# Estimation of Signals and Systems

**Prof. S. Mukhopadhyay**
**Department of Electrical Engineering**
**Indian Institute of Technology, Kharagpur**

## Lecture - 04

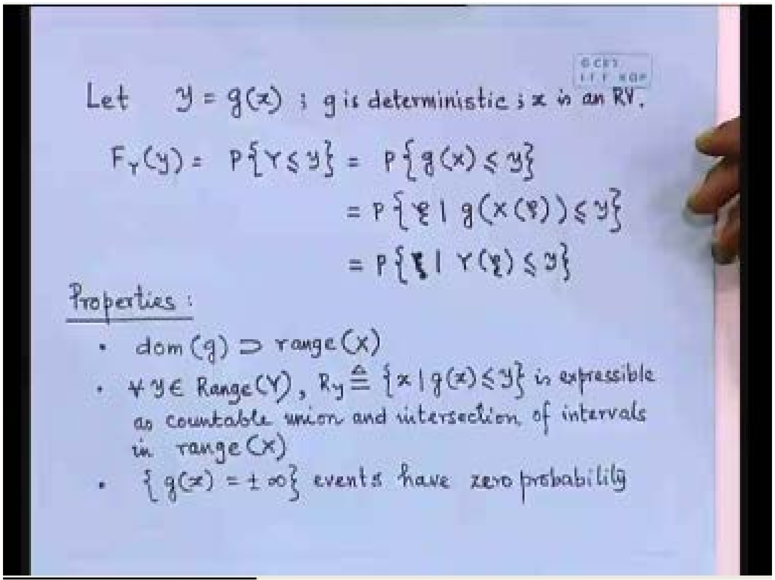
## Function of Random Variable Joint Density/Distribution

What we have done is that we have defined the random variable as a mapping from the set of events to the... to the real line so we got a real variable. From events we have now got a real variable which we call random variable which takes values between minus infinity and plus which takes values between minus infinity and plus infinity.

Now now the moment we have a real variable we can define other real variables which are functions of that variable. The moment we have a random variable x we can define another random variable which would be let us say x square. So the question is that if we know the what we are going to see today mainly is if we know the properties of x as a random variable we know by by property we basically mean if we know its distribution function if we know its density function can we calculate the density and distribution functions of the variable y right? And we can extend this concept saying that if we have a... if we have one random variable x why not have two random variables x and y then what is their what is their joint statistics that is what is the probability that this and this occurs at... that is the x is less than 2 and y is less than 3 that is a varied problem; we can we can in general extend it to N variables. See, the moment we can do that we can talk about vectors that is random vectors because each component is going to be random variable. So just so that we are able to talk about things like you know states and multivariable things so we should know that if we have more than one random variable how to handle that. Then once we have more than one random variable we can also define a function z which is a function of x and y; why only a function of x. So so we can we can also talk about such variables.

So we will so as we are gradually advancing we are we are trying to see that from from the basic idea how we can get the other various other cases which are also going to be of interest in this course right. So so we will talk about these things today.

(Refer Slide Time: 00:03:17 min)




So the first thing is that let y equal to g x; x is a random variable, remember that that g is deterministic, g is there is there is no uncertainty about g so y is equal to 1/2 x square that that operation of 1/2 and that squaring random variables these operations are not uncertain, it is x it is the value it is the argument of the function which is uncertain right.

Now for example if you have y is equal to x plus c where that is also a function of x where c is a random variable not x. In that case the function itself is random, there is a subtle difference, we are not talking about random functions we are talking about deterministic functions of random variables right.

So so now the question is obviously these are our objects of interest. The moment we define another new random variable we would like to know what is this what is its distribution function. So we will we will always go back to the go back to the definition.

For any random variable its distribution function is actually defined as the probability that Y is less than or equal to y, that is how the distribution function is defined. So what is y nothing but g X so it is the it is the probability of having an X now we are transforming from the set of Y's to the set of X. We are now... it is the same as the probability of finding out an X or all those X's for which g x is less than or equal to y because y is less than or equal to g X rather y is equal to g X correct?

Or you can even go back to the events themselves or you can see that it is finally because finally the probabilities are coming out of experiments or outcomes of the events that is the basis. So you can also say that finally it is a probability of those events that corresponding to those events if you have a random if you have the random variables A X which are functions of psi now and corresponding to those X's whatever is g g X psi that is less than or equal to y.

Now if you have g X psi this this g X these two mappings you can club and you can say that it's a new mapping from the set of events to y you can also say that. So you can say that it is Y is a new mapping which is which gives me the new random variable; any random variable is a is a mapping from the set of outcomes to the set of reals right. So you can define a new Y Y. This is just you know to to bring down the physical meaning of the quantity. It is the physical meaning of y remains the same here, it is the probability that that value y comes from... now for example, suppose you are saying that you are you have connected a voltage source to a resistor; now if you arbitrarily pick up a resistor let us say 5k nominal written it is not going to be 5k it is going to be random.

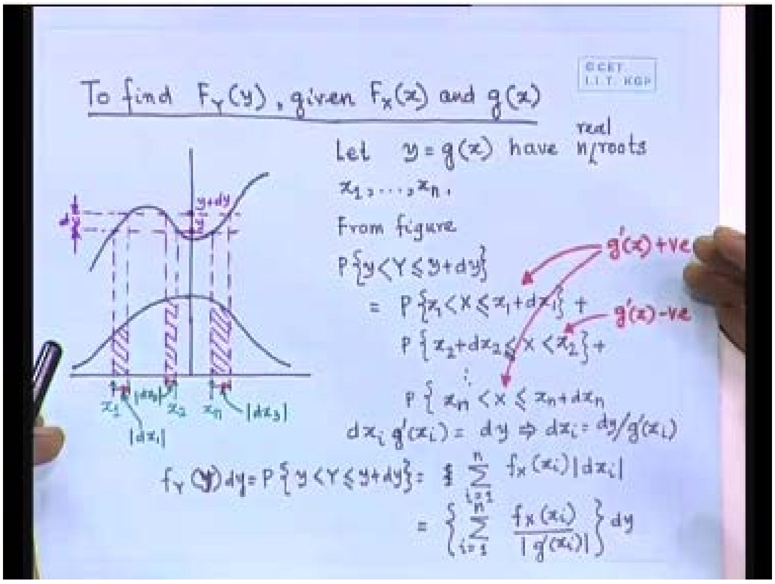
Suppose we imagine that the value of the resistance is a uniformly uniformly distributed variable so then the then the R is the R is the random variable which is which here works as x. Now you can define you can ask so many other question you can say what is... obviously if you connect a random resistor to a to a constant voltage source which for the time being I am I am assuming not to be random in general that itself is going to be random then you can also ask what is what is the current, that is going to be random variable and it is going to be a function of the R, you can ask what is a power. So so this is why you see once you have a once you have something random at the basis you need to determine several derived quantity this is a this is a this is a very common requirement that is why we are studying this property right and there are certain certain properties that this function that this function g should should satisfy otherwise it will not be... if you just take any function g then it will not be giving you a standard variable y right?

For example, if you take... for example suppose you take g as a root function then root over of negative quantity is not a real so therefore corresponding to all negative values of x you will not get a y. Now the point is that those negative values of x we have some probability associated with it. So then the all the probabilities of y will now not add to one. So there are so there will be some probabilities which will be left unexplained right? So so therefore you need that the domain of this function must include the range of X right. Similarly this is that that standard technical requirement of the the underlined field being a being what is called a ((00:08:35 min)) field. These are these are high some high mathematics, we do not need to understand it. But it says that the all the basically the set of X's which which which constitute a valid event you should be able to always form by countable union and intersection of intervals on the real line. This is a this is the requirement because if you... in in some very very pathological cases you can you can dig up a function which will not do this and then you can show that you cannot define a probability measure mathematically consistently on that kind of a set. So we will assume that it happens it will always be satisfied in our cases.

