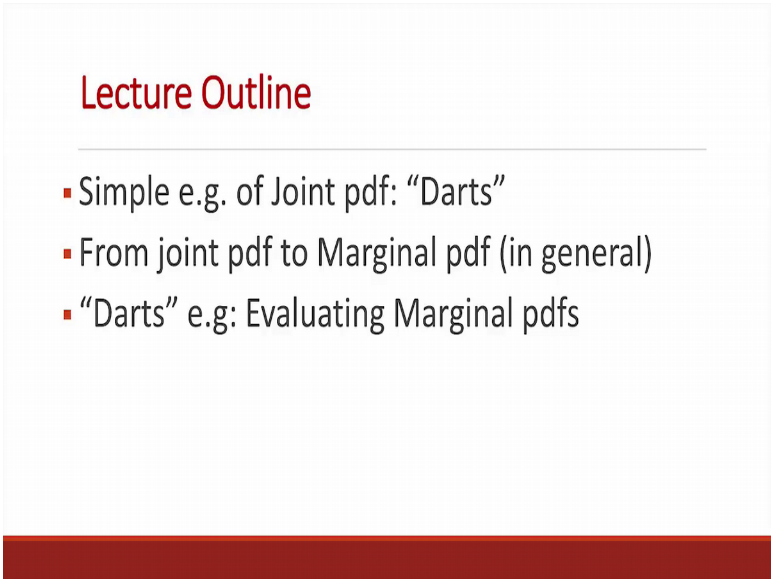
Probability Foundations for Electrical Engineers
Prof. Aravind R
Department of Electrical Engineering
Indian Institute of Technology, Madras

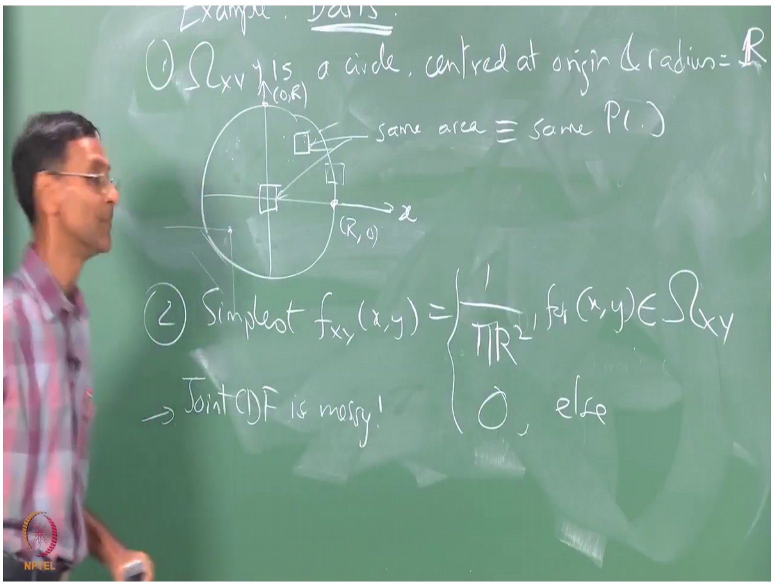
Lecture – 54
Darts Example and Marginal pdfs

(Refer Slide Time: 00:16)



The most popular example is darts. Why is darts so popular? Darts is very popular because of the simplicity of the model itself all right.

(Refer Slide Time: 00:39)



So, we are going to assume that omega xy is some circle and for lack of anything better say if we do not have any other circles any other anything else to use as yardstick you might as well have origin center and radius equal to 1, unless no there is no other I said no other measuring devices use along like this, is not it.

So, you can always said this is the only thing going on which is what we are going to assume here centered at the origin and radius equals 1, if you so want you can generalize it to have some arbitrary radius R as a parameter, but let me not I do it now or maybe not, let me say radius is R why not; that means, a radius equals to R. Just for, because otherwise you know it is some formulas will use the R and that will be more like the role was the R will become clear let it be R, some capital R so that you know it is a constant.

