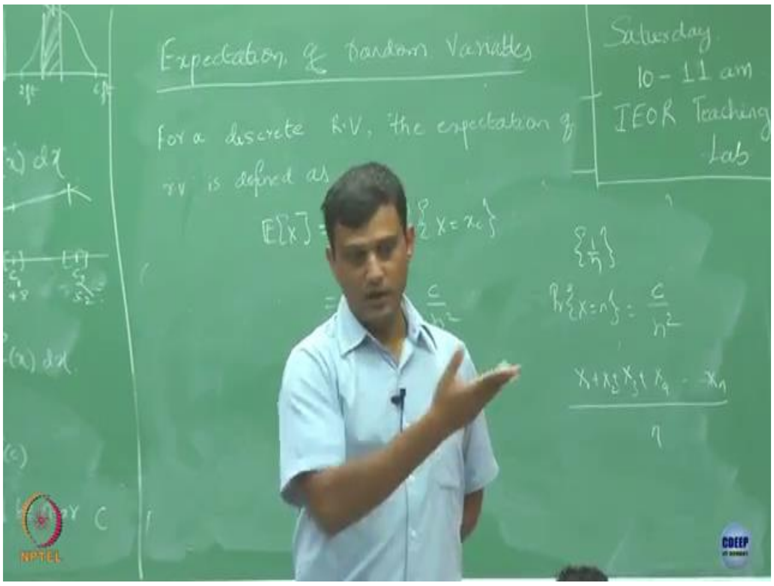
Professor: Yeah.

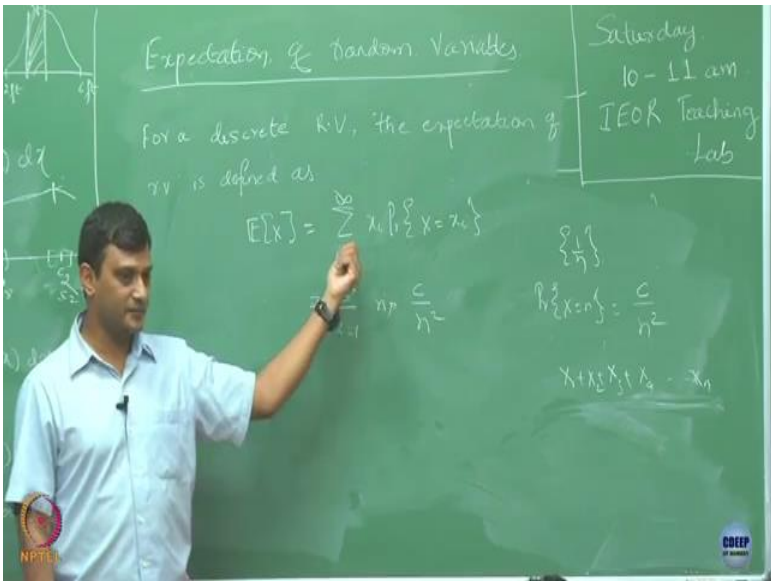
Student: I am not going to get 3.5.

Professor: No, averaged over all values you got that done. So, I am saying you are a gambler, you want to casino first day, you got some, you lost or gained whatever you got. But you are now addicted to this, you go there every day, every day and you are going to win, lose whatever every day. But I do not care, you do not care what happens unless you go bankrupt or whatever. So, what I care is to throw this what is the value I got, averaged overall.

I mean on an average, like okay, so so other way of interpreting is you are going to do this experiment, let us say you do one, two, three times, every times you did this. Every time you are going to get a different value and if you take average of all this, what is the value would I have gotten? That is the value that is the interpretation with which we are defining this expectation. So again, coming back to this gambling one, if you are going to whatever the money you are going to put for whatever number of days you are going to do it, what are the total wealth you are accumulated?

