**Introduction To Adaptive Signal Processing**

**Prof. Mrityunjoy Chakraborty**

**Department of Electronics and Electrical Communication Engineering**

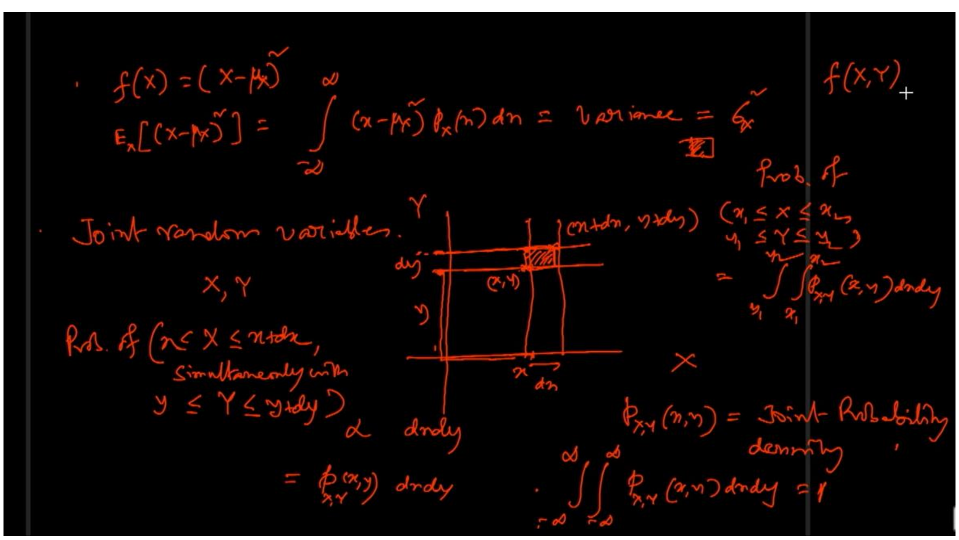
**Indian Institute of Technology, Kharagpur**

**Lecture No # 02**

**Probability and Random Variables**

So, we start from where we stopped last time to recall we are considering joint random variables. Now suppose it is basically nothing but what I am doing is nothing but generalization of the one variable case to two variable cases. So, I have certainty that is one also probability of that is one ok involving two variables now. Now like earlier is that like the case of single random variable now for the joint case also we can assume that we are given a function $f(X,Y)$ that is every time I observe x and y say temperature humidity this function I in this function I plug in those values I get something next time again I conduct the experiment I plug in the value of temperature humidity that is x and y I get another value and so on and so forth which means the function also is random. So, if so, what is this average value. So, what I will do I will multiply by the chance factor that is if see $f(X,Y)$ I will assume when it is within this box its value does not change because the widths are infinitely small.

**(Refer Slide Time: 01:25)**



So, its value will be same as its value at the corner that is $f(x,y)$ alright. So, that value, but it is not guaranteed that I will get that value always. So, I have to multiply by its chance, chance is chance of $X$ and $Y$ following here. So, I multiply by the chance again the same thing I do on another box, on another box, on another box, I go on adding ok.

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\, P_{X,Y}(x,y)\, dx\, dy$$

So, basically multiplying this by a weight. So, I get a weighted sum.

$$\frac{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y) P_{X,Y}(x,y)\, dx\, dy}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} P_{X,Y}(x,y) dx\, dy}$$

So, I will have the weighted average weighted average will be if I go to the next page will be f of this value you are multiplying by the chance, chance is this much as we have seen and this you do for very small boxes all over the plane and add the we add this ok. So, you get a weighted sum and it is a summation now over the entire range and then divide by the sum of the weights there is a weighted average which are called expectation. These are the weights to sum them, but this is equal to 1 because this is certainty probability of certainty because I have my range is from $-\infty\ to\ \infty$ capital $X, Y$ falling within this range is always certain.

