Probability and Statistics

Prof. Dr. Somesh Kumar
Department of Mathematics

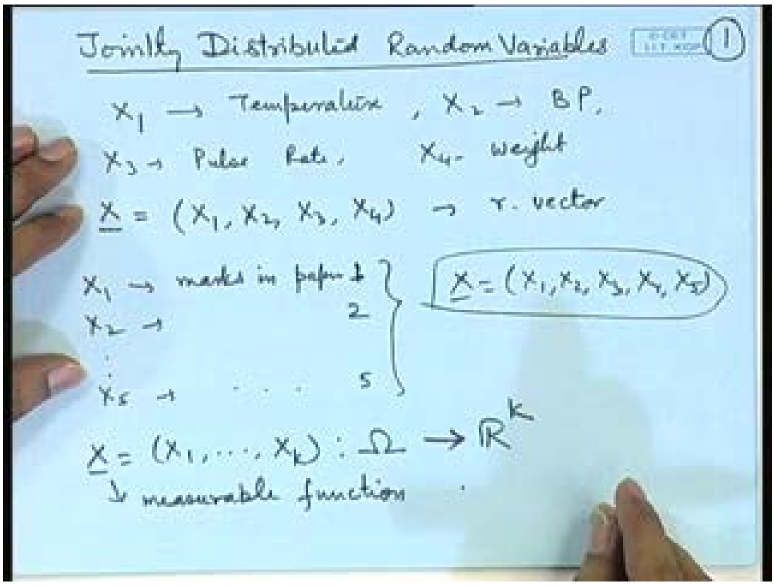
Indian Institute of Technology, Kharagpur

Lecture No. #18

Joint Distributions-I

So far we have considered the phenomena where if a sample space is given,we are considering function ofathe function which isthe mapping sample space to the real line. So, we are considering one characteristic at a time.For example, it may be heights of the students, it may be marks of a student in a test paper, it may be the life of an electronic equipment. However, sothis constitutes the values of a single random variable X. So, we have sofar concentrated on the distribution of the random variable X.Many times we are not havingthe luxury of considering only one characteristic, but several characteristics such asa patient goes to a doctor for a medical checkup, the doctor takes his say temperature, he may take his blood pressure, he may take his pulse rate and he may record his say weight.

(Refer Slide Time: 01:20)

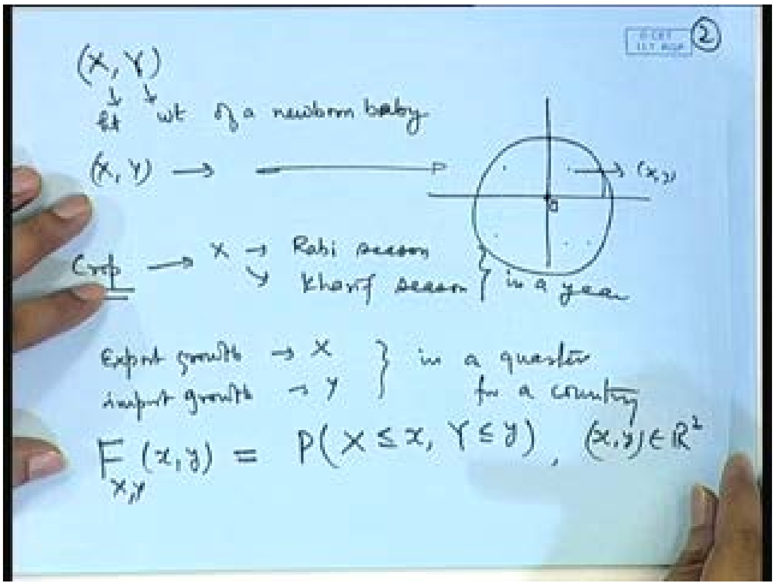


So, for different patient's four quantities X 1, X 2, X 3, X 4 are recorded. So, here this constitute say random vectoror jointly distributed random variable X 1, X 2, X 3, X 4.

So, X is called a random vector.We may be looking at say X 1 as marks in paper 1, X 2 as marks in paper2, say X 5 marks in paper 5. So, students are studying five subjects in a particular semester and each of them will get different marks in different test papers. So, for each student, if we record X is equal to X 1, X 2, X 3, X 4 and X 5, then this is a random vector or a jointly distributed random variables X 1, X 2, X 3, X 4 and X 5.

So, in general if I am considering X as X 1, X 2, Xk, then this is afunction from omega into the k dimensional Euclideanspace R k.The random variable X was a measurable function from omega into R. That is one-dimensional Euclidean space. So, a higher order random variable or a random vector is a function from -it is a measurable function from omega into R k. So, X is measurable function that we have to ensure in order that the probability functions of X are well defined. For convenience we will restrict attention to two-dimensional case in the beginning.

