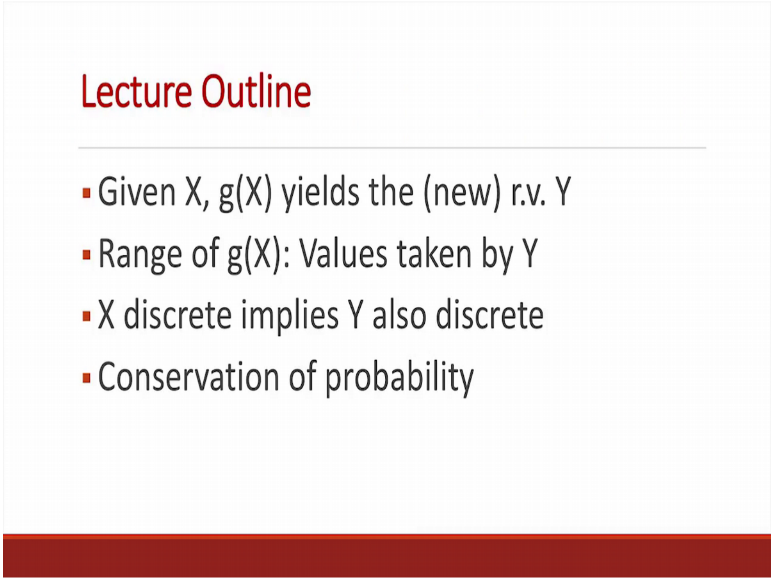
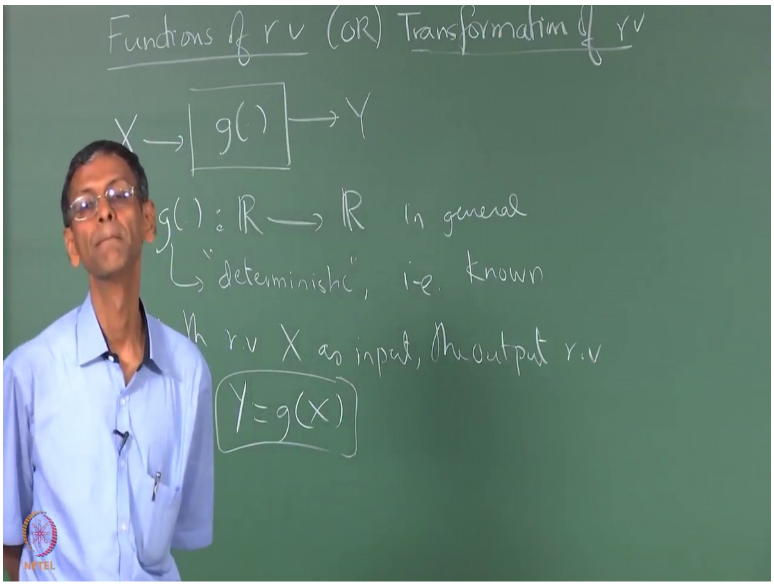
Probability Foundations for Electrical Engineers
Prof. Aravind R
Department of Electrical Engineering
Indian Institute of Technology, Madras

Lecture – 58
Transformations of Random Variables

(Refer Slide Time: 00:15)



(Refer Slide Time: 00:24)



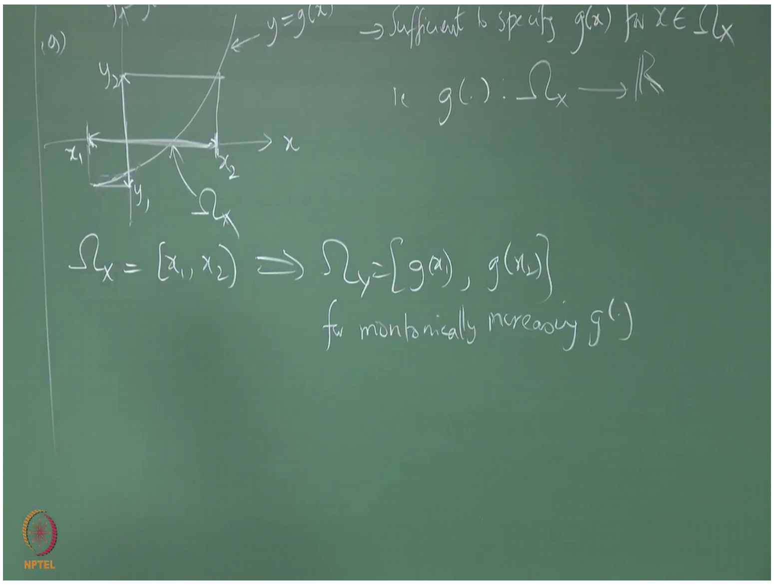
Today we are going to start a new topic which is transformation or functions of random variables. I think it is the shorter name is better it is just functions of random variables, but you have to view these functions as transforming one random variable into another in the sense that right, so the sense of this diagram. So, I have an input output relationship where the transformation is a simple function it is not a complicated operation which transforms one sequence into another or one function of time into another function of time vary in such thing it is just a simple memory less mapping between one number and into another which is basically it is standard functional mapping that you studied way back when 10th standard, 11th standard type material that is all this is.

