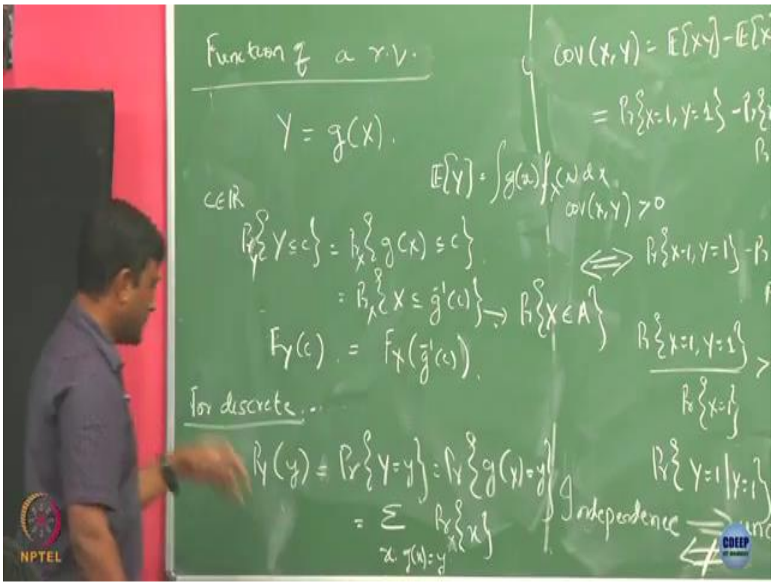
**Introduction to Stochastic Processes**
**Professor Manjesh Hanawal**
**Industrial Engineering and Operations Research**
**Indian Institute of Technology, Bombay**
**Lecture No. 14**
**Transformation of random vectors**

Okay, so let us move on now, I want to discuss, how to find distribution of a function of random variable. So, often what happens you have some outcome? Let us say your outcome of experiment, I am going to define it by a random variable x. But most of the time it is not necessary that you will be just interested in x but some function of this x.

So, maybe like if you have a outcome, which can take values both positive and negative. That is, so your random variable x is such that it can take positive or negative values, but you are not interested in the sign of the outcome. So, you may want to define a new random variable y which is simply the absolute value of x. That is what matters to, the absolute value the magnitude not the sign.

So, then what is happened is? You have basically interested in a function of this random variable. Or alternatively, let us say you have a outcome, you do not want to just look at outcome but you want, you are interested in the squared value of the outcomes. Now, how to find the distribution of such a random variables? So, let us start with a single random, function on single random variable, then we will move to how to find distribution of the set of random variables.

(Refer Slide Time: 1:53)



Suppose, let us say our random variable x, but I am into, what I am interested in some function of this random variable x and let us call that function to be g. Now, what I want is I will tell you, what is the distribution of x?  Now, I want you to give me the distribution of y. How you are going to do that? If I am simply interested in the expected value of y, what I am going to do? What will be expected value of y? g of?

Student: x.

Professor: So, here if you just tell me the PDF, let us assume it is a continuous random variable and if you give me g function, then I just find it. Now, I am asking you more than this, but here you find you give me, you can find the expectation like this, but I do not want you to just give me an expectation. I have asked you to give me the actual distribution of y itself. So, how you are going to do that? So 1 thing to do is simply, you want to find distribution of y's.

So to do that we start with our CDF, so I am for some some see let us say something up. I will start with what is that probability that is equal to c, this is going to be simply probability of g of x just another less than or equals to c. And support let say you can this function is invertible then you are going to find g is equals to g inverse of c.

So, since c is a constant here, you know g function and suppose if you can invert it, you will end up with another constant here. And if you already know the CDF of this, you can find a CDF for this function. And then what you can do is? Now, this is nothing but f of g this, now to find CDF of y what you need to do? Sorry, this is going to be, so this is going to be F of y at c.

So, in general what we are going to do, we have basically translated this condition which this I can basically write as something like I have translated this into x belongs to some set A. So, this x being less than or equals to g inverse of c, saying that x belong to some A, where A is now defined in terms of this constant c and my function g.