(Refer Slide Time: 13:18)

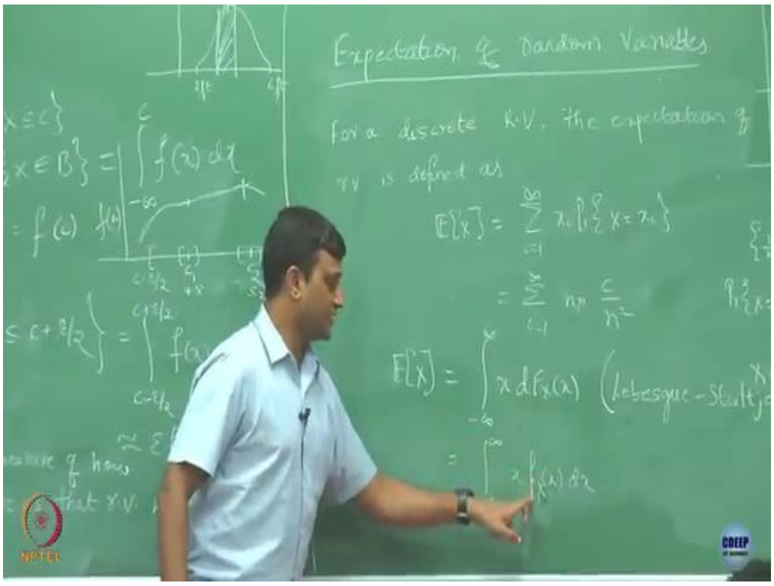


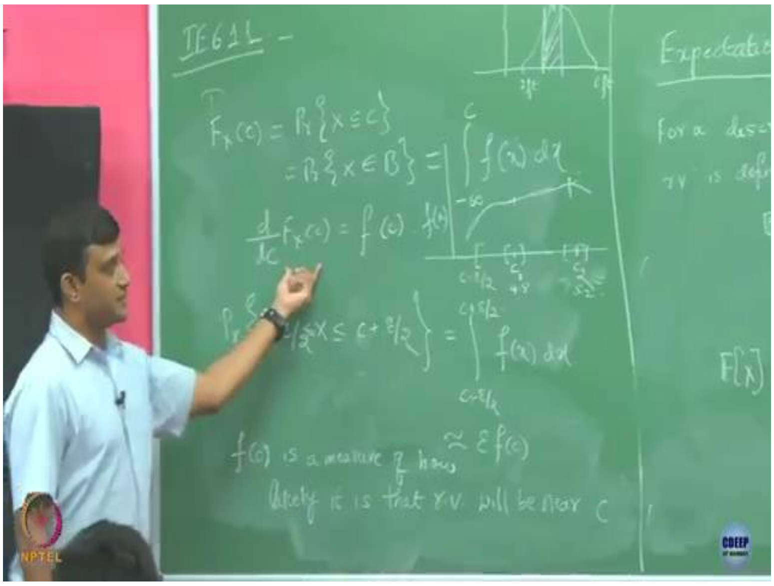


Now I want to average If I average it, then I can kind of in a sense I can interpret this is what I would have got on an average in on each day because like so, suppose so let us say on day 1 you got this, day 2 you got this, day 3 like this, day 4 like this, you add all of them and then n. You did this for n days and you got this ratio. Now, other way of interpreting it instead of this looking at over for all these days, on an average this is what I would have got every day. Instead of looking at the value you got per day, instead of looking at the accumulation and then dividing it by n, I am, I, we can think of this average as per day maybe this is what I would have got?

That is what the interpretation like we want to come, I mean this is what the interpretation, this is what, this is what that leads to this definition, this kind of interpretation that we want to derive that should be naturally coming from this formula. So, fine this expectation need not be the final point is this expectation need not be one of the possible outcomes, this could be any value in the real number.