This is a... this is a you know this is some crooked thing which the mathematician think up. They are finicky about every detail; they will not leave any possibility anywhere. But these are will always be satisfied so we do not need to so much to bother about it. And we are saying that the function should also be so defined that if x goes to I mean this this g x equal to plus minus infinity events whenever you will get such events they they must have zero probability why? Because again we we want that suppose g x equal to minus g x equal to minus infinity events must have zero probability because they because nothing can be smaller than minus infinity and g x equal to plus infinity events must also have zero probability because otherwise suppose at infinity you have some some discrete probability suppose that is then the the probability that g x is less than less than infinity is not going to be one. So the moment you can go on go on having increasing and increasing values of y and still not get a probability of 1. So these are you know technical things but they... I mean basically what you have to remember is that when you are defining a function of a random variable I mean you cannot take ideally you cannot take any any arbitrary function they have to obey certain rules, you have to you have to look at that function and verify whether it is a valid function.

Now the question is, now we come to the main problem that is to find... given given the distribution function of x and giving g and given g x how do you find you F y; how do you actually compute it right.

(Refer Slide Time: 00:11:01 min)




Suppose this is the function g x, this side is x and this side is y and this is f x of x okay; I have drawn two graphs the upper side is y so this curve is the is the g x curve; y equal to g x and this is the probability density curve of x; x is a random variable so it has a density function assumed to be and this is that.

Now we want to know what is the density function of y right? So I mean how do you calculate that. So first of all note that for any given value of x right if you take any value of y suppose this is y and this is y plus dy you have to find out what is the density function you have to find out that what is the probability that y lies between y and y plus dy and then divided that by dy that is going to give me the density function, it is the it is the derivative of the distribution function.

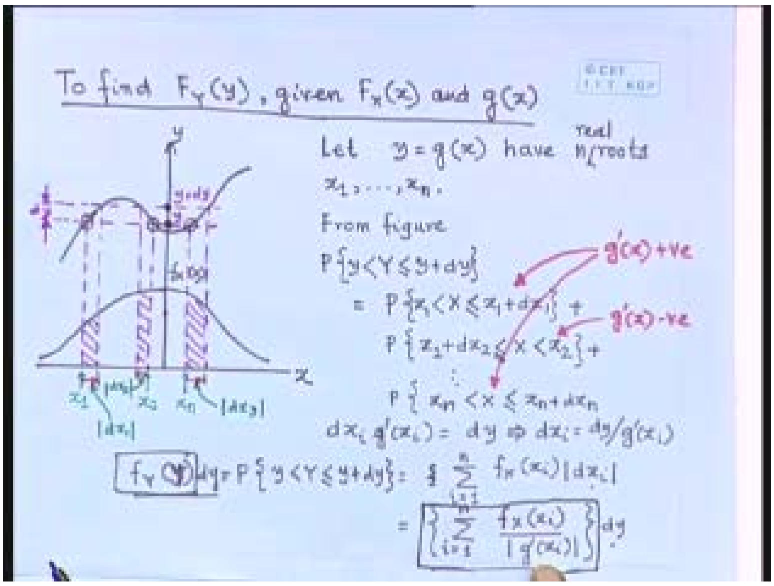
So first thing you have to note is that for a given value of y there may be many values of x. See my final first probability must come from x because I have the disruption for x. Now it turns out that for a value of y there are many values of x: this is one value (Refer Slide Time: 12:22), this is another value, this is another value. So there are three values of x which satisfies y is equal to g x for this value of y. So in general if you put any value of y you may get a number of real roots right. Now that means there that now if you change y around that if you change y to y plus dy then all these points will now contribute the probability so then there will be a change on x 1 to x plus d x 1 plus d x 1, x 2 to x 2 plus dx 2 and x n in general up to x n to d to x n plus dx n in this case 3 and here I have written n.

So basically what is the probability of y plus dy? It is it is this hatched areas this pink hatched area (Refer Slide Time: 13:13) that is the probability that y lies between y and y plus dy. Now so that is what I have written that probability that y is between capital Y is between small y and y plus dy is basically the sum of these probability sum of these hatched areas.

Now it turns out that... see you here we have taken a positive increment. Now it turns out that whether dx 1 there is a there is a small sign problem which you have to take care. It turns out that if you have if you have if you have the gradient as positive as it is here, here the gradient is positive at x 1, here the gradient is positive but here the gradient is negative (Refer Slide Time: 13:59) so therefore similarly you have dx 1 is positive, dx n is positive but dx two is negative right; why? That is because dy is equal to that is simply because dy equal to dx i into g dash x i and and dy is positive, so if this is negative this must be negative.

Now when you are calculating probability you do not see positive negative so your probability is actually is sigma i is equal to 1 to n fx x i into mod dx i so it is this way or that way you are going to have probability. And what is mod dx i? Mod dx i is nothing but dy by mod g dash to x i. If I take mod here I can also write this as mod, this into mod this because this is going to be positive, I have taken a positive increment. So it turns out that s that the probability that y is between y and y plus dy which will obviously be by definition this is given by this this into dy which means that this part is equal to this part.

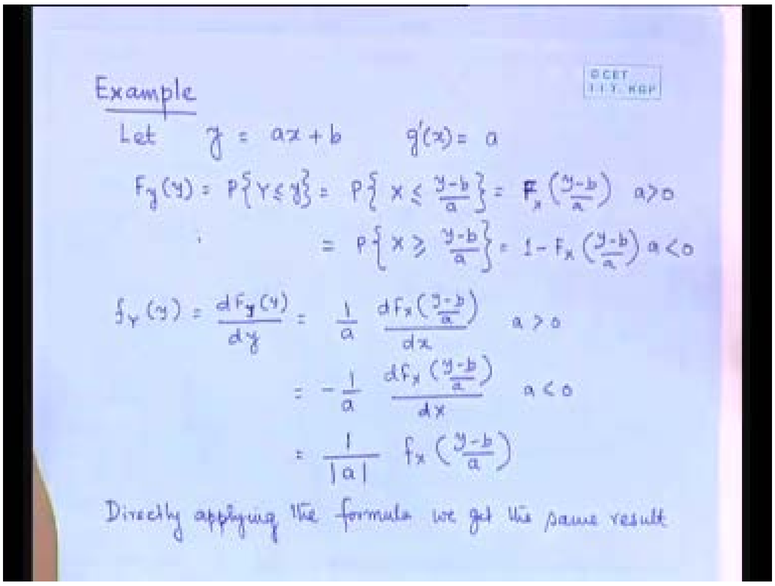
(Refer Slide Time: 15:20)



So now you find the see that given the distribution function of X and the function g if you are given g you will calculate g dash; g dash is a derivative of g, you can calculate f y. So at every point y solve g equal to that y find out the roots then you have to sum at at at all those roots you have to you have to compute this then you will get the density function of y right.

There is a simple this is a this is a very very simple example, in the book you will find more much more complicated examples. One of the one of the one of the simplest maps that you can that you can think of is a linear map. I mean strictly speaking mathematically this is not linear in the sense that it is it is affine that is superposition cannot be applied on this so in that sense it is not linear but it is an equation of a line, it is a fine map.