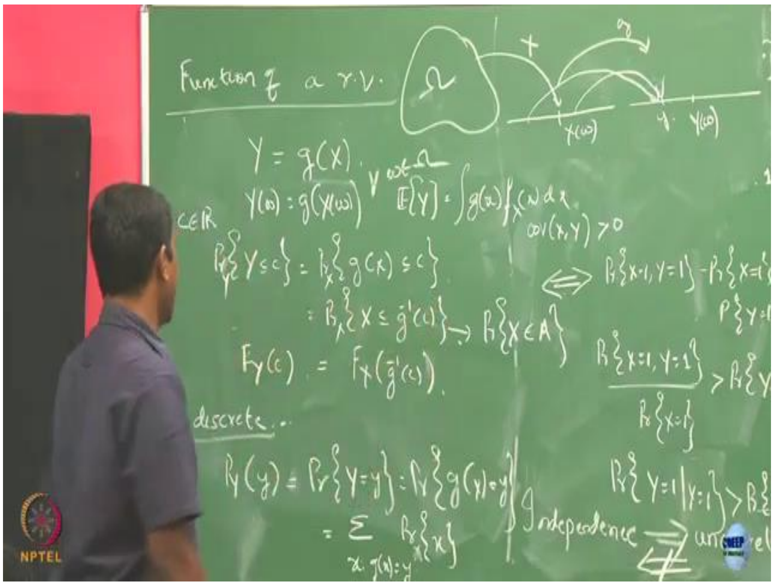
So far, we head CDF of y, now how you are going to find the PDF differentiate with respect to what? C, so suppose time being assumed that this guy is continuous and differentiable, then you can differentiate with respect to c and you will you will going to get right away the PDF of y, what you really wanted?

Now, so this is for the case when you have, you are y to be sorry x to be continuous random variable. But what have what if you are x is discretes random available. So in that case, you want to basically interested in first the PMF of your random variable y and for discrete, we are going to say, P of y if my x is already, so this is nothing but P of y equals to y but now this probability that g of x equals y. And this is going to be set of all x, such that g of x equals to y of probability of x.

And now, see, like when I am applying probability, we have to be also careful. I am computing this probability with respect to random variable x or y or on computing this probability jointly. So, in this case when I write it here I am computing probability that y is less than c. So, here the probability is with respect to distribution of y. And here now I have replaced y by g of x. Now, here probability with respect to distribution of x, so often to be more precise, we can subscript this values like this. And here it is x and all.

But when it is obvious, we just drop it. So, now, let us come back here. Probability of y that it takes value small y is nothing but this right like and then I have just replaced Y by g of x that is the definition of y. Now, what I am interested is? This is nothing but here, set of all x such that g of x goes to y, it is not necessary that 1 point goes to, when you apply g function on that 1 point, it goes to 1 point that there could be multiple points that can go to y.

(Refer Slide Time: 8:55)




So, for example, let us say this is our omega space and this is your random variable x, y. So, what is x doing is? For each value of omega that lies in this big omega, it assigns some value on the real line. That is the meaning of x, x is basically giving values to each point in this.

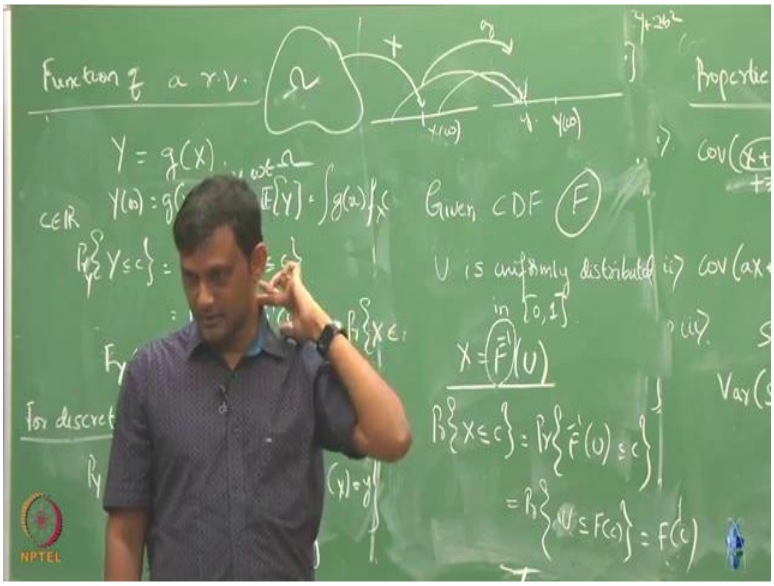
Now further, what we are doing is we are doing is? We are further g on this to get y and that is further give me some for the same omega, the new value y of omega. So just to be clear, what we mean by this is? Y of omega is equals to g of x omega for all. So, now when I apply this g function, let us say this in some y, let us say I am interested in some y here. That is my small I, when I apply this g function, it may happen that many function can fall to the same y.