So, in other words what we are saying is that we are going to pick points at random only from this circle this is the point R comma 0, this is 0 comma R we are never going to pick any points outside the circle I am going to pick points only from here. How I going to pick points from here? You are going to stand up there and throw a dart and hopefully you will never go outside we will go only inside. Are you good enough for this? We do not know fine that only is omega is.

What about the f xy? What is the simplest model that you can think of? A uniform value for all of omega xy, the simplest f xy I have to say it is one by pi R square let me just say why do I need one by pi R squared this circle right. So, in this case I could have used omega itself, but I am as I as I just said somewhere out there where we are going to keep the generality in mind and always use only omega xy rather than say you know it is only R squared. So, it is enough for me to. So, maybe I do not want to make exceptions like that. So, I going to uniformly use this omega xy full for the region in which they take problem values.

So, everybody agrees to the simplest f xy, you cannot have anything simpler than this. What does it mean physically? It means that you people when you throw your likely do pick any point on the inside the circle, arbitrary close to the origin and arbitrarily far close to the edge also. So, this model has been used to describe a novice player where the player has not much control over where the dart lands. So, regions around the origin do not have higher probability compared to regions closer to the edge.

You can debate the usefulness of this model, but as I said it is just an example at this point. If you want to use a more complicated model fine do it, right, (Refer Time: 04:44). Another byproduct of using this model is what? A certain region of some area has the same probability no matter where that particular region is located inside the circle. So, if I take a circle, let us say a small square out here of some area and I take a small square out here or whatever of the same area. So, same area means same probability. And what is the probability? It is just area multiplied by pdf because pdf is constant the integral reduces to just multiplication of the pdf value at the area of the region D. So, that is true in every case of constant joint pdf, in every example of joint pdf the probability calculation cannot get any simpler because its regardless of the overall shape omega xy for any region which is entirely inside the omega xy the probability of that region is very simple to calculate. If this region lies partly in and out well you only have to take the region the portion that is inside that is also obvious.

So, what are we going to say here f xy is not only is it 0 by domain extension you have to put this, at all the last time I am going to write it, but let me just write it at least once so that you know that this is part of the definition. So, you can you know put some square out there and say well I only this part will count in this part f xy will be 0. So, it will not give you any property, fine.

So, you can see in this example how messy rest to for to specify joint CDF. What would be the you know; if you would pick some point out here say for example, you this point is probably not one of the more difficult ones to look to find a joint CDF, but if I pick something out here or even here you could see that the dependence on this area, on this point is very complicated in general it is not easy, even here it is not so easy.

So, joint CDF is automatically I do not have to emphasizes anymore very messy. Only some values you can easily specify the joints CDF. For example, what is the joint CDF with the origin?

Student: 1 by 4.

Ah?

Student: 1 by 4.

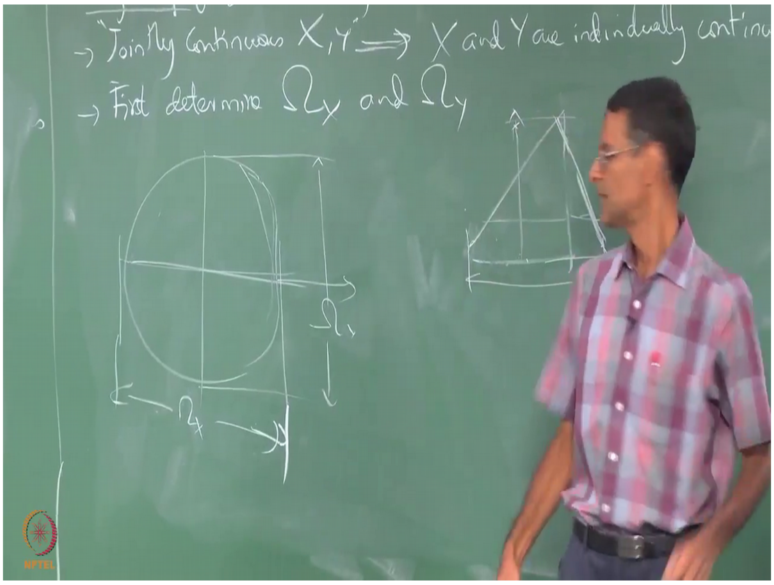
1 by 4, and the joint CDF here is what?