(Refer Slide Time: 03:48



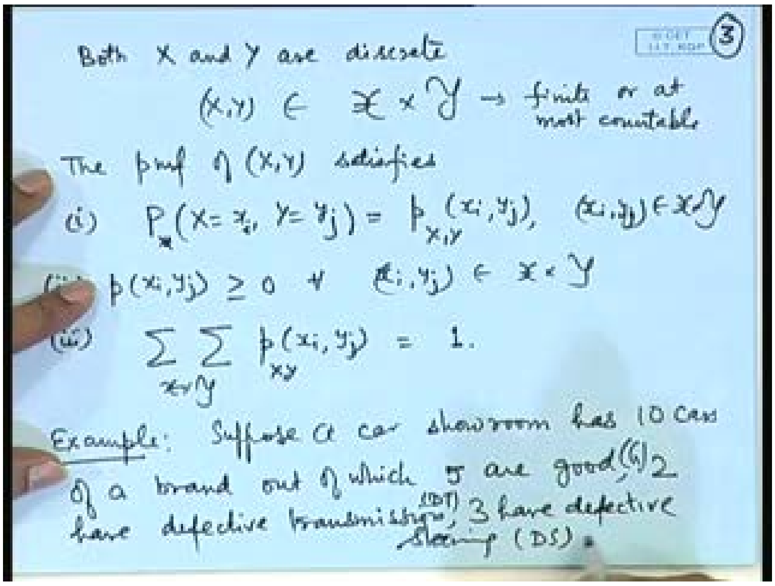
So, let us consider jointly distributed random variables(X, Y). So, here X could denote the height of a new born and Y could denote the weight of a new bornbaby. (x,y) could denote the coordinates of a dart hitting a target. So, suppose this is the target, we consider it as origin and the dart may hit anywhere. So, the coordinates(x,y) of the dart hitting may be considered as a jointly distributed random variable.Suppose we are considering crop in a area. So, crop in a say Rabi season and in a Kharifseason in a year.

We may consider, say export growth, say x and sayimport growth y in a quarter of year for a country.So, these are all examples of jointly distributed random variables.

Now, naturally the question arises that how do we evaluate the probability distribution of(X, Y)or how do we define the probability distribution of jointly distributed random variables.So, we have seen in the case of one variable, the description of the probability distribution depends upon whether the random variable is discrete or continuous. So, if the random variable is discrete, we have a probability mass function, if the random variable is continuous, we have a probability density function.Of course, for any type of random variable, we have thefacility of using a cumulative distribution function, but that is not useful all the time.

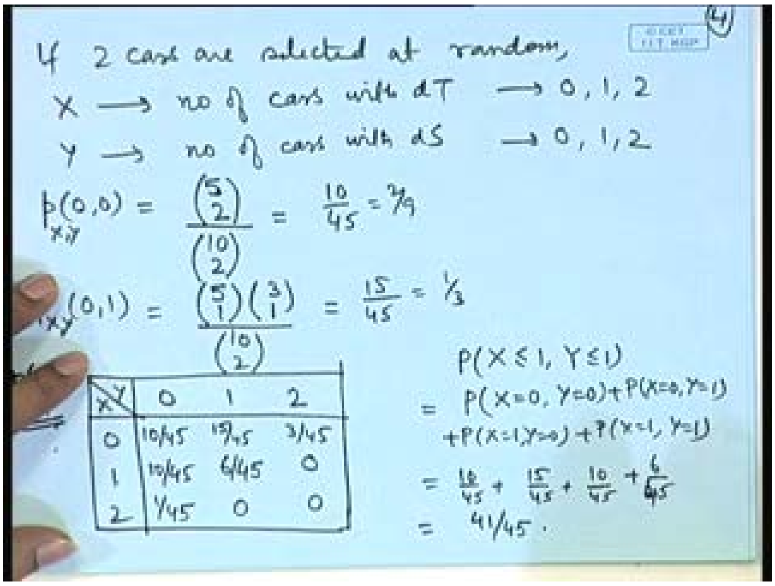
So, when we have a jointly distributed random variable then there can be several cases; Xmay be discrete, Y may be continuous, both may be discrete, both may be continuous, X may be discrete Y may a mixture, X may be a mixture random variable Y may be continuous, etcetera.So, in each case the description of the distribution may be of different nature. However, one may make use of the joint CDF that is probability of X less than or equal to x, Y less than or equal to y. This is defined for all x y in the two-dimensional plane. So, we will look atthe properties of this CDF later on. Firstly, let us consider the case, when both X, Y may be discrete or both X, Y may be continuous.

(Refer Slide Time: 07:00)



So, if both X and Y are say discrete. That means, the values that X, Ytake in a space say xcross y; this is finite or at most countable. So, in this case, the probability mass function of X, Y this will satisfy. So, P X - probability X equal to x, Y is equal to y. Let me put here X equal to x i, Y is equal to y j, because the values are at most countable. So, we can innumerate them. This is P X Y(x i, y j) for allx y x iy j. Then this function is always non negative for all x i y j and the sum over all the values of the probability mass function is 1. Let us take one example here. Suppose a car showroom has 10 cars of a brand out of which 5 are good, 2 are having say defective transmission and 3 have say defective saysteerings. So, we call it say DT and D S, and good are G.

(Refer Slide Time: 09:32)

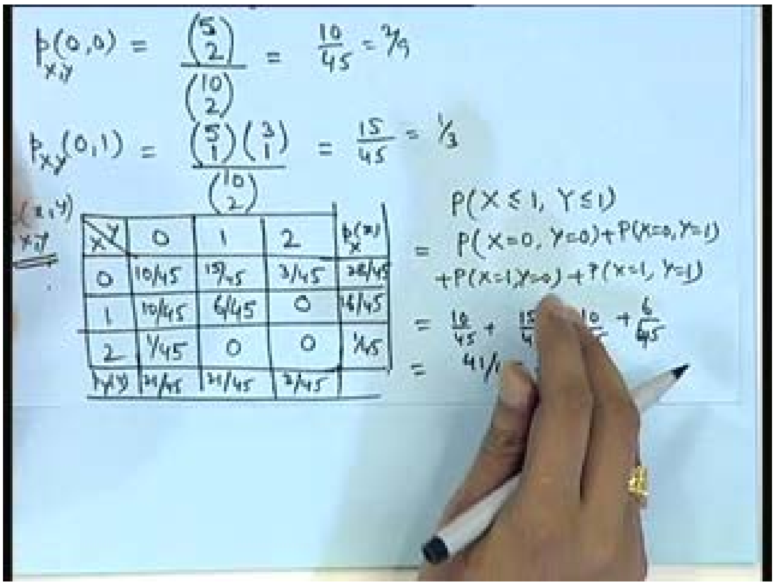


Now, if 2 cars are selected at random, let X denote the number of cars with defective transmission and Y the number of cars with defective steering mechanism. So, hereX, Y is a discrete random vector both X and Y can take values 0, 1, 2. So, the probability distribution ofX, Y can be described, what is the probability that X is equal to 0, Y is equal to 0.Now, this is possible if both of the selections are made from the good 5 cars out of the total selections from 10. So, this is equal to 10 by 45or 2 by 9. Similarly, we can calculate probability X is equal to 0 and Y is equal to 1. So, here it means that out of1 good, out of 5 good, 1 good car has been selected and out of 3 cars with defectivesteering mechanism 1 has been selected, out of total 2 selections of 10. So, that is equal to 15 by 45 or 1 by 3.