So, that is why I have to look for all x that takes up value y and then add all of the probability, that will give me probability of y at the value of that due to probability that y takes value small y, understand this? So, this is a process, if I just give you x with a certain PDF or just CDF for its PMF, If you want to ask find a PMF or PDF of y, this is the general step. You have to just basically translate that value in terms of x itself for some known set and write it in this form.

Now, where is this is this is useful? Of course, this is useful whenever you want to find the PDF of a function of a random variable. But most often it comes to use when you want to simulate a random variable or samples according to a given distribution. So, many of the time you want to do a priori some analysis. You feel that your system has this kind of CDF a but you cannot always work on your system.

For example, your system could be very costly aircraft carrier or like a big aircraft, fighter aircraft or something. So, you cannot like directly work on this, you have to basically simulate some of the things in your room. So, to and suppose, so like some critical component or whatever component if you feel that it is going to behave in this form, then in the in your laboratory on your computer, you want to generate samples according to the distributions. So how you are going to generate that distributions? So here this function of random variable come handy. Suppose let us say you have a system which produces outcomes according some distribution F.

(Refer Slide Time: 12:04)




Now, I want to generate samples according that is that obeys the CDF, or come up with a random variable of characters around a variable that has this area. So suppose if I can set up an experiment that gives outcomes, which has a CDF, as per your requirement, that means I am basically simulating whatever the real thing that you wanted to characterize through this F. So, now how to generate a random variable which has this F? So often, we will be given x and will say this is the CDF of it. Now, I am asking the reverse question, given the CDF generate me a random variable who would outcomes will have this CDF, how we are going to do that?

So, 1 simple thing to do is we you know, uniform random variables. So, the claim is that using uniform random variable and description of this function F, you can generate a random variable, whose CDF is exactly this. So suppose let us say u is uniformly distributed in the interval 0 1. My claim is that if I am going to construct a random variable like this, this random variable x, will have the CDF F. So, let us first see why that is true, and then let us try to make it a bit more precise.

F is been given to you, when I say F is given to you that is a CDF and I expect it to satisfy all the three properties of a CDF, what are those? Monotonicity, the limits being 0 and 1 at the 2 extremes, and then.

Student: Right continuity properties.

Professor: Right continuity properties. So, let say it already has those things. Now my, my task is to generate a random variable that has the CDF. So, now if I define this Maclaurin, if this random variable has exactly that CDF, why is that? Suppose, let us say. And here F inverse, that is the job of g. And u is uniform random variable, which I already know how it behaves and how to, let us say, generate it. But now my goal is from that generate a random variable, whose CDF is going to be the given F. Now, let us say I am going to do this, what is this? Probability that F inverse of u less than or equals to this.

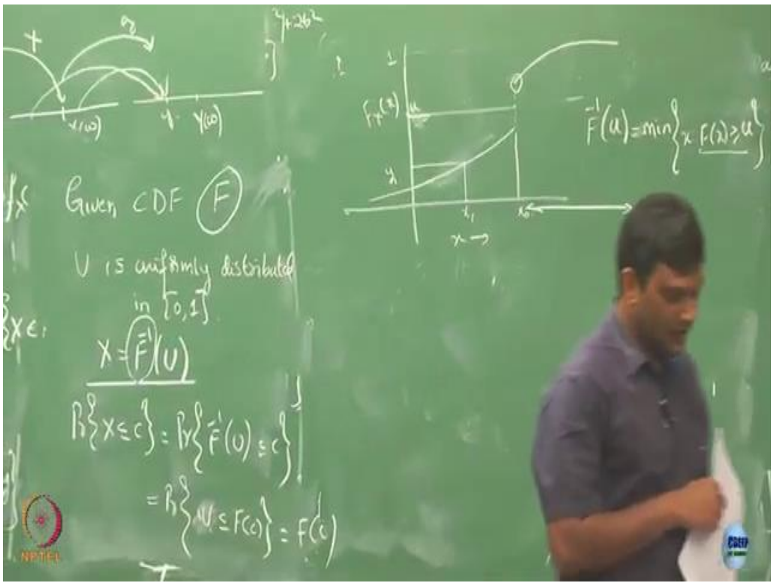
And now if you simplify this, this is going to be u is less than or equals to F of c. So is that? F is a CDF, what is the range of F, F, it is going to take value only between 0 to 1. Now, I am asking the question after simplifying this, this uniform random variable taking value less than or equals to F of c which is between 0 1. What is this value? It is simply going to be F of c.