(Refer Slide Time: 14:56)

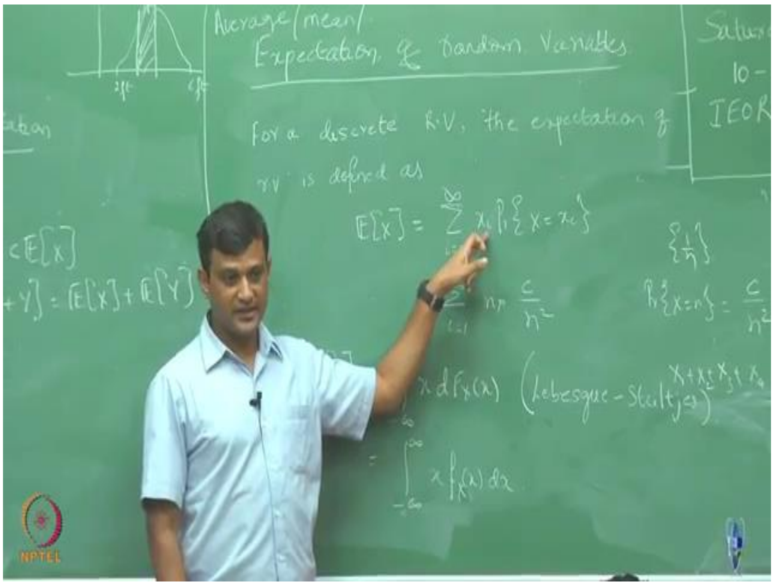


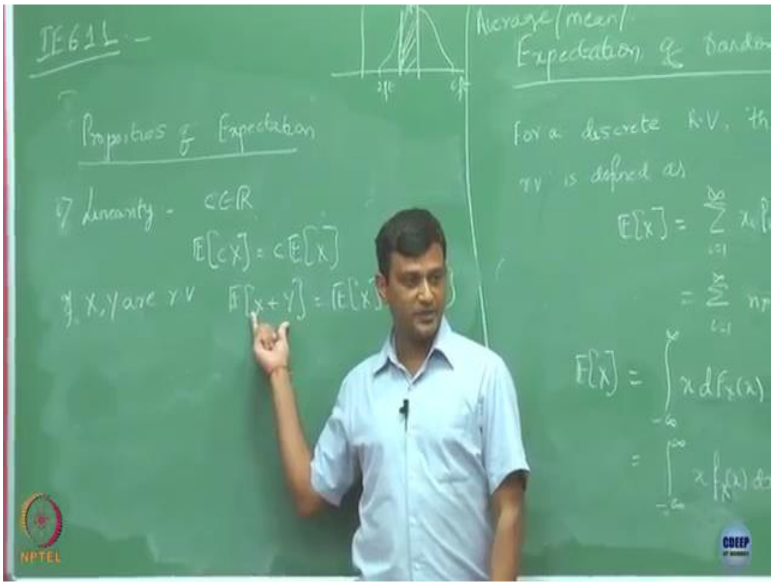


Now, how to define it for a continuous random variable? We are going to define expectation of a continuous random variable as what is called as the lebesgue. So, as we expect in the continuous version, to define integration, actually the integration will come into picture. And there are different notions of integration, Riemann's integration, lebesgue's, definition of integration and this lebesgue is this kind of integration. So, what will take, we will go with this, it is called a Lebesgue Stieltjes Integration and this is going to be defined as x of derivative of my CDF function, which I can write it as, we have already shown that d f of x is nothing but this from this definition here.

So, through this definition of expectation of random variable or mean value of random variable, we are trying to characterize the kind of global observation about that one can make about this random variable x, not about the individual value it is going to get, but what we are saying as the value we are going to get on an average or the expected value here. Now, we are going to define it like this maybe what are the properties this character has here, like when we defined our CDF, we looked at what are the properties that satisfies. So, what properties this expectation satisfies?

(Refer Slide Time: 17:21)





So, so for this we are going to also call it the average of random variable, or we are going to call it as mean of random variable or expectation of a random variable, all them mean the same thing. So, first thing what we call Linearity property, so take a C, then expectation of C of X is going to be C of expectation of X.