In a similar way, we can calculate all other probabilities and we can describe the probability distribution in the form of a tabular representation;the probability X equal to 0, Y is equal to 0 is 10 by 45; the probability X equal to 0, Y is equal to 1 is 15 by 45; the probability that X equal to 0,Y is equal to 2 that can be calculated to be 3 by 45; the probability that X equal to 1, Y is equal to 0can be calculated to be 10 by 45; X is equal to 1, Y is equal to 1, 6 by 45; X is equal to 1, Y is equal to 2 is not possible,because total selections are only 2. So, this is 0. Probability thatX equal to 0, Y is equal to… X equal to 2, Y is equal to 0 is 1 by 45; again 2 1 and 2 2 are not possible. So, this gives the joint distribution P x y of the random variables X, Y.

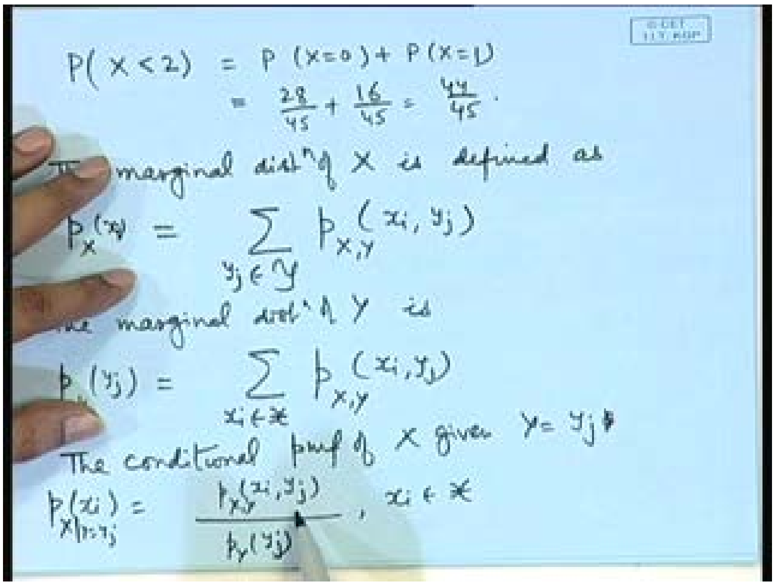
You can look at various probabilities related to random variables X, Y from here. For example, if we ask a function what is the probability that X is say less than or equal to 1, Y isless than or equal to 1, then this is corresponding to probability of X is equal to 0, Y is equal to 0; probability X is equal to 0, Y is equal to 1; probability X equal to 1, Y is equal to 0and probability X equal to 1, Y is equal to 1. So, this is equal to 10 by 45 plus 15 by 45 plus 10 by 45 plus 6 by 45 which is equal to41 by 45.

(Refer Slide Time: 09:31)



We may answer some other questions from here, regarding the probabilities.For example, if we ask what is the probability X is say less than 2 that is probability X equal to 0 plus probability X equal to 1. Now, this relates to the probabilities of one random variable when the joint distribution is given.Now, noticehere that if Isum row wise, then this will give probability X equal to 0, Y is equal to 0; probability X equal to 0, Y is equal to 1; probability X equal to 0, Y is equal to 2. That means,this will give the distribution of X. That is equal to 28 by 45, 16 by 45 and 1by 45.Similarly, if I add the columns here, I will get the distribution of Y. So, that is 21 by 45,21 by 45 and 3 by 45.

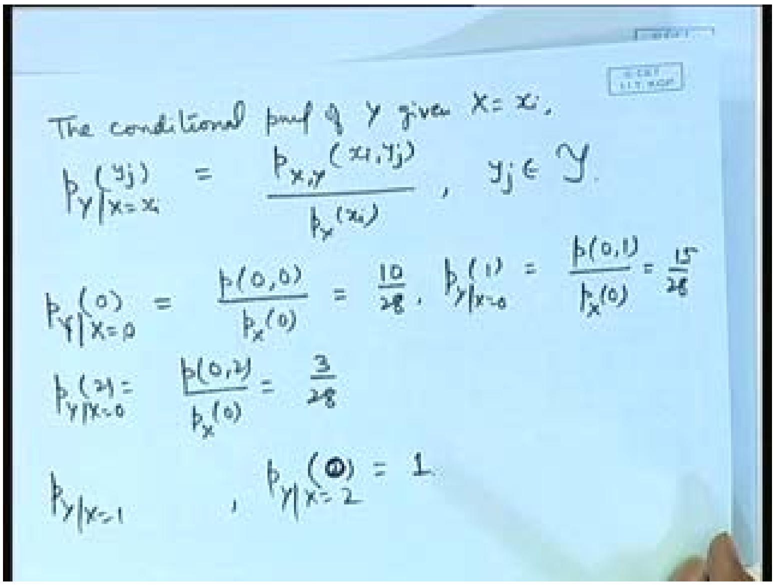
(Refer Slide Time: 14:46)



So… Now, probability X equal to 0 and probability X equal to 1 can be easily obtain as 28 by 45 plus 16 by 45 that is equal to44 by 45. This process of adding row wiseor column wise, this gives rise to the individual distributions of X and Y, they are known as marginal distributions. So, in general the marginal distribution of X is definedas… This is obtained by adding the values of y over the range of this.Similarly, the marginal distribution of Y is obtained by adding the joint probability mass function with respect to x i. So, we can answer all the probability statements regarding individual distributions of X and Y also from the joint probability mass function.

