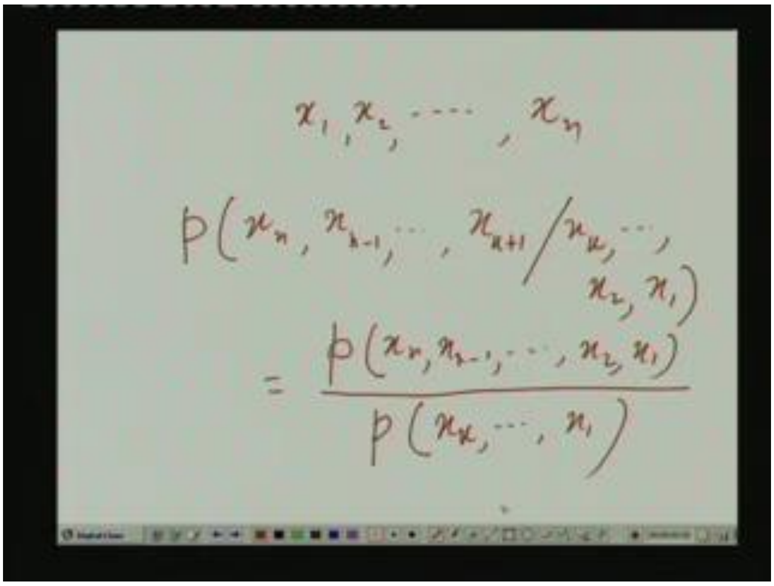
Probability and Random Variables
Prof. Dr. M. Chakraborty
Department of Electronics and Electrical Communication Engineering
Indian Institute of Technology, Kharagpur

Lecture - 25
Conditional Densities of Random Vectors

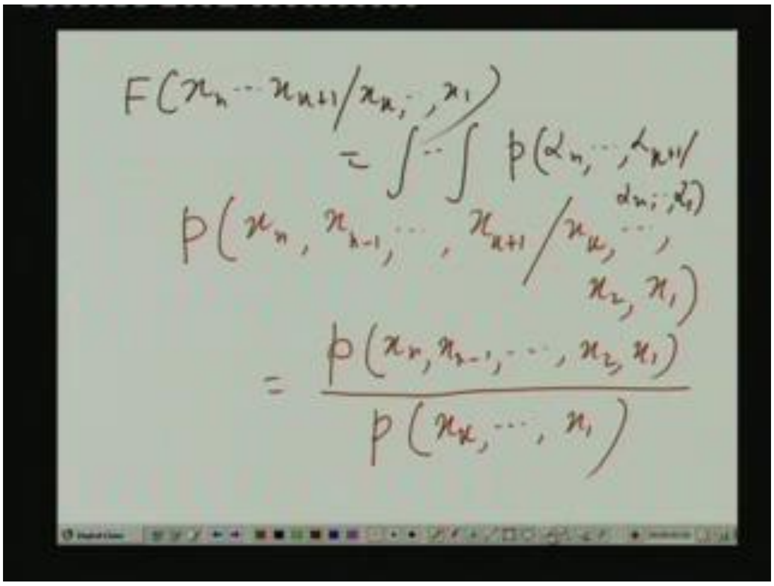
So, in the last class, we had been considering these random sequences or maybe you can call a random vector. If the sequence has just a finite number of terms, you can put them in a vector form. So, today we continue from that. We consider this topic of conditional problem density of a random vector.

(Refer Slide Time: 01:19)



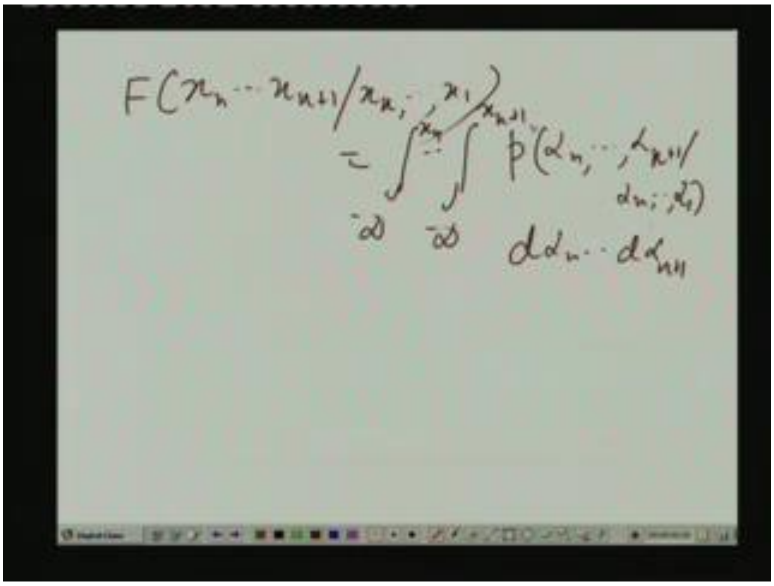
Suppose, we have got the sequence, then this thing that is given the values of x 1 x 2 up to x k condition to that what is the probability density of x n x n minus 1 up to x k plus 1? This before is nothing but the ratio of two joint densities. One is there is a total joint density of all the n variables divided by the joint density of the variables which are conditioned here, that is x k down to x 1 that you can easily define using this.

(Refer Slide Time: 02:50)



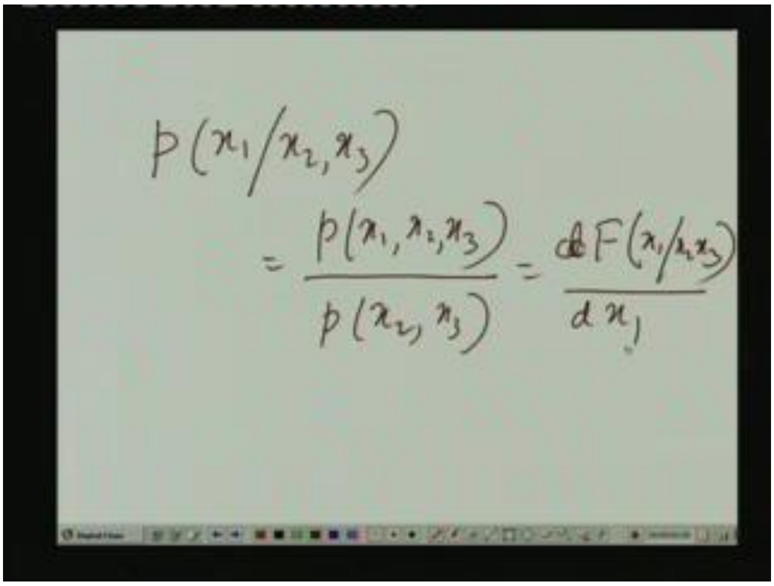
What is the conditional distribution? After all you have to just integrate that is the conditional distribution is nothing but just a integral that is you have to just integrate this conditional density function with respect to these variables x n, x n minus 1 down to x k plus 1 over their entire range. No sorry from minus infinity up to the value x n to x k plus 1. So, you just change these variables, may be you say alpha n. Thus, give them new names to alpha k plus 1 conditioned to alpha k dot dot dot alpha 1 an integral. So, I am running short of space here.

(Refer Slide Time: 04:20)



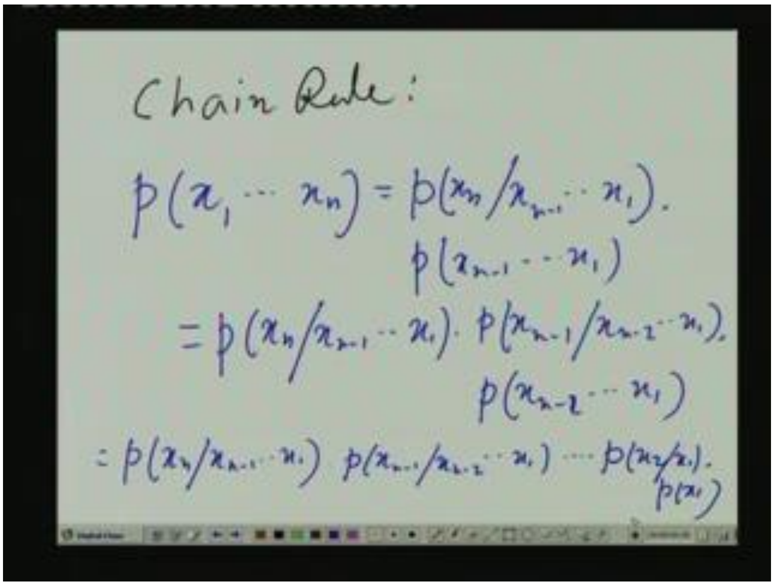
So, let me erase some part. D alpha n down to d alpha k plus 1 and the limits are minus infinity to x n, that is for alpha n dot dot dot minus infinity to x k plus 1. That is we are doing nothing new. We are only extending our concept of conditional probability density and distribution function from one variable and then two variables to in general a set of n variables is nothing basically principle remains same.