So, do you think this is true? This property? So, what I am saying is, you give me c whatever the initial random variable x, I had? I am going to scale it by c. And then if I look at the expectation of that scaled random variable, that is nothing but scaling the expectation of the original random variable. And further, if x and y are random variable, then expectation of x plus expectation of y is expectation of x plus expectation of y, and this should be very much straight forward from this formula. So, when I say c x that means I have to replace xi by what?

Student: cxi.

Professor: cxi, so then it is already true that you can just pull out c and (()(19:28) remains. And that same thing holds in your continuous random variable also. And now, expectation of x plus y how we are going to verify this?

Student: (())(19:47).

Professor: We are going to replace xi by xi plus yi and then what is this probability?

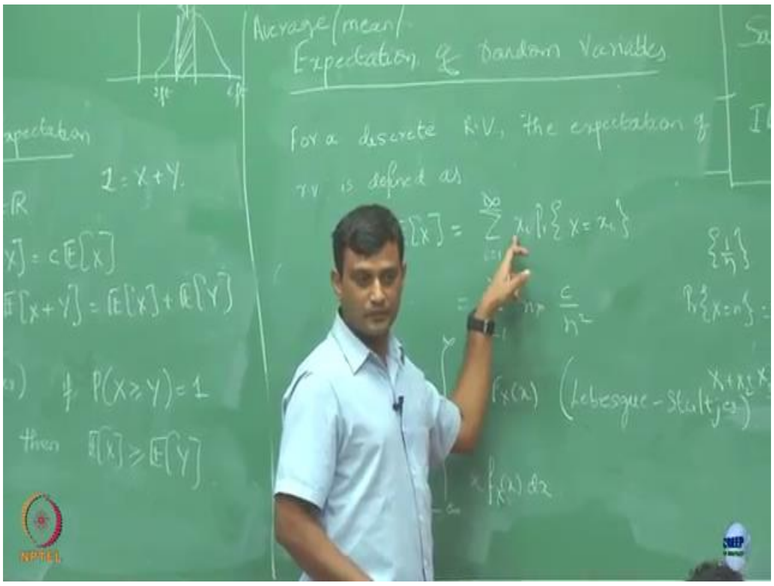
Student: It is distributed in both of them, so either become expectations of that (())(20:07).

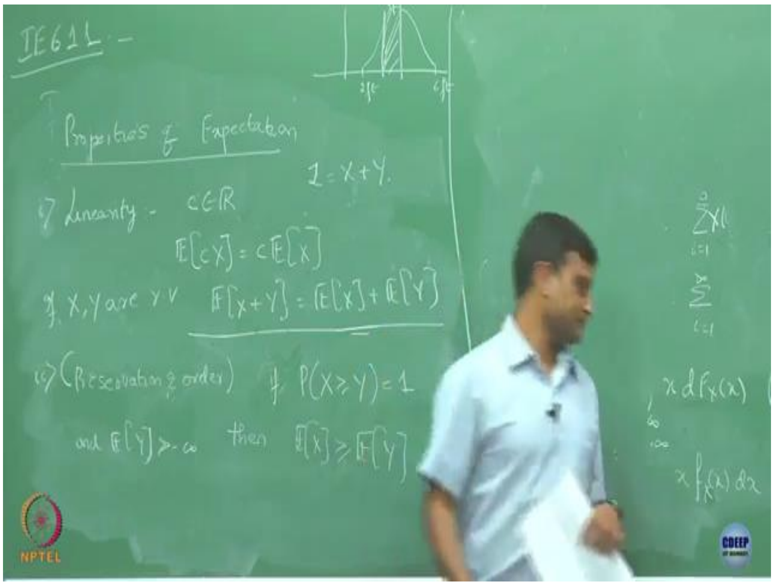
Professor: They will be different.

Student: (())(20:09).

Professor: No, I am not saying independent, I did not say that, just I say x and y are random variable, so, how we will show that? So, what is this?