Now, there is one more thing; when we have two random variables, we may also look at the conditioning events.For example, if Y is equal to 2is given, if Y is equal to 1is given, what happens to the probability distribution of X. This is known as the conditional distribution.So, we may define the conditional probability mass function of X given a value say Y is equal to y j. So, this is given by the joint distribution of X and Y at x i, y j divided by the marginal of thiswhere x ivaries is over X.To see that it is a valid probability distribution, let us sum over all the values of x i. So, if you sum over all the values of x i, the numerator becomes the marginal distribution of y j which is the denominator. So, this will become 1. So, this is well defined.

(Refer Slide Time: 17:23)

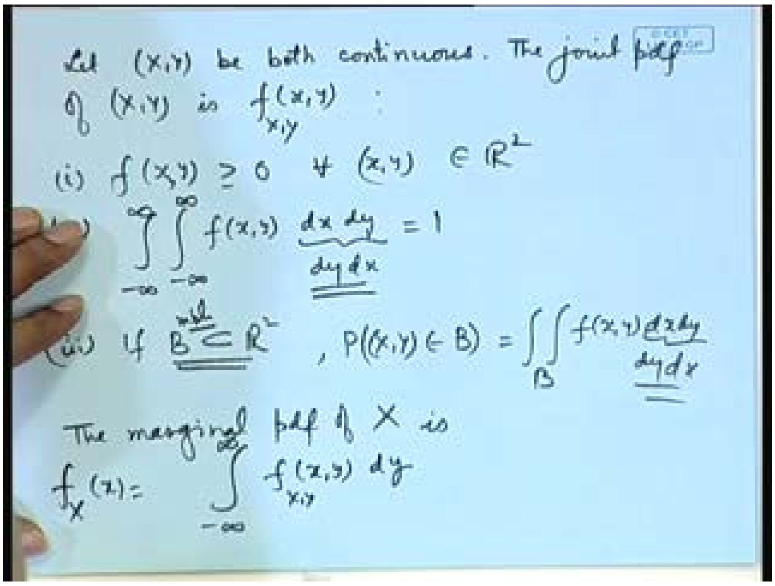


Similarly, the conditional probability mass function of Y given X is equal to x i. So, that is equal to p Y given X equal to x i is add the value y j, it is defined to be the joint distribution divided by the marginal distribution of x i for values of y j over Y. Once again you can see that, it is a valid probability distribution, if we sum over the values of y j, the numerator here will become p x, x i which issame as the denominator. So, it will give the value 1.

So, in thisgiven problem, let us look at the conditional distribution,say probability of Y given X is equal to 0. So, the values of Y are 0, 1 and 2. What is the probability that Y is equal to 0 given X equal to 0. So, it is p(0, 0) divided by p X(0). Now, p(0, 0) is 10 by 45 and P X(0) is 28 by 45. So, this becomes 10 by 28. What is p Y 1 given X equal to 0? This is p 0, 1 divided by p X 0. So, that is equal to 15 by 28. Similarly, p Y X equal to 0at 2, this is p(0, 2) divided by p X 0; that is 3 by 28. So, you can see here, the sum of the three probabilities 10 by 28, 15 by 28 and 3 by 28 gives 1.

In a similar way, 1 may calculate probability distribution of say Y given X equal to 1, probability distribution of Y given X equal to 2.Notice here that if I say X is equal to 2 then only Y is equal to 0 is possible. So, probability of Y is equal to 2 given X equal to 2 that will be 1. This is a degenerate distribution. This is p(2, 2)divided by probability X equal to 2that is 1 by 45divided by 1 by 45, because if X equal to 2 is fixed, then Y cannot take any other value except 0sorry this is not 2, this is 0,because the total number of cars that we are purchasing is 2,therefore, if one of them is having the value 2 then the other value has to be 0.

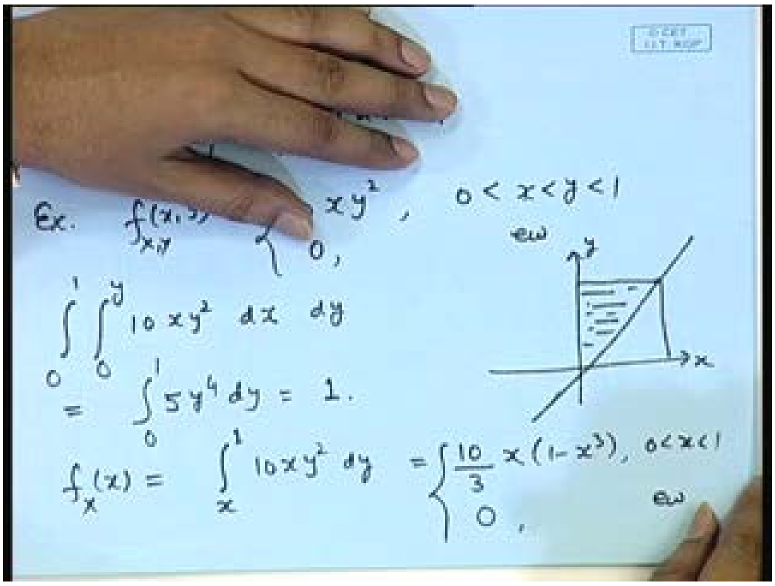
(Refer Slide Time: 20:33)



Now, in a…if both the random variables are continuous then we have joint probability density function. So, let us consider, let X and Y be both continuous. So, the joint probability density function (X, Y) is f (x, y); this is satisfying that it has to be a non negative function.The integral over the whole range with respect to both the random variables is 1. So, herein place of dx dy1 may writedy dx also. And if B is a measurable subset of R 2, then probability of X, Y belonging to B will be given by the integral of the density over the set B integration may be in any order.