(Refer Slide Time: 05:14)



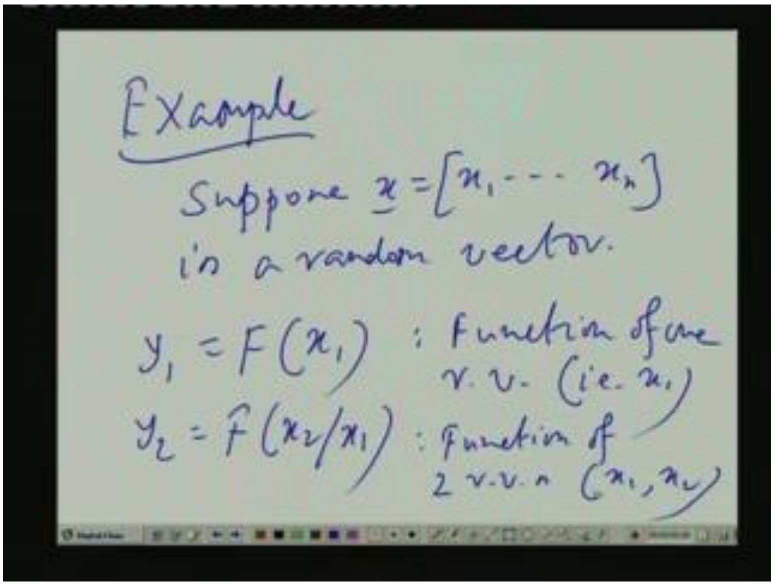
Just for an example, you can say that p x 1 by x 2 x 3, this is nothing but p x 1 x 2 x 3 divided by p x 2 x 3, and this probability density is nothing but like its integral with respect to x 1 was giving you the conditional distribution. That means, this density is nothing but the delivery of the distribution with respect to x 1, that is you can also write this as dF d F x 1 by x 2 x 3 d x y. We can form chain rule using this.

(Refer Slide Time: 06:15)



We can now write like this. First isolate xn p of xn subject to or condition to the other ones multiplied by the corresponding joint density of these remaining ones. Then, this again you express like this that is these remains as before. The second term, you express like this conditional density of x n minus 1 subject to the rest that is xn minus 2 downward to x 1 times the joint density of xn minus 2 up to or down to x 1. Likewise if you continue, finally you get this dot dot dot p of x 2 by x 1 into p x 1. This is the chain rule. It is very useful at times.

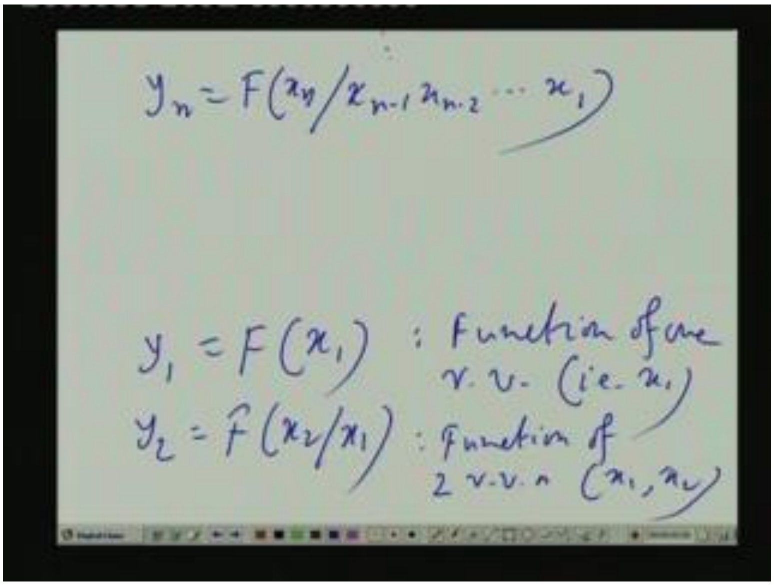
(Refer Slide Time: 08:32)



We will take an example to show how this rule sometimes becomes useful. Suppose it is a random vector. So, it has got some joint density and then, conditional density is conditional distribution and all those things as defined just a while back. Suppose we form this quantity first y 1 which is nothing but the distribution of x 1. Now, my claim is y 1 is a random variable. After all F of x 1 is nothing but a function of x 1 or whenever x 1 takes a value, you get a value of this function F x 1 and that is assigned to y 1. So, if x i mean anytime x takes the value x 1 F of x 1, that means what is the total probability of the random variable x 1 taking values from minus infinity up to that. So, you get some value. Next time, x 1 value changes. So, F of x 1 also changes so on and so forth. Obviously, y 1 is a random variable.

What is the function of only one random variable? Random variable is x 1. Then, define y 2 as conditional distribution of x 2 given x 1. Again sum for any two values for any specific values of x 1 and x 2, you will get some value for this distribution function and give it to y 2, but x a at x 1 and x 2 changes with respect to experiment. The value y 2 also changes. That means, y 2 is a random variable, but this term is a random variable of the function of two random variables x 1 and x 2.

(Refer Slide Time: 11:02)

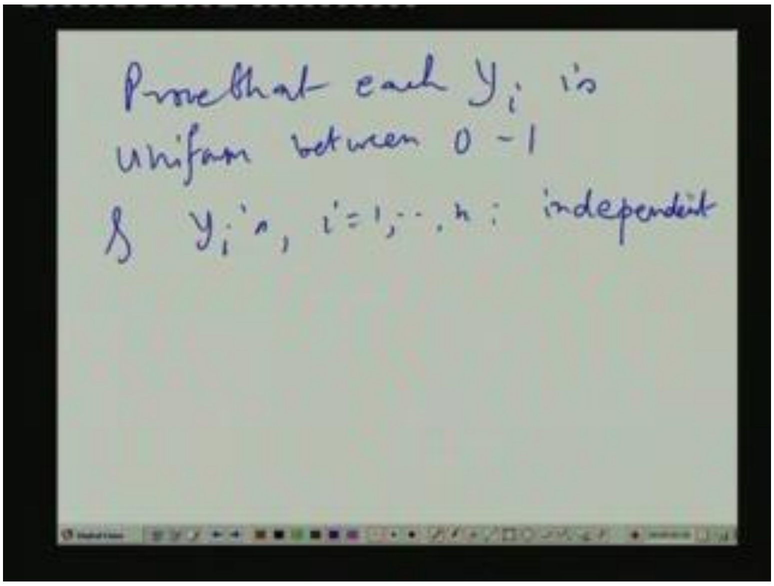


Likewise we define dot dot dot y n. So, you get n random variables. First one is a function of only one random variable that is x 1. Second one is a function of two random variables that is x 1, x 2 dot dot dot. Last one is a function of all the n random variables x

1 up to x n. We will now show that these random variables are usually independent, statistical independent and each is uniformly distributed between 0 and 1.