(Refer Slide Time: 00:16:12 min)




So if you have y equal to ax plus b then what is g dash states? It is a. So now see that what is the probability that y is less than or equal to y? It is a probability that x is less than or equal to y minus b by a simply solve form here. See this is a one to one map so that corresponding to every y you only have one x. So in this case they are they are going to be only one root there is no point where you will get n roots right. So so is so it is a simple case, you need to consider only one one point you do not need to sum any, this is a this is a very simple case.

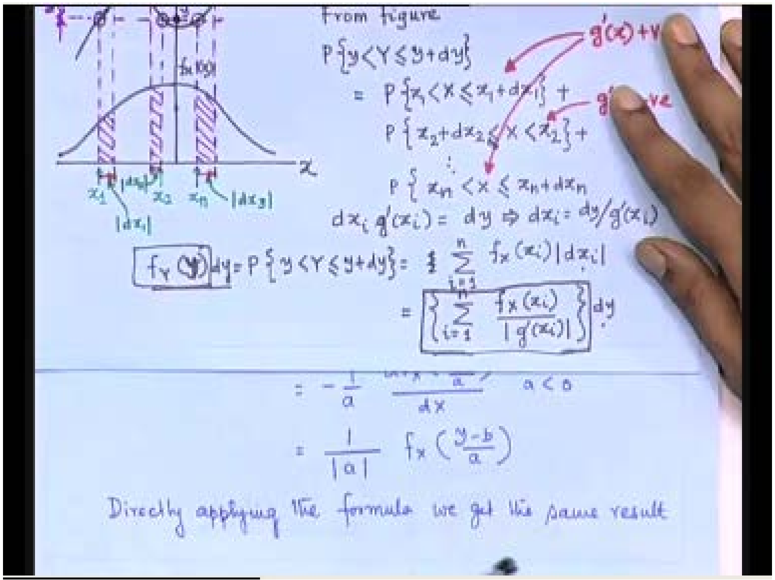
So obviously probability that x is less than or equal to this is nothing but this; in terms of the distribution of x it is this and if a is less than 0 then there is a problem then probability that y is less than or equal to y translates to x is greater than or equal to y minus b by a because a is negative which means that it will become 1 minus F x into y minus b by a if if a is less than 0.

So now what is F y y?

It is... we know that it is nothing but we have to differentiate this so we define it now what is dy? dy is adx. So when you when you are differentiating with respect to y it is like differentiating with respect to 1 by a dx and then dfy we have put this one we have taken from here (Refer Slide Time: 18:00) so this is like this if a is greater than 0 and this is like this minus when a is less than 0 because you are now differentiating this one so here is a minus. So obviously when a is less than 0 minus and a is greater than 0 plus means you can club these two formulae and say 1 by mod a. So this is the... this is the probability probability density function of y given the probability density function of x.

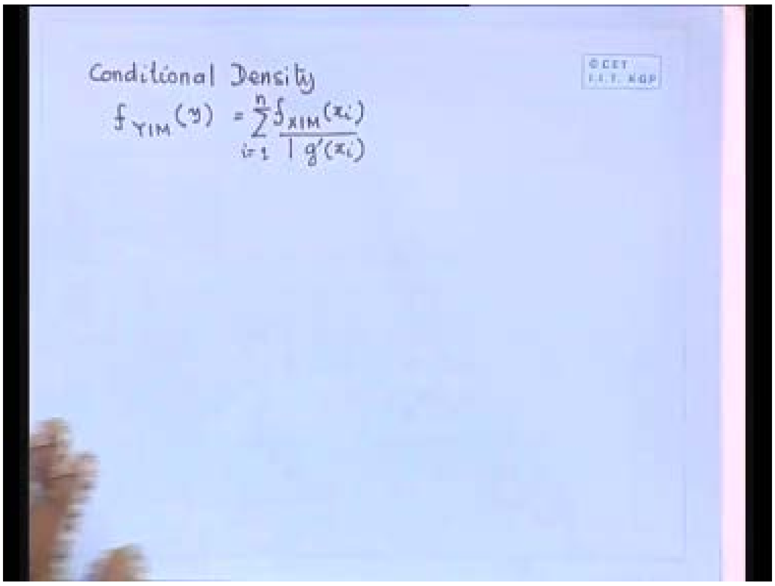
And you can do the... this we sort of you know obtain from common sense principles. Now you can put it in the formula.

(Refer Slide Time: 00:18:43 min)




If you put it in the formula you will get the same thing. See we are getting this mod g dash x i is is this is mod a and and and f x x i is this f x y minus b by a so we are getting the same thing and and we have only one root because of that function. So it is the same thing.

(Refer Slide Time: 00:19:08 min)




Now you can define you know in exactly the same way the conditional density functions can also be defined. It is basically the same thing because the because you can follow exactly the same arguments (Refer Slide Time: 00:19:22 min) if instead of calling this as f x given x if you call this as f if you call this as f x given m if if this is a conditional density function of x given the event M then also the same thing, exactly the similar argument will happen so the same formula will apply. So, f of y given m is is going to be related to is going to be related to f of x given M in in exactly the same way. So so you have so there is no change here and you have the same formula and only here I have put y given M and here I have put x no no no change (Refer Slide Time: 20:00).

[Conversation between student and professor: 20:03 – 20:06] Sir x... what is the... no no no x given M means that what is the probability that what is the probability of having a certain x given that some event has occurred. Now M is an event right okay... [Conversation between student and professor: 20:22 – 20:26] So derivative already defined there cannot... derivative yeah yeah that that is our a-priori knowledge, we know that such a thing has occurred right. For example, we know if we if we cast a die then somebody tells me that you know that an even number has occurred. What is the probability now that a that a 2 has occurred, this is probability of x given M okay.

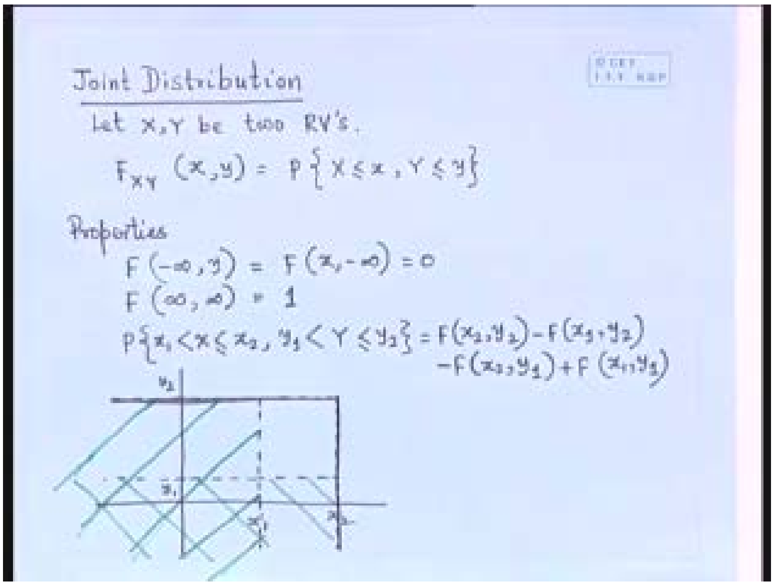
Now we talk of more than one random variable right. So so now let x and y be two random variables. They may be related right. The thing is that they may be related so so that is what makes it interesting, if they are not related they become we will see that they become simple products, relations.

So now let x and y be two random variables. Now the question is again everywhere we are interested to compute density and distribution function because once we know density and distribution functions we know everything about that variable. But I mean as far as probability is concerned that constitutes total information right. But I mean it it actually turns out that it that is very difficult and expensive, not only very difficult, it is very very expensive to find out probability distribution functions of of real objects simply because you have to perform a large number of experiments so that is why people I mean that is why a lot of things have been done where people try to estimate it using small samples or people do not I mean try to try not to require the exact knowledge of the density function and things like that.