If the probability density function is given, we may be able to answer any probability statement related to the distribution of X, Y, the joint distribution ormarginal distributions of X and Y. 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. 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.

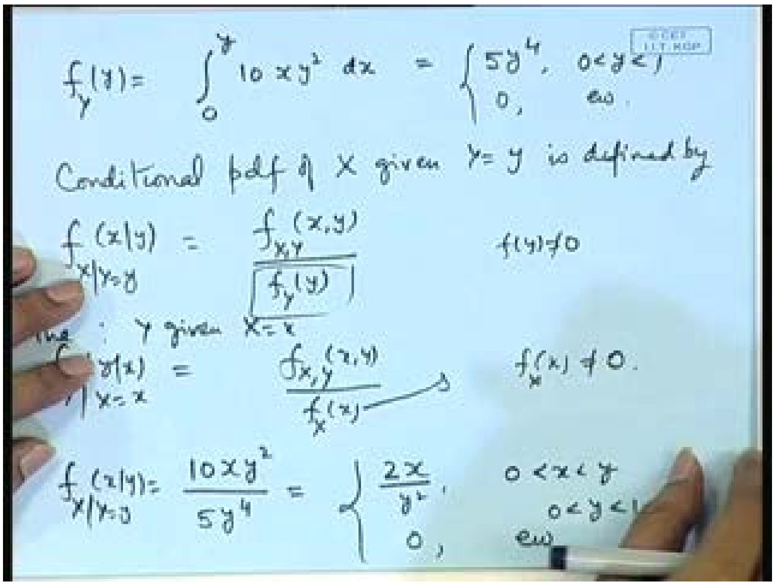
(Refer Slide Time: 22:44)



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 validcdf. 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 the distribution is defined over the interval 0 to 1 on the half side; that is x less than y. So, if this is x axis, this is y axis. So, we are on this side.So, you can easily see that - this is 5 y to the power 4 dy which 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 toy,the range of x y will be from x to 1.This gives 10 by 3x into y cube from 1 tofrom x to 1 that is 1 minus x cube. So, the marginal distribution of x is given by 10 by 3x into 1 minus x cube.

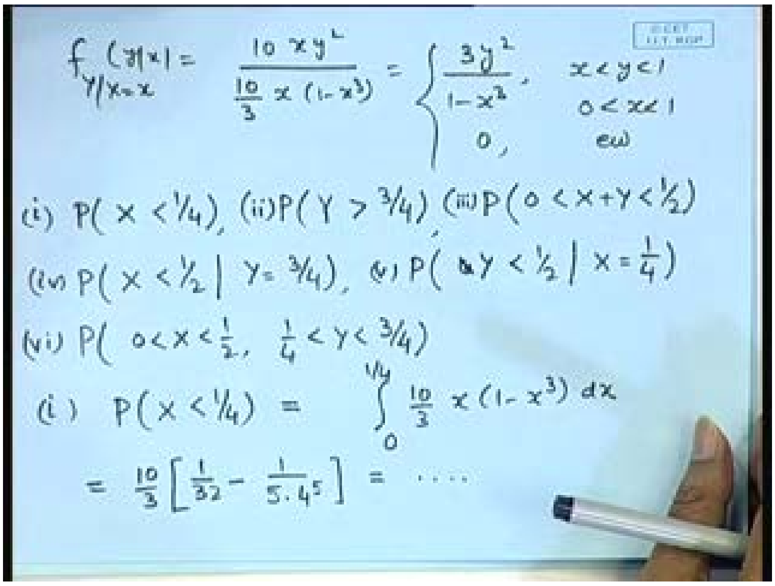
(Refer Slide Time: 24:54)



Similarly, we can obtain the marginal distribution of Y.That is integralof 10 x y square d x. Now, here the range of x will be from 0 to y; for a given y the range of x is 0 to y; so, 0 to y. So, this is equal to 5 y to the power 4 for 0less than y less than 1, and 0 elsewhere. So, the marginal distributions of X and Y are easily obtain.

The conditional pdfof X given a value of Y is equal to y is defined by f X given Y is equal to y; that is the joint distribution of x y divided by the marginal distribution of y.Of course, this is defined for f ynon 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 square divided by 5 y to the power 4; that is equal to 2x 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 is the valid distribution, if we integrate from 0 to y, we will get y square. So, y square by y square will be 1

(Refer Slide Time: 27:16)