Student: Half.

Half. Joint CDF here is also half these are the only 3 points, I think where you can write the joint CDF exactly without any calculations at all. And of course, the joint CDF will become once that we that semi infinite region completely includes the omega xy. So, this we will come back to later, just keep this example around. Maybe I will keep it here for the rest of this today's lecture.

Now, there is not much more we can do now without covering a little more of theory, so let us continue with the marginal pdfs in general.

(Refer Slide Time: 08:37)



We already said that x and y are continuous random variables to begin with. So, they must have an independent existence independent pdf on their own without having worry about the join all the time. So, how are you going to go from joint to marginal in this case, the jointly continuous case?

So, what term I am going to use when you when you are defining omega x y and f xy and so on we are going to use the term jointly continuous to describe them, which means automatically they are individually continuous as well right. So, jointly continuous X Y implies that X and Y are also individually continuous and therefore, their individual statistics are very important, in the one the discrete case what did we do we added we drew a straight line for some x we looked at all the points that lie on that line and we added all those probabilities because the what do we call it called marginal pmf is it not.

So, what do you think we should do here?

Student: (Refer Time: 10:06).

What should we do here?

Student: Integrate about the (Refer Time: 10:15).

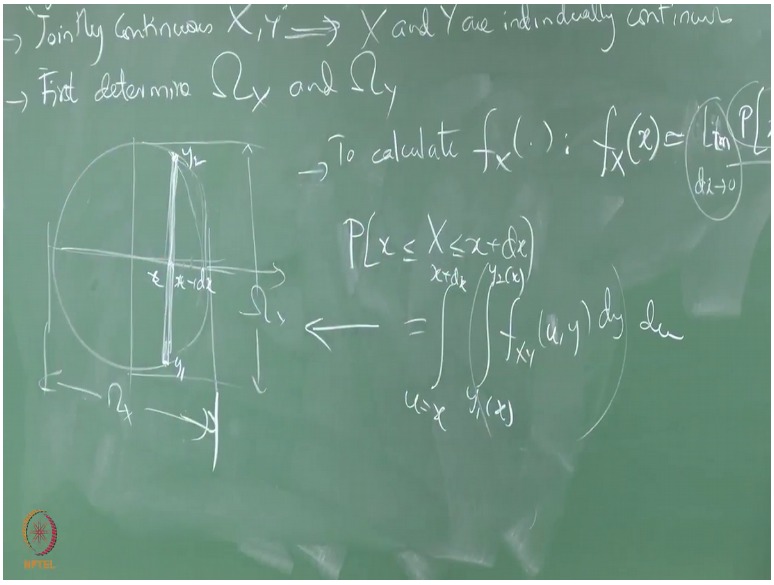
Integrate. But before that before doing integration and you have to freeze what is omega x and omega y where do these random variables live in the individually. So, what do I mean by? So, first determine, if I have some arbitrary singly connected region similar to that, but not exactly that. Omega x and omega y will be the maximum. How do you call this? The maximum spread of x and y which includes all of omega xy.

Why is this important? Because here for example, you know the capital x both capital x and capital y we will take values in what, what interval? Only minus R to R that you have to determine that it need not be as symmetric as this is a very nice symmetric situation because we put the origin plumb in the middle, if we did not put origin plumb in the middle it is not going to be, so nice is not it.

