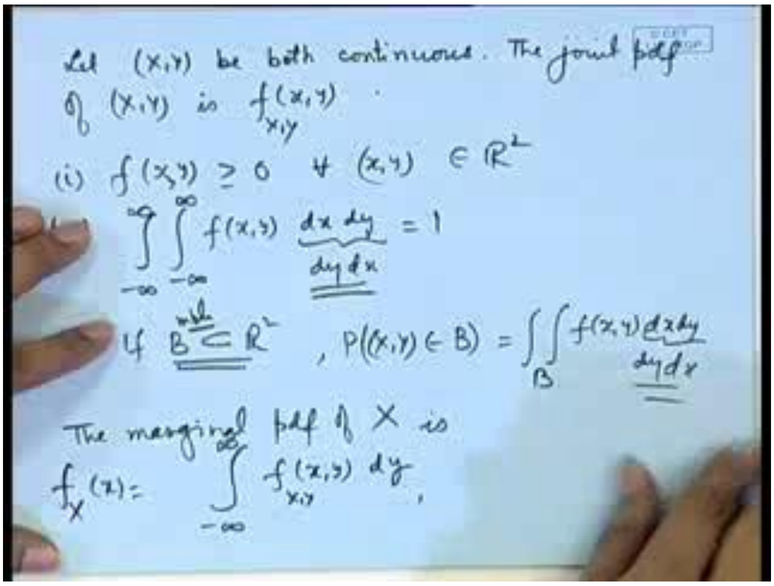
Probability and Statistics
Prof. Somesh Kumar
Department of Mathematics
Indian Institute Technology, Kharagpur

Lecture - 36
Joint Distributions – II

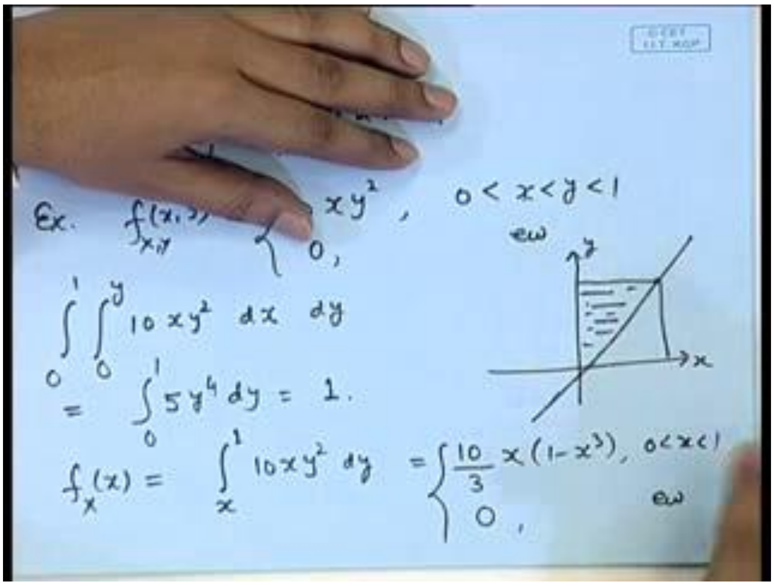
So, like in the case of the discrete random variable one may talk about the marginal distributions of x and y. So, in the case of discrete we had summed over the other variable to get the marginal distribution of one variable. In the case of continuous random variable we will have to integrate.

(Refer Slide Time: 00:41)



So the marginal distribution or marginal pdf of X, so we will denote it by say f x x that is equal to integral of f x y with respect to y over the appropriate range.