(Refer Slide Time: 20:23)





Let us say my new, z is my new random variable x plus y. Now, this z may take values different then what this x and y would have taken. Let us say x is throw of my one die, y is through of another die? So x tooks value 1, 2, 3, 4, 5, 6, y also took value 1, 2, 3, 4, 5, 6, what z took? z started from 2, 3 all the way up to 12? So, it is, that is going to be different

And now to find the expected value of z, you have a different probability mass function. But do you think that is going to separate out and you can get this. Why this should be true?

Expectation of x plus y is equals to? Think about it. Let me complete the other properties. So, what is the meaning of this? Probability that x is going to be take greater than or equals to y is 1. What does this means? Now, what do you mean by that?

Student: (())(22:45).

Professor: So, if you take any, so x and y are the two random variable, they are defined on the same sample space. On any sample point x is going to give a higher value than y and the collection of those sample points is such that, that has a probability of 1. If x is going to give a higher value for every possible value of sample point, they expect this expectation to hold then in expectation also they should be higher.

So, basically what you did here is? When I wrote this quantity, so you can replace this by x of omega i. xi is what? xi is one of the realization taken. And I can say that realization taken is for some sample point, omega. And this is what my expectation, if my y, if my y is such that, on the same sample point, it is going to assign a smaller value then you expect that this whole of the expectation is going to be smaller.

So and also and also have assumed that this expectation is strictly greater than minus infinity. This is because if this expectation was not strictly greater than minus infinity, but it is minus infinity then this trivially holds. Because anything is going to be larger than minus infinity.

Student: That means x and y’s are, the same omega.

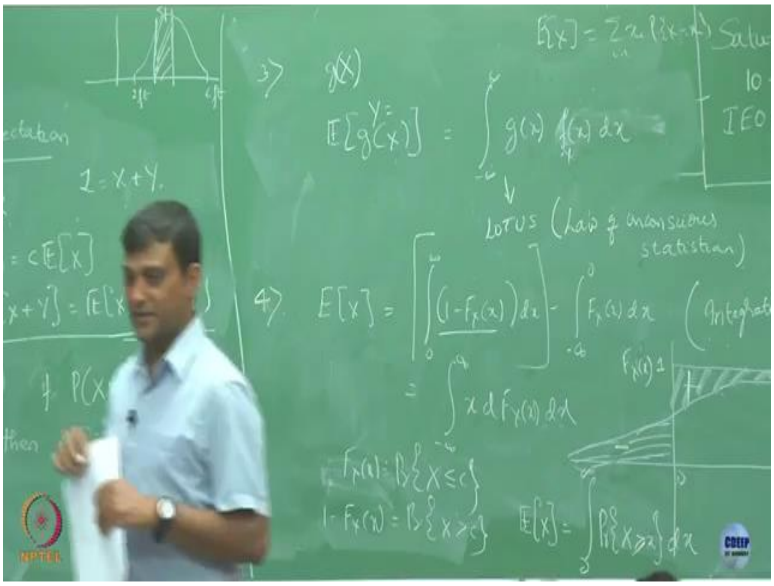
Professor: Right.

Student: So x and y should be like two random variables taken from a same (())(24:41) or.

Professor: On the same, on the same probability space, so this is what I taken here. And by the way and this property is bit more than that like this, you can see that this not necessarily that x and y here should be coming from the should be defined on the same probability space even if they are possibly defined on different this should this should be satisfied. But if you had not checked that we have just we just showed that this mean we are not actually showed anything but we are just saying that if x and y are two random variables, this linearity property holds that, I am going to take this expectation of the sum of random variable as nothing but the sum of the expectation of the random variables.

And you can extend this argument not just to random variables, you can extend it for any n number of random variables.

(Refer Slide Time: 25:53)



Okay third, suppose you have a random variable x and you have a function, you apply a function g on that. So, what does this mean? What are the outcome of a random variable? You are going to further apply another function on this. For example, you want to like you, whatever outcome of the dice, you just do not want it you are going to take twice the value of that, then the g is that function which is making it twice.