Although we are you know very simply talking about density and distribution functions they are they are practically obtaining density and distribution functions of physical real variables are sometimes very very difficult. So but anyway we will we will we will cross those bridges when we come to them. So let us so here we have... now I am trying to define what is known as Joint Distribution function of two random variables x and y and the and and this is the meaning.

So the meaning of F x y x y is the probability that simultaneously x is less than or equal to x and y is less than or equal to y simultaneously that is important right. So obviously what are the properties that F is minus infinity and y or F is x of minus infinity is 0 because y cannot be less than... either x or y none of them can be less than minus infinity. See both have to be simultaneously satisfied. So it is a probability that X is less than or equal to minus infinity and Y is less than or equal to y because x cannot be less than or minus infinity so the probability is 0 and similarly here (Refer Slide Time: 23:23). And similarly F of infinity infinity is 1 because X must be less than infinity and plus infinity and y must be less than plus infinity. But F x infinity and or or F infinity y is not one; in fact it will be the marginal distribution right. So here we can substitute one infinity and get a nice number, here we have to substitute both, here we cannot substitute 1 then we will not get 1 here. This I hope you realized.

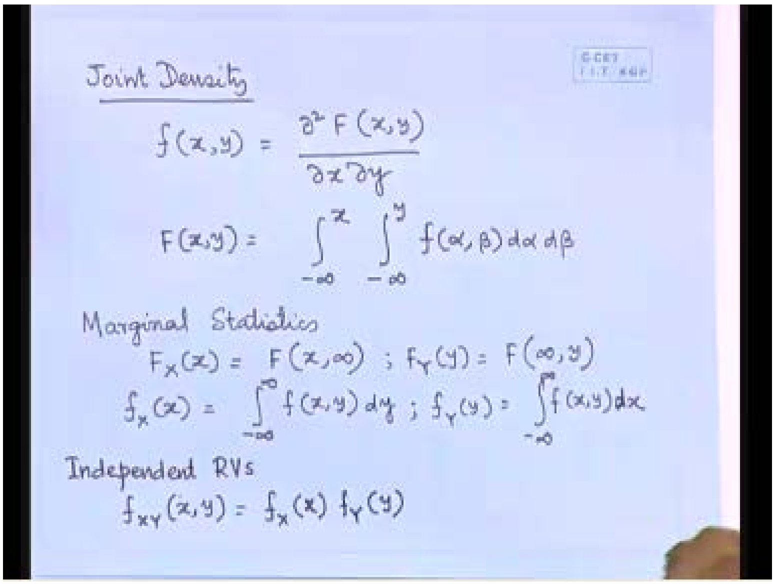
(Refer Slide Time: 23:55)




So now now let us see many times we would like to know what is the probability that x lies between x 1 and x 2 and and y lies between y 1 and y 2 given the distribution functions right. So you see this is this is y 2 and this is y 1 suppose these two lines and this is x 1 and this is x 2. So what is the probability that that that simultaneously first of all what are we talking about we are actually we want to find out the probability of this part right naturally so that you have to find out in terms of the distribution. So we first say that it is this area which is F x 2 y 2 I am assuming that x 2 is greater than it minus F x 1 y 2 what is x 1 y 2 that is this green hatched part (Refer Slide Time: 24:52) this part minus x 2 y 1 which is this part.

So now if you have done that you will actually subtract it this part twice. This part you have subtracted twice so you should you should not subtract twice. So you see you have to make it up that is why plus F x 1 by 1. So then what is left is this square right? Very simple.

(Refer Slide Time: 00:25:29 min)




Now we talk about Joint Density. So far we have seen regarding distribution. So how do you get Joint Density? It is nothing but double differentiation del x del y because now I have two two random variables. Now it is a plain right; so so it is del square F x y by del x del y. That is given this one you can find this one in this way and given this one (Refer Slide Time: 25:56) you have to integrate the density function in this way that is it will be a double integral now because here you have to have minus infinity to y because it has a probability that beta is less than or equal to y so it is from minus infinity to y all possible values then from minus infinity to x all possible values, so it is a double integral, that will give you F of x y correct?
Now actually what happens is that suppose there are there are two random variables and you are asking question about...

Suppose you have a red ball and a green ball and you are saying what happens to the red ball, you are not asking any question about the green ball which means that the green ball can be anything. So all possibilities about the green ball whatever its diameter big, large, weighty all cases are being considered you are you are you are not keeping your eyes closed about the green ball.

You are only asking questions with about the red ball when you are trying to of... when there are two random variables which are related and you are trying to characterize the the the probability of only one not talking about the others that is called Marginal Statistics i mean that is the name given to it so that means that for for all the other random variables which you are not considering you have to consider all possible combinations for them. So therefore if you have a distribution function for the other argument you have to put infinity because they can take any value.

Therefore this is what I was saying sometime back that if you put 1 plus infinity then F x infinity will be will give you the just the statistics of x and similarly F infinity y will give you the statistics of y. So from so so so given the joint statistics of two variables by putting by by putting an infinity into one you can always reduce it to one random variable just simply by ignoring the other; basically it is ignoring the other and if you have... if you have two random variables which are independent. That is suppose the two random variables are actually coming at coming from two different experiments so the value of one random variable is in no way affected by the other. So if if if that is so then their joint statistics is nothing but this into this.

Because this event, because the events corresponding to y are unrelated to the events corresponding with these. So when you have two mutually exclusive events we know that the probability of a and b occurring is equal to probability a into probability b and when we have Joint Density we are talking about a and b occurring together so it will be nothing but the product of densities right? Okay.

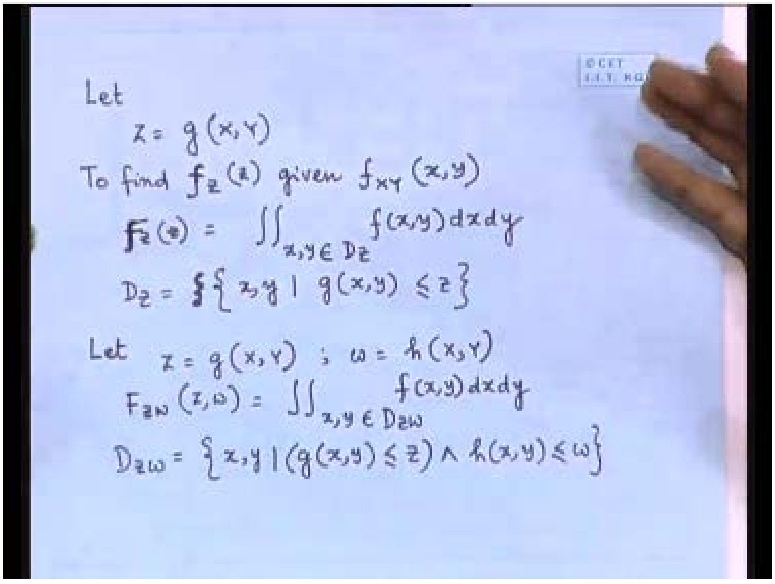
[Conversation between student and professor: 28:51 - 29:06] Sir which is the order of integration... which one? That integral in minus infinity to x integral minus infinity to y is simply... or the other way is different... obviously it will be it will be d beta dr and... because alpha actually I am I am integrating this is supposed to be the x argument this is supposed to be the y argument; see (Refer Slide Time: 29:14) if this is the y argument I am integrating this first so it should be d beta dl.