(Refer Slide Time: 01:02)

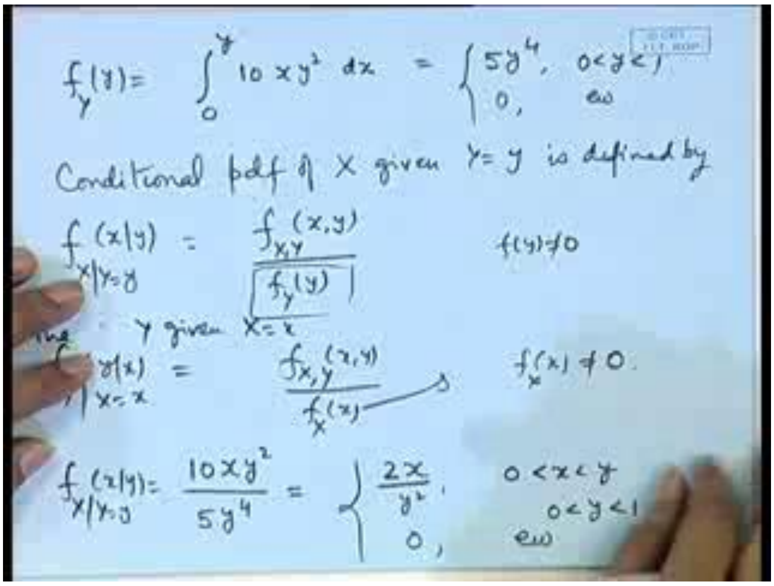


Similarly, the marginal pdf of y; f y y is equal to integral f x y, x, y d x, let us take an example say f x, y is equal to say 10 x y square, 0 less than x, less than y, less than 1, it is 0 elsewhere; let us analyze this cdf.

First of all is it a valid cdf. So, you can see that the values are non negative and if I take the integral over the full range. So, here if I integrate with respect to x first then it will be 0 to y and then the range of y will be 0 to 1. Basically the range of the distribution is defined over the interval 0 to 1 on the half side there is x less than y. So, if this is x axis this is y axis. So, we are on this side, you can easily see that this is 5 y to the power 4 d y this is equal to 1.

So, it is a valid probability density function, we may look at say marginal distribution of x, so that is obtained by integrating with respect to y. So, if we integrate with respect to y, the range of, y will be from x to 1, this gives 10 by 3 x into y cube from 1 2 a from x to 1, that is 1 minus x cube. So, the marginal distribution of x is given by 10 by 3 x into 1 minus x cube.

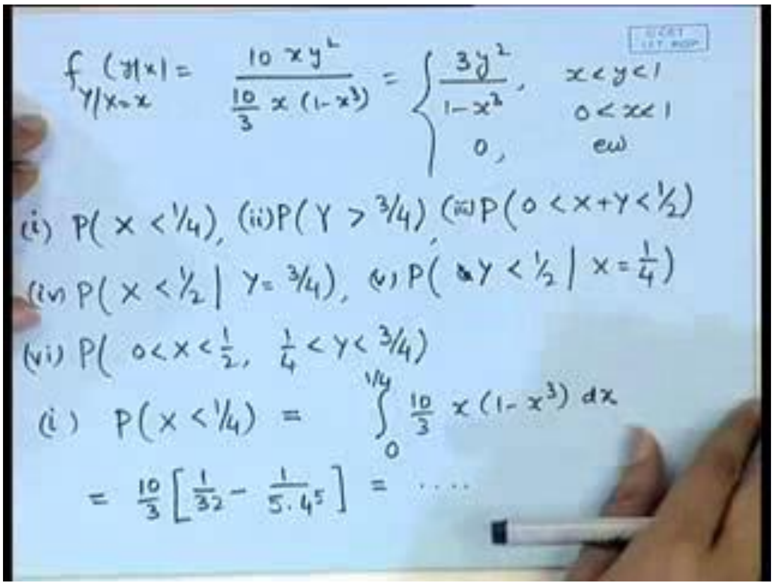
(Refer Slide Time: 03:10)



Similarly, we can obtain the marginal distribution of y, that is integral of 10 x y square d x, now here the range of the x will be from 0 to y, for a given y the range of x is 0 to y. So, 0 to y, this is equal to 5 y to the power 4 for 0 less than y less than 1 and 0 elsewhere. So, the marginal distribution of x and y are easily obtained; the conditional pdf of x given a value of y is equal to y is defined by f x given y is equal to y that is the join distribution of x y, divided by the marginal distribution of y of course, this is defined for f y non 0.