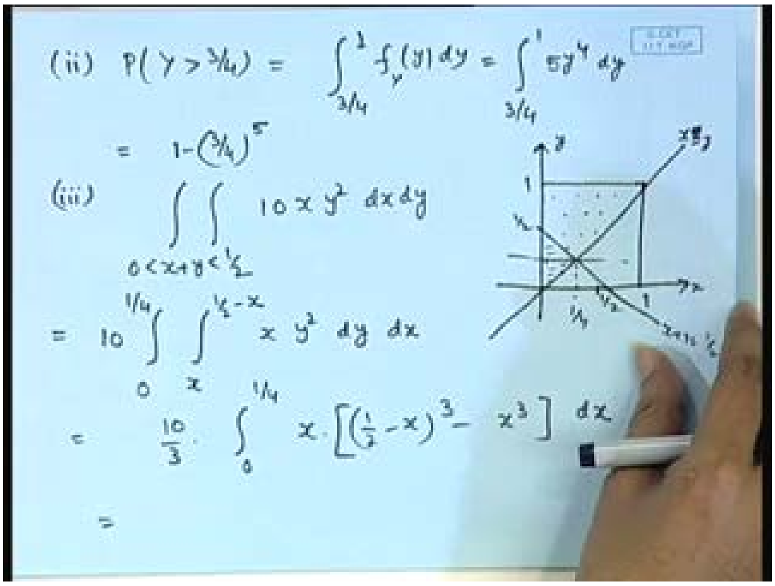


Similarly,we can define the conditional distribution of Y given X here. So, that is 10 x y square divided by the marginal distribution of X which we obtained as 10 by 3 x into 1 minus x cube. So, divided by 10 by 3x into 1minus x cube;that is equal to 3 y square by 1 minus x cube. Here the rangeof y is from x to 1 for a value of x between 0and 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 probability statements regarding the joint probability of X, Y, the individual probability related to random variable X and Yor the conditional probabilities related to X and Y. So, in this particular case, let us look at some of the such questions.

So, suppose I say what is probability X is less than 1 by 4; what is the probability that 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 givenX 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 probability X less than 1 by4. So, this is related to the marginal distribution of X which we calculated as 10 by 3x into 1 minus x cube for x lying in the range 0 to 1. So, this is 0 to 1 by 4, 10 by 3x 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 isx to x to the power 4 that is x to the power 5 by 5; so, 1 by 5into 4 to the power 5. So, this can be simplified.

(Refer Slide Time: 31:05)

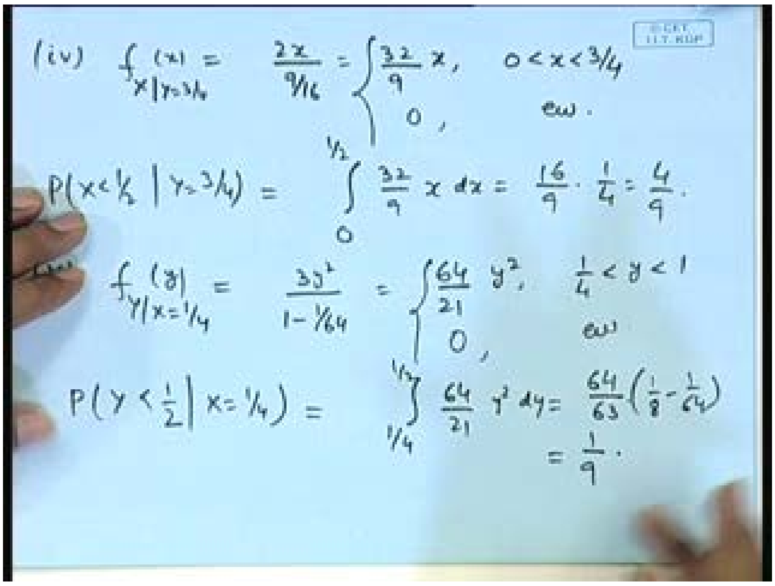


Similarly, if we look at probability Y greater than 3 by 4, then this can be obtained from themarginal distribution of Y. So, if we answer this question, probability Y greater than 3 by 4. That is integral f y from 3 by 4 to 1, because the range of y is from 0 to 1. So, this is equal to integral 3 by 4 to 1, 5 y to the power 4 dy. 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 yliesbetween 0 to half of 10 x y square dxdy.Of course, this can be dy dx 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 1. 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 firstlywith respect to y, then the range of integration is from xto… So, thison this line, this is x plus y is equal to half. So, on this line y is equal to half minusx. And the range of x is from 0 to 1 by 4. So, this is 10 x y square dy dx. 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 to evaluate this quickly. So, here you can observe that, if we want to determine certain probability related to the joint distribution, we should determine the regionof integration from the description of the distribution that is given there. So, if we had if we wanted to integrate in a reverse way, then firstlyyou willsplited into two 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 the carefully these region, then we might have integrated fromsay 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.

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 determine. So, here if we substitute Y is equal to 3 by 4, we get the appropriate distribution.

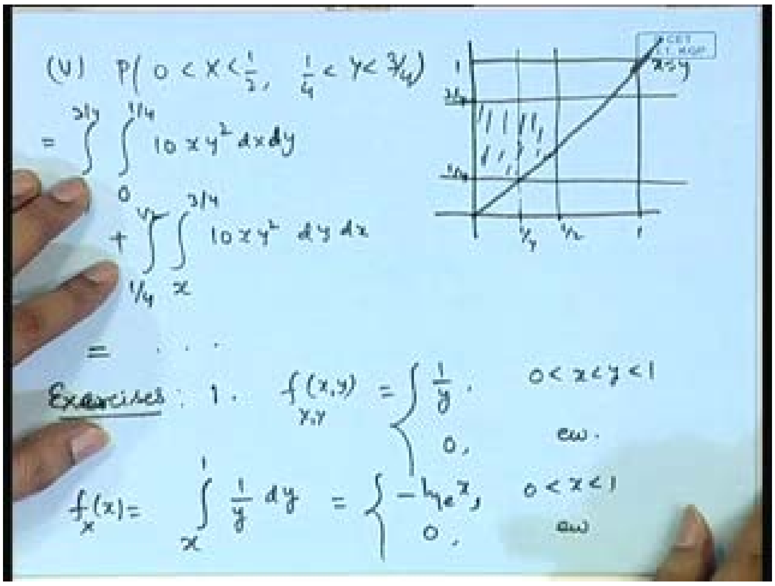
(Refer Slide Time: 35:27)



So, firstlywe write down the conditional distribution of X given Y is equal to 3 by 4. So, this is obtainable from here; this is 2 x divided by in place of y we put 3 by 4. So, we get 9 by 16 that is32 by 9 x for 0 less than x less than 3 by 4, and 0 elsewhere. So, nowif we want to calculate probability of Xless than half given Y is equal to 3 by 4, it is the integral from 0 to half,32 by 9 x dx. 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 equal to 1 by 4, then we use the conditional distribution of Y given X and then we substitute X is equal to 1 by 4 here. So, the conditional distribution of Y given X equal to 1 by 4. That is obtained as 3y square divided by 1 minus 1 by64. So, that is64 by say21 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 sayY less than half given X 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 is64 by 21 y square dy; that is equal to 1 by 8 minus 1 by64; that is equal to 1 by 9.