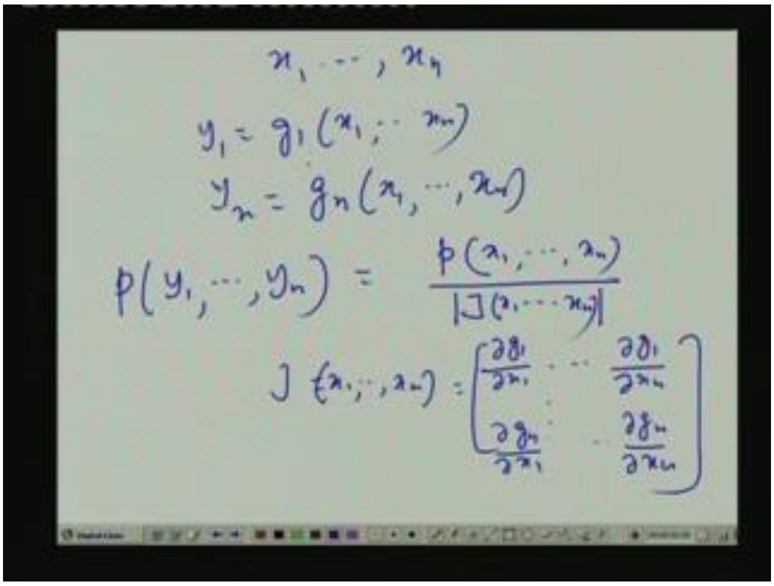
Now, obviously any distribution function, its minimum value is 0 and maximum value is 1. So, they are contained within the range 0 to 1. There is no doubt about it, but what is important is that irrespective of I mean the probability density or the distribution function for this random vector, it could be Gaussian. It could be anything irrespective of that. These random variables will always be statistically independent and each one is uniform between 0 to 1. This is what we have to prove. This is very interesting and useful result.

(Refer Slide Time: 12:31)



We will prove that each F i, sorry uniform means uniformly dependent between 0 to 1, and y i is i equal to 1 to n are independent. This is what we will prove. Now, we can just recap a little bit. One particular technique that is given a set of random variable and a function of the random variable, say g of x 1 to x n which I can call it y. What is the probability density of y given a set of functions like that?

(Refer Slide Time: 13:56)

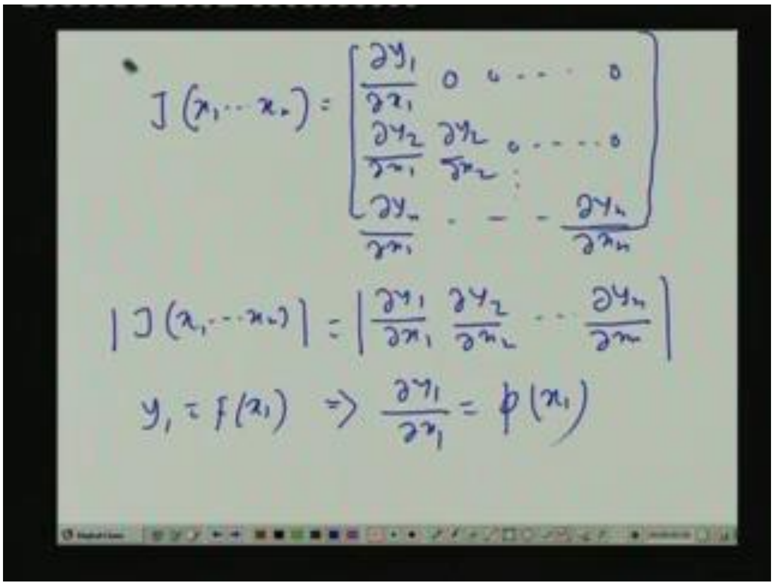


Suppose you are given x 1 to x n and y 1 as some g 1 x 1 to x n y 2 y n is say g n. So, it is a n functions. So, what is the probability density of joint density of y 1 to y n? First we said that you want to find out the joint density of these at a particular value of say y 1, particular value of y 2 and particular value of y n. First step is you put those values in these equations for y 1 up to y n and solve. If you get a solution x 1, if you get any solution, if you get no solution say if suppose you cannot find out any x 1 to x n for with these equations are simultaneously satisfied. That means, this is an impossible case because you cannot really get that kind of solution for y 1 to that kind of result for y 1 to y n in practice because no x 1 to x n will give you that output that combination of y 1 to y n. So, the probability density will be 0.

On the other hand, suppose by putting your specific values of y 1 to y n in these equations, simultaneous equations, you can find and you get a unique solution. Suppose solution is x 1 to x n, then the probability density will be p. This is the probability density joint density of this variable x 1 to x n divided by determinant and not only determinant, magnitude of the determinant that is with the plus sign of the Jacobean, just a minute. J x 1 to x n, where J x 1 to x n is this matrix del g 1 del x 1 dot dot del g 1 del x n dot dot dot dot del g n del x 1 dot dot dot del g n del x n.

So, order of the rows is not important because your determinant is invariant to the order and if you have got many solutions for a given set of y 1 to y n, if you get many solutions, then you simply have to add terms like this that is the density function at one set of solution divided by the corresponding Jacobean determinant positive determinant for the Jacobean. Again same thing at another solution and like that add them. This is just for the recap. Now, we use this as an example that we considered.

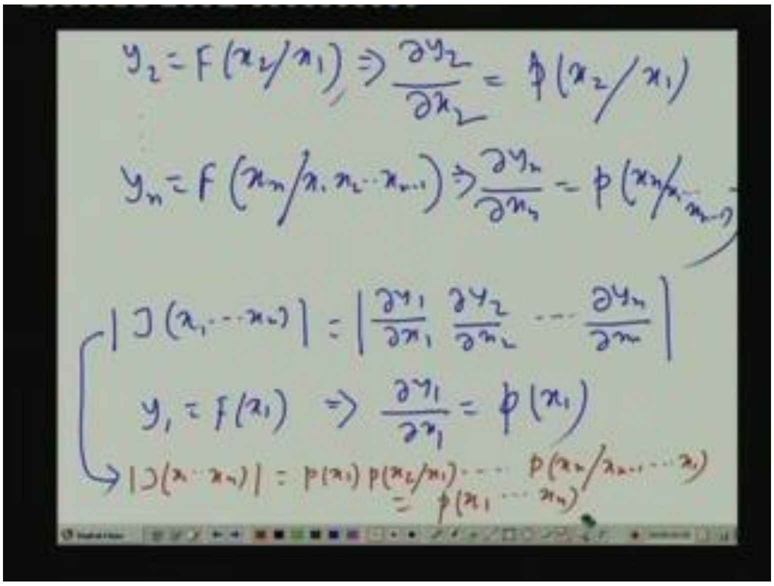
(Refer Slide Time: 17:33)



Firstly, we quickly write down y 1 was only a function of x 1. It was not a function of x 2 or 2 x n. So, therefore, that J first is del y 1 del x 1 and then, zeros because y 1 does not depend on x 2 to x n. So, corresponding derivatives are 0. Then, y 2 del y 2. It is a function of two random variables x 1 and x 2. So, del x 1 del y 2 del x 2 and then, zeros dot dot dot dot finally, del y n del x 1. It goes to the last term del y n del x n. So, it is an lower triangular matrix. So, determinant is given by just the product of the diagonal elements because upper half is 0. So, that means the in all derivatives, no derivative is negative. Also, we can easily see because every function, every y 1 y 2, they have distribution functions and distribution functions are usually positive.

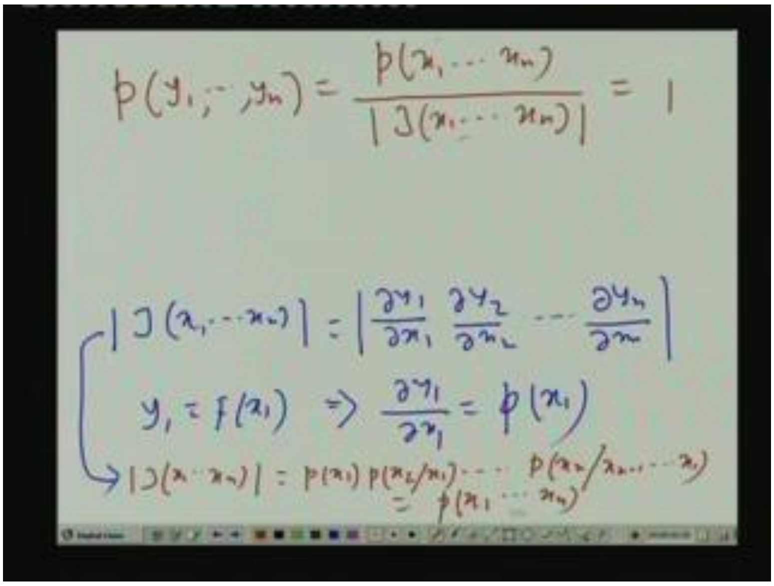
In any case, we will take up determinant and put a positive sign. So, sign is not important here. So, this will be given because of the magnitude of this product. Now, see what is y 1 or y 1 was F x 1. That means, what is del y 1 del x 1. If you differentiate it with respect to x 1, you simply get the density. Now, from distribution we will go to density. So, you get p x 1.