So, in general your omega x omega y is a maximum spread if you will, it is a unscientific way of saying it, but I do not think I put on a better way of saying it out here, anyway. The picture explains it. If you have some unfortunately this figure also is very symmetric, but if I take took some took some non symmetric thing like, took some non symmetric, but it is bounded by straight lines if this was my right. So, this would be spread of x and this would be the spread of y and so on. I do not think we need to spend too much time on this.

Essentially omega x will determine both of these without any reference to probability you know where x and y jointly take values you know where they also individually take values that simple as that, right.

(Refer Slide Time: 12:32)



Now, what you do is you freeze let us say you want to calculate f x. What we do? You pick some x in omega x just like we did in discrete case you draw a vertical line. So, let us say you want to calculate the pdf of x at this point. So, you draw a vertical line with delta x.

So, how are we going to, we are going to use the fact that f x of x is basically a limit well we will first write probability can I hope this is I think it should be you can always do this and then this limit delta x is, this is not a strong mathematical statement it is just a review you know what will happen here is this dx will cancel and in a way its I mean we can write this it should be it should written properly may put l in here. So, it is a limit as delta x goes to 0 of this ratio, that is f x of x, just by standard arguments that is the meaning of a density function is it not. Exactly as we had in the 1D case; there is no difference is; note that this small x is a fixed number as far as all that operation is concerned. So, the entire operation on in there small x is a fixed number delta x is also some fixed interval.

So, what is varying here? You are varying y from here to here. The limits of the variation of y are determined by omega xy, but all books we will say integrate from minus infinity to infinity if you see in books. Unfortunately that does not tell you much about how to apply it in any particular case. So, I would term it as some kind of intellectual laziness, intellectually lazy to write y varying from minus infinity to infinity, but that is the way in which is written, but anyway.

So, what is this? So, to calculate this numerator this numerator it turns out is a nice aligned rectangle if you do not worry about the starting and stopping points. So, what is this p of x? The numerator p of this probability that x is in, x takes values only in some x to x plus dx all right. So, is what? Is basically if you use that formula which I erased the integral basically you integrate you start with f xy you assuming of course, you have the joint pdf you integrate over y from some y 1 to y 2 which depends. So, we should write though y 1 and y 2 we should not be intellectually lazy we should not run away from the fact that they are there is always some y 1 and y 2 and that y 1 and y 2 will depend in general on the value of x you pick.

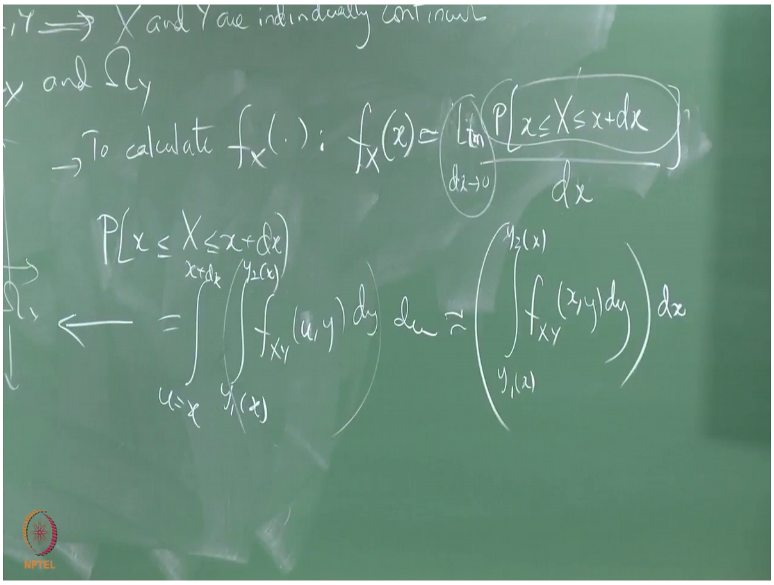
So, I should in general be prepared to write y 1 of x here y 2. This is just one example, I mean it I can have multiple such intervals also its supposing I have a bunch of you disconnected regions for omega x y. This is where I have a single region, that is why the books run away from the saw from this they are talking about this all these possibilities and immediately write minus infinity to infinity very convenient right.