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} P_{X,Y}(x,y) dx\, dy = 1$$

So, probability of that which is given by this is double integral that is the probability of certainty which is a global phenomenon a constant universally constant there is 1. Therefor E is this one Take the function instead of capital XY substitute by small x small y because they are the integral integration variables multiplied by the joint density. So, is a function of this small x small y and dx dy. one thing expectation, this called expectation, expected value. Expectation is a linear operation I will explain what I mean.

$$E_{X,Y}[f(X,Y)] = \frac{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)P_{X,Y}(x,y)\ dx\ dy}{\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} P_{X,Y}(x,y)dx\ dy}$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)P_{X,Y}(x,y)dx\ dy$$

Suppose consider single variable case and then we can easily extend to this case also I will do that. Suppose you are giving two functions $f_1(X), f_2(X)$. Now if you create a function $f(X)$ as a linear combination of the two, where $c_1$ and $c_2$ are some scalar numbers. $c_1$ numbers 2, 3, 4, 5.3, 7.9 anything ok.

$$f(X) = c_1 f_1(X) + c_2 f_2(X)$$

So, again this is a random function because every time I observe $X$ is a single variable case I going back to and can be easily generalized to this also. So, every time I observe $X$ this takes a value this takes a value. So, linear combination takes a value next time I conduct another experiment this takes a different value, this takes a different value, these two functions again the combination takes another value. So, this function also random. So, therefore,

$$E_X[f(X)] = \int_{-\infty}^{\infty} f(x)\ P_X(x)\ dx$$

I have to take this there is this function $f(x)$ as a chance factor all right.

$$E_X[f(X)] = \int_{-\infty}^{\infty} [c_1\ f_1\ (x) + c_2\ f_2\ (x)]P_X\ (x)\ dx$$

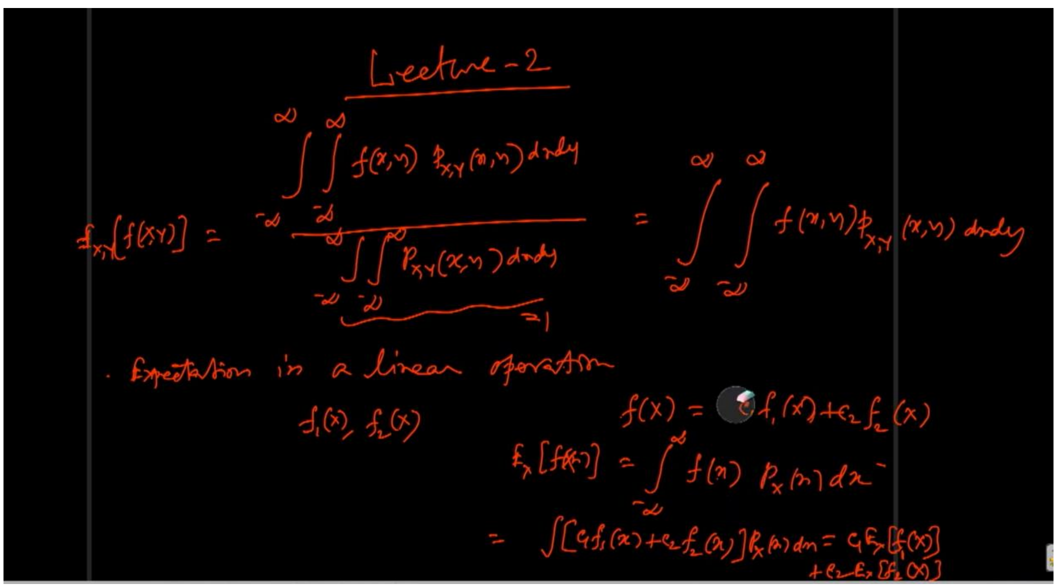
And now if you replace this by this $c_1$, x will become $x$. Now you know integral is linear, you can separate out the integral into two integral one is on $f_1(x)$. So, $c_1\ f_1(x)P_X\ dx$, one integral another is $c_2\ f_2(x)P_X\ dx$ another integral. If you do that you will have $c_1$ times

$f_1(x)P_X\, dx$ which is expected value of $f_1$ by definition. Expected value of $f_1(X)$ is what you replace $f_1(X)$ by $f_1(x)$ multiplied by the chance factor $P_X\,(x)\,dx$ integrate.