So, in a similar way the distribution of y given x that is defined as the joint distribution divided by the marginal distribution of x and of course, this f x should not be 0. So, in this problem we can talk about say the conditional distribution of x given y. So, the marginal distribution is 5 y to the power 4. So, 10 x y squared divided by 5 y to the power 4 that is equal to 2 x by y square and the range of x is 0 less than x less than y, for a value of y between 0 and 1 it is 0 elsewhere; you can check that it is a valid distribution if we integrate from 0 to y, you will get y square, so y square by y square will be 1.

(Refer Slide Time: 05:33)



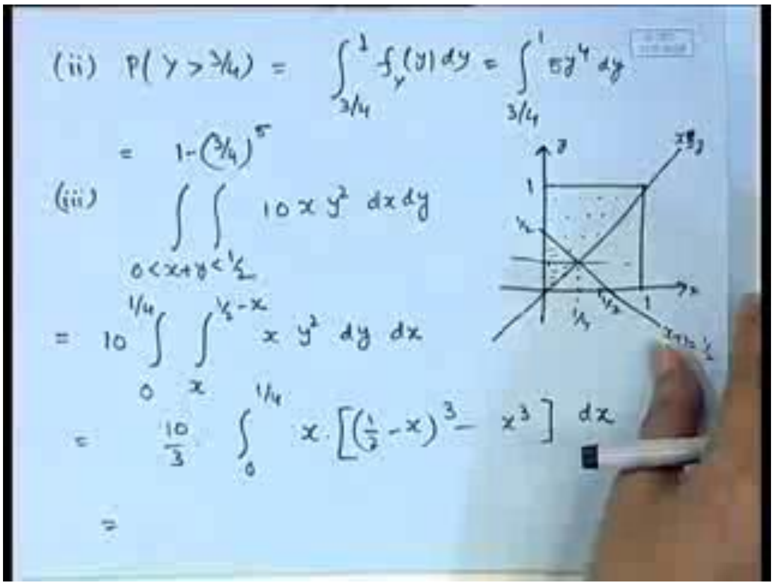
Similarly, we can define the conditional distribution of y given x here. So, that is 10 x y square divided by the marginal distribution of x, which we obtained as 10 by 3 x into 1 minus x cube. So, divided by 10 by 3 x into 1 minus x cube, that is equal to 3 y square by 1 minus x cube; here the range of y is from x to 1 for a value of x between 0 and 1 and 0 elsewhere. So, given the joint probability density function of x and y, we are able to obtain the marginal distributions of x y, the conditional distributions of x given y and y given x and therefore, we can answer all the probabilities statements regarding the joint probabilities of x y, the individual probability is related to random variables x and y are the conditional probabilities related to x and y.

So, in this particular case, let us look at some of such questions. So, suppose I say what is probability x is less than 1 by 4; what is the probability that is say y is greater than 3 by 4, what is the probability that say x plus y lies between 0 and half, what is the probability that x is less than half given y is equal to 3 by 4, what is the probability that y is less than half given x is equal to say 1 by 4, what is the probability that say 0 less than x is less than half and 1 by 4 less than y less than 3 by 4. So, these are various statements regarding the marginal conditional or the joint probabilities of x and y.

So, in the context of this particular problem, let us answer these questions. So, let us number them 1 2 3 4 5 and 6. So, what is the probability x less than 1 by 4? So, this is related to the marginal distribution of x which we calculated as 10 by 3 x into 1 minus x cube, for x lying in the range 0 to 1. So, this is 0 to 1 by 4, 10 by 3, x into 1 minus x cube d x. So, this can be evaluated easily 10 by 3, now integral of x is x square by 2, so this is giving you 1 by 32 minus integral of x to the power 5 that is x to the power 4 that is x to the power 5 by 5, so 1 by 5 into 4 to the power 5. So, this can be simplified.