So you see that the CDF of this random variable is exactly what you wanted. It is, it is the CDF of, it is exactly the F function. Now, the question is to do this, I should be able to invert my function F, whatever my function F that is given to me. So, is it possible like if you take a CDF which has all these three, three properties, I should be always able to invert it.

Student: Yes.

Professor: Yes, what? That is fair enough. So let us look a simple.

(Refer Slide Time: 17:10)




Let us take an arbitrary CDF. Let us say it has this shape and let say jump jumps here and then something goes like this. There is a jump for us here and let say this is this is the right continuous function at this point, x and this is. Now, it is clear that how the invert look invert point looks, so if I want to find inverse of my function F at this point, let us say this is let say y here. So, how you are going to find the value of x that gives you this value y on y axis? So you just draw a line there and then come back here.

So at this point, let us say this is, let me call this y one, and this 1 x one. So this x 1 is the inversion of this y one. But how we are going to do at a point here, let us call this y 2. So, what this where any value in this way is going to map is it possible to clearly define what is the inversion here? What is the inverse of a y to here?

Student: This will not be (())(18:48).

Student: It does not take the value (())(18:54)

Professor: What?

Student: It does not take the value (())(18:51).

Student: It will not be the domain of a (())(18:56).

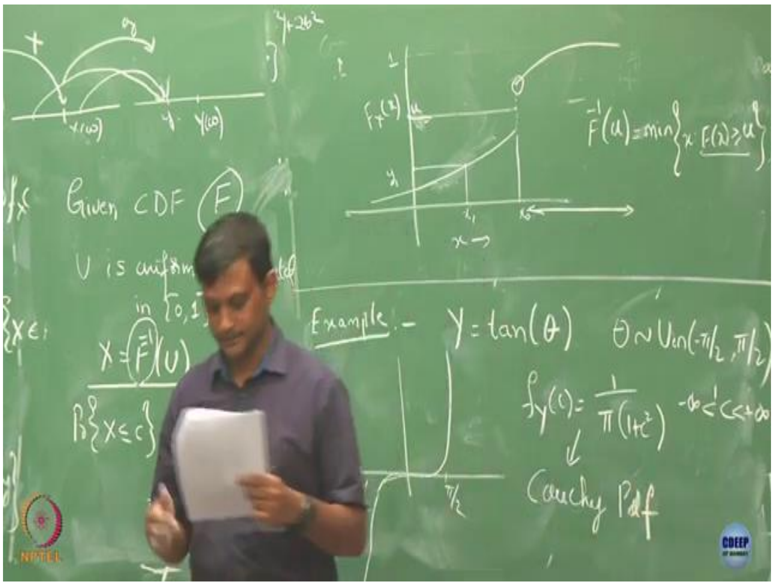
Professor: Right, but to make this work I have to up assign some values to F inverse function, I need to define what is my F inverse function. So, 1 possible way to work on this, what we have to do is? We have to appropriately define F universe u, what F inverse u you can define is suppose, let us call this as u in this case, which is in this. We can define it as min of x at F of x. So, you look at all the values of F of x which is going to be greater than this point and in that you take the smallest one.

So, it is going to be this. So, this is monotonically increasing. So, all the points above this, you take but from that, so all this region will be covered in that take the smallest one. And if you do this this is still well defined and from that you can continue with this definition of F inverse function in this and you can see that like if you just define F inverse like this everything goes through.

So, this is way you see that right like where how this function of random variables comes to help at least if you want to simulate a 1 random variable which is complicated random variables, if I can define a function appropriately g function, then using a simple random variable in this case uniform random variable, I should be able to generate a random variable which satisfies the more complex CDF characteristics.

So, fine, so we will in the in the you look into the textbook there are so many examples about how to derive CDF of 1 function which is expressed as a function of another random variables.