But, supposing here we have a y 1 and y 2 a single y one and a single y 1 which is smallest and single y 2 which is the largest then. So, this in this only worries about this then you have to integrate over x from and I should not put I should its better do not put x here I should put some u something integrate du. So, this is the inner integral the outer integral is from what u equal to let us say x plus dx never reuse the variables, I hope you are all aware of that you should never reuse the variables inside and on the limits you have enough letters in the English and Greek alphabets to at your disposes. So, do not ever put x dx here and put x and x plus dx out here.

So, if you do all this what do you get? If you assume that there is f xy does not vary much as you go from x to x plus dx. So, what will they this is, this integral will stay as it is except that it will become if this is constant over the variation with x, so this will be basically roughly equal to this integral y 1 x to y 2 x of f xy x comma y dy the whole thing multiplied by dx. Can I say this or not?

Student: (Refer Time: 17:41).

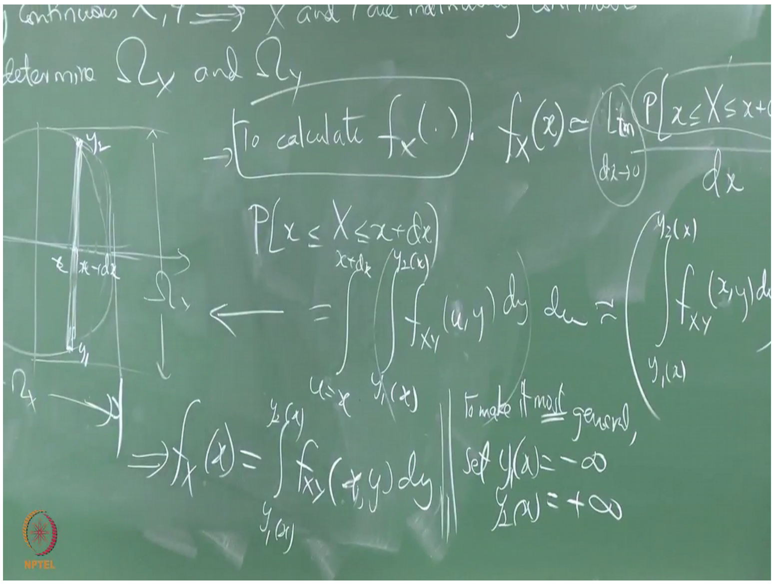
(Refer Slide Time: 17:26)



Now, I can put x in because I am integrating over only this I am putting that value of x in here. This number as a number does not change for any let us say for all values of u it is if it is the same number that is what I am saying, this number does not change when I vary u from x to x plus dx. So, it sort of comes out and all I have is integral du from a and what is the value you know, what u am I using; I am using x of course, I could say I can use any number between x and x plus dx here, but that makes it unnecessarily complicated and ultimately I am going to take delta x going to 0. So, I might as well put x there is not it. And the limit as delta x goes to 0 or sorry this d x goes to 0 I have to put x only, I cannot put anything else here.

So, just this is just going back to how double integrals work, is not it. So, what is the effects of x in this find the limit? So, you do all the, if you divide of by dx and take then the dx goes away. So, the limit can be peacefully taken without any repercussions whatsoever it becomes what?

(Refer Slide Time: 19:05)



So, the bottom line is you will get this very important expression the marginal pdf of x to calculate this write the marginal pdf. So, it is basically the integral y 1 of x 2, y 2 of x, I am keeping it here of f xy the x that you are interested in comma and you are integrating over y.