(Refer Slide Time: 20:00)



What y 2 is, F of x 2 given x 1. That means, del y 2 del y 2 del x 2. As I said all are I mean all these derivatives are positive for 0. It is because they are not negative because there distribution, I mean they correspond to distribution function of some variables and differentiation is with respect to that variable like x 2 here. Then, if you differentiate it with respect to x 2, you get the conditional density p x 2 by x 1 so on and so forth. Finally, we know y n leads to the conditional density of x n given x 1 to x n minus 1 and that product will be in this case is x 1 to x n. This Jacobean determinant will be what? P x 1 p x 2 by x 1 dot dot dot up to p x n by x n minus 1 up to x 1 and this is the chain rule that we discussed previously. So, this is nothing but the joint density of x 1 up to x n.

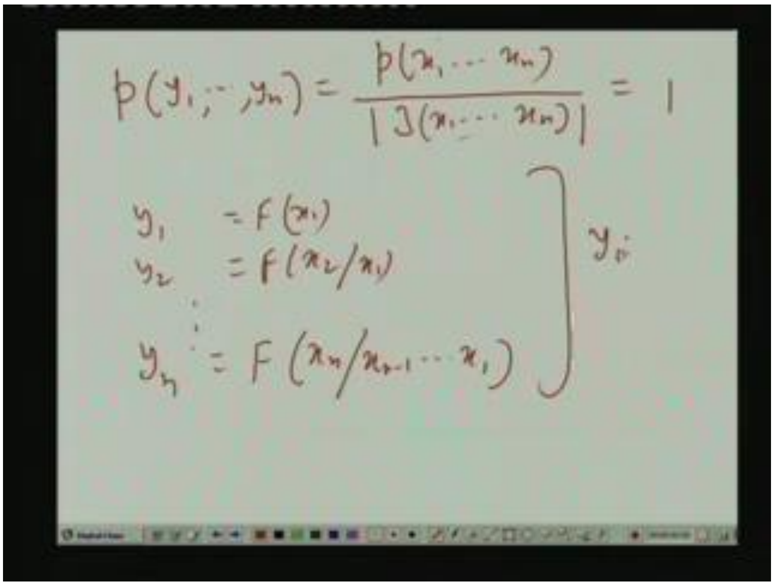
(Refer Slide Time: 22:24)



That means, what is p y 1 up to y n is nothing but p 2. P’s are not same. This is the probability density of y 1 to y n. It is a different function, and when I say p x 1 to x n, it is again a different function though same symbol p is used. Please do not think that function p is same and once put in y 1 to y n and next time, put in x 1 to x n, the two functions are different. In fact, ideally I should put a subscript here x and subscript here y, so that p x p y indicated different function, but assume that by now we have attained sufficient maturity. So, we will not be thinking that we have got only one probability function, sometimes having the variable y 1 to y n, sometimes x 1 to x n. That is not the case.

The two probability density function or the joint density function, they are different function altogether. Anyway, this was the formula. By the way first let us evaluate what it is. So, if you substitute for this determinant of this Jacobean, you will get the joint density function is positive and the ratio cancels and you get 1.

(Refer Slide Time: 23:56)



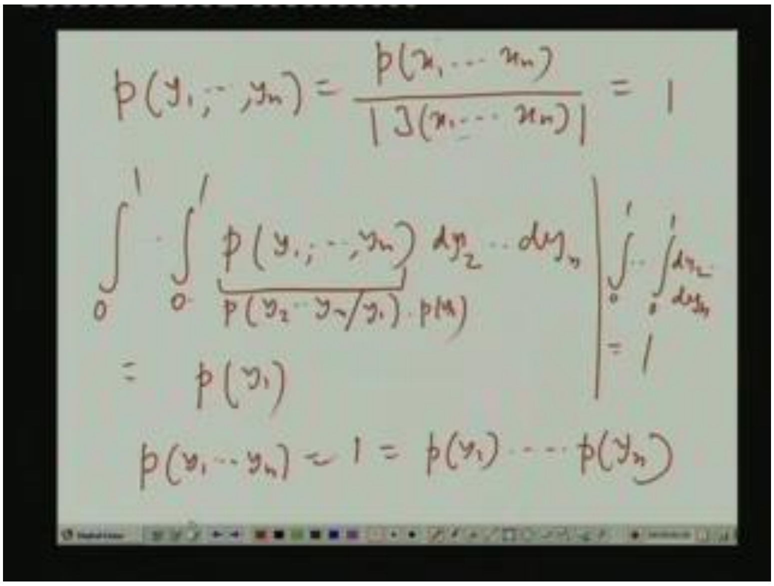
Now, I forgot to mention that consider y 1 n as before that was F x 1 y to n that was F x 2 by x 1 dot dot dot dot y. What y? Sorry there is no n. Now, each of these functions mind you y 1 to y n, they are random variables. Now, consider any y i. Y i can take value between 0 to 1. So, outside that region 0 to 1, its probability is 0, that is y i taking value greater than 1 or less than 0 has probability 0. Let us group up all; that is y 1 y 2 up to y n. So, you can assume n dimensional hyper cube, where each axis, one axis corresponds to y 1 and another axis corresponds to y 2 dot dot dot up to y n, and within each axis, you take out a segment 0 to 1. So, you get a hyper cube. Each side is from 0 to 1 n dimensional. So, each axis, one axis corresponds to y 1, another axis corresponds to y 2 dot dot dot y n.

So, this y 1 to y n, it remains. It takes values within that hyper cube with some probability and they are taking values outside the hyper cube that is 0. Further, if you then consider in a particular value of y 1, particular value of y 2, particular value of y n within that hyper cube, it corresponds to an unique x 1, unique x 2 up to unique x n obviously because consider this equation y 1 equal to F x 1. For a particular y 1 x 1 is fixed that is from the nature of the distribution because for any x 1 F x 1 cannot take two values. If F x 1 gives the total probability of the random variable, x 1 taking values from minus infinity up to x 1. So, that has got only one value. So, if that value is given corresponding x 1 is known, and then if x 1 is known and y 2 is F of x 2 giving x 1, there is that x 1, then again x 2 is fixed. If y 2 is known, there cannot be 2 x 2 giving rise to the same y 2 so on and so forth.

So, within that hyper cube each point y 1 up to y n that corresponds to a unique choice of x 1 to x n, if you give some specific values of y 1 to y n and solve for corresponding x 1 to x n, you get only unique solution within the hyper cube outside the hyper cube. No solution. So, their probability is 0. Inside the hyper cube, you get a unique solution. So, therefore, this is I mean you just have to do this once probability density of x 1 to x n divided by the Jacobean determinant and that is equal to 1. So, I repeat that joint density is 0 outside the hyper cube and within the hyper cube at any point within the hyper cube, it is 1. That is interesting, right.

So, there is a hyper cube in a dimensional plain, in dimensional space. Each axis is given by the random variable ith axis given by the random variable y i i equal to 1 to n. Hyper cube each side is from 0 to 1. In that hyper cube at any point if you go inside the hyper cube, the corresponding joint density of y 1 to y n that is equal to 1 and that any point outside the hyper cube joint density is 0.