Similarly, if we look at probability y greater than 3 by 4, then this can be obtained from the marginal distribution of y.

(Refer Slide Time: 09:23)



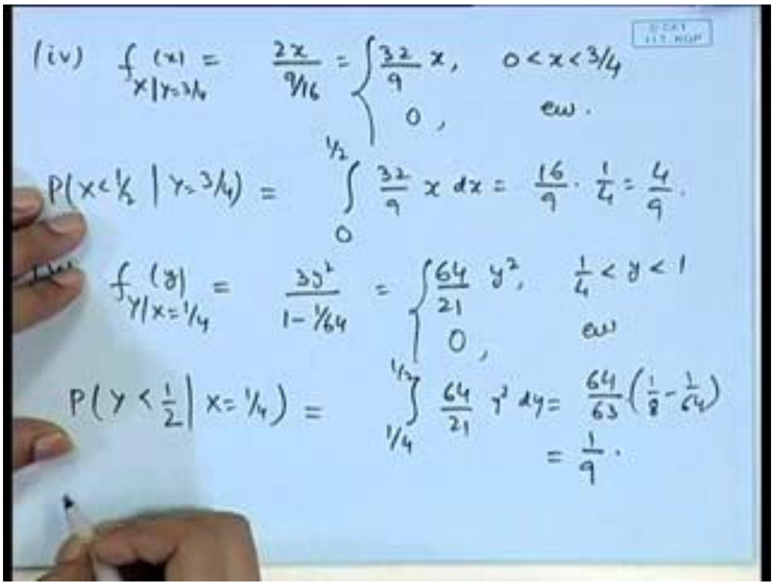
So, if we answer this question probability y greater than 3 by 4, that is integral f y from 3 by 4 to 1, because the range of y is from 0 to 1. So, this is equal to integral 3 by 4 to 1, 5 y to the power 4 d y. So, this is y to the power 5 that is 1 minus 3 by 4 to the power 5. If we look at the third problem, here we need a joint probability statement regarding the distributions of x and y. So, this can be calculated from the joint distribution of x and y. So, this is integral where x plus y lies between 0 to half of 10 x y square d x d y of course, this can be d y d x also depending upon the order in which we integrate.

So, let us determine the range of integration; the density is defined over now here we are saying x plus y is less than half, so the line x plus y is equal to half that is this one. So, the region of integration is reduced to this, so this is half, this is half. So, this point is actually 1 by 4. So, if we integrate, firstly with respect to y, then the range of integration is from x to. So, this on this line this is x plus y is equal to half. So, on this line y is equal to half minus x, and the range of x is from 0 to 1 by 4. So, this is 10, x y square d y d x. So, this is 10 by 3, 0 to 1 by 4 x.

Now this is y cube so it is half minus x cube minus x cube. So, this is a simple integral and one may be able evaluate this quickly. So, here you can observe that if you want to determine certain probability related to the joint distribution, we should determine the region of integration from the description of the distribution that is given there.

So, if we wanted to integrate in a reverse way then firstly, you would it split in it 2 portions: in this portion x is from 0 to y and y is from 0 to 1 by 4; in this portion x is from 0 to half minus y and y is from 1 by 4 to half. If we do not see be carefully this region, then we might have integrated from say 0 to half minus y for x and then for y between 0 to 1 etcetera so that would have been a wrong region here.