So, what does this g? g is just in general is a function which takes you from R to R and it is a deterministic function. So, g is determining. When we say deterministic what do we mean we mean its fully known, ie it is specified or full or known it is a known mapping, there is nothing random about g. Standard examples are squaring linear transform like ax plus b and so on. I do not have to elaborate that point, but what is interesting in this context is that its input is not just any real number x, but a random variable capital X.

So, in otherwise you are going to feed into g some you connect it going to connect it to the output of a random experiment the random experiment will throw I do some realization of some value of taken by capital x and that will be transformed by g into y. So, if you look at all the possible values you get out here as a result of this mapping as a result of the random variable x going into this mapping g. So, what do we get? So, can we call this a output random variable it turns out, yes because you are assuming that there is some statistical regularity or behave like say some statistical description of what is coming in at the input and that will transform into the statistical description of the output. So, with the random variable X as input we have we have Y equal to g of X right

So, this output random variable is simply this. So, if you take for example, simple picture eg supposing I take some non-linear.

(Refer Slide Time: 03:51)



Now, we are interested in what are we interested in; obviously, we are interested in some non-linear function g just linear functions are not very if you cannot restrict yourself only to them right, but let us assume that we have a monotonically increasing g just for. So, if you draw something like this means it is. So, what am I plotting here? I am plotting the domain of g along the x axis and the range or the values taken by g or the y axis there is nothing random about this portion this is just a specification a graphical specification of g and it turns out that. So, this is y equal to g of x it turns out that I in general I do not need to worry about j for all of R because in many cases that the omega x the ranged of x is finite is bounded.

So, there is you really need to specify this g only for omega x. So, I make that point to clear may a first (Refer Time: 04:59) size is efficient to specify g or g of x for x n omega x you do not have to go around specifying it for every real number under the sum right.

So, ie, you can write g as map going from omega x and it of course, the output could be any real number we are not going to put any restriction on the on the output of g. So, this may be in general, but, so there is one thing. So, here like supposing this is my omega x let us say my random variable x is only going to take values in that interval what does this mean on why first of all what are the set of values I can I will observe any output. Obviously, I will in this particularly with respect to this particular function I am only going to observe this if this is my omega x let us say from x 1 to x 2 say then I will only observe values in this case as monotonically increasing or in this example I will only observe omega y 1 to y 2 of course, in general also even if it is not monotonically increasing you can still identify the range y 1 to y 2.

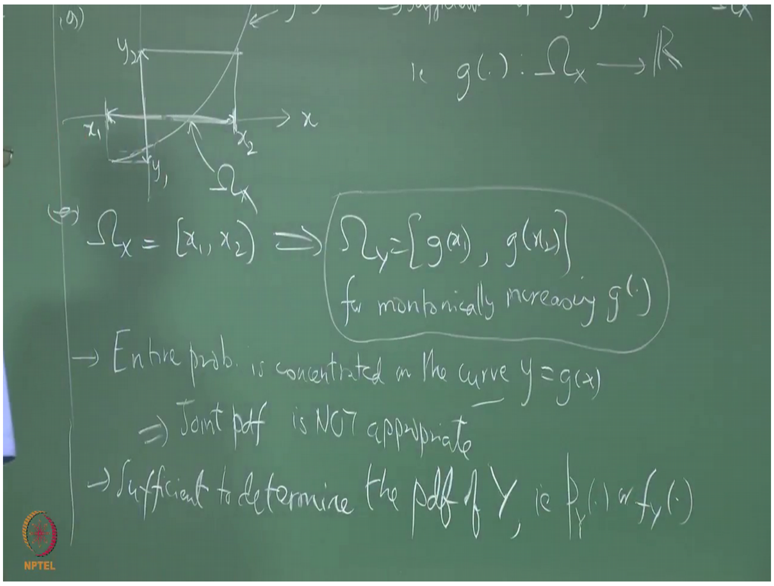
This is easy to draw. So, I am drawing it. So, omega x equal to let us say closed interval x 1 x 2 does not matter implies that omega y will be in this case g of x 1 to g of x 2 comma g of x 2 for monotonically increasing g. This is just an, but this is a true statement in general. So, I do not mind writing it down. If it were monotonically decreasing your put g of x before g of x 1 it is ok. If it is going up and down I am now going to draw all that because I am I i do not have to be enough space on this board to do all the (Refer Time: 07:06) that normally do in dsp on 27 or whatever right. So, if it going up and down you would of course, I you would do the overlay the same thing I do you just have to look at this x 1 to x 2 and apply to g and see what is the range of the output. So, that I will leave it to do you do anyway. We will do enough example specific examples that will illustrate the point. So, there is no point in drawing pictures now, right.

