Probability & Random Variables

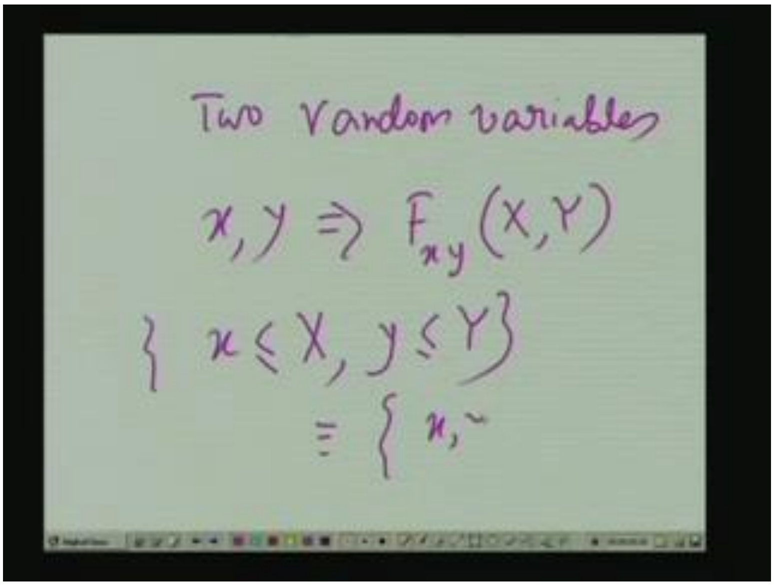
Prof. M. Chakraborty
Department of Electronics & Electrical Communication Engineering
Indian Institute of Technology, Kharagpur

Lecture - 12
Two Random Variables

So long we have discussed just one random variable and function of one random variable. We so far did not discuss what is called joint variation; that is earlier we considered random variable x; its probability density, probability distribution and a function of this random variable and all that. Similarly, there could be a separate random variable y; we discussed its properties and all that. But, there is an important topic actually; where, x and y – they are two random variables, which are mutually related; that is, the underlying process is such that x and y has got some kind of interconnection – I mean interdependence rather. Then, the question of not individual variation, but rather joint variation and joint statistics comes.

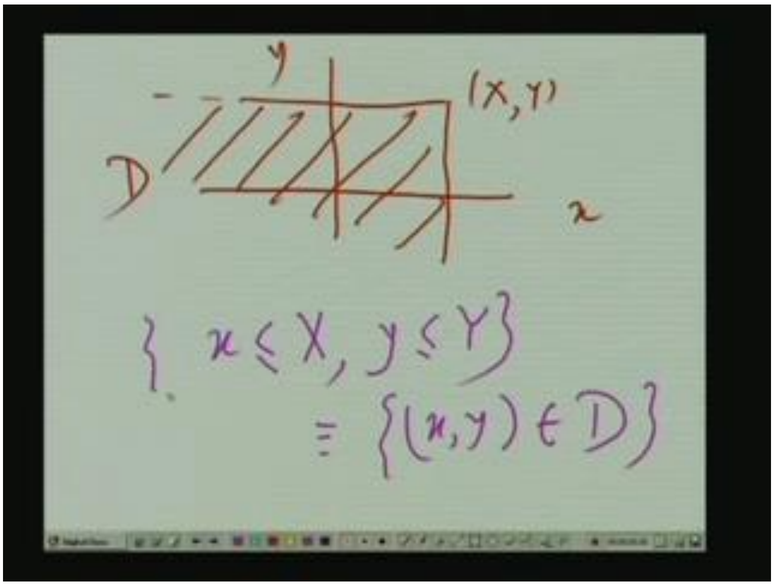
For instance, you can consider these two variables: one is say temperature of the environment and humidity in the environment. Now, you know temperature could be a variable T; humidity could be H. Now, you all can visualize that, the underlying physical process is such that definitely one depends on the other. Surely if the temperature is high, there will be some kind of one type of humidity; the temperature is less, another type of humidity and vice versa. Now, we do not go into this physical modeling, because that is very difficult. But, we treat both T and H as two separate random variables, which are jointly related – mutually related; obviously, because if T is high, maybe H will be less or H will be high. Again if H is less, T could be high, less, because that is determined by the mutual interdependence. So, we define…

(Refer Slide Time: 02:46)



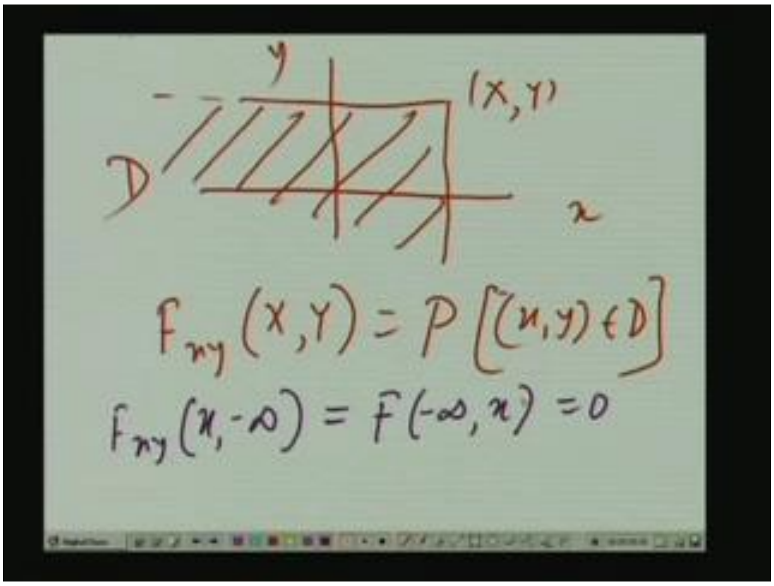
So, our topic today is… In fact, now, we are considering two random variables; but in general, it will be say n number of random variables. So, we will extend this concept to more than two number of random variables. So, we now, consider always a pair – under this topic, always a pair. x is a random variable y is a random variable. But, since they are related, we consider xy. In that case, we have to define the joint probability distribution of x and y. Now, what does it mean? Let us give a definition actually. We consider the event. This random variable x taking value less than equal to some given number capital X; y taking value less than equal to some given number capital Y; which actually means this pair is taking value for a particular region D in the xy plane. What is that region?

(Refer Slide Time: 04:18)



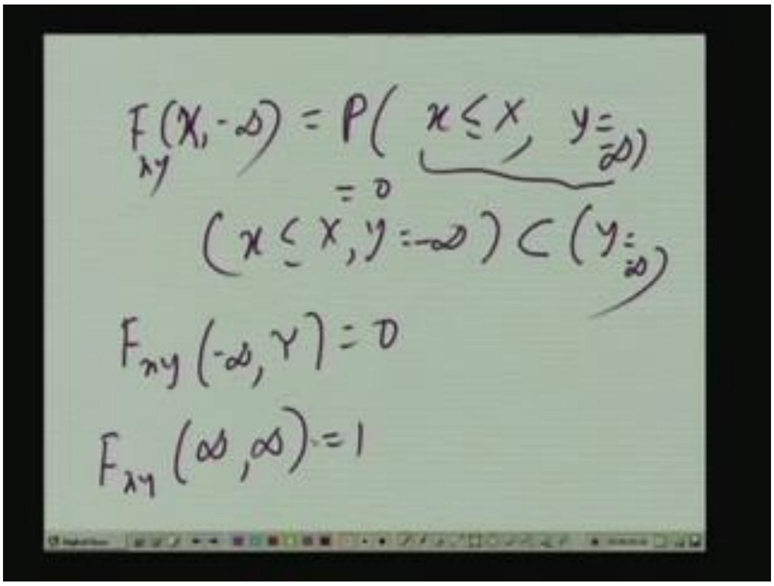
Suppose this is xy plane – x-y; and this point is capital X, capital Y; that if you draw this line, this area is D. Obviously, within this area, both these are jointly satisfied; both x less than equal to y; x capital X; y less than equal to capital Y. See if x at y, they are not considered together; then we consider only x individually. In that case, probability distribution of x was what? We have simply taken this vertical line; and the entire plane to the left corresponds to points; where, x is less than equal to capital X – independent of whether y is less equal to capital Y or higher. But, since we are considering the joint variation, that is, under tighter condition; that not only x is less than equal to capital X or y less than equal to capital Y; but they are satisfied simultaneously. This is the simultaneous variation or joint variation. So, probability distribution will be actually the probability of this event; that is…