(Refer Slide Time: 21:10)



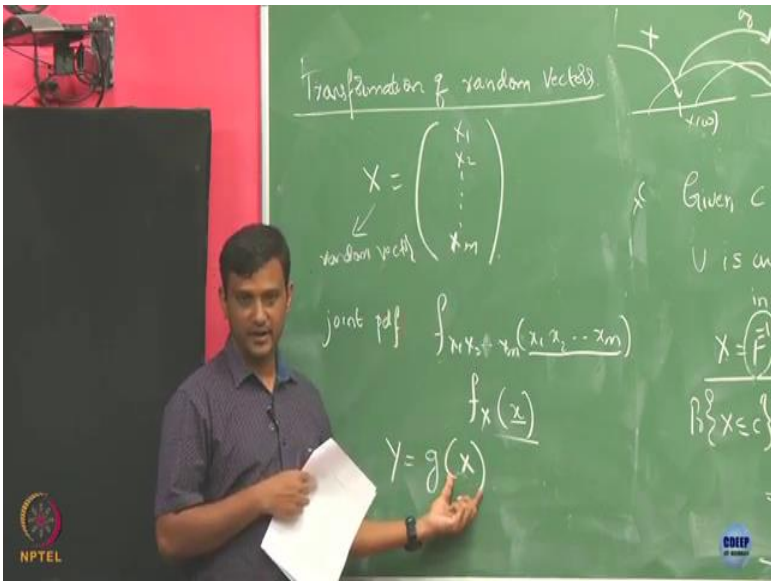

So, just like 1 example, which you can work out is suppose let us say you have a random variable which is defined as y equals to tan of theta, where theta is distributed uniformly from minus pi by 2 to pi by 2. You understand this? Theta is uniformly distributed random variable that takes value between minus pi by 2 to pi by 2. Now, we are applying tan function on that and you are defining the new random variable as y. So, how this function look like?

So, let us say this is my pi by 2 and this pi by 2 minus infinity 2, so it goes likes maybe something like, thing like this. Now, what is the range of y? Minus infinity to plus infinity, whereas the range of theta is just minus pi by 2 pi by 2. Now, how to find a CDF of y here? So, just like using this method you can work out and you will see that f of y of c is equals to 1 by pi, 1 plus 3 square 4 c into (())(22:55) plus infinity.

So, this is 1 upon pi, 1 plus c square and this PDF is in the literature it is called a Cauchy PDF. So, this is function of one kind of variables. It is not necessary that most of the time you will be dealing with one random variable like as I said, if you are interested in whether you may be interested in temperature, humidity, density of whatever the clouds whatever, so there are so many random bunch of random variables you have to deal with. And then you may want to look at transformation of these random variables.

So, here it is the transformation of single random variable, but you may be interested in transformation of a set of variables. So, how to do that?

(Refer Slide Time: 23:58)




Now, when we have a set of random variables, all are defined on the same probability space. I put it as them as vector. Suppose let us say I have n random values, x one, x 2 all the way up to x m. I can treat them as a vector having m components. So, I am just going to represent that a random vector now. So, a collection of random variables I have, I am just going to take them as a vector and denote that as x here. Now, what is the x here is a random vector here.

So, I am just going to use the same notation for random vector and also sometimes in a single random variable I am going to write it as x. So, you should be clear from, from the context that I am talking about x, which is just like single random variable or it is a random vector. Now, how is the distribution of this random vector? When we have a set of random variables, we had something called joint probability density function.

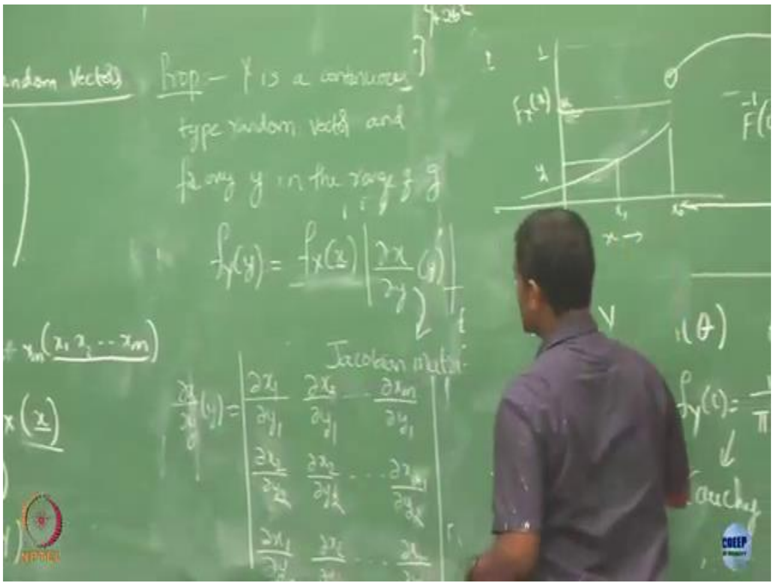
We are going to take that joint probability density function as the distribution of this random vector. So the joint PDF, so I already told you that when I have m random variables, they are completely characterized by this joint probability density functions. I am going to now take this itself as the probability density function of this vector and that makes sense. Like now in this case, I will simply write it as x, here x is this vector and this x is this vector.