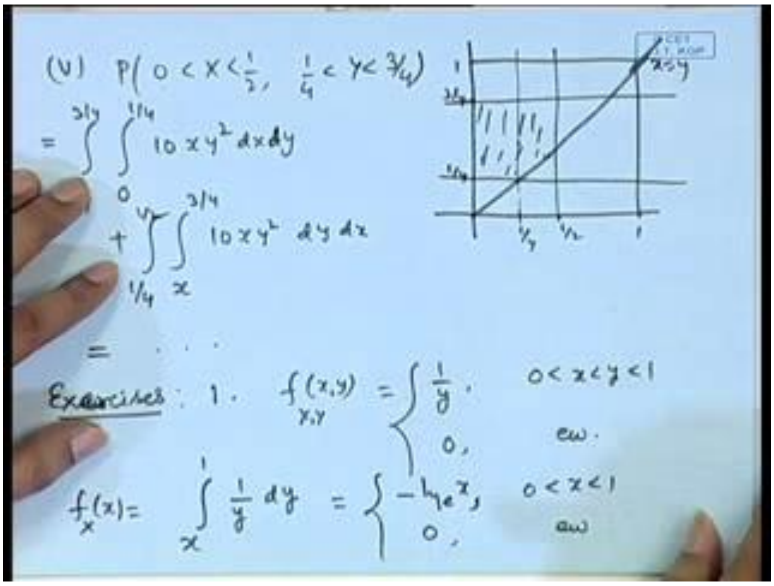
(Refer Slide Time: 13:45)



Let us look at the conditional probabilities also. So, if we are calculating probability of x less than half given y is equal to 3 by 4; now this can be evaluated from the conditional distribution of x given y is equal to 3 by 4. Now conditional distribution of x given y is equal to y we have already determined. So, here if we substitute y is equal to 3 by 4 we get the appropriate distribution. So firstly, we write down the conditional distribution of x given y is equal to 3 by 4; this is obtainable from here, this is 2 x divided by in place y we put 3 by 4. So, we get 9 by 16 that is 32 by 9 x, for 0 less than x less than 3 by 4 and 0 elsewhere.

So, now if we want to calculate probability of x less than half given y is equal to 3 by 4, it is the integral from 0 to half 32 by 9 x d x. So, which is 16 by 9, 1 by 4 that is 4 by 9. In a similar way if we want to calculate probability of y less than half given x is equal to 1 by 4, then we use the conditional distribution y given x and then we substitute x is equal to 1 by 4 here. So, the condition distribution of y given x is equal to 1 by 4 that is obtained as 3 y square divided by 1 minus 1 by 64. So, that is 64 by say 21 y square, where the range of y is from x to 1. So, it will be from 1 by 4 to 1 and 0 elsewhere. So, if we are looking at the probability of say y less than half given x is equal to 1 by 4, then this will be integral from 1 by 4 to half, because the range of y is from 1 by 4 to 1. So, it cannot be from 0 to half. So, this is 64 by 21 y square d y, that is equal to 1 by 8 minus 1 by 64 that is equal to 1 by 9.

(Refer Slide Time: 16:08)



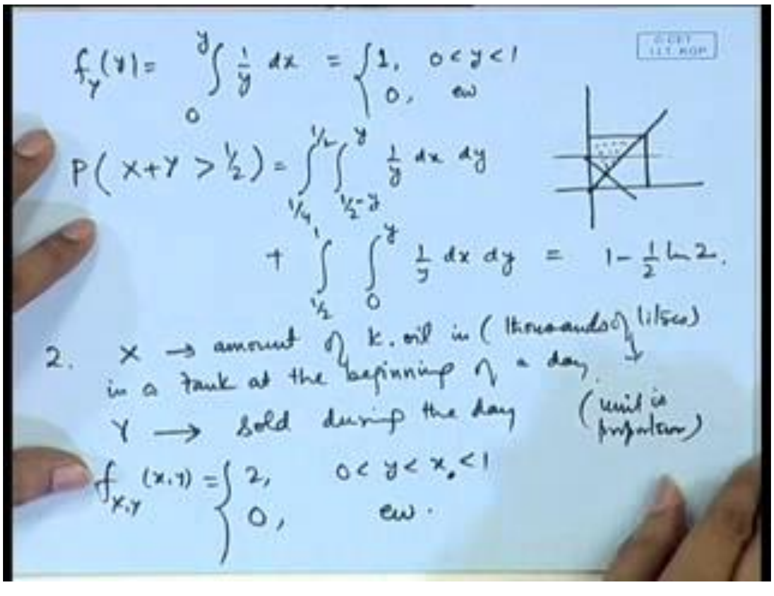
Lastly, probability of 0 less than x less than half, 1 by 4 less than y less than 3 by 4. So, once again we look at the region; here this is the region of the density, so if we say x is between 0 to half, we consider this line and here y is from 1 by 4 to 3 by 4. So, basically we have this particular region; let us determinate separately because this will continue the region with the earlier one.