(Refer Slide Time 28:35)

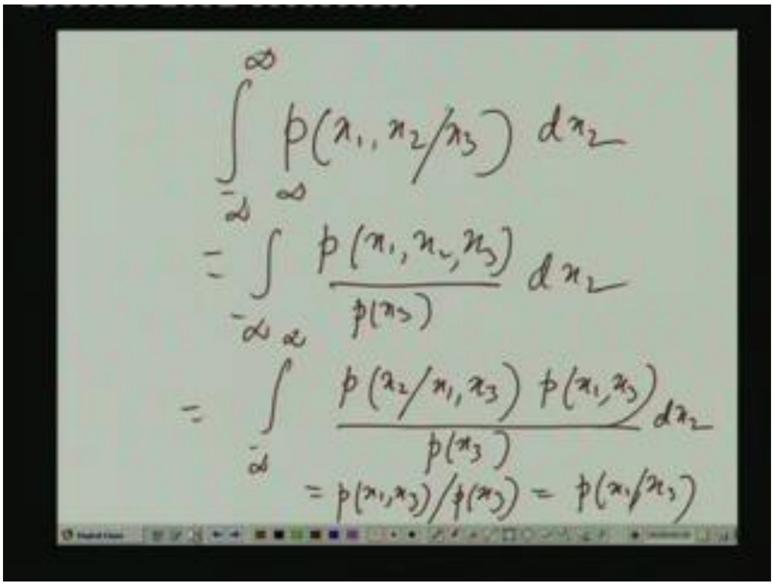


Now, what does that mean? That means suppose we consider this function integrated with respect to say y n y 2. Initially you can integrate from minus infinity to infinity, but since outside the range 0 to 1, probability density is 0. It is enough that we integrate from 0 to 1. What does this give? Actually you can write it. You can write this thing as p of y

2 up to y n condition to y 1 times p y 1. So, p y 1 comes out of the integrals and the remaining thing when integrated that will give rise to 1 because total probability is 1. So, this give rise to p y 1, but if i do the same think on the right hand side that is that we have got one. So, if you integrate this one, how many times if you integrate this 1 from 0 to 1 up to 0 to 1, and with respect to the same variable d y 2 dot dot dot d y n, every integral gives rise to 1. So, 1 times 1 times 1 times dot dot dot n minus 1 times is 1.

So, you see p y 1 equal to 1. So, y 1 is a random variable. That is unique and that has value 1 within the range 0 to 1 and outside the range, it is 0. It is similarly for y 2 and similarly for y 3. So, each of them, this part we have proved that each random variable here y 1 to y n is uniformly distributed random variable within the range 0 to 1 and outside that it is 0. The fact that the independent is very simple, you can write down easily since p y 1 up to y n is equal to 1. You can also call it, you can write y 1 like this. After all each individual density is 1. So, 1 times 1 times 1 times dot dot dot dot 1 which is equal to 1. That means joint density is nothing but product of individual density. So, they are statistical independent. So, that is the proof. Some interesting things we can observe, now because these are very useful in practice.

(Refer Slide Time 32:16)

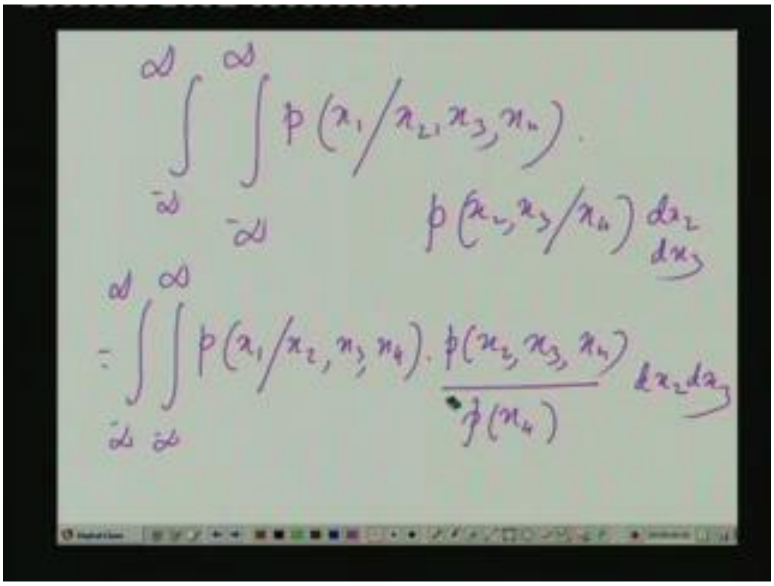


Suppose you got just three random variables, x 1 x 2 x 3 and you are doing some integration like this. You are given suppose we first integrate and then we will see. This is given and you are integrating with respect to x 2 from minus infinity to infinity. What do you get? Well, we can always write it like this and then, this we can write as. So, only this quantity depends on x 2. So, when this is integrated with respect to x 2 from minus infinity to infinity, that is equal to 1 because it is a conditional density of x 2 subject to some x 1 and x 3, but if x 2 is moved from minus infinity to infinity, and this probability density is integrated, you obviously get 1.

So, you get this ratio p of x 1 x 3 divided by p x 3 that is you get p of which is nothing but p of x 1 by x 3. That means, suppose in the beginning you are given a conditional density like this p of x 1, x 2, stroke x 3 that is given x 3, the joint density of conditional joint density of x 1 x suppose it is given and you want one variable to be eliminated. We want say x 2 to go. What did you have? You have to integrate this with respect to that variable. So, you just took that density function p of x 1 x 2 given x 3, but integrate it with respect to x 2 only over the entire range, immediately you get p of just x 1 by x 3.

I repeat again if you have got a density function conditional density function and this is a conditional line, the slash, to the left of the slash, there are some variables and you want to remove say one of them or we want to remove say a few of them. Then, you just integrate this density function with respect to those variables for over their entire range say minus infinity to infinity those variables which we want to remove. You then get the net thing as a conditional density of the remaining variable.

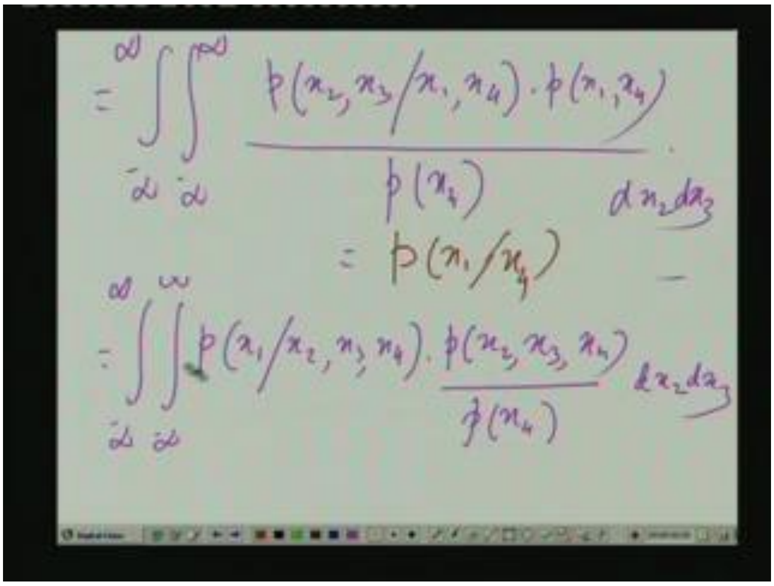
(Refer Slide Time: 35:33)



On the other hand, suppose you have got a thing like this. Now, suppose four random variables x 1, x 2, x 3, x 4. Now, suppose you are doing this integral x 1 stroke. There is a slash x 2, x 3, x 4 into p of and integrate with respect to x 2 and x 3. What do you get?

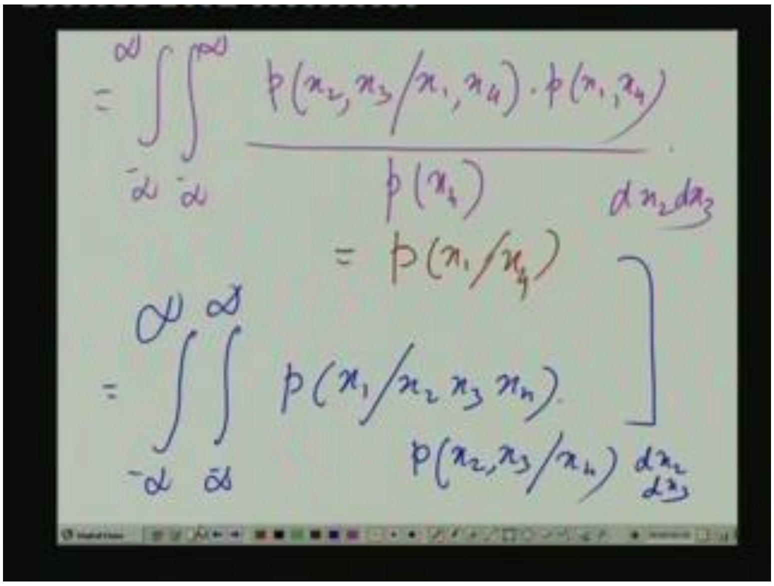
Think for a minute what do you get. Its purpose I will tell later, but suppose you carry out the integral like this. Well, you can write like this, that is this conditional density can be written as p of joint density of x 2, x 3, x 4 divided by the probability density of x 4. You combine the two. This conditional density of x 1 giving x 2 x 3 x 4 multiplied by the conditional density of x 2 x 3 x 4. So, that will give rise to the joint density of all the upper variables x 1 x 2 x 3 x 4. I will delete this part.