(Refer Slide Time: 05:52)



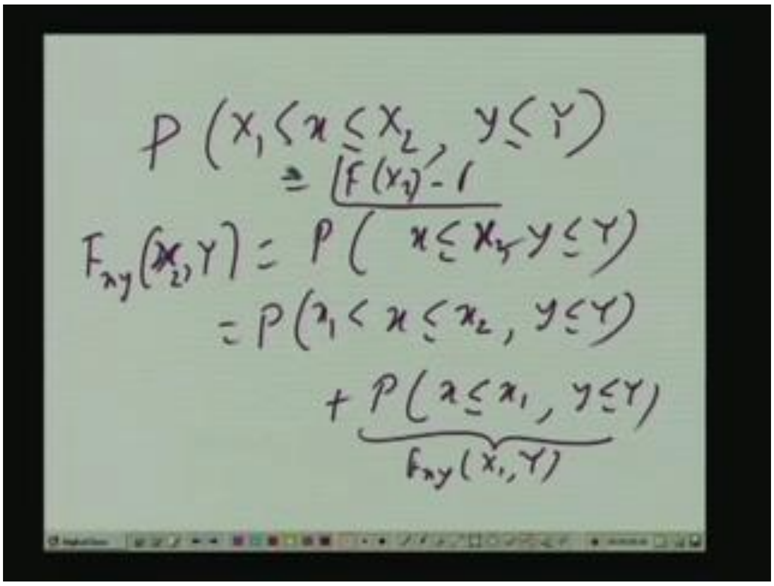
It is nothing but the probability of the event that I have shown; that is, the pair X comma Y falling within this zone. There is a probability of this event. Now, some of the definition – some properties follow immediately from this; what are the properties? First, this is equal to 0. Why? It is very simple actually.

(Refer Slide Time: 06:54)



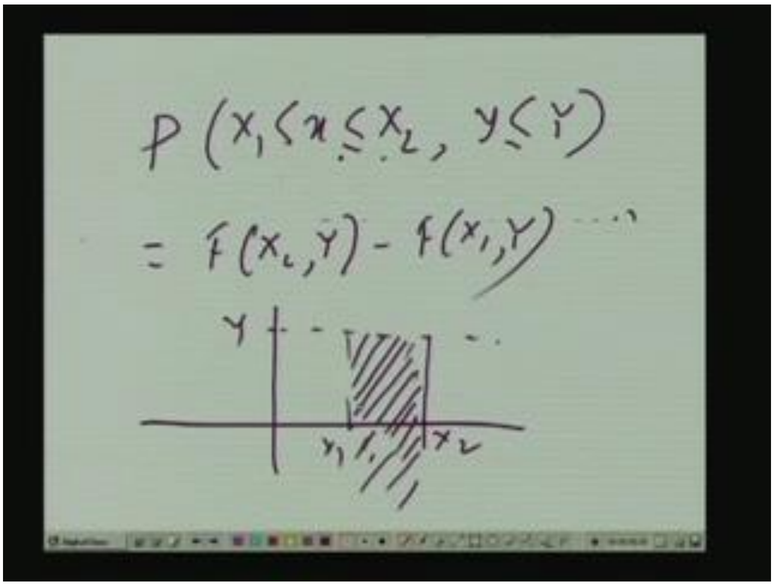
What is F x minus infinity? It is actually the probability of this event. See you can… I am writing like this; actually, I should write in a different way; I mean I should say that, it is x y. The probability that x less than equal to capital X and Y. There is nothing equal less than equal to minus infinity; so… But, this is the event; this event is basically a subset of this event. But, probability of this event as we know from our earlier definition is 0; which means this probability is 0. So, this is equal to 0. By the same way, we can have – equal to 0. For that, this should be equal to 1 obviously, because this covers that entire plane. All points on the xy plane are covered by this; that is, the event actually consists of all pairs for all points of the xy plane. And obviously, the corresponding probability distribution will be 1, because it considers that total set.

(Refer Slide Time: 08:53)



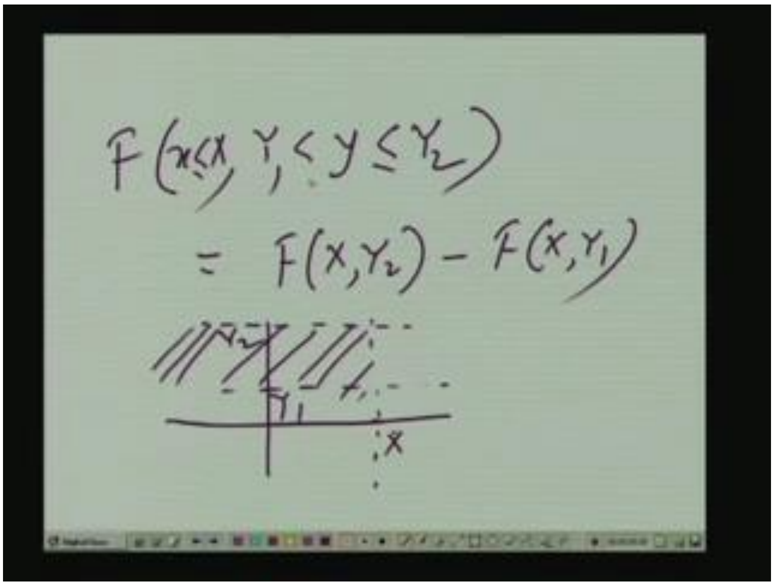
Then, consider this – what is the probability of x taking value less than equal to some capital X 2 greater than X 1 and y taking value less than equal to sum Y. This means we can always write it like this. See the event consists of this thing x should be lying within this zone – less than equal to capital X 2, but greater than X 1 and y less than equal to capital Y. Now, this is the event. Now, this set can be written as a combination of two non-overlapping sets – non-intersecting sets. One is this event actually. Let me make one small correction. We know that… Instead of capital X, we put x 2 – small x 2 you can say. Now, this set or this event can be written as union of two non-intersecting sets – two non-overlapping events: one is this; another is… So, y ((Refer Slide Time: 10:42)) less than equal to y; but this zone is separated into two zones. One is for x lying, that is, less than equal to x 2, but greater than x 1. In another case, x lies – x is less than equal to x 1. Now, this is what I wanted. And what is this? This is simply F x y x 1, Y; which means this probability is nothing but – this is nothing but F x 2 minus… In fact, I can erase this.