So, the line x is equal to y is this. So, x is between 0 and half, and y is from 1 by 4; so that will be somewhere here 2 3 by 4 so this is the region. Now here we will have to integrate the density into 2 parts. So, we may split between 2 this type of portion. So, this point is 1 by 4, this point is1 by 2, this point is 1 by 4, this point is 3 by 4. So, if we integrate the density 10 x y square d x d y, then in this region x is from 0 to 1 by 4 and y is from 1 by 4 to 3 by 4, plus in this portion we may consider it with respect to y first y is from x to 3 by 4 and x is from 1 by 4 to half. So, this integral can be evaluated.

So, this way the questions regarding the joint probabilities of x y, the marginal probabilities of x and y are the conditional probabilities of x given some value of y are y given some value of x can be determined from the joint distributions. So, I give certain exercises here, consider say the joint distribution is equal to say 1 by y, 0 less than x less than y less than 1, it is equal to 0 elsewhere.

So, let us look at the marginal distribution say f x that is equal to integral 1 by y with respect to y; now the range of y is from x to 1 so that will give us minus log of x, because it is log x so log 1 is 0 and here the range of x is from 0 to 1 we should not think that this is a negative value actually x between 0 to 1. So, this will be a log of x is a negative so minus log will be positive value.

(Refer Slide Time: 19:59)

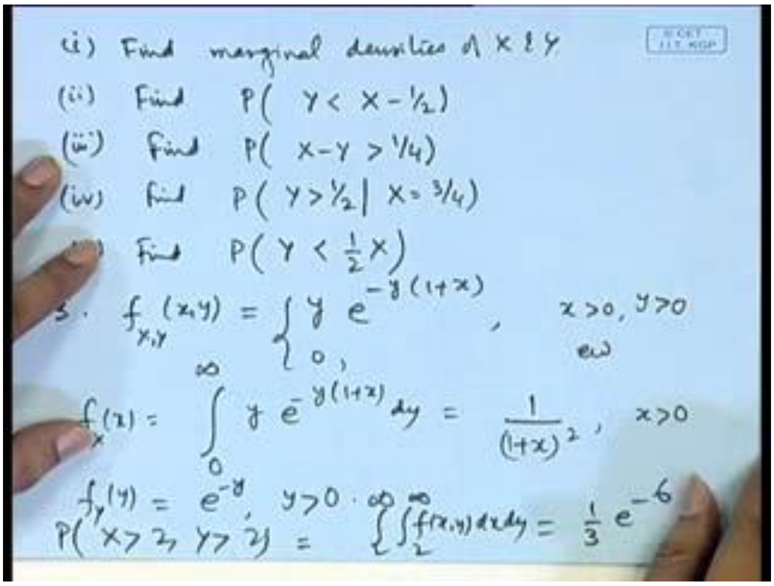


Similarly the marginal distribution y can be obtained. So, that is simply x and therefore, we will get y minus 0 that is equal to 1.

So, the distribution of y is simply uniform distribution on the interval 0 to 1; suppose we want to answer a question regarding say probability of x plus y greater than half. So, once again we look at the region of integration. So, x plus y is equal to half is this line so the reason of integration for this part is this, which we can again split into 2 portions; here x is from half minus y to y, and y is from 1 by 4 to half plus in this portion x is from 0 to y and y is from half to 1. So, this can be evaluated and the some of the integrals says 1 minus half log 2.