[Conversation between student and professor: 29:20 – 29:24] Sir it does not that that is a question of notation. It does not matter so much. Obviously if you I mean the variable you are integrating you have to take the corresponding difference naturally, okay.

Now take take this case that z is a function of it is a function of two random variables x and y. It can happen. Very very common example is z equal to x plus y okay. General example is sum of n variables, weighted sum of n variables these are very common quantities which we all the time all the time encounter. For example; we we regularly average four things, so if you take three things and if you compute an average what you are doing is x 1 plus x 2 plus x 3 by 3 so so what are going to be the randomness properties of this average if you know the randomness properties of the sample that is a very common question right.

Suppose we have we have this can we can we talk about the density function of this given the given the joint density of x and y.

(Refer Slide Time: 00:30:24 min)




See in general if you are given x and y, if you have two random variables x and y you need to know their joint density because they may be deducted they they they they may not be related right so so so the joint statistics is the complete information about those two random variables right. So obviously it is very simple that... what is the distribution D that is what is the probability that z is less than or equal to z? Obviously you will have to integrate this over a region which I am calling Dz what is that region? that region is that this g(x, y) this region consists of all those x y's such that this g(x, y) is less than or equal to z. so you have to basically you have to characterize that region and then integrate the joint density function over that region then you will get the distribution function of F z we will see this naturally.

Are you able to follow this?

See basically what, you have to find out if you get have to get the you have to get the distribution function if you have to get the distribution function you have to integrate the density function. now integrate the density function over which points over all those points such that capital Z is less than or equal to z. now what is capital Z capital Z is g(x, y) so you have to consider only integrate about that region of x and y such that g of (x, y) is the is less than or equal to z. Only over this region this probability is applying is it not? Right? And this is another case that is if you have a now... gradually you know one once we have something you will have you would like to have something more.

So once you have defined a z is equal to g(x, y) then you say okay let z be equal to g(x, y) and let another random variable w be another function of x, y which is h(x, y). Now obviously once I have z and w, I can define their joint statistics because because they are also two two random variables and since they have both functions of x and y so this joint statistics must be related to the joint statistics of x and y because x and y are the basic entities, z and w are basically defined from x and y so their joint statistics must be computable from the joint statistics of x and y right? How to do that?

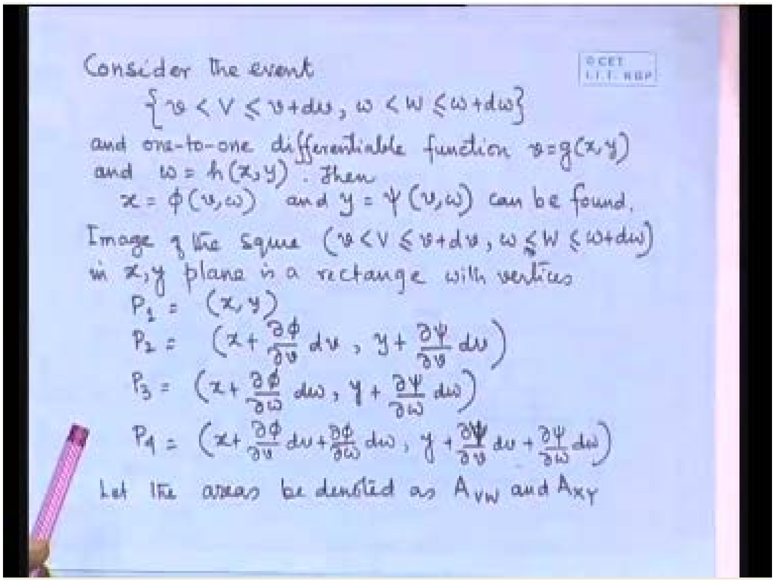
So again same argument that is that the basic basic reasoning is very simple basic reasoning is that what is going to be this distribution function?
It is the probability that capital Z is is less than or equal to small z and capital W is less than or equal to small w right which means that we have to... now the probability is actually coming from the random variables x and y because they are the basic random variables; z and w are derived random variables right so we have to now integrate this function over again somebody here just like here.

Now what is this region now? It is now going to be a different region. Previous region was same that g(x, y) is is less than or equal to z. Now I have to choose my region such that both these conditions have satisfied, naturally. Because this means that both z must be less than or equal to z and and w has been less than or equal to w. here there was only one condition. So now I have to... so basically my region will become smaller right but anyways still it is a region. So now I have to integrate it over that region.

See we are... this is sounding abstract but but the moment you solve the problem go to a problem as we will go next week you will find particular examples of g, h. So immediately by by inverse solving you will you will have to compute this region and then if f(x, y) is given then you have to evaluate that integral; in evaluating that integral sometimes maybe you know it is difficult, it may involve beta gamma functions those things are separate. But the basic idea is that the integral must be evaluated over a region, over the appropriate region that is all. So that way it is uniform.

Now here, here what they are doing is that here they are... now the question is what what you get is actually by doing this integral what you get is the density function the distribution function. Now the density function if you want to get then you will have to differentiate it right? So rather than that here is a general way of getting the density function right?

(Refer Slide Time: 00:35:34 min)



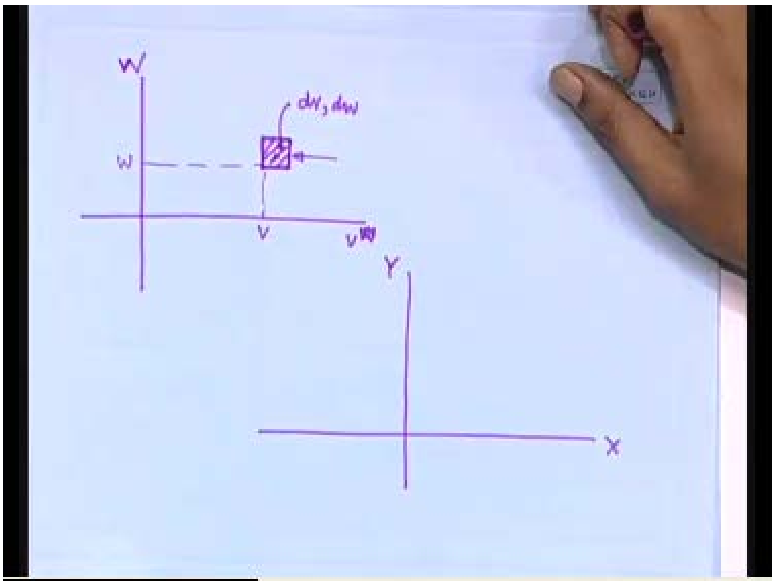

So what does it say?

This is again principal wise it is not complicated. So what is the density function that is f let us say... previously we have taken z and w here and actually I have it is a human error because I have taken this part from a from a different book I mean without recognizing I have gone to a different notation. So here we are talking about f of (v, w) (v, w) is equal to the probability there. Actually we should say dv dw is it not? That is what is the probability that capital V is between v plus dv and v and capital w is between w plus dw and w what is that probability? That probability is nothing but this joint density into dv dw; dv dw is the area; density into area will give you mass a mass area or volume right?

So basically what you have to characterize is this is this probability that is what you have to characterize and you have to get it in terms of x and y, that is how that is the thing right?

So now what do we do? Now you see that we now have two spaces okay.

(Refer Slide Time: 00:37:06 min)




For example you have a space called v w and you have a corresponding space of x y. This is y, this is x, this is v, this is w, or maybe this is v, this is w. So what are you considering, you are considering an elementary volume which is dv dw and this is v and this is w. You are considering the probability... you are basically integrating the function over these.