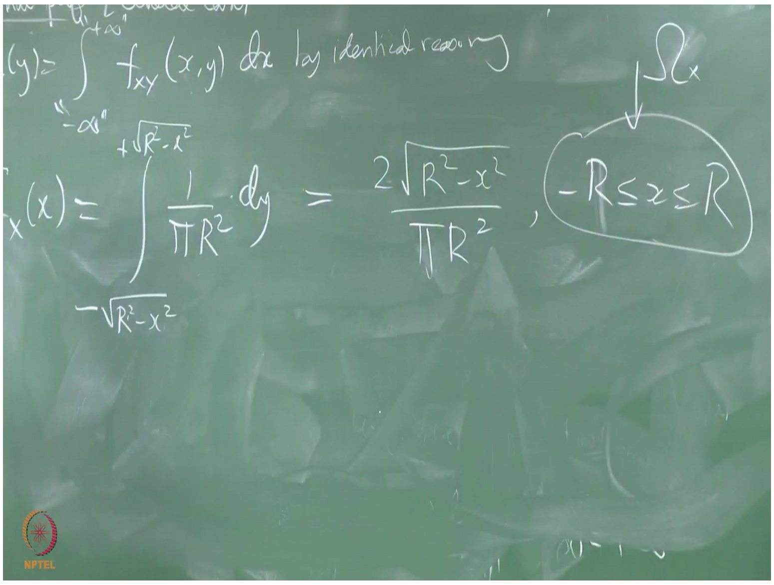
So, it is very important here the notation the x is the value for which you are evaluating the marginal p pdf the y is a variable of integration this y 1 to y 2 is what in books you find written as minus infinity and plus infinity by domain extension. So, to make to make the formula most general make it most general most general said y one of x equal to minus infinity and y 2 of x equal to plus infinity which is also said. So, hide all the details under the carpet that is another way of thinking about this, is not it.

Now, none of them I see details about what omega x y and is there anything just simply hide it all, this is the way you find in books. So, you have to understand what that is when you go look at that formula in the bookends what is it really saying it is basically exactly saying this and this is the idea which will keep climbing up a lot when you D look continuous random variables. So, even 1D not just 2D.

Because what is the ultimate meaning of a density function is exactly that, you are evaluating the probability in a small interval divided by the length of that interval and take the limit as that interval goes to 0 that is the meaning of a density function with one variable. Is this clear, everybody you all understand this at least over as here I cannot worry about who is not here, but at least ones who come.

So, what about y? Say exactly the same thing except you are going to think of horizontal strips instead of vertical strips and I will keep that there and maybe I will move it write it here its implies that f y of y is basically again I will put this now minus infinity to infinity in cloth in quotes because; obviously, this cannot be low small as minus infinity that cannot be larger right.

(Refer Slide Time: 21:49)



So, you pick the most the big the case that includes everything and anything and put it there F x y this unfortunately the thing is this looks very innocuous, but it is you have to realize which is a variable you are integrating over now. What should we put here?

Student: (Refer Time: 22:22) dx.

Dx you are integrating on the horizontal direction keeping y fixed. So, this is very analogous to what we did in the.

Student: Discrete.

In the discrete case, you in the discrete case if you get the pmf of y you draw a horizontal line and you looked at all the points on that line and you added all those probabilities. So, here you have to do an integration by identical reasoning.

Now, the reason the nice things to have this up there right, I applied this all of this to this, I will remove this. So, once again I think the importance of first determining omega x and omega omega y cannot be under emphasized you need to write first thing about a random variable please think about what values it can take any derived or what otherwise there which is not straight away given to you first thing always should be what values it takes. So, when you apply this marginal pdf calculation to that example what do you get; for the darts case; the simple model. I guess I said all about and I do not want to put stuff about it novice player and all that it is, you can write it down of course. So, in future let me also say we will use novice darts as a simple short hand for all of this. So, I darts pdf for a novice player n o v i c novice. So, I cannot keep writing this all the time ok.

So, for this example what is f x and what is fy? Somebody do the calculation tell me what is f x? So, if I pick an x here between minus R and R what will be the integral. So, this is 1 by pi R square into dy. What are y 1 and y 2?