Now, you want to complete the expectation of not x but g of x. How does expectation of g of x look like? The x has some underlying let us say if it is a continued, continuous random variable, it has some underlying a PDF. And now, if I apply g of x on that x, it is going to take different values. And g of x may have a different outcomes done x itself. And because of that its PDF may change. And its CDF may also change.

Is it necessary that I should first compute the new CDF? And then with respect to that I should find this expectation. So, one natural way to, so let us call this y, whatever is y, if I can find the PDF or CDF for this y, I know already you know how to find the expectation. If I know what are the possible values of y and associated CDF, PDF, I already have a formula. But without computing the new CDF or PDF, can I write this with the original CDF, PDF of my x? It looks like that is possible and the this is going to be simply g of x of. So, I really do not need to find the, so this is the PDF of my x, earlier if I instead of g of x if I just write x here, what it would have been?

Student: g of x.

Professor: It is just the expectation of x. But then like even if I want to find expectation of g of x, all I need to do is replace this x by g of x, I still returning my PDF function up the original x here. That means in a way then I am doing this computation the x, the random variable x is still having this property, only the value you are looking at these values. And but, so we are just like so you are it is that your value outcome of the random variable has changed.

But the underline x's are still today generated at the same PDF. So that is why we can use this formula. And this formula is often called what they call Lotus or what they call law of unconscious satisfaction. So, it is called unconscious here because like even though you are computing the expected value of g of x, you still not caring about changing that is PDF or this you are still working with a original PDF. And one formula which comes very handy in some cases, the expectation of x can be written in terms of only CDF like here we have written it in terms of the PDF. But it can be written only express in terms of the CDF, the area under the CDF, how it looks like? 0 to infinity, so you see that, so how I get this formula?

This is I think we just need to apply integration by parts here (())(31:30) my formula, so then just check that like we know that expectation of x is nothing but x d of F of x dx. We know this is the case, just use integration by part formula and manipulate it, it should, you should be able to get this. Now, what did this saying is? If I have some CDF like this, let us say this is 1 here. In the interval 0 to infinity, what I am looking at, I want to integrate the area under

1 minus F of x. So, what is 1 minus F of x? So, this my F function, 1 minus F of x is this portion.

From this, saying that you just subtract the area under my F of x but in the negative half of the interval and what is this region? This region is going to be just this. So, you find the region total area of this and from this you subtract the total area from this, then this is going to be still the expected value of expected value of your random variable x. Suppose, let us say your random variable x is such that it is only going to take positive values, it is not going to take negative values.

So, then this part will not be there. And then in this case, your expectation is simply going to be this part, which is just a much simpler version, right it is just like you are going to integrate the area when the complimentary area of your CDF in this interval.

Student: Sir, can we write expectation in terms of CDF when discrete randomly?

Professor: Discrete random variable, so you can also show that, so by the way, what is 1 minus F of x? So, what is F of x? F of x is.

Student: (())(34:27).

Professor: F of x is I know this is probability that x less than or equals to c. What is 1 minus F of x?

Student: Greater than c.

Professor: Is just probability that x greater c, so then this simplifies that your expectation of x is nothing but you can write it only in terms of your probability that probability x less then c into or d of x, this is provided your x is only taking positive value that mean that case you can ignore this negative values. And also, if your x is a continuous random variable, even if I include equality here nothing changes. Is that true?

Because probability x exactly taking value x is 0. That is not going to change anything here. And so what was your question? Can we do write a similar formula for the discrete ones. So based on this, you can, do think you should be able to write your expectation in a discrete case also similarly? Let us see, let us see what you have in mind. So, in what way you want to convert it? We have just PMF, you want to remove this PMF and bring in?

Student: CDF.

Professor: CDF, but we already have said that the relation between PDF and CDF, so we know that this is nothing but we said that a probability of x equals to taking Xi's nothing but the jump of my CDF at that point. So, then we have you can write this expectation in the discrete case in terms of the CDF directly in this fashion. So there is nothing fancy about this. Okay, fine then let us stop here.