(Refer Slide Time: 11:41)



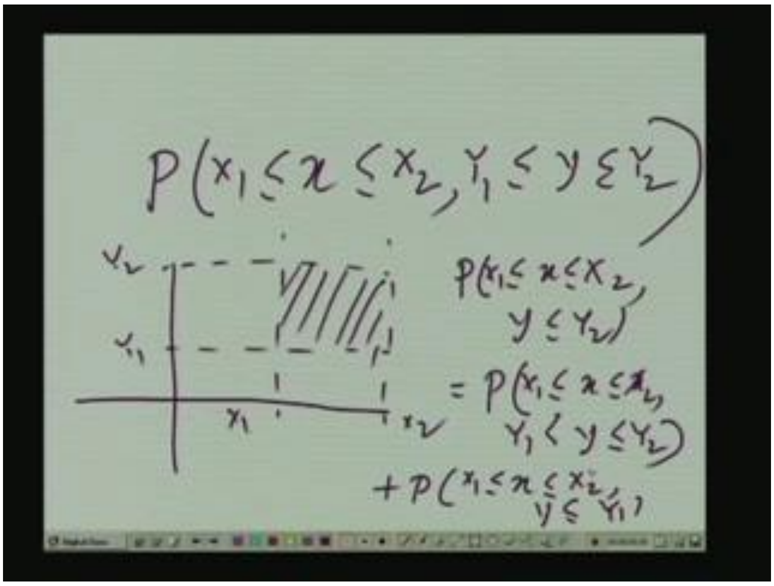
It is a 2D plane. It means what? Suppose X 2 is here; X 1 here and this is Y; that meets; this is the area of interest; where, X is less than equal to X 2 and goes up to X 1, but does not touch X 1; this is given that X 1. So, this was the event. This probability is nothing the pair X comma Y; taking values from within this; I mean it is falling within this zone. What is the total probability of that? That answer to be this. On the other hand, by the same logic, we can say… Just give me a minute please.

(Refer Slide Time: 13:12)



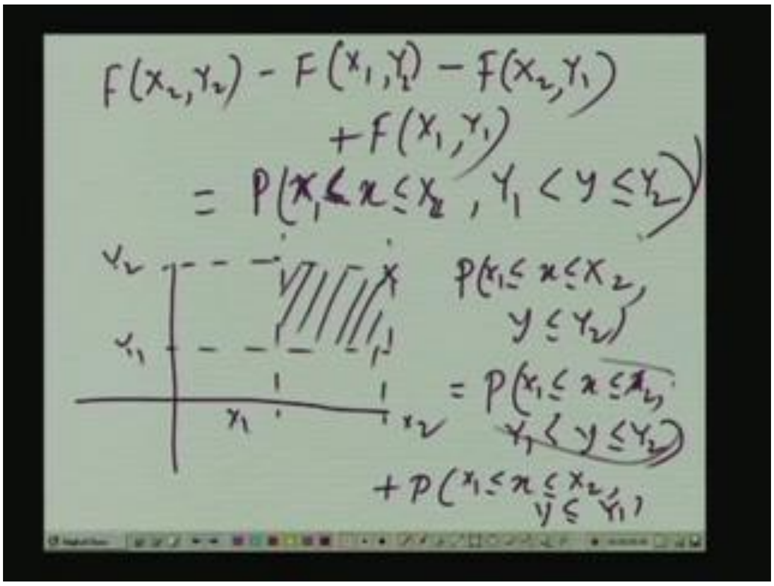
By the same logic we can say that, take some… This probability y lie within two zones within two limits Y 1 and x less than equal to some capital X. I do not think I have to repeat anything; you can just apply the same logic and get this. This will be what? Y 2; this is the larger one minus Y 1. In the 2-D plane, xy plane – it means what? That suppose this is X and there are two values; this value is Y 2; this value is Y 1. Then, this is the region. What is the probability of the pair XY? It is falling within this region. That is given by this. This now we can generalize further.

(Refer Slide Time: 14:40)



How about both general thing; that is, probability of x lying within two limits and y also lying two limits; what is this? You can easily see this if I draw the 2-D plane; this means that, if this is Y 2; this is Y 1; if this is X 1; if this is X 2; so this is the region. This probability means what is the probability of xy pair taking values falling within this region; that xy pair taking value within this region. What is it? And you can see one thing; if I do not limit myself to this box only, but go down fully; then what I get is nothing but what I already obtained; that is, if I am interested in this that, attributes as it is; but y is only less than equal to y 2. So, I go down fully. This event… I have already known these probabilities; it was calculated in the previous page. But, this event is nothing but… I mean this set is nothing but the union of two non-intersecting sets. One is what? One is this box; another is the box below; that is… But, in another case, y 2 greater than equal to y 1 and P – first component events as it is; and y is less than equal to y 1; obviously. Now, this is my quantity of interest. I have to find out what it is. I know what it is; I know what it is. It was derived earlier. This is what F of x 2 comma y 2 minus F of x 1 comma y 2. And this is what F of x 2 comma y 1 minus F of x 1 comma y 1.

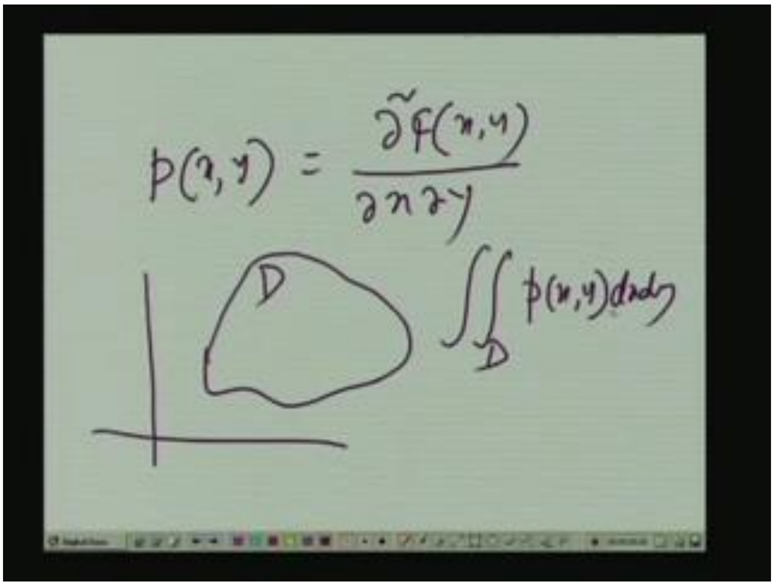
(Refer Slide Time: 17:30)



So, in fact, we can write down… First component is what? F of X 2, Y 2; then this point – minus F of – the other component is… I mean here F of X 2, Y 2 minus F of X 1, Y 2. What is X 1, Y 2? This point. And take this quantity to the left. So, sign gets reversed. Originally, this was F of X 2 comma Y 1. Now, that comes with a minus sign. And the other quantity, which was present with negative sign; what was that quantity? F of X 1 comma Y 1; that comes with the positive sign. So, this is equal to what we wanted – this quantity less than… Remember this; that I need this expression later. So, ((Refer Slide Time: 18:58)) probability distribution.

Next thing that comes is probability density as before. You remember what I did last time. In the case of single variable, we had F of x as the probability distribution function of a single random variable; then the derivative of that was what? That was the probability density; like that is, if F of x denotes the probability distribution function; then del F del x was P of x, which is the probability density. What is the physical meaning of it? That if at any point x and between a; I mean if I consider a point x, then what is the total probability of x falling between x to x plus dx? That was px into dx. All that was fine for the one dimensional, one variable case; we will just generalize that concept or that notion to the two variable case.