So, this is the first important thing you have understand, you have to see, what is a range of they output random variable. In every situation this is the first thing that you have to always determine and we have done this so many times now that I again I am sounding like a broken record repeating myself again and again they write any conditioning any extra stuff you throw on it always changes the range and that is so the new range is what you have to calculate right.

So, this particular thing is very easy, but if it is going up and down it is not so easy you have to be a little more careful. And also one other thing this omega y as in all other previous examples of determining this new range is a, has to be done on a case by case basis, there is no magic equation or formula that one can write down except maybe here.

So, now, having write said you know dealt with this issue of finding this range. Let us look at how you might want to write the probabilistic description the joint description of x and y. Here now what happens is the entire probability is concentrated on that curve as opposed to the earlier treatment of to joint x and y where you had a region of nonzero area. So, in this case remember that no matter what g is going up and down or not or going monotonically. The entire probability is concentrated on the curve y equal to g of x.

(Refer Slide Time: 09:14)



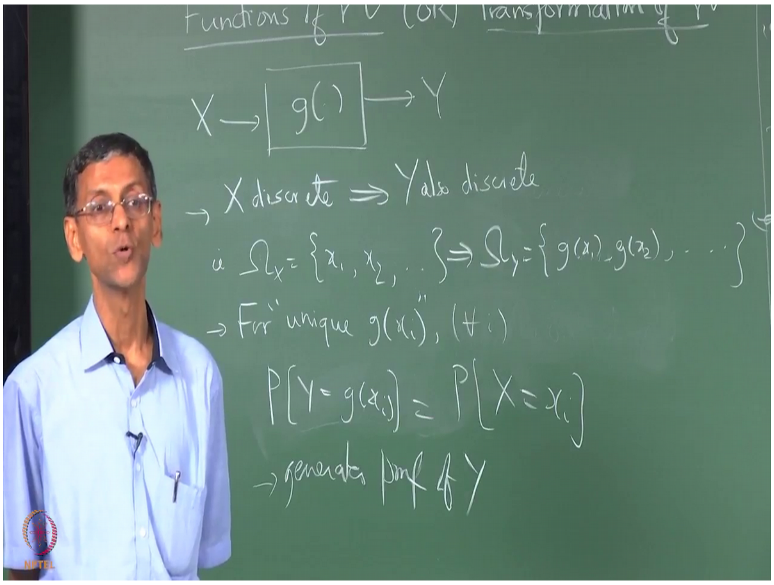
Now, remember this being a function there up and down this it means that for a given value of x there is only one value of y, but for a given value of y there could be more than one value of x that that is that is understood by this. No matter what which irrespective of the kind of g the entire probability; obviously, cannot has to be only on this one curve. So, this means that the joint pdf is not the thing to be looking for you cannot write a joint pdf in such situations. The joint pdf can be written properly only if it exists only if it can be defined or, for a certain area nonzero area curves always have 0 area in the xy plane right.

So, joint pdf is not appropriate, it is not the appropriate tool for the situation that is not so. But it turns out that what you can do is treat y as any as a separately existing random variable and ask for the pdf of y separately right. So, that is what we will that is the approach we are going to look at we are not going to look at the; I am. One more point I can we can; obviously, say is given that x is some given that x is taking some value in its range then there is no more randomness in y you know exactly what is the value of y, y has to take the value g of whatever g of x naught what. So, that can, so the conditional if you look at it the conditional pmf, so y becomes a deterministic quantity. So, with probability one it will only take the value g of x naught when x takes the value x naught.

So, you also have that going for you in the in such a situation. So, it is sufficient in this case to just to look at the standard pdf which we started looking at long back in this course. So, it is sufficient to determine the pdf of Y, determine namely now I am here I am going to have to make a small distinction like need write the discrete versus continuous. So, what kind of a random variable is Y going to be? So, before we proceed any with any calculations we have to understand that right.

So, what kind of a random variable is Y going to be? Supposing X is a discrete random variable I think maybe I do not need this supposing X is discrete which means that you can take only a countable set of values here then Y also has to be discrete.