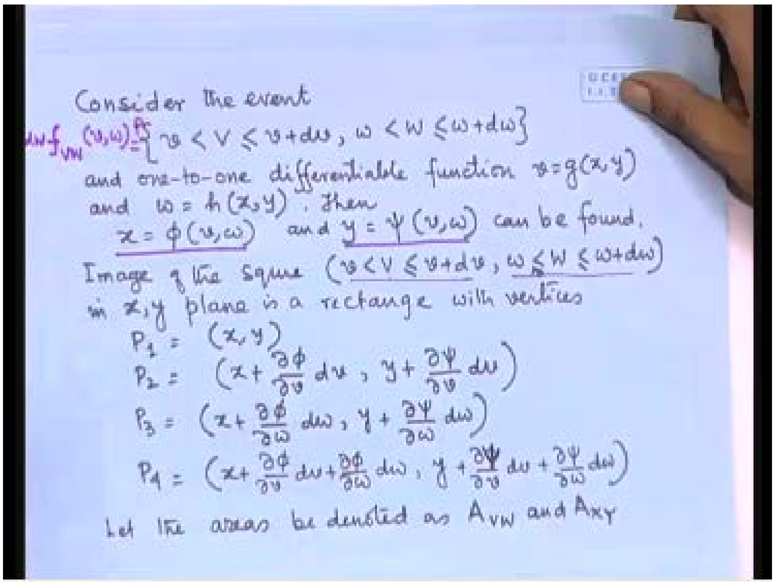
Now if you have to find out the... now this corresponding to this... where from this has come? This has come from x and y. So you have to identify that that which is this region in x and y which has given rise to this region in v is it not? You have to find out the corresponding region in the x, y plane and you have to integrate the f(x, y) function over that region agree.
So the... so in general what will happen? Now that you can do.

For example this this square (Refer Slide Time: 00:38:24) that is where v is less than or equal to this and w is less than or equal to this, this square in general maps into what, in general maps into a parallelogram.

Actually I have drawn it.... I should not draw it like this, I should draw it..... in a... I mean in a general case it will be arbitrary, it will be an arbitrary parallelogram.

[Conversation between Student and Professor – Not audible ((00:38:54 min))] That is a parallelogram, it might be it maybe it maybe a... actually we are we are we are considering it as a parallelogram because we are considering differentials because they are so small then that curvature that that amount of difference is actually an actually a higher order differential so we can ignore it, okay?
So now what are the... what are the coordinates of this? That you have to find out.

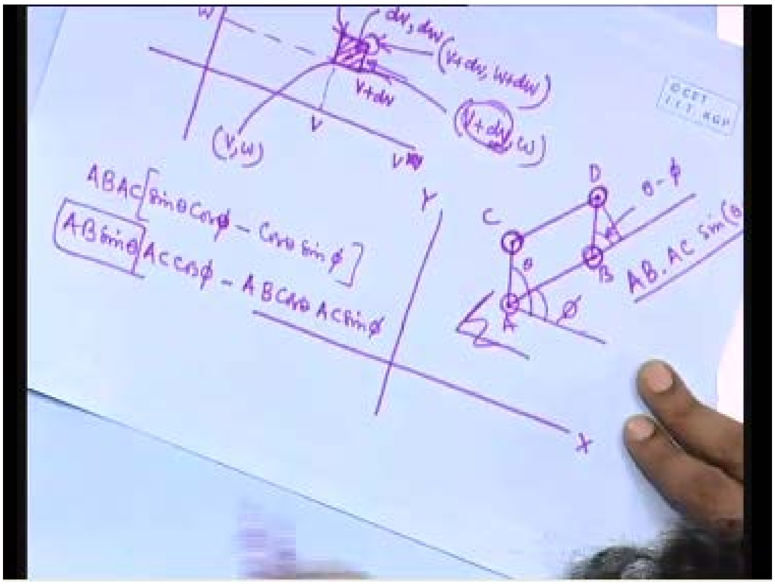
(Refer Slide Time: 00:39:27 min)




So the coordinates of this are... obviously one of them is going to be x, y. Basically what we are doing is this function if this function is one to one and differentiable when can functions be inverted. That is if you have y is equal to g x when can you write x equal to h y? This is the same relationship when they are one to one otherwise you cannot. If it... if it is many to one then you cannot invert it right. So if it is one to one and differentiable as you mean we are taking that in case then then you can define the same function maps and you can invert and you can write x equal to phi v w and y is equal to psi v w that is I am doing different. Previously I was saying v is equal to g(x, y) and w equal to h(x, y). So I can invert them, right?

So now now it is now simple because once... why I have come because because I am going from dv dw to the corresponding increments in the x, y plane I want to identify them. So now once you have defined these maps this parallelogram coordinates are going to be... one will be x and y that is given v and w one one coordinate is going to be x and y that is phi (v, w) and psi (v, w); the other coordinates are going to be this, it will be x plus del phi del v dv and del psi del v dv.

(Refer Slide Time: 00:40:56 min)




Basically what are these coordinates? This coordinate is V, this is V plus dv. So what is the coordinate... so this coordinate is (v, w), this coordinate is (v plus dv, w). So corresponding to this v plus dv both x and y will change. So by so by how much will x change because of this dv? x will change by del phi del v into dv because x is phi (v, w) right. By by by how much will y change because of this dv? That will be (del psi, del v into dv) because y is defined by the psi function.

Similarly, if you take this coordinate (Refer Slide Time: 41:42) that is (v, w plus dw) so now w is changing, so now if if if w changes then by how much is x and y changes? Now it is del phi del w into w and del psi del w into dw so these are the corresponding coordinates we are getting. And finally here you get both v plus dv and w plus dw so both are changed so you have to add their corresponding things.

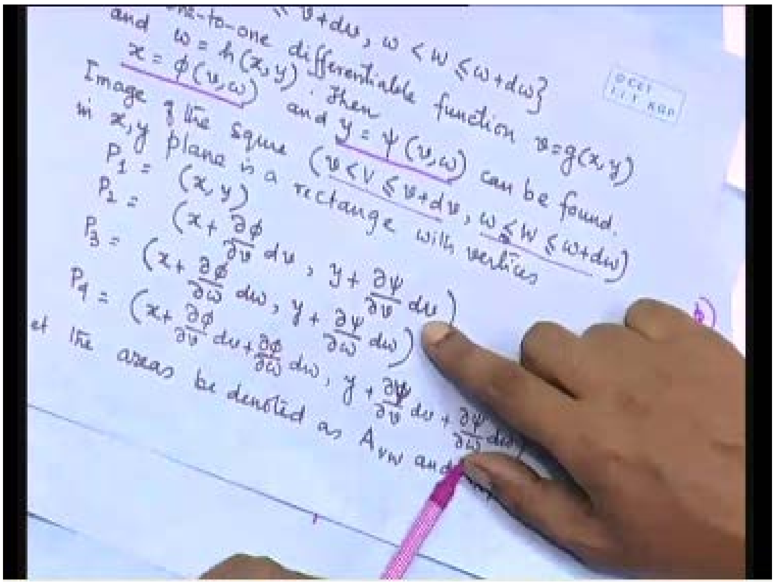
So now you got the four coordinate points on the x y plane, now you have to find its area. So what is the area of this? The the area of a parallelogram is... suppose this is A B C and D so and... suppose this is say phi and this is say theta then what is what is the area of this parallelogram it is AB it is AB I hope you can see it, it is AB into this one (Refer Slide Time:

43:00) so this angle is theta minus phi, this is AC so it is AB into AC into sin of theta minus phi that is the area of the parallelogram. Agree?