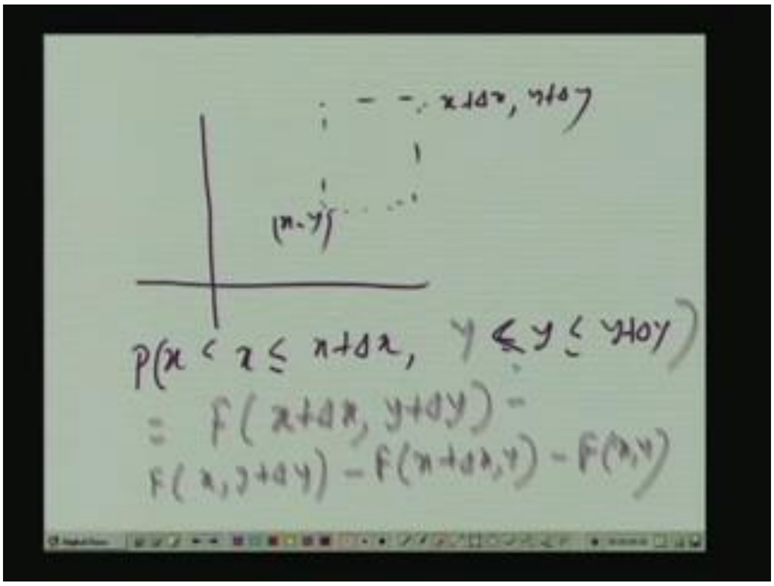
(Refer Slide Time: 19:54)



So, here we will define the joint probability density function. Since we have got two variables, these are all partial derivatives. Remember this is partial derivative and sometimes I do not write in the general form, where F; then we have subscript xy; then the values of x, y; where, I just write F as a function of x and y; I just put as unknowns – no known quantity given. This is general form of the function. I simply put x variable and y variable here. This is the probability density. Why this is probability density?

Reason is if I consider any region say D in the xy plane, then what is the probability of xy falling within this region? We will show that, that will be nothing but this integral about this zone. I put double integral, because dx dy over this zone. And what is p x comma y dx dy? Let us see what it is?

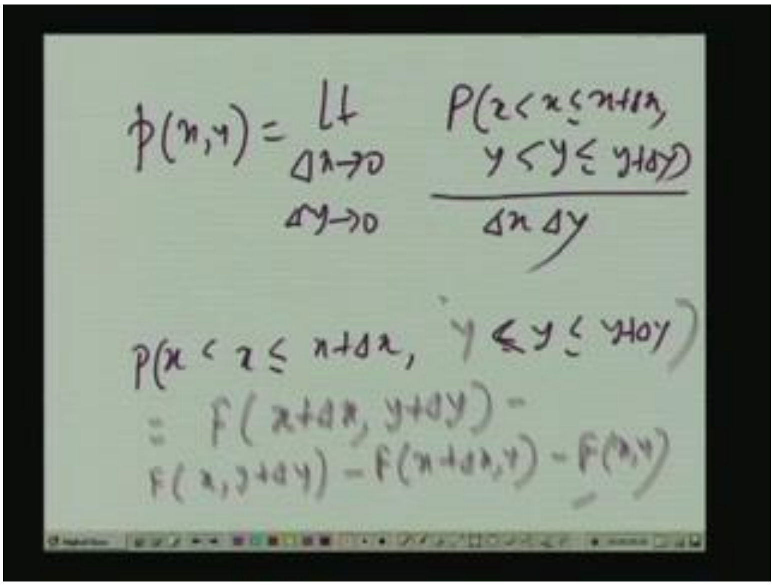
(Refer Slide Time: 21:33)



Consider this point x plus delta x, y plus delta y. And this is x, y. Of course, in the figure, it appears to be very big area; but this is actually infinitesimally small area, because this is delta x, delta y. Just for our convenience, I have drawn reasonably big sized area. But, actually, the area is very small – infinitesimally is small, because delta x and delta y are small; and they will be in fact tending to 0. Anyway, what is the probability of x comma y falling within this zone? We have already seen that probability; that is, what is the probability of x lying within x plus dx and y? That we have already seen on the previous example. That was equal to what? Minus… This point. So… Then, this point… This is the probability.

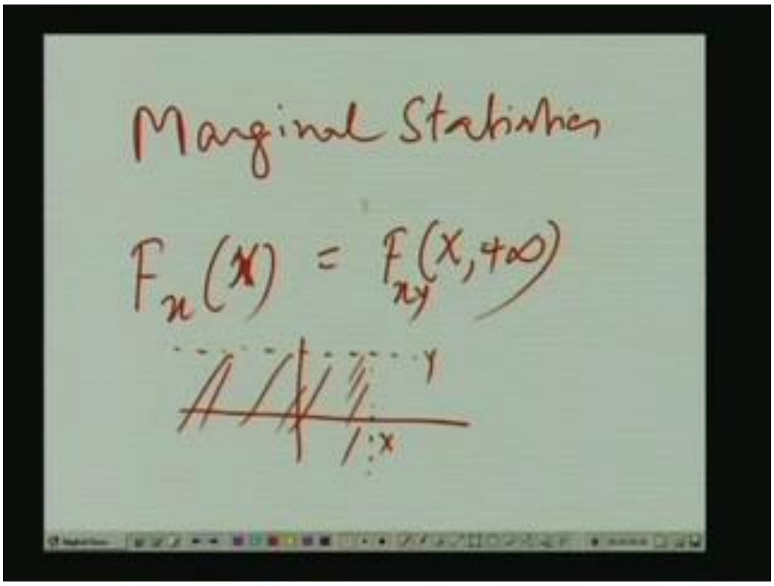
And if I divide this, total probability of x, y falling within this zone by delta x times delta y; and let delta x approach 0, delta y approach 0. Then, what we get is nothing but the probability density. That is the meaning of this partial derivative actually; that is, where I say the del square F, del x del y – it means find out in a rectangular area of sides delta x delta y located at point x, y. What is the probability? What is this differential? That divided by delta x delta y. That is why the probability of xy falling within this zone; that is, this divided by delta x delta y. And let delta x tend to 0, delta y tend to 0. That will be equal to probability density; that is I can erase this, that is…

(Refer Slide Time: 25:35)



That means if I now take a very small infinitesimally small rectangular area, where the rectangular area delta x delta y – meaning the sides are delta x and delta y. Then, what is the probability of x, y falling in that area? that is take p x comma y times delta x delta y. I repeat again what is the probability of xy pair falling within the infinitesimally small rectangular area delta x, delta y located at the point x, y. What is the probability? Obviously, that will be equal to p of x comma y times delta x delta y. So, what is the probability of xy pair falling within the total region D? It will be nothing but a summation of this infinitesimally small probabilities over the entire region; that is, I can break up the entire region as a union such infinitesimally small rectangular areas. So, total probability will be nothing but a summation of these individual probabilities; and it is a continuous summation. So, we get that integral relation, that is, P of xy, dx dy integral; one is entire region – D; that will give you this thing. There is a probability of x y falling within this specify region D.

(Refer Slide Time: 27:43)

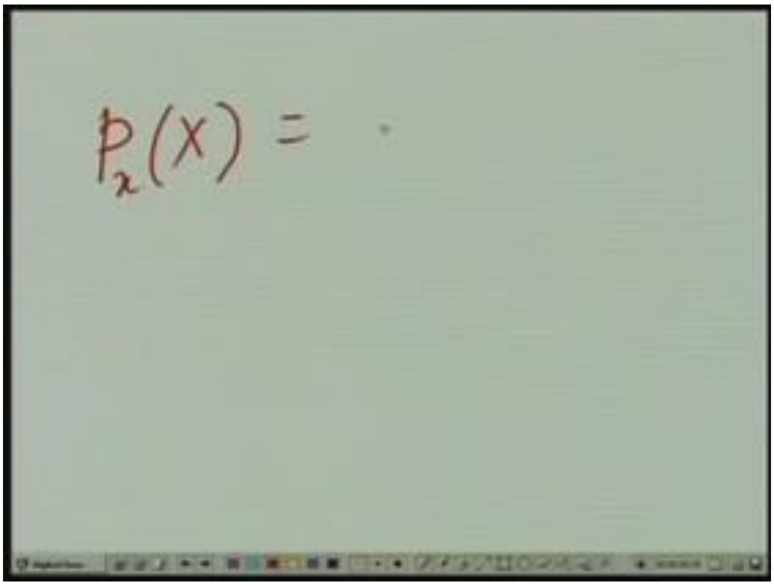


Next, we move to another definition, which is called marginal statistics. Suppose it is given that, two variables x and y; they are joint jointly random variables. So, we are giving the joint statistics like joint probability density, a joint probability distribution and all that. In this scenario, whether I consider just one random variable separately and takes its individual statistics – like its own probability density function probability distribution function – things like that; then that will be called marginal statistics of the random variable x. You can easily see that, if I say this; this means as usual the probability distribution of x, that is, small x taking values less than equal to capital X; where, capital X is very specified number. This is the probability distribution. This will be called the marginal distribution of x; where, x, y are jointly random.