So x is this vector of random variable, this x is the values, the vector realizations they are taking. Now, again, so again notice that I am simply writing it as x to denote this entire vector. And it should be again, clear from the context I am talking about single random variable or a multiple random variables.

Now, on this multiple lambda variables or a random vector, now I may be interested in the same. If I have a PDF of this random vector to be x and I will be interested in another random vector which is a function of this random variable. Now, how to find how to find the distribution of this random variable y? So, we already discussed when this x is a random vector is just like similar values like just as 1 component we already computed how to do this but now this x is a vector how to do this?

So, for this, I am just going to write a formula, which is expressed in terms of the Jacobean matrix, we will not derive it. But we will take it and you will make yourself familiar with that and you are going to do some exercise, how to use that?

(Refer Slide Time: 28:20) 01, 32:15,



Now, this is like, let say y is a continuous, so notice that once you have here, I can also write x to g inverse of y where y is again a vector. For any y in the range of g, we have that f of y, f y is equals to f of x, x divided by delta of x, delta of y computed edge y. So, when I wrote write this is the Jacobean matrix? And when I say this to 2 vertical block, verticalized this is in limited that determinants of this Jacobean matrix.

Now, what is this Jacobean matrix? It basically looks how each component varies each, dependent component, independent component varies as a function of the, independent component varies as a function of the dependent component. So let us say you have x 1, here y is what? Dependent component, it is dependent on x, x 1 delta as y one, x 1 by 1, all the way up to the let say x m to let the bar 1 and then delta x 2, delta y 1. So if you are going to solve this, so you have to construct a matrix like this and then so this is going to be a square matrix.

So they are also going to assume that this y is also of the same dimension of x. So, then you are going to find the determinant of this. So now here, this is you are going to compute with for a given y. And know but for that you have to first find x here is nothing but g inverse of y. And then you are going to compute this matrix at y.

Student: So, (())(32:34) x one, there also, first row last column.

Professor: Last column what?

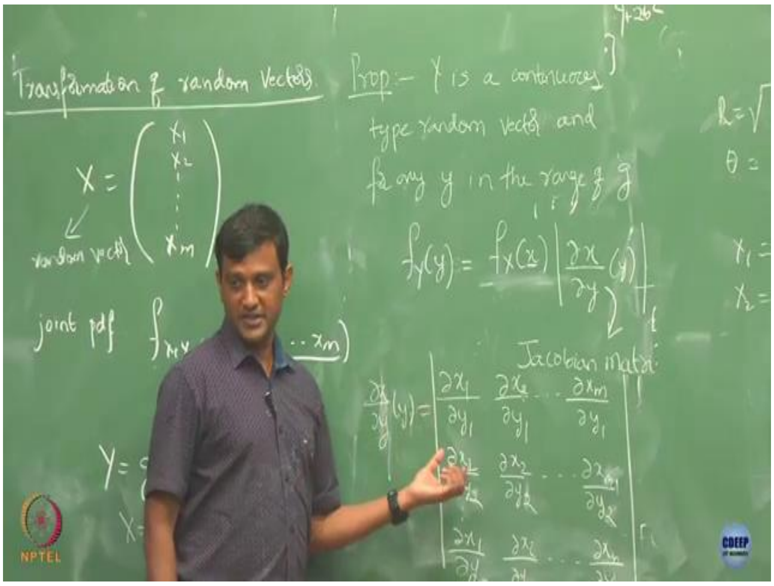
Student: (())(32:54)