So now sin theta minus phi is what?
So it is AB into AC into sin theta cos phi minus cos theta sin phi. So now you can say that it is AB sin theta into AC cos phi minus AB cos theta into AC sin phi, like this you can... now if you find out these now what is rather rather rather once second one second.... rather we will say we will combine them differently, no, but we can do that same thing.

Now you see what is what is AB sin theta? AB sin theta is this one, AB sin theta is we should we should we should we should do it the other way. We should combine AB with cos phi and AC with sin theta. Then AB cos phi is what AB cos phi is is basically the x coordinate change between this point and this point. So what is the x coordinate change between this point and this point? That we know. Suppose this is P 1 this is P 2 then the x coordinate change is between these two points is del phi by del v into dv.

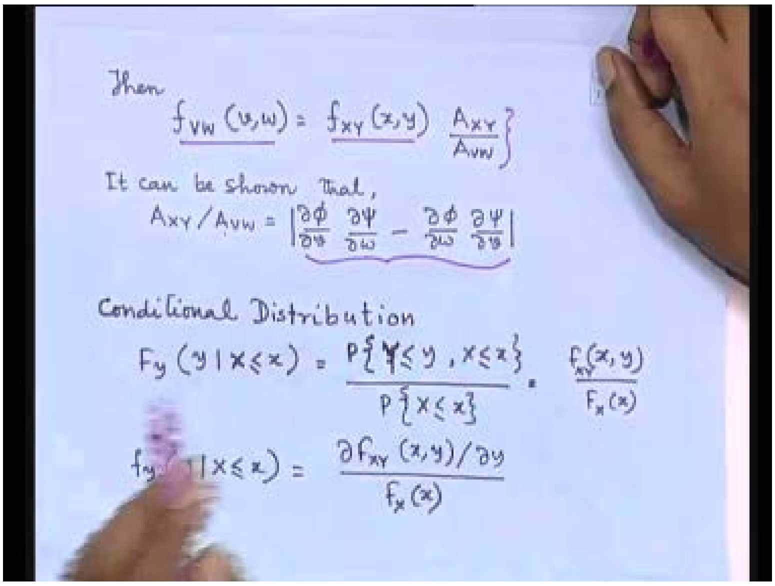
(Refer Slide Time: 00:44:37 min)




See I am trying to find out a general thing. Of course it is not fully general in the sense that it is not in n dimension but in 2 dimensions I am doing a general thing. So if you if you now simply simply calculate these terms what you will get eventually is this. What you will get is the... where is that? Okay. What you will get is the area is let let us see... so now what what

what you can what you should do is... just a moment, let me think otherwise I may miss something, it should have it somewhere here.

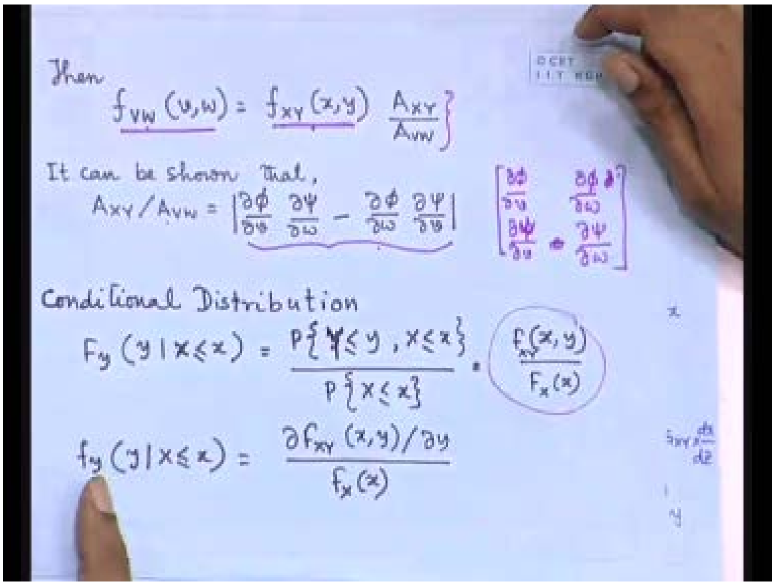
(Refer Slide Time: 00:45:46 min)




So now so what we have is that... finally the probability is must be the same whether you are computing it in the v w space or whether you are computing them in the in the x y space probability should not differ. So therefore f w v into this small area dv dw which I am given in the name area in the v w plane must be the same as f x y into x y into into A xy the corresponding area in the x y plane. So so now we can compute this one in terms of this one and this area ratio.

Now you can you can just workout using the principles that we have said just now and you will get that A x y by A v w is this quantity (Refer Slide Time: 46:36) which turns out to be noting but the determinant of this matrix del psi by del v you take this matrix, if you take its determinant then you will get this into this this into this minus this into this that is this one, right.

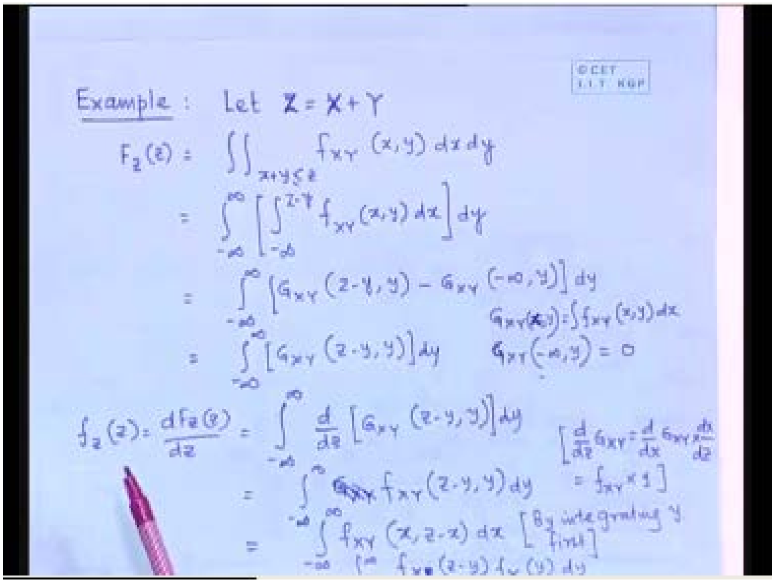
(Refer Slide Time: 00:46:42 min)




So this formulation actually gives us... actually from this you can you can really generalize it in the sense that if you go to n dimensions it turns out that you will you will get an nth order determinant but simply this matrix will become nth order that can be proved; of course the proof is much more complex but the but the idea of putting it in this form was to extend it for n dimensions so that you can even compute the interrelations between n random variables. So this is so this is how you can calculate directly calculate that is given the phi's and the psi's and given the density function... see here we are not going for the distribution function, directly from the density function and the phi's and the psi's you can calculate the density function of the derived variables.

And then conditional distribution is the same thing for example what is the probability that of that of y given that x is less than or equal to x; such things. Obviously it is this this is Bayes formula. So this Y is less than or equal to y and X is less than equal to x divided by X is... so this is this. So we have... here is a small example which is which is one of the very main cases that one should see.