Now, this is nothing but… Why? It will be plus infinity, not minus; minus means we will make it 0. We have seen already. Obviously, it will be entering 2-D plane. As such suppose we take at point x; we come back to this previous figure. Suppose this was capital Y. And this is x, y of course. And this region – we are considering this region earlier. That was the total probability of xy falling within this region was F xy capital X comma capital Y. But, if this capital Y moves to infinity; then this will be nothing but the probability of just x taking value to the left of this line; that is, x taking values less than equal to capital X irrespective of whatever y; or, all possible Y's were considered. So… In fact, that region will give you all possibilities by which small x can values less than equal to capital X. So; obviously, when it is infinity, that is, when this goes up to infinity;

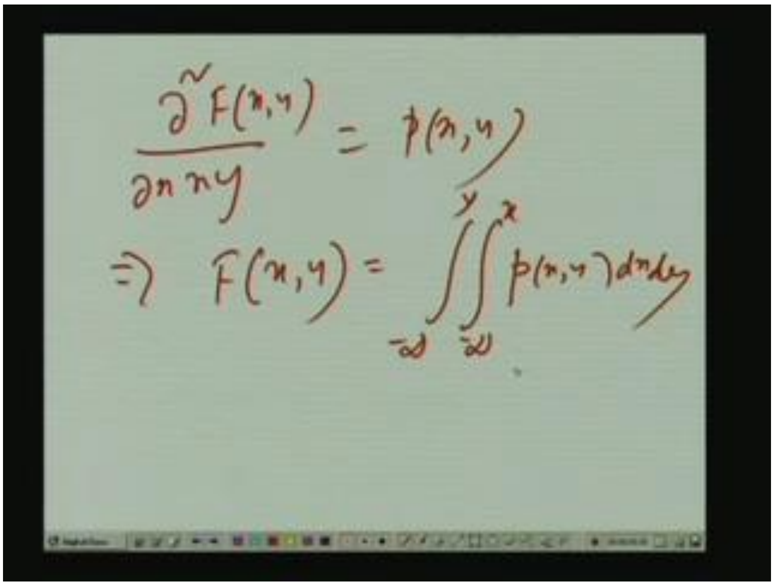
that is, I have the total plane to the left of capital X line – x equal to capital X line. That up-plane did indicates what? I mean that up-plane altogether considers all cases, where small x takes values less than equal to capital X. Now, there is no condition on y, because y equal to infinity means all y's are covered – all y's are possible. So, I am left with X only. So, I get this. So, by the same token, I will not repeat it. This will be nothing but…

(Refer Slide Time: 31:35)



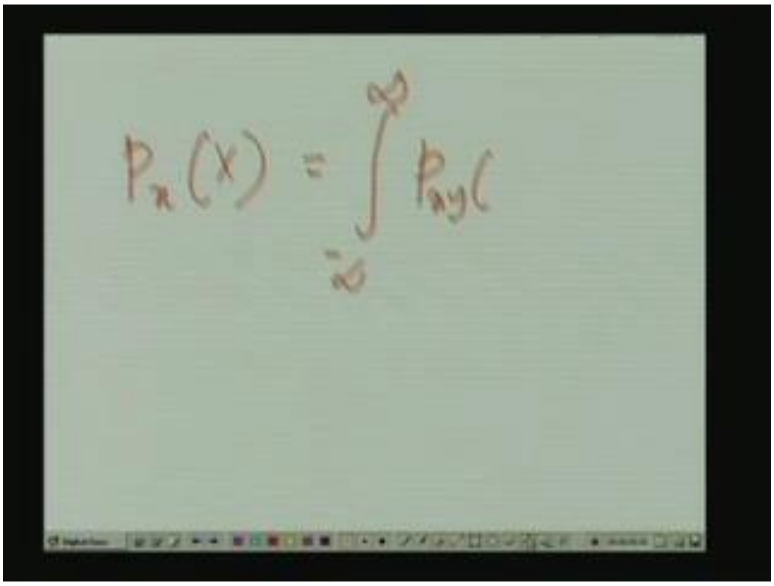
And finally, this is very important P… I write formally here P X – capital X – that will be nothing but… By the way I forgot to mention one thing. So, it is very clear; we should always at least write it out.

(Refer Slide Time: 32:04)



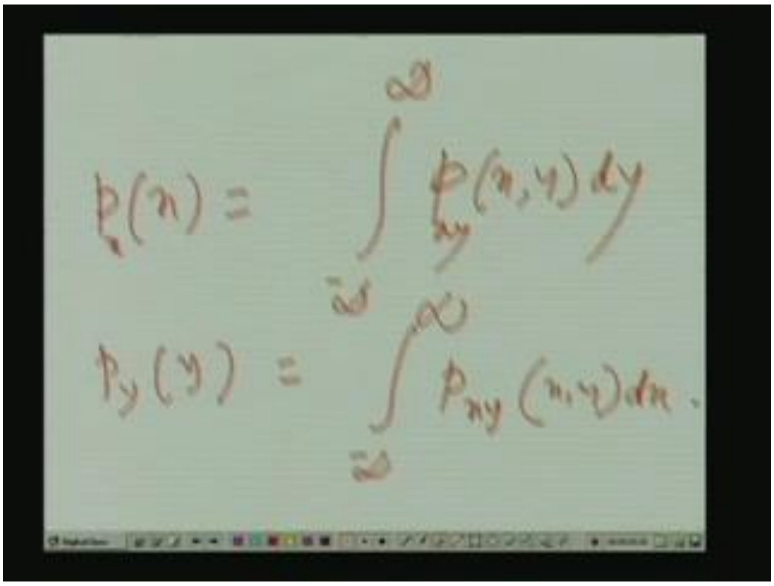
We have seen that del square F by del x del y – this was p x, y; which means where x and y are general variables now. So, that is why I am not putting them in the subscript and putting some specified value here. This will be nothing but double integral of… You can put limits x and minus infinity. This comes from ordinary calculus. I cannot go back to that. P x, y dx dy – it is derivative. So, this is an integral relation reverse. So, I come to what I was doing.

(Refer Slide Time: 32:58)



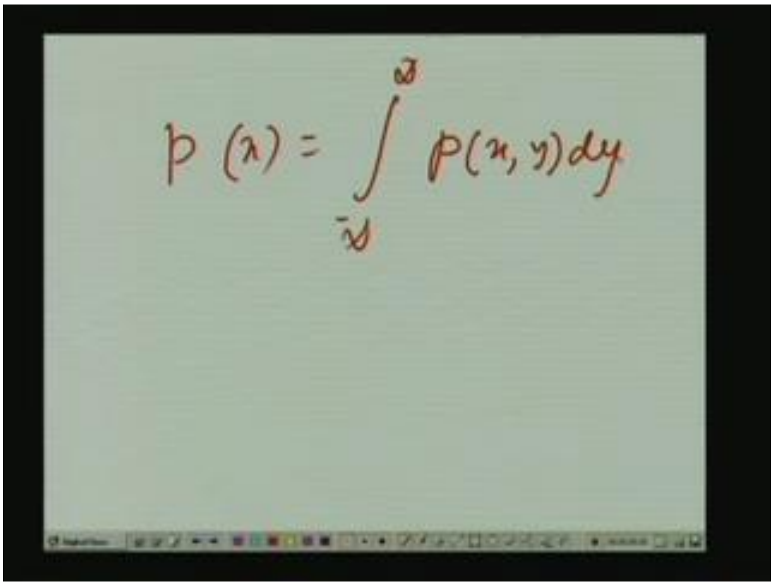
This will be P x y… Or let me write it in a more simplified way, because this is becoming complicated.