And $c_1$ constant goes out of integral plus $c_2$ times, again $c_2$ goes out of integral $f_2\,(x)$ and this which is again $E_X$ of this function $f_2\,(x)$. That is how you see linear, expected value of this is nothing, but $c_1$ times expected value of this, plus $c_2$ times expected value of this.

$$E_X[f(X)] = c_1\, E_X\,[f_1(X)] + c_2 E_X[f_2(X)]$$

**(Refer Slide Time: 08:04)**




So, expectation from outside can be applied directly on the constituent functions all right. Expectation of the overall $c_1$ if for the x plus $c_2 f_2(x)$ will be nothing, but $c_1$ times expectation of this and $c_2$ times expectation of this same thing we can we have here. If we have two functions $f(X,Y)$ is nothing, but $c_1\, f_1(x,y)$, plus $c_2 f_2(x,y)$ ok, here you replace this by that.

So, one double integral for $f_1$, another for double $f_2$, $c_1$ goes out, $c_2$ goes out and you get the same thing the expectation is linear all right this is the average. So, now I consider some case that we already have one random variable $X$ and we know its mean ok that is expected value of $x$ itself we have done it $X \rightarrow \mu_X$, plus and y is expected value is mu y $Y \rightarrow \mu_Y$ they are constant they are not random. Now, we are given a function f is

$$f(X,Y) = (X - \mu_X)(Y - \mu_Y)$$

So, what is the expected value I have to take this double integral instead of $X$ it will be now $x$ within integral, again instead of $Y$ inside the integral it will be $y$, because those are the integral variable all right sorry one term is missing. The joint density most important term is missed out all right this is expected value of this thing, is a very important thing it is called covariance, I will give you the physical meaning of this of $X$ and $Y$, you denote it as $C_{X,Y}$.

$$E_{X,Y}[f(X,Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x - \mu_x)(y - \mu_y) P_{X,Y}(x,y) dx\, dy$$

$$= E_{X,Y}[(X - \mu_X)(Y - \mu_Y)]$$

$$= C_{X,Y}$$

C for covariance. What it means is this it actually gives you an idea about how well the $X$ and $Y$ are related to each other internally, are they very uncorrelated, no relation between them or there is some relation like temperature humidity has some relation ok, but temperature and you see rupee value of dollar, they are absolutely two different things no connection. So, there is no correlation both are real, both are continuous ok. So, you see $X$ is a random variable, $X$ minus its mean. So, what is mean, I am going this is my $\mu_X$ and $X$ can be here to here, any you know it is fluctuating every time.

So, suppose it has taken this much value and again this is your $\mu_Y$ and $Y$ is going like this. So, by $\mu_Y$ has taken this much value. So, this is that increment $X - \mu_X$ in a particular experiment this is the increment $Y - \mu_Y$ in a particular experiment I am multiplying the two and this product I am averaging. Now, suppose $X$ and $Y$ are highly related to each other. So, when this goes up either this goes up always.

So, difference, they are positive. So, multiply you get a positive number or they go down it goes down. So, it also goes down. So, negative, but multiplication is positive. So, if they are highly correlated then what will happen if this is positive this is positive or this is negative, this is negative, product is always positive.

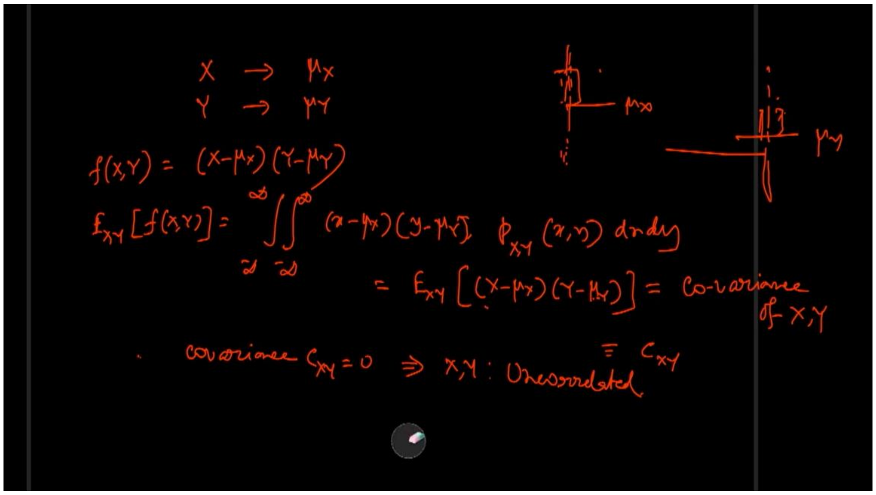
So, average will give a good value or it can so happen that the relation is such that when it is going up relation is such it will go down. So, positive increment negative increment. So, multiply you get a negative value and when it goes down it should go up ok, but there is this relation deterministic kind of thing if it goes down it ok it influences this or this influences this, either both positive or both negative or when this is positive this is negative and vice versa. When this is positive, this is negative product is minus again product is minus and we average you get a sizable good negative number. So, magnitude of this average will be high that means, if this average is high in magnitude, I will assume that they are influencing each other very well ok.