Suppose X denotes the amount of kerosene oil in say thousands of liters in a tank at the beginning of a day and y is the amount which is sold during the day. So, f x y is suppose given by say 2, 0 less than y less than x less than 1. So, if we are considering say proportion here the unit is proportion and 0 elsewhere.

(Refer Slide Time: 22:48)

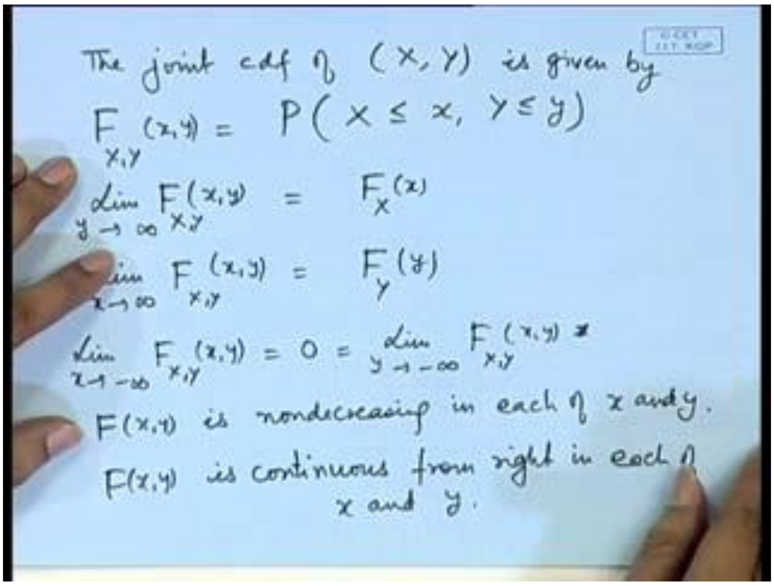


So, find marginal densities of x and y, find probability y is less than say x minus half, find probability say x minus y is say greater than 1 by 4, find probability y is greater than 1 by 2 given x is equal to 3 by 4, find probability say y is less than half x etcetera.

Suppose x and y denote component lies of certain equipment consisting of 2 components; let us look at say marginal distributions here, then this is y e to the power minus y 1 plus x d y so it is equal to 1 by 1 plus x whole square. If we consider say marginal of y then we integrate with respect to x, then that is equal to e to the power minus y y is greater than 0. Suppose I say find probability of x greater than 2, y greater than 2. So, it will be obtained by integrating the joint density from 2 to infinity.

So, after integration of this term we get 1 by 3 e to the power minus 6. In the case of unvariate random variables, we considered various characteristics of the random variables such as it is mean, variance, higher order moments, the measures of estimates kurtosis the median quantiles etcetera. So, in a similar way we can talk about the characteristics of the joint distributions also.

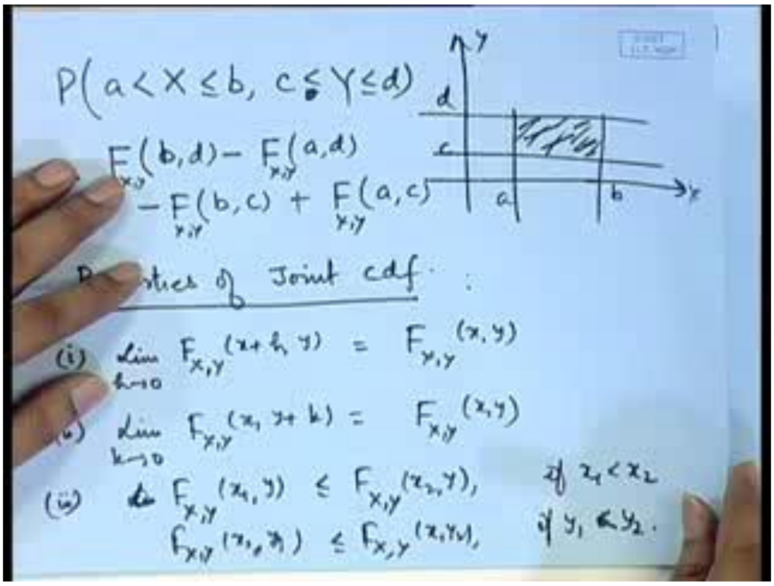
(Refer Slide Time: 25:44)



So, first of all we introduce the joint cdf. So, the joint cdf of a bivariate random vector x y is given by F x, y as probability of. So, we can see here that this function gives information about the type of the random variables that x and y are, as well as it will yield the individual cdf’s of x and y also.