(Refer Slide Time: 33:31)



Just P x and put it in the general variable x. This is nothing but… So, integrate it out with respect to y, because it is a function of only x. And that will be p x. When I write p here and this p – these two are different functions. In fact, that is why I should have put p subscript x and this is subscript x y. These two are two different functions. This is the probability density function of x; this is a joint probability density function of x and y though I use the same symbol P. These are actually two different functions. And similarly…

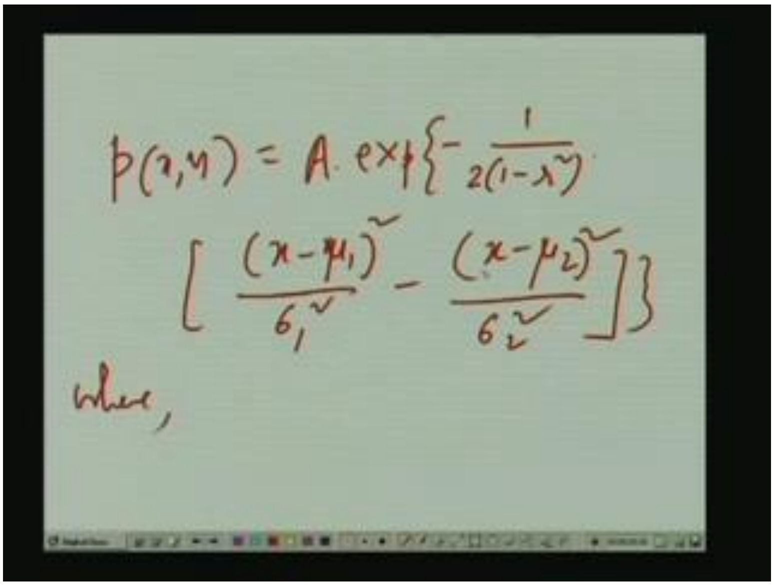
(Refer Slide Time: 34:50)



It is not very difficult to show that; probably you have seen that… We have just a while back seen… You put some… Actually to be very correct mathematically, I should not use the same symbol x, y here, because they are coming in the limits. Let me put alfa, beta. ((Refer Time: 35:51)) later. So; obviously, if I differentiate it with respect to x, this goes; I am left with only this. Convert this F – F of x y. Now, if I put y equal to infinity; then in this integral, this will become infinity of course. And here… But, we have seen what is F x comma infinity. That is nothing but F x. And what is del F x del x? Which is in fact dF dx now, because it is a function of just one variable. So, del F del x actually means the dF dx; that is nothing but the probability density of x. That gives the proof, that is… That gives the proof, that is, p x is nothing but… We can put now x y here also; does not matter. There is the other one also for that very easily.

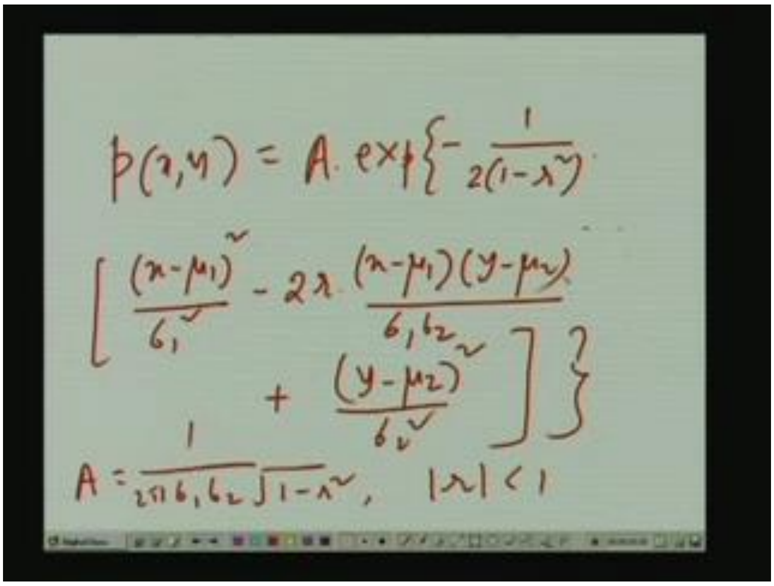
Now, take an example. Earlier you have seen that x is a random variable and x was taking values; like x was following some distribution density like normal; that is, Gaussian or uniform. Or, maybe we took Binominal or Rayleigh or things like that; Poisson of course. That time x was only individual random variable – single random variable. Now, we consider similar example for joint random variables like what is meant by two variables, which are jointly Gaussian.

(Refer Slide Time: 39:03)



Now, two variables are jointly Gaussian if that probability density always is this function. Some A times exponential… Just a minute – times we can write it; where, you see this has to qualify for a probability density function; it should always be non-negative and its integral over the entire plane xy plane for this will be y. No, this I made a mistake. This is a very special case. So, this is a mistake; let me correct it first.

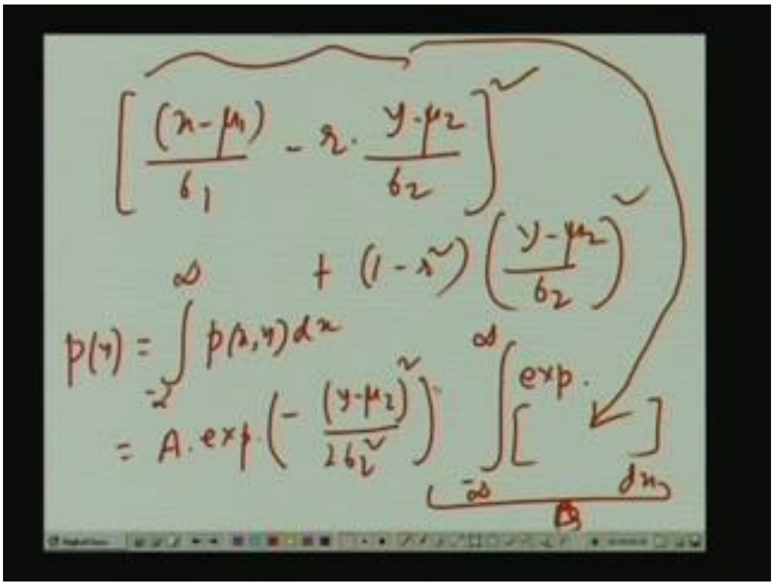
(Refer Slide Time: 40:58)



So, this is a very big expression actually. So, this is a function of two variables: x and y. So, in order to qualify for… In order to be qualified as a probability density function, this should be non-negative everywhere and its total integral what the entire xy plane should be equal to 1. That will be satisfied provided A is – A should be equal to… r square; and r should be such a quantity, whose value is less than 1. So, A should be such – A should be this; and mod value of r should be less than 1. In that case, if you integrate it, what this entire xy plane? This should be equal to 1 and this should always be non-negative. So, this is just a condition. This will go jointly Gaussian. In fact, what is mu 1, what is mu 2 – you can now see. We can find out the marginal statistics of x and y; we will see that x and y – they are individually also Gaussian random variables. And mu 1 is the mean of x and sigma 1 is the… Sigma 1 square is the variance of x. Similarly, mu 2 is the mean of y; sigma 2 square is the variance of y. That is the marginal statistics we can obtain from this.