Student: Minus root over R square minus x square.

Minus.

Student: Root over R square.

R squared minus x square to.

Student: Plus root R square.

What is the value of this integral?

Student: 2 times root over R square.

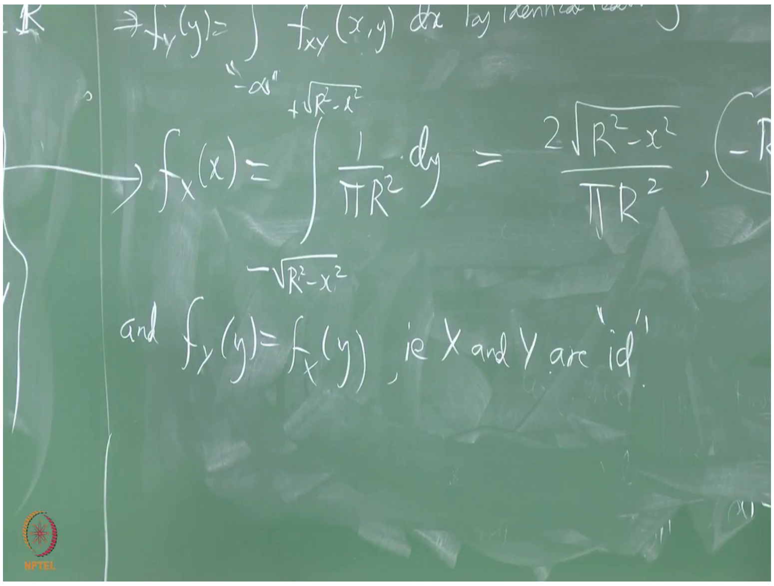
And automatically 0 outside. So, this is your omega x.

Student: (Refer Time: 25:17).

Note that this is not a uniform pdf anymore automatically. So, it has higher probabilities of taking values closer to the origin than at the extremities, that should be obvious from here. If you say any point is equally likely; obviously, the x coordinate of that point is more likely to be close to the origin than minus R and plus R right. So, this outer boundary of the circle is a very artificial construct in this cases somehow suddenly the probability of dart lining here is something and then outside is 0.

So, in a way its unfortunately is not a great physical model maybe if you take R extremely R's it might become better, but then the if you say R is a similar lies in this is not going to be that physically meaningful, but anyway some marriage of not so, physically meaningful ideas, but you have we are living with it because it is so simple.

(Refer Slide Time: 26:30)



By symmetry what can you say about the marginal pdf of y it is exactly the same, and f y of y, I can just say its f x of y without any without rewriting anything.

Note that you can use any argument do I do not have to say there is no I know there is nothing there is no rule to say that this must be the uppercase that this must be exactly the same lower case and nothing of the sort. So, what is the meaning of this statement? This statement says that x and y are identically distributed their ID with the same statistics the same range the same pdf.

Just summarize what we did. So, we started with omega x y, f xy, then we have shown how that model can be used to calculate first of all any probability by integration remember always, that the double integral I mean the joint pdf has to have unit volume otherwise it is not properly normalized it should right. So, that again I did not write it out in the beginning of this class, but remember joint pdf always has to be normalized to unit volume.

So, in this case one more thing to remember is probability is volume, one more way of remembering this joint pdf thing you have a volume here. Then that is what we did first and then we took up this example and be showed how in general you can go from joint to marginal. Very similar to analogous way to what we did for discrete, but you cannot go from in general from marginal to.

Student: Joint.

Joint. Supposing I gave you those two pdfs how would you write in your wildest day, you know you just this is mathematically impossible to do you cannot say that just given this pdf here of x and the similar pdf of identical pdf of y that this must be the joint pdf it is not possible. Somebody we can say no and they would be right.

So, when is it possible to that the marginal and joint are equivalent ways of describing situation that this requires independence. So, in general random variables are never independent they are always dependent in independence is a very very special case, remember.