So, for example, if I take limit of F x y as say y tends to infinity, this gives the marginal cdf of x; if we take limit as x tends to infinity of the joint cdf this gives us the marginal cdf of y, if we take either of x tending to minus infinity or y tending to minus infinity we get 0. F x y is non-decreasing in each of x and y. F x y is continuous from right in each of x and y, we can also consider say sell in 2 dimensional space.

(Refer Slide Time: 27:45)

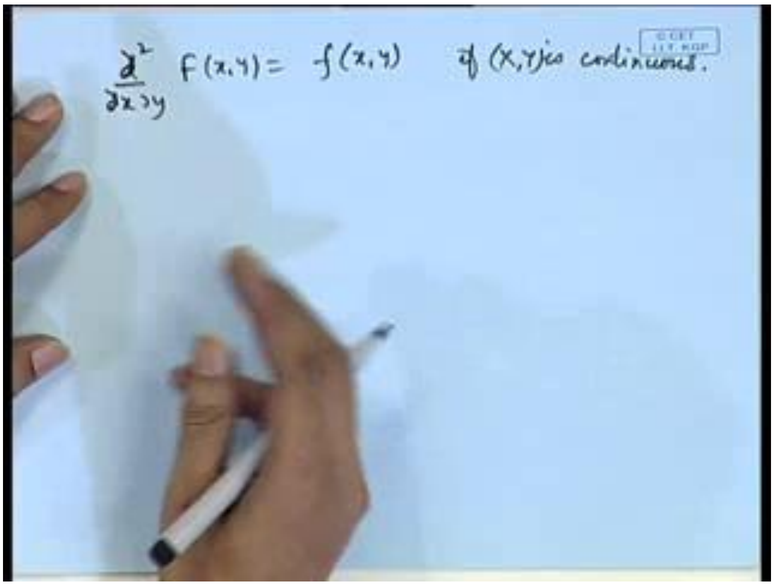


Suppose there is a point a, b this is x, this is y and this is a c, d. So, if you want to look at the probability of this region. So, probability of a less than x less than or equal to b, c less than y, less than or equal to d; then this is equal to F of b, d minus F of a, d minus F of b, c plus F of a, c. So, the probability of a rectangular region can be evaluated in terms of the cdf. Since we are able to obtain the individual distributions from the joint cdf, we can find the individual nature of the random variables that is whether they are discrete continuous or mixture random variables sum here.

Moreover if the random variable x y is continuous throughout, then capital F x, y will be a absolute, continuous in both of them both of the arguments and the derivatives with respect to x and y will give you the probability density function of x and y. So, this joint cdf is a quite useful function for calculating various characteristics of the random variable, the joint distribution x and y.

So, let us look at the other features a. So, this is the right continuous behavior in both the arguments, the non decreasing nature in both the arguments.

(Refer Slide Time: 31:15)



In case of the discrete random variable, the probability of a point can be calculated in terms of the differences taken by the cdf. So, if you are looking at say probability of x is equal to x i y is equal to y j, then we can consider it as probability of x less than or equal to x j y less than or equal to y j minus probability of x is less than or equal to x j minus 1, y is equal to y j etcetera and we can express in terms of the joint cdf at these points.

So, in general the joint cdf gives information about the complete information about a jointly distributed random variable. So, we will discuss about other property such as product movements, the correlation coefficient between the random variables in the coming lectures.

Thank you.