But, this is the formula for… This is the expression for joint Gaussian density; where, A is this and r is studied by this condition. Actually the meaning of r will later. It is the cross correlation or correlation you say between x and y. Now, how do you show this? How do you show that, under such case, x and y – both are also individually Gaussian random variables with means mu 1 and mu 2 respectively; variance is sigma 1 square and sigma 2 square respectively. For that I just consider this quantity. This quantity can be written as very easily in this manner.

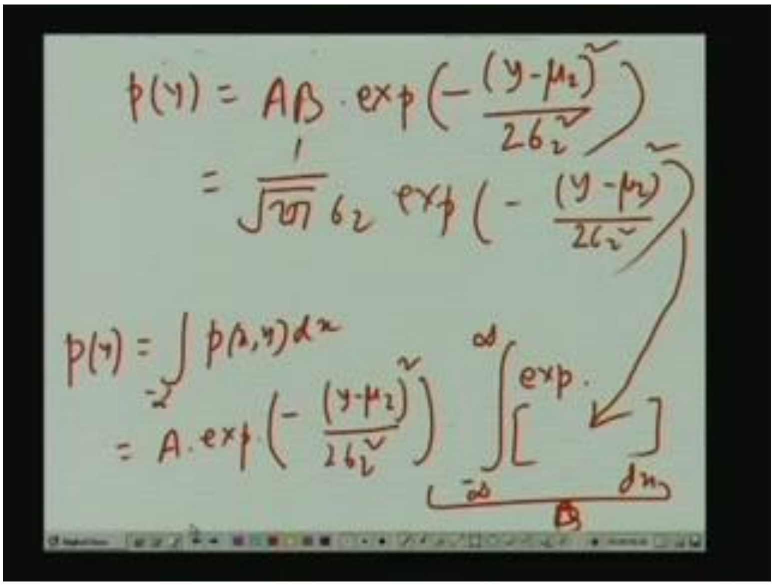
(Refer Slide Time: 43:58)



We can write that quantity simply as minus r times whole square. So, I am bringing an additional term. So, this has to go out – 1 minus r square to y minus same thing whole square. So, suppose… Then, find out P y. Then, it will be nothing but dx. If we integrate see this part; it has no x; it will come out. So, we will be getting A times exponential… This gets… Some cancellation takes place with this at the other terms, which lies outside, that is, 1 by 2 – within bracket, 1 minus r square. 1 minus r square, 1 minus r square cancels; you will get this thing. There is a minus sign also. So, you get… And within this, there is an integral – exponential; I have no space. So, I will just say that this quantity – this comes over here – dx. This integral – even though there is a y present, because it is very interesting that, this integral does not depend on y; simply because we are integrating with respect to x. Suppose this entire quantity x minus mu 1 by sigma 1 minus r minus r times y minus mu 2 by sigma 2 – you call it x prime. I am just telling given the hint.

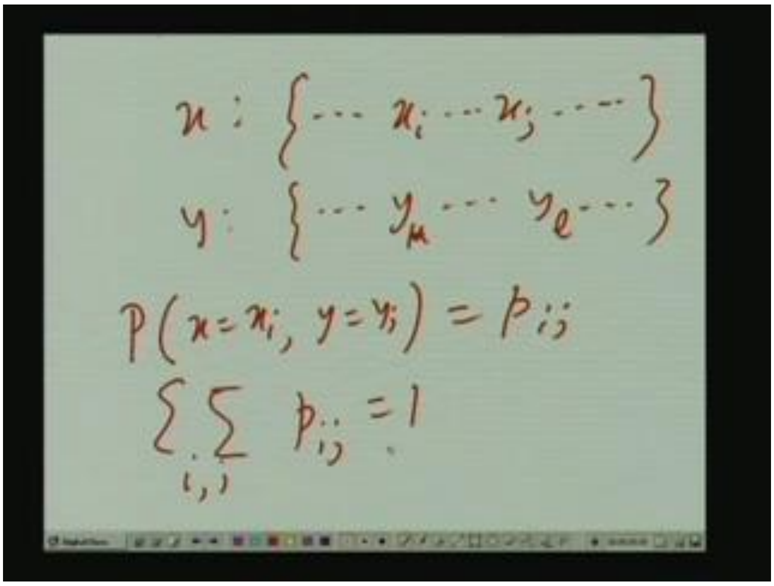
In that case, dx by sigma 1 will be dx prime. So, sigma 1 will go out. So, dx will be sigma 1 times dx prime. So, sigma 1 will out come out. And what will be the limits? When x is minus infinity, x prime also minus infinity; so this limit will remain. When x is plus infinity, x prime also is plus infinity; this limit will remain. So, you will have some sigma 1 times exponential. Some quantity called x prime whole square dx prime. And this integral exponential x prime whole square dx prime. This is a constant. If you take that sigma 1 also – sigma 1 or sigma 1 square whatever along with this, you can call this integral to be then just a B, because it is independent of y. So, you get AB times this quantity. But, this has to be a probability. And we know this already takes the Gaussian shape – y minus mu 2 square whole square by twice sigma 2 square.

(Refer Slide Time: 48:15)



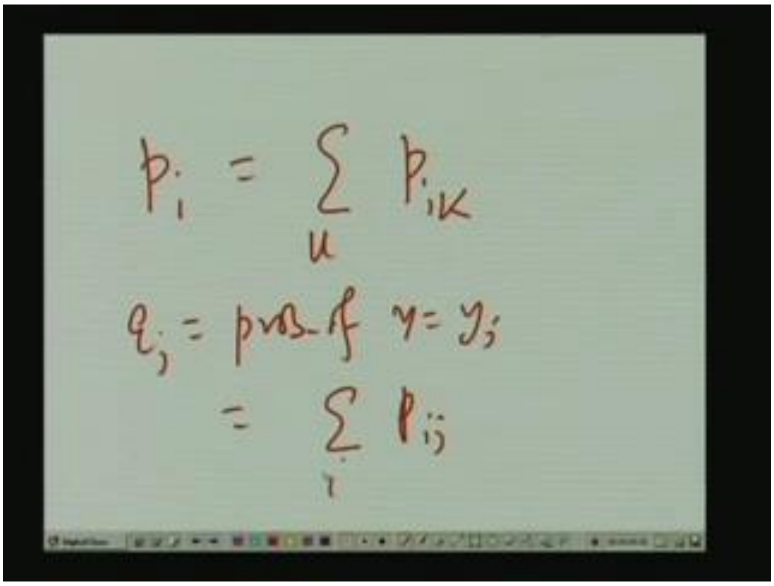
So, that means… Let me repeat again P y can become AB… And obviously, you know this takes the Gaussian form because of this kernel. And as a result, AB has to be equal to 1. So, that means… The integral has to be equal to 1 from the Gaussian kernel; obviously, AB should be equal to 1 by root 2 pi sigma 2 and this. So, it is clear that, mu 2 is a mean. It is clear that, P y, that is, the individual – the marginal statistics – marginal probability density of y is Gaussian with mean. Then, y has a mean mu 2 variance sigma 2 square. By the same token, by the same analysis, you can also prove that, marginal density – P x also is Gaussian. And in that case, x has a mean – mu 1 variance sigma whole square. This was for continuous variables. So, we can just repeat these things for discrete variables also.

(Refer Slide Time: 50:04)



Suppose x and y – they are two discrete random variables. x takes values from this set; I mean… While I mean… It may not be finite. So, let me not put it this way. x i, x j and like that. It could be a finite set; it could be infinite set. Similarly, y k, y l dot dot dot like that. In that case, the joint probability that, x takes the value x i; y takes the value y j; it is denoted by P i j. Obviously, i and j have put a double summation. That should be equal to 1.