(Refer Slide Time: 37:29)



That join density you can then write as p of x 2 x 3 that is the variables with respect to which we are doing the integration that is x 2 and x 3. So, x 2 x 3 slash x 1 x 4 multiplied by p of x 1 x 4 that will give rise to joint density, and this is divided by p of x 4 as before d x 2 d x 3. Now, in this integral you see the first function p of x 2, x 3 subject to x 1 x 4. Only that depends on x 2 x 3. Now, if that is integrated with respect to x 2 and x 3 from minus infinity, you get 1 because this conditional density of x 2 and x 3 x 1 x 4 are conditioned, but then x 2 and x 3 are moved over all possibilities from minus infinity to infinity. So, the total value has to be 1. So, 1 times the remaining things that is p of x 1, x 4 divided by p of x 4, and p of x 1 x 4 divided by p of x 4 is nothing but p of may be sorry. This is then equal to what we rewrite, what we started with. We started with this.

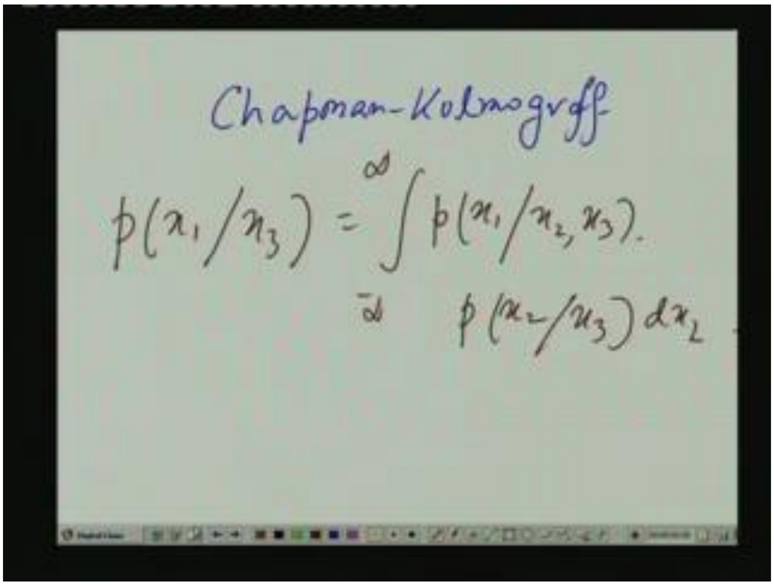
(Refer Slide Time: 39:33)



We were given x 1 probability density conditional probability density x 1, given subject to the other three variables. If from this suppose this conditional density is given p of x 1 slash the other variable and if here you want to eliminate some of the variables to the right of the slash that is the variables which are conditioned to the right of the slash. If we want to remove some of these variables in this case x 2 and x 3, then first you multiply this function by the conditional density of these variables which you want to eliminate subject to the remaining ones. In this case, x 4 and integrate with respect to these two variables this is what we started with. So, these two things are same.

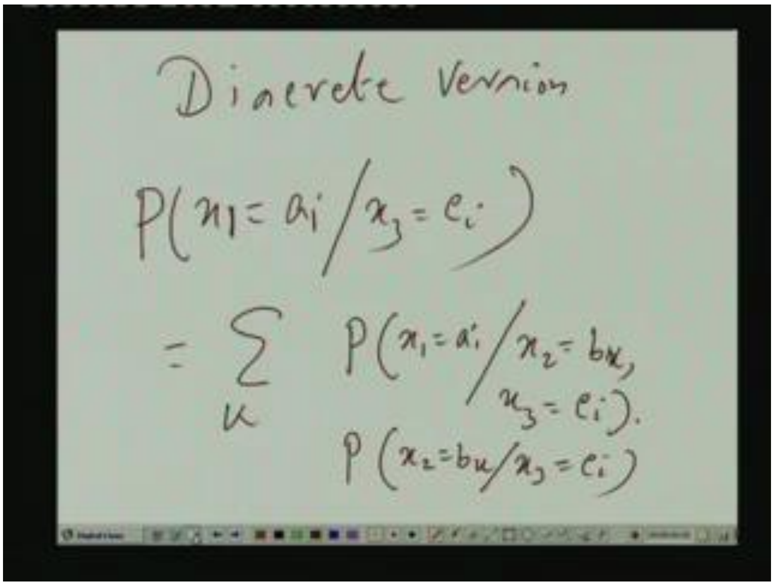
So, I repeat if you are given conditional density function, where within bracket you have got a slash sign like this to the left, there are some variables to right, and from the right you want to eliminate some of the variables as in this case x 2 and x 3. Then what we do? We multiply this function by the conditional density of those variables to the right of the slash x 2 x 3 which you want to eliminate condition to the remaining one again to the right of this slash, in this case only x 4. After doing that integrate it with respect to these variables which you want to eliminate, that is x 2 x 3, then you get p of x 1, x 4. This is a very interesting variable, because this comes very handy in practice.

(Refer Slide Time: 41:39)



In fact, a particular case arises often in practice and that is called Chapman Kolmogroff. In this case, often this is written as stroke x 2 x 3 multiplied by p of x 2 by x 3 d x 2. Now, we have to consider its discrete version.

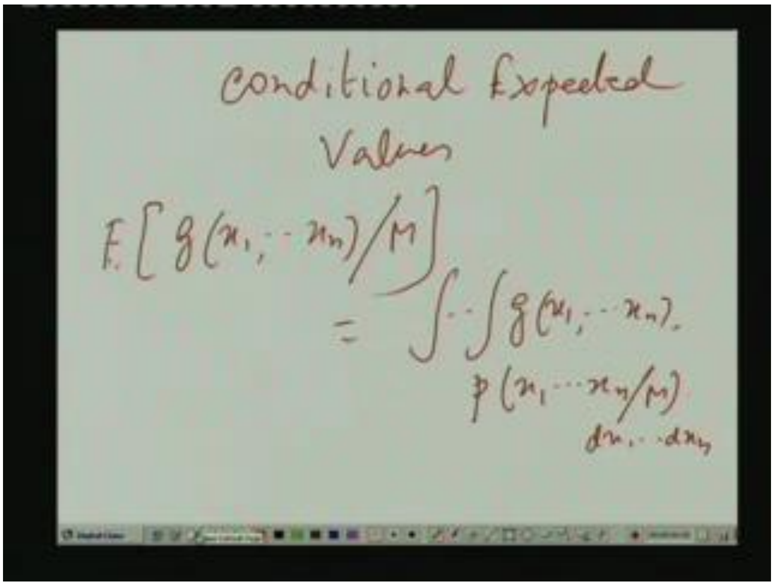
(Refer Slide Time: 42:27)



In the discrete version, the random variables x 1 to x n, they are continuous. They take discrete values, some finite set of values that it can take, but discrete. In that case, you can usually extend this logic and write that. Suppose this conditional probability that x 1 takes may be say xi take some value ai x 1, say x 1 take say a particular value a 1 or ai, all right. X 1 for x 1, there is a set we call it alphabet a 1, a 2, a 3 dot dot dot dot out of which x 1 is taking value a i conditioned to another random value say x 3. It takes values from another set c 1, c 2, c 3 dot dot dot, but right now it is conditioned to take the values say c i.