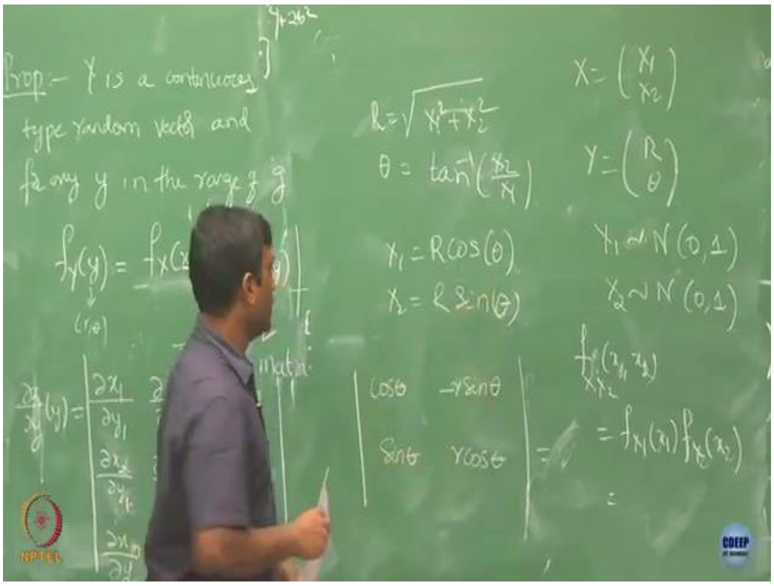
Professor: So, let us quickly look into this not going to see how this will come? But let us just quickly get a grasp. So let us say you want to, for example, as an example, let us say you are in a Cartesian, in a Euclidean space where the, how you are going to denote a given component point? Through its coordinates, so like if you are in a 2 dimensional space any point you are going to denote as x 1 comma x 2. Now, let us say you want to shift to polar coordinates. So in polar coordinates, what are the components you need?

Student: R and theta.

Professor: R and theta, so r and theta can be expressed in terms of x 1 and x 2. So that is the map. So, let us say you are already in the Euclidean space, and now you want to basically go to polar coordinates. So, that you can do through some function and now if this Euclidean points have some distribution, you want to understand, what are the distribution of the polar coordinates? Now how we are going to do this? This method is going to help you.

(Refer Slide Time: 34:31)




So let us take a simple case, where my r 1 r is equals to x 1 square plus x 2 square radius and my theta is equal to how it is going to be x 2 by 2 by x one. So here my x is going to be x one, x 2 and my y is r, n theta and my function g is such that it is being this kind of mapping. It is taking r 1 and r 2 and giving you sorry x 1 and x 2 and it is giving you r radius and it is giving you theta. So, now you have expressed independent kind of variables. So, these are independent here. So, how the dependent variable depends on them? So, I could also write in a reverse form. So, how can I represent x one, x 2 in terms of r and theta?

Student: R cos theta (())(35:48).

Professor: R and, so, now in this case, my, this matrix is going to be simply through the 2 cos 2 matrix why? Because I have only 2 components in this. So, can you quickly compute this and tell me what is this Jacobean matrix is going to look like? So, take y 1 to be r and y 2 to be theta, cos theta.

Student: Minus r sin theta. Sin theta or cos theta.

Professor: I will just check this is what the computation I have here I am just writing that.

Student: (())(36:52).

Professor: Going to be transpose of this matrix, why is that?

Student: To do Del x 1 by del R, it can be cos theta.

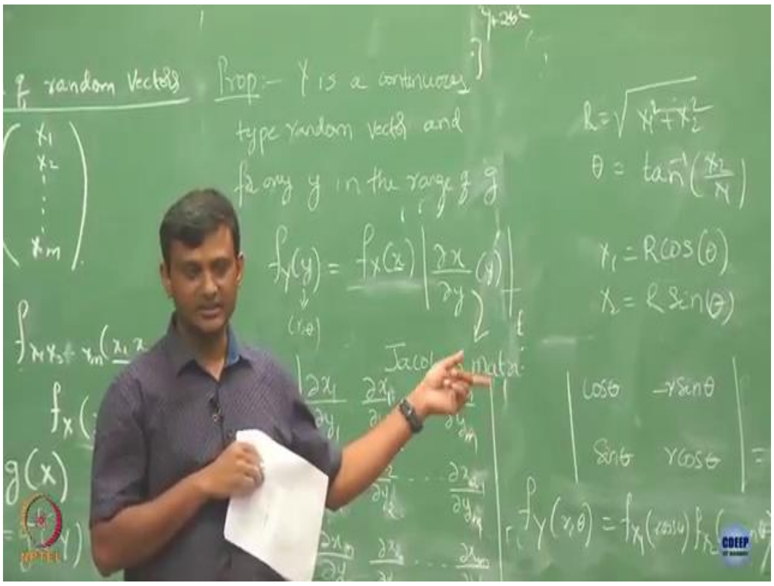
Professor: x 1 I am just going to do R, this is simply going to be cos theta and then this 1 I am going to do with R. Did I mess up? Sorry, I think I messed up this, this should be now, but let me just define this correctly determinant will be same but whether are you if you take transpose or not is the worry. Now it is fine. Now, what is this function, so let us now talking this, what is the determinant of this? It is going to be R. So, now what is this quantity? So for a given y what is this? You do not know it, like I have not told you what is this f of x.