(Refer Slide Time: 37:51)

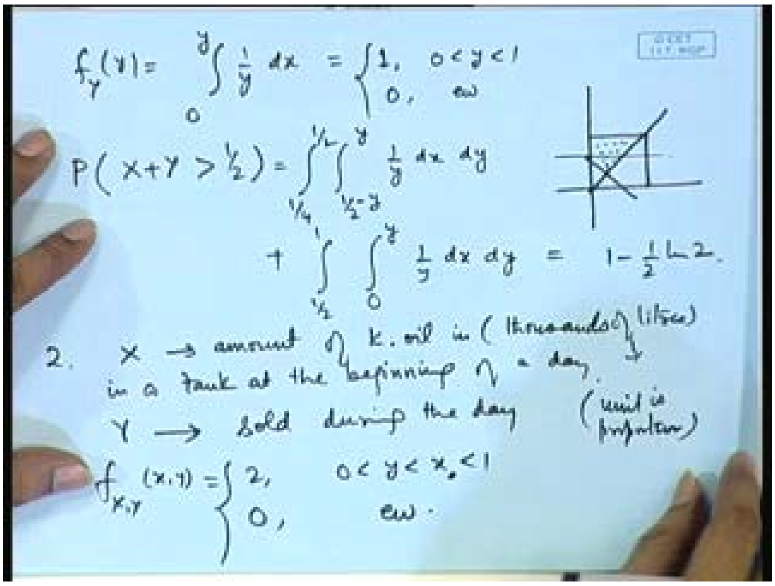


Lastly, probability of 0 less than X less than half, 1 by 4 less than Y less than 3 by 4. So, once again we look at the region here.This is the region of the density. So, if we say X is between 0 to half,we consider this line and here Y is from 1 by 4 to 3 by 4. So, basically we have this particular region. Let us determine it separately, because this will confuse the region with the earlier one. So, the line x 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 to 3 by 4. So, this is the region.Now, here we will have to integrate the density into two parts. So, we may split it into this type of portion. So, this point is 1 by 4, this point is 1 by 2, this point is 1 by

4,this point is 3 by 4. So, if we integrate the density 10 x y square dxdy, 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.

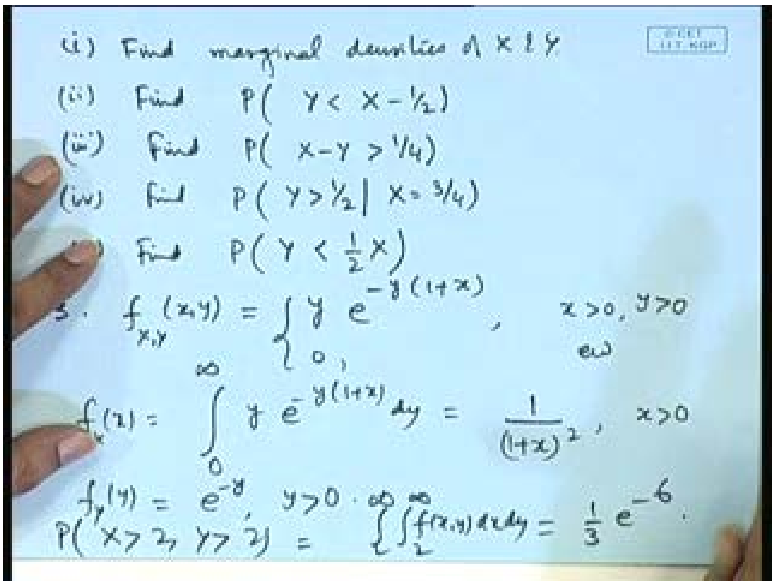
So, this way the questions regarding the joint probabilities of X Y,the marginal probabilities of X and Y or the conditional probabilities of X given some value of Y or Y given some value of X can be determine from the joint distributions. (No audio from 40:23 to 40:30) So, I give certain exercises here. Consider say the joint distribution is equal to say 1 by y, 0 less than x less than y less than1, it is equal to 0 elsewhere. So, let us look at the marginal distributions 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 y, solog 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 is between 0 to 1. So, this will be a log of x is a negative value. So, minus log will be a positive value.

(Refer Slide Time: 41:42)



Similarly, the marginal distribution of Y can be obtained. So, that is simply x and therefore, we will get y minus 0that is equal to 1. So, the distribution of y is simplyuniform distributionson 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 region of integration for this part is this;which we can again split into two 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 fromhalf to 1. So, this can be evaluated and the sum of the two integral is 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 0is here.