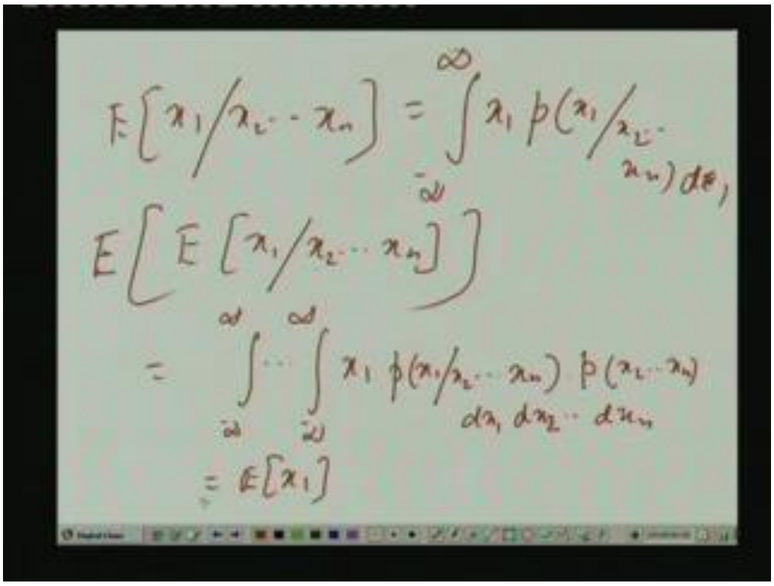
This analogously you can write and you can prove also analogously as p of x 1 equal to a i subject to two random variables x 2 taking say some b k and x 3 taking ci. X 2 is another random variable which takes values from a set b 1 b 2 b 3 dot dot dot. So, it is b k here. This multiplied by probability of x 2 taking b k divided by x 3 taking c i and like we did integration with respect to x 2 here, it is summation of k. So, x 2 can take b 1 b 2 b 3. For each case, this product is evaluated and added. A very simple expression of this is the conditional expected values.

(Refer Slide Time: 44:39)



Suppose there is some condition M, and we want to find out expected value of a function of random variable say g x 1 to x n subject to M. Obviously this is nothing but integral of this function multiplied by the conditional density of this subject to M integral. This is expected. Obviously this follows from definition, but as a special case we can now consider g to be x 1 itself.

(Refer Slide Time: 45:38)



So, excepted value of say x 1 given x 2 to say x n. Obviously, that will be multiplied by this conditional density of x 1 subject to integral with respect to x 1. Now, you see one thing here, we have assumed some specific value of x 2 to x n subject to that x 1 can take any value, and we are taking the mean. This is given by this conditional expected value.

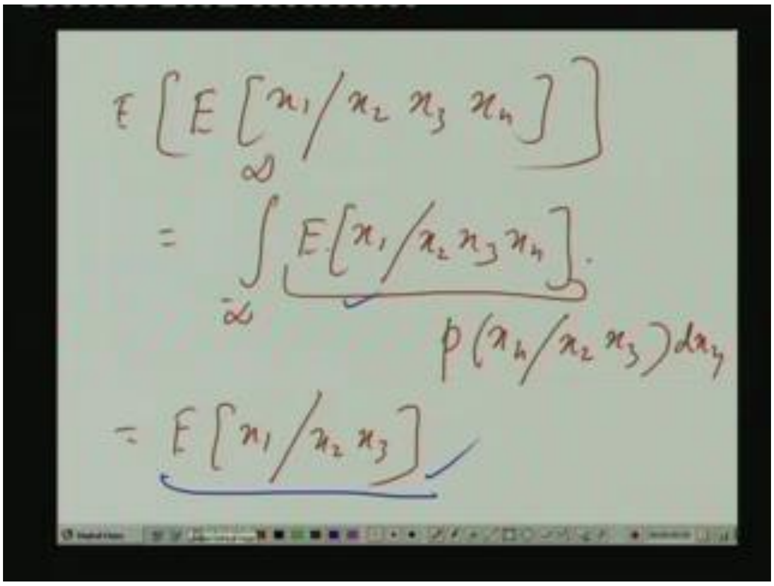
Next time again we can change the values for x 2 x n. We can give some other value here and again evaluate this. So, you will get another value for this overall mean and likewise. So, that means if you now consider both x 1 and also x 2 to x n as variables here, not constants, then this entire thing, entire expected value, conditional expected value is also a random variable, but variable. I mean random variable, its function of x 2 to x n, it is a random variable and its function of x 2 to x n. Is it not?

Here what we have done? We chose a particular value of x 2, chose a particular value of x 3. There is constant. Add another constant value for x n and put them here subject to that find out the expected value of x 1. So, integrate that x 1. I mean multiply x 1 by the corresponding conditional density and integrate. Next time suppose the values for x 2 to x n change, you give some other value. So, obviously you get another value of this mean so on and so forth. So, that means that this entire thing is a random variable. If you now allow x 2 to x n to vary, that is if you now take it as a function of x 2 to x n, then this random variable.

So, what is then expected value with respect to x 2 to x n of this mean? That is first you found out the mean and then, we are saying that it is a function of the other random variables x 2 to x n. You give them some specific values, you get one value of the mean, you give another specific values, another set of specific values for x 2 to x n and you get another mean and likewise. So, only the expected value of that is if I do not want to take the expectation over this x 2 to x n, that means this entire thing now is to be multiplied by the joint density of x 2 to x n and integrated, but if you multiply this integral with respect to p of x 2 to x n, you can easily see that it becomes a joint density of x 1 to x n because after all it is conditional density of x 1 subject to x 2 to x n.

This if you multiply by the joint density of x 2 to x n, then the total product becomes just the joint density p of x 1 x 2 dot dot dot dot x n that is multiplying x 1 and integrating will give rise to just a mean of x 1 or you will have d x 1, now d x 2 to d x n. So, this product has two probability densities. One is conditional and is joint, that is the joint density of product becomes equal to the joint density of x 1 to x n. Multiplying x 1 and if you integrate it with respect to all the random variables, obviously that will give you the mean of x 1 because joint density can be written as p of x 2 to x n subject to x 1 multiplied by p of x 1 and the first one when integrate with respect to the respective variables will become 1. So, we will be left with x 1 multiplied by p of x 1 integrated from minus infinity to infinity with respect to x 1. That will give rise to the expected value of x 1. Since we have done similar things in the past, I am not getting into those lines.

(Refer Slide Time: 51:00)

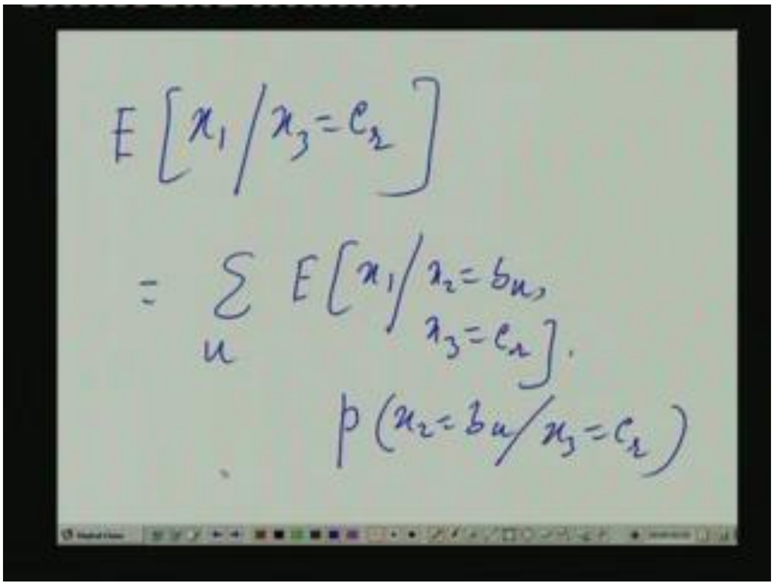


Similarly, just as an example. Suppose we are considering a case of four random variables, and you found out E of x 1 subject to x 2 x 3 x 4. Now, suppose here when you just write this, it means that we are giving some specific constant values for x 2, specific constant value for x 3, same for x 4 and finding out this mean. Now, suppose here I am not changing the values for x 2 and x 3, but I am changing only x 4 from case to case, and obviously the overall mean changes. So, this mean is then a function of x 4. So, what is the expected value of this mean with respect to x 4?