(Refer Slide Time: 51:18)



And p i; that is obviously, p i k – summed over all possible k's – w p i. If q j is the corresponding probability of y equal to y j, this will be again equal to… p i is the probability of x equal to x i; that is, this summation. p i k is the joint probability. And here it is… q j is nothing but again p i j, but summed over i. This is pretty obvious. So, today's class we conclude here. We stop here. We continue to expand on this – on this theme getting it to what is called probability mass. There is a very important concept called statistical independence, then joint characteristic functions and things like that. That will be for the next time.

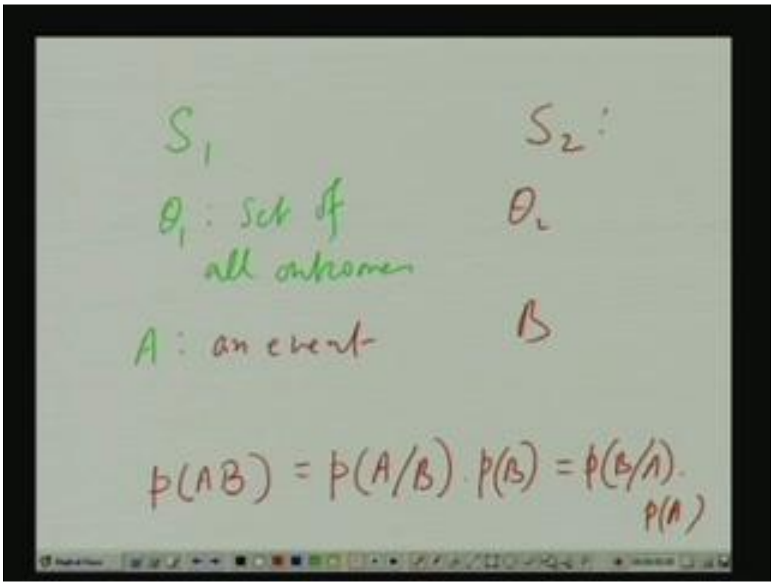
Thank you very much.

Lecture - 13

Function of Two Random Variables

So, last time, we were discussing two random variables, that is, the joint statistics, joint probability density, joint probability distribution and things like that. Then, a question that – a very important question, a very important point – a property is rather is statistical independence of two random variables; that is, when can we say that the two random variables x and y – they are statistically independent. Now, to answer that, let us go back to some of our earlier discussions.

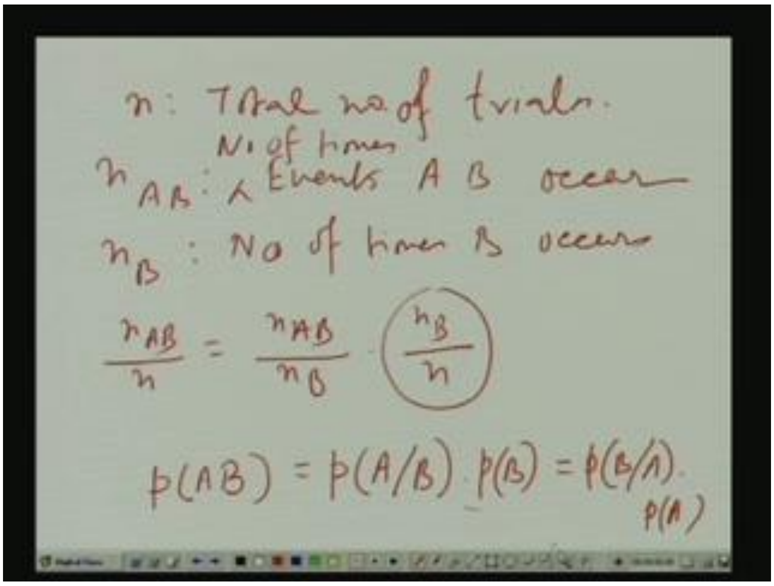
(Refer Slide Time: 53:24)



Suppose we have got two probability spaces. One is S 1. That has the outcomes say theta 1, that is, set of all outcomes. And A be a particular outcome. Similarly, suppose S 2; given theta 2 and B; theta 2 is the set of all outcomes here. This is another probability space and B is a particular outcome. In that case, we have seen one result that, the joint probability of this event; or, maybe instead of calling A to be an outcome, it will be better or more generally we say it is an event; that is, a set of outcomes. And. B also is an event. So, let me erase it. So, what is AB? This is joint event; that is, we have got both these experiments going on – theta 1 is the set of all outcomes; theta 2 is a set of all outcomes. A is a particular event; B – a particular event; AB means joint event – A occurring and B occurring simultaneously. What was this probability of this joint event?

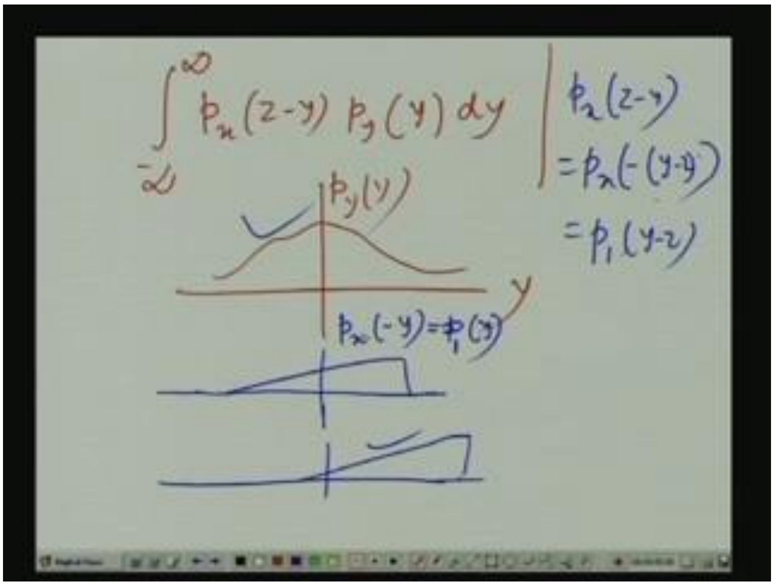
That you had seen was given by either like this p A by B into p B was equal to also p B by A into p A. In case you have forgotten, we can also see how it comes.

(Refer Slide Time: 55:34)



That is, suppose n is the total number of trials; out of these total trials, actually we are observing both… I mean outcomes from or events from S 1 and S 2 simultaneously – parallely. Suppose on n A B occasions, events A B occur. And on n A… What is an n B occasion? Number of times… This also number of times; number of times the events A B occur, that is, A and B occurring simultaneously; that is, n A B. And n B is the number of times B occurs.

(Refer Slide Time: 56:36)



p x z minus y – you can write it as p x minus of y minus z. So, first plot this function p x minus y. You can call it… Give it even the name p say 1 y; you call it as p 1 y. You get some function. Maybe I should not draw the same shape for both the functions. Maybe suppose it is like this. But, you want to evaluate this function at z. So, it should be p 1 y. This is… So, it will be p 1 y minus z. So, you have to shift it to the right by z if z is positive; or, shift it to the left by z, if z is negative. So, you shift it by z and then you multiply this function, then this function. See the overlapping area. Multiply into an integral – integrate. That will give rise to this integral. This is a typical way of very standard way rather of carrying out convolutional integral.

So, in the next class, we will carry out some examples of this; that is, two random variables – x and y. And then having some particular probability densities for each, you form a random variable z as a summation of x and y. What sort of shape? What sort of probability density function you will have for z? That we will evaluate by carrying out this convolution integral. So, that is all for today. We start from here in the next class.

Thank you very much.