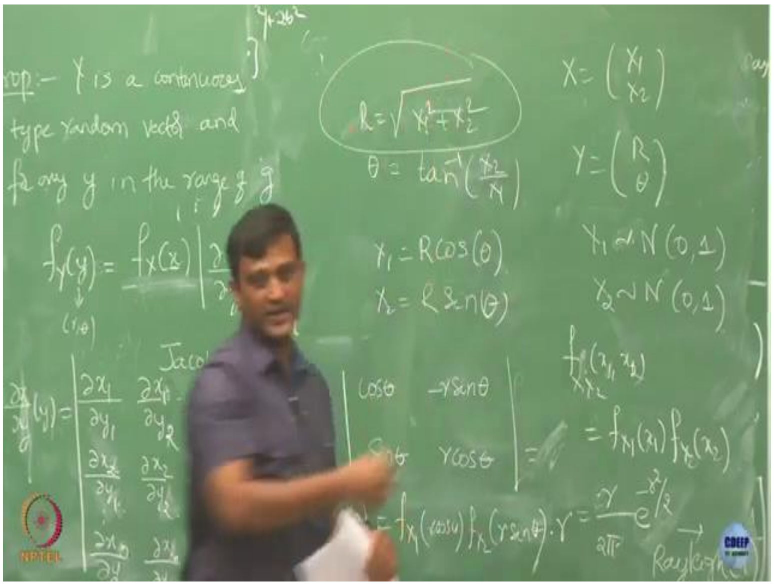
Okay, for time being, so now let us assume for to do, to do this I need to know f of x distribution of the independent random variables here. Suppose assume that x 1 is Gaussian sorry, x 1 Gaussian with zero mean and let us say variance 1 and also as will same, x 2 is also the same thing and assume that they are independent, x 1 and x 2 are independence. Then what is f of x, x 1 x, let us say x 1 x 2. So, what is this distribution? If I am saying independent, it is going to split. It is going to be x one, x one, f of 2 and x 2, what is this distribution f of x 1 of x 1?

Student: Gaussian.

Professor: Gaussian, we know its formula and for this also we know the formula. Now, this quantity, now, what is x 1 and x 2 here? If I tell you y, y is now for me what? R and theta, if I tell you that what is this x 1 is going to be? I have already written what is that?

(Refer Slide Time: 39:51)




Now, so finally, what is the formula is going to look like for f of y for some pair r, theta what is this? F of x 1 r cos theta, f of x 2 r sin theta and what is the determinant? The fourth x 1 of the fourth textbook, r sin theta and what is the determinant? R, Can you quickly plug in, the Gaussian distribution here and see what is that you are going to get?

Student: R by e to the power...

Professor: R by 2 pi e to the power, so it is going to be r square cos square theta plus r squared plus sine square theta. That is simply going to be r square by 2.

Student: Simply I was saying r square by 2.

Professor: Excuse me. So it is fine, so if you notice this we have already come across this distribution, where is where was that and we gave a name for this. So if you look your last distribution in the continuous, so this is relay distribution. But there you will have another term that we have a rated for more general case but if you set sigma square equals to one, this is what you are going to get. So this is relay with parameter what sigma square equals to?

Student: One.

Professor: One, so that that time when we discussed about this relay distribution, we said that this is going to be the envelope of a some of random variables. So if you now look into this, this is exactly what we mean by envelope that time you just take the squared sum, and look at the square root of that this is what envelope. So, this is how like the sum of the Gaussian squared and if you take the square root, we recover the relay distribution. So we will say again more things can be computed, depending on what is our applications. So let us stop here.