(Refer Slide Time: 12:35)



This is just you do the function the nature of the function g itself given that x equal to x 1 or x equal to x 2 you y can only take of a accountable number of values g of x 1 g of x 2 and so on. So, in this case we are going to say ie, omega x if omega x equal to some numb some countable set like this implies what omega y now of course, some of these may be the same g of x 1 could be the same as g of x 2, does not matter if these are the same then you have to recognize that and avoid application, but that is besides the point let me write we cannot you know. So, we are assuming that these are all distinct, but they need not be distinct right.

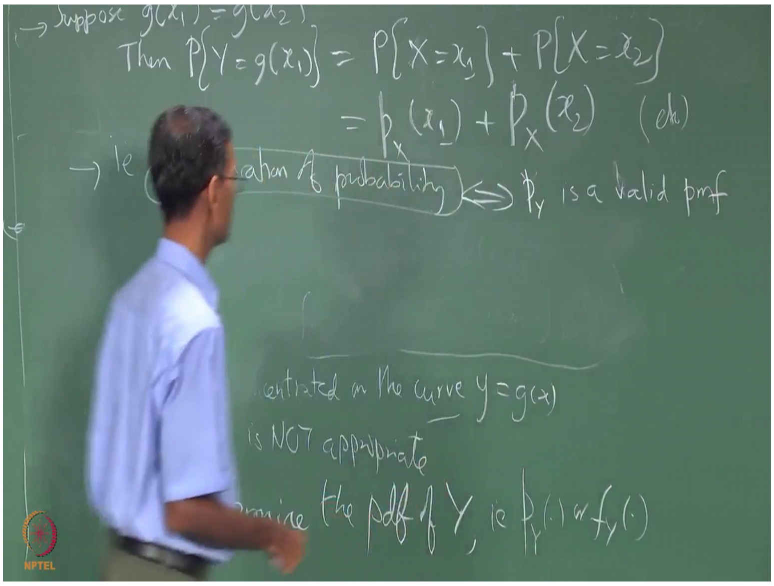
So, in this pmf in this discrete situation it turns out that it is not much pending too much time on this, why is it not spending too much time on this because there is the pmf of y is

I think extremely easily written down you should be able to do it at this point. What is for example, supposing the mapping is g is monotonic in that for x 1 there is only one value for any x i there is only one value g x i. So, if what do you call such functions is, one to one on the set a on this set this is as the name for it injective, surjective, one to one I (Refer Time: 14:05) so many I did not document all those things, anyway.

So, for anyway one to one supposing g for unique let us say g x i in the sense that this is g x, g x i is not equal to g of x j, for any i not equal to j that is what I mean. If g x i is unique for all i, in other words then P of Y equal to g x i will be identical to what? P of X equal to x i. So, this determines the pmf of Y. If you rule this for all I automatically generate the pmf of Y. So, this generates.

But supposing you do not have unique g x i supposing there are two numbers x 1 and x 2 let us say which map to the same g x 1 then what is that probability y equal to g x 1 it will be the sum of the two numbers P of Y x equal to x 1 plus P of X equal to x 2 because if you get either number in X, you going to get the same number in Y. Again I am not going through all write, all of that do not know maybe, I will write it down for the sake of the camera right.

(Refer Slide Time: 15:58)



You know suppose let us say g x 1 is equal to g x 2 and write x 1. How did we write this? We wrote this as small p of x 1 plus this function p x we used write, so etcetera. So, all this procedure can be repeated as needed for all points in omega y. Once you repeat it for or do it for all points in omega y you generate the pmf remember you have to do it for every single point in omega y, is not it. You do not completely specify the pmf of random variable will use to nail down all the probabilities of all the points or values it can take.

But I am assuming that that is not a problem. So, this discrete or discrete mapping through this function although it is important in its own its not considered to be sufficiently significant to devote a lot of space. So, we will not come back to this anymore it is you can do it by just common sense. Only thing I wish to emphasize is this kind of a calculation is basically it is hard of what we are going to do in the continuous case also and it is basically called conservation of probability.

What is conservation of probability mean? That the unit probability in omega x somehow has to be translated to unit probability in omega y. So, probability has to be conserved you cannot have a different overall probability here, this probability on this set must be identical to one, the whole patrol probability. So, you have this important principle of conservation of probabilities, among all the other conservation laws you have this one also. And this is not the first time we are seeing it all the conditional pdf, pmfs, everything that we wrote in the past to a week or so, every one of those has to be a valid pmf of pdf what means it must have unit probability, whether you write it as an unconditional pmf or a conditional pmf or a pdf does not matter. As long as you call it a pdf without any (Refer Time: 18:32) anything in front of it, it has to have unit area unit sum. So, here also obviously, you need to have write the same, this means what there is a p y is a valid pmf.