(Refer Slide Time: 00:48:47 min)

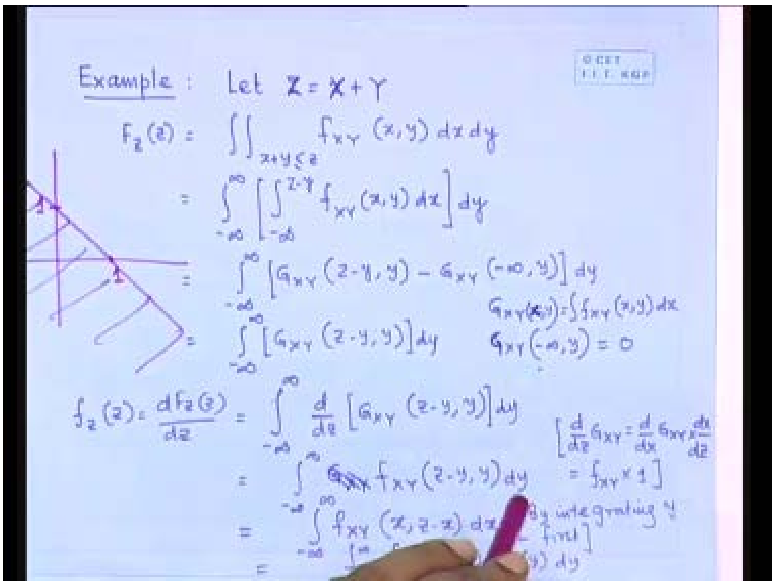



For example; I mean Z is equal to X plus Y because this is.... and one of the most common way of combining into random variables is that you may... that people would like to do simply add them, you can extend this result easily to adding with weights and adding n numbers some things like that that is very simple.

So here what they are doing is again... if Z defined in this way, so what is going to be, what is going to be the distribution function? It will be the integral over in region over which region? What is now my dz? I was saying that you have to integrate over a certain region dz now.

Now what is the dz? It is those set of points x and y such that x plus y is less than or equal to z.

(Refer Slide Time: 00:49:32 min)




So basically it is... if you have the x y plane so this is your 1 and... suppose this is your 1 and suppose you are having z is equal to 1 then this is your x plus y is equal to one line so you should integrate over this region (Refer Slide Time: 49:53); for all points on this x plus y is going to be less than or equal to 1. So like this for different values of z you are going to take different set of straight lines and then integrate over the corresponding region.

So now what is going to be this?

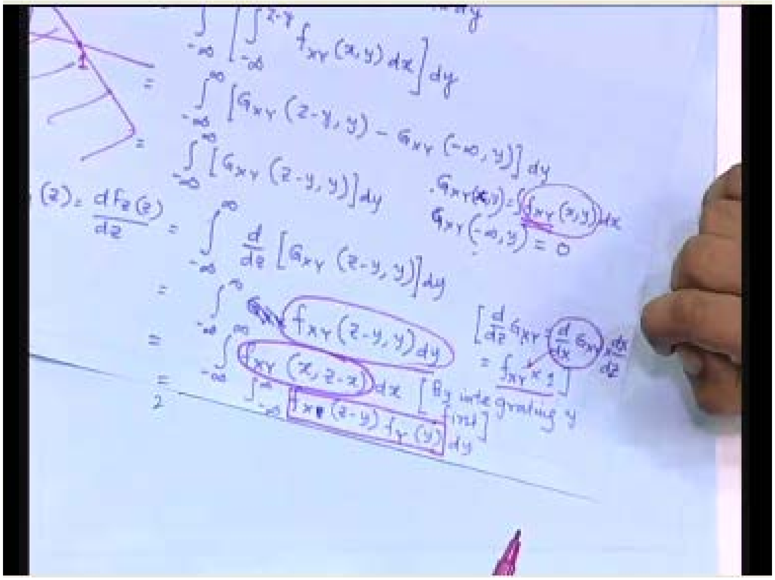
Obviously if x plus y z then what is x? x x varies between minus infinity and and z minus y; I am I am actually eliminating one of the variables, three variables are not needed because they are actually related by one equation so you can use one equation to eliminate one variable so I am now eliminating it, so so the limit of x become z minus 1.

So what is going to be this?

This is going to be... if if you take the take the indefinite indefinite integral of f(x, y) if you denote by this G xy then it becomes z minus y y and minus G x of minus infinity y, simply putting in place of x these two limits. If if this G function is the indefinite integral corresponding this then how do you get a definite integral? Just by putting the limits. So you have... this is this.

Now G xy of minus infinity y must be zero because no number can be less than minus infinity so this term is gone so you have only this term. Now if you want to if you are interested in the density function then you have to differentiate this so you put it in and finally obviously because this is this is supposed to be the supposed to be the indefinite integral of this so so what is going to happen d dz of G xy is equal to d dx of G xy into into dx dz. now d dx of G xy is what? It is nothing but f xy because f xy is the is the indefinite integral of I mean G xy is the indefinite integral f xy with respect to x.

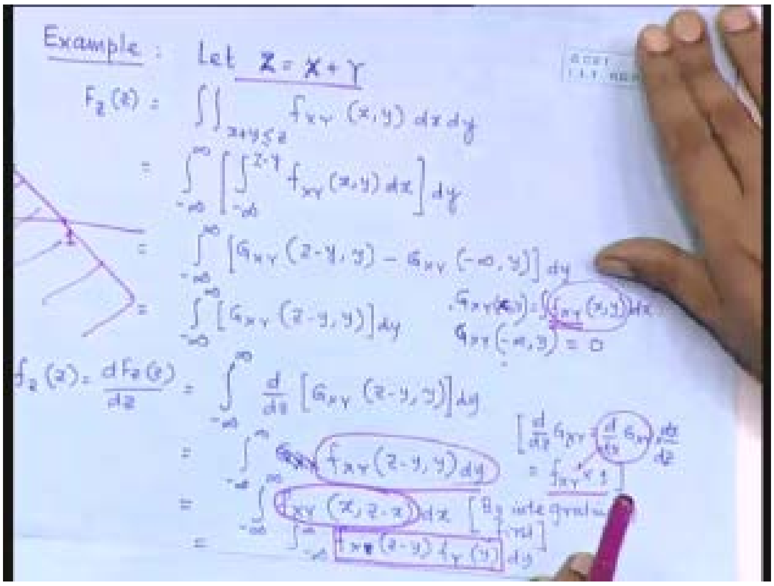
(Refer Slide Time: 00:51:57 min)



So if you differentiate this with respect to x you will get back this so this part becomes this; and what is x? x is z minus y so so d of z minus y dz is 1 so you get 1 so you get only this term if you get the density function.

Now here what we did is we we we first integrated x then integrated y. Nobody tells you to do that, you can also do it in the reverse order it is all the same. So then if you do it in the reverse order you will get you will get nothing but f xy x and z minus x. Here you have eliminated the variable x by integrating first. You can also integrate the... you can also eliminate the variable y, you can always do that. So then it will be this and and if these variables are independent that is x and y if they are independent variables then their joint distribution will be replaced by this part, it will simply be the product of the individual distribution individual densities and this is what is known as a convolution integral sometimes. So so so this is a special case, just to demonstrate that how you can get a get a special case when this function is Z is equal to X plus Y which is a common function.

(Refer Slide Time: 53:37)




So, as I was telling we will we will we will stop here today. As I was telling that next week we will have tutorials and though I am teaching all kinds of theories and which may seems a bit abstract but what is expect is that you will apply the theory; I am not expecting that you will you are going to memorize these theories, it is not at all needed though the theories I am teaching. What I expect from you is that you will read the theory and you will be able to apply it in a situation, in other words it means that you will solve problems.

So the set of problems is very simple. I have got two books here, they have got nice exercises, large number of solved examples and exercises, so I expect that you solve them and in the next class, next Tuesday we are going to have tutorials for 2 hours where I will solve I will try to solve only those problems that in general you could not solve, so the ball will be in your court, you will have to point out to be saying that, Sir I could not solve this problem; it is not that I will come and solve another set of... because there are... this already contains a number of solved examples so so I am not going to redo another set of solved examples okay. So, if we... so if you do not point out problems I will assume that you could solve everything and then we will have a nice half........... alright?