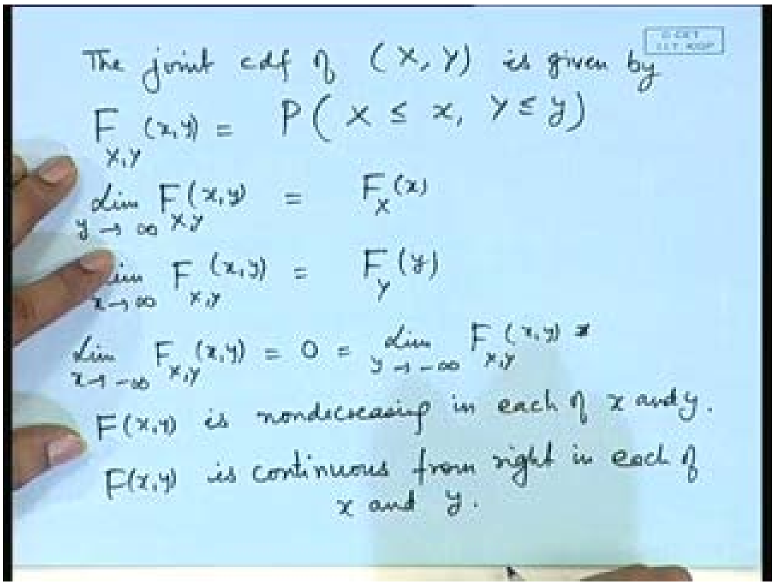
(Refer Slide Time: 44:30)



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 Ydenote component lives of certain equipment consisting of two components.Let us look at say marginal distributions here then this is y e to the power minus y 1 plus xdy. So, it is equal to 1 by 1 plus x whole square. If we consider saymarginal of Y then we integrate with respect to x, then that is equal to e to the power minus y, y greater than 0.Suppose I say find probability 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.

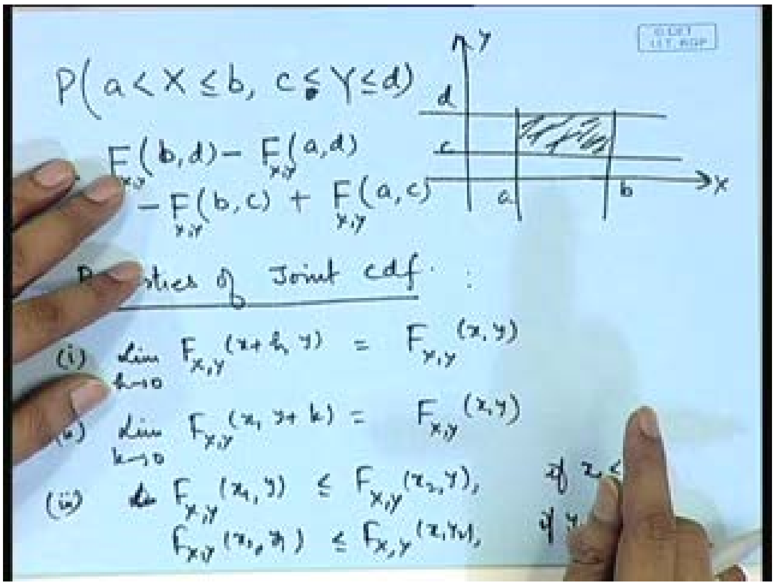
In the case ofunivariate random variables, we considered various characteristicsof the random variables such as its mean,variance,higher order moments, the measures of skewers kurtosis, the median quintiles', etcetera.So, in similar way,we can talk about the characteristics of the joint distributions also. So, first of all we introduce the joint cdf.

(Refer Slide Time: 47:27)



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 itwill yield the individual cdfs of X and Y also. So, for example, if I take limit of F(x, y) as say ytends 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.

(Refer Slide Time: 49:28)



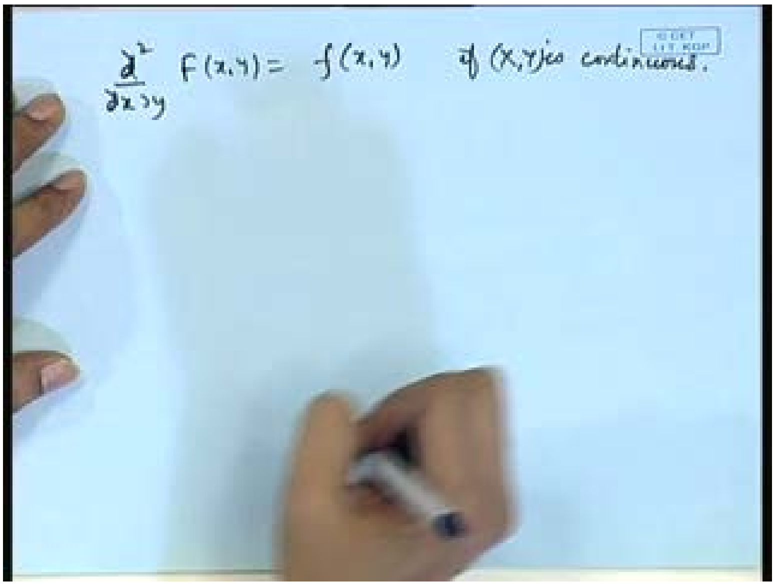
We can also consider say a cell intwo-dimensional space. Suppose this is the point a b this is X, this is Y and this is say c, d. So, if we want to look at the probability of this region. So, probability of a less than X less than or equal to b, c less than Y less than or equal to d, then this is equal to Fc(b,d) minus F(a, d) minus F(b, c) plus F(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 jointcdf, we can find the individual nature of the random variables that is whether they are discrete, continuous or mixture random variables from here.

Moreover, if the random variableX, Y is continuous throughout, then capital F(x, y)will be absolute continuous in both of them both of the arguments and thederivatives 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 therandom variable, the joint distribution X and Y.So, let us look atthe other features of a...

(No audio from 51:32 to 52:47)

So, this is the right continuous behavior in both the arguments, the non-decreasing nature in both the arguments.

(Refer Slide Time: 52:57)



(No audio from 52:55 to 53:15)

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 we are looking at say probability of x 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 thejoint cdf at these pointsSo... So, in general the joint cdf gives information aboutthe complete information about a jointly distributed random variable. So, we will discuss aboutother properties such as product moments,the correlationcoefficient between the random variables in the coming lectures.Thank you.