So, that means I have to take this, I have to take this, I have to take this and multiply by here. It is interesting. Multiply by what? Will it be just the probability density of x 4? The answer is no because in this entire business, I am keeping x 2 and x 3 fixed subject to that I am varying x 4, right. That means I have to multiply by this and then, integrate with respect to x 4. What will this give rise to? You consider this expression. It was E of x 1, given x 2 x 3 x 4 and there I took the further average over x 4. So, I am left with nothing but just E of x 1, given x 2 x 3. So, that means given a conditional expected value, this say consider this and this, this one and this one. Given a conditional expected value of x 1 subject to a set of random variables, that is there is a slash to the right of the slash, there is a set of random variables.

If you want to eliminate say a particular random variable or a few or more random variables like in this case x 4, what you have to do is you take this conditional expected value, but multiply it with respect to the conditional probability density of those variables which you want to eliminate. In this case, x 4 subject to the remaining ones to the right of this slash, that is a x 2 x 3 and integrate with respect to those variables like x 4.

(Refer Slide Time: 53:59)



It can be easily extended to the discrete case, that is E of say x 1 subject to say x 3 taking a value say c r because x 3 takes values from a set c 1 c 2 dot dot dot dot. So, this is specific now. With respect to this, this is called the condition. What is the mean of x 1? You can easily extend this logic. You can make a summation over k E of x 1, given x 2 equal to b k, x 3 is equal to c r multiplied by the probability of x 2 is equal to b k divided by x 3 is equal to c r and summation over k like earlier, we are integral. Now, you have summation.

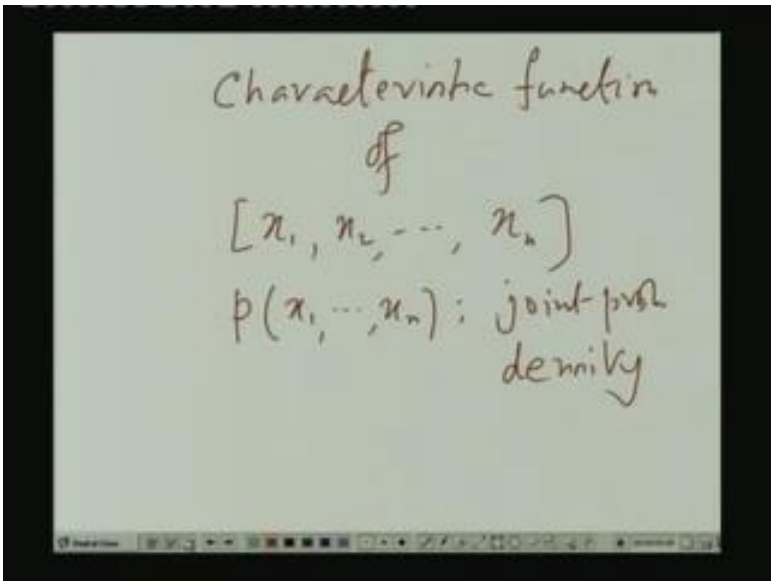
So, that is all for today. So, in the next class, we will be considering the characteristic functions for this general case of random vectors. We will derive some interesting properties. We will march towards what is called central limit theorem from that, so all that is for next class. Thank you very much.

Preview of Next Lecture
Transcriber's Name: Mahesh
Probability and Random Variables
Prof. Dr. M. Chakraborty
Department of Electronics and Electrical Communication Engineering
Indian Institute of Technology, Kharagpur

Lecture - 26
Characteristic Functions and Normality of a Random Vector

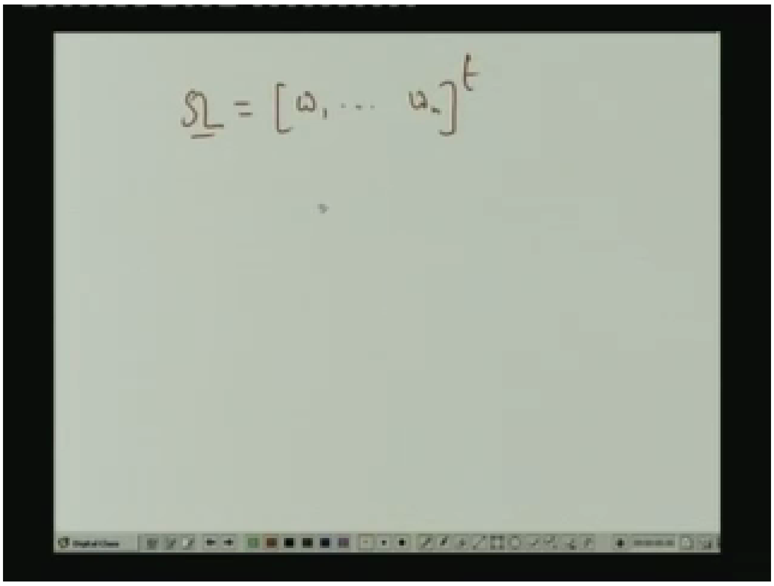
So far we have been, I mean in the last class, we have been considering random vectors. In fact, for last few lectures only we have been on this topic. So, today we will be considering characteristic functions for random vectors. See you remember I mean earlier we considered only a single random variable and with respect to a single random variable, we considered a characteristic function. At that time, it was a function of just one frequency variable omega 1. Then, we extended that to the domain of two random variables. So, that time also we had a characteristic function. It was a function of two variables, two frequency variables omega 1 and omega 2. So, now that whole approach will be generalized to a random vector that has got say n number of random variables, right.

(Refer Slide Time: 56:42)



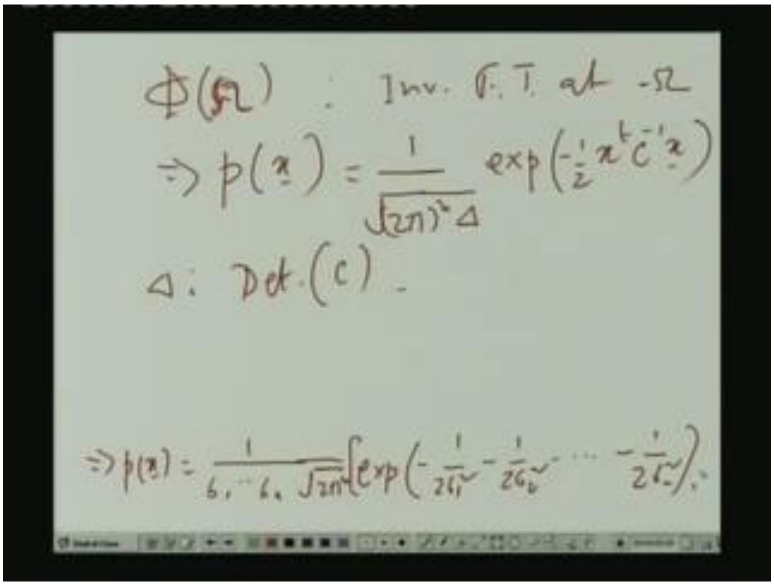
So, we will be considering random x 1, x 2 dot dot dot say x n. There are n random variables, there are jointly random, they have got a joint density that is you can say the joint density joint probability density, it is for probability. Obviously you can understand that since there are n random variables, we have n frequency variables. Now, omega 1 associated with x 1, omega 2 associated with x 2 dot dot dot dot omega n associated with x n.

(Refer Slide Time: 57:43)



So, the characteristic function I mean if I define a vector, in my case all vectors are actually column vectors, whereas in the book by Papoulis, he normally takes vectors as row vectors. There is a difference you may find, this is row. If you put a transpose, it becomes a column vector.

(Refer Slide Time: 58:09)



So, the characteristic function actually is dot sigma n 2 pi to the power n square root in to exponential minus 1 by 2 sigma 1 square minus 1 by 2 sigma 2 square dot dot dot minus 1 by 2 sigma n square. It amounts to just multiplying n individual Gaussian density functions for n random Gaussian random variables that have zero mean and variance is sigma 1 square sigma 2 square dot dot dot sigma n square. So, we stop here today. In the next class, we will talk about stochastic conversions and we got to the central limit theorem.

Thank you very much.