So, they are correlated on the other hand if there is no correlation as I told you one is temperature and another is the you know rupee value of US dollar, no correlation. So, if they go up ok, but while it goes up it can sometimes go down sometimes go up. So, go up go down. So, positive into positive and sometimes positive into negative or sometimes this is negative this can go up go down. So, negative into positive, negative into negative.

If you average out 4 cases it will be 0 or close to 0. If they are not correlated to each other, they are not influencing each other. So, if they are uncorrelated then this value covariance will be 0 or close to 0 ideally 0. So, implies x y uncorrelated alright x y uncorrelated.

$$C_{X,Y} = 0 \implies X, Y \ are \ uncorrelated$$

**(Refer Slide Time: 14:22)**

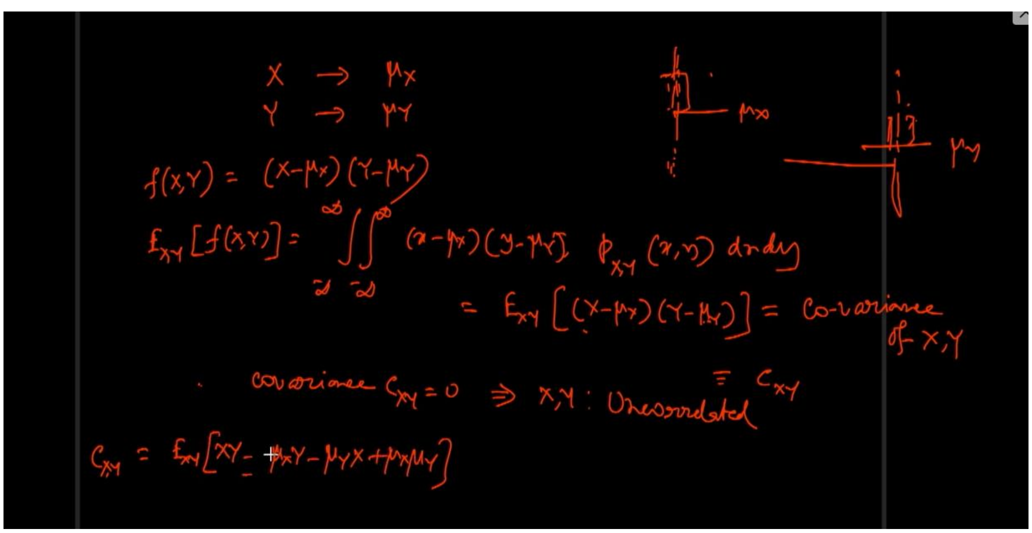



Anyway, now we come to this formula, that was this now this I multiply $(X - \mu_X)$ and $(Y - \mu_y)$, simple multiplication will be $XY - \ \mu_X Y - \mu_Y X + \mu_X \mu_Y$ and I told you E is a linear operator.

$$C_{X,Y} = \ E_{X,Y}[XY - \ \mu_X Y - \mu_Y X + \mu_X \mu_Y]$$

$E$ on summation of this functions is same as $E$ on this, $E$ on this, $E$ on this, $E$ on this, because if you multiply this by the joint density of $X, Y$ and then integrate the double integral you can separate out one on this, one on this, one on this, one on this.

**(Refer Slide Time: 15:31)**



So, you get average of this, average of this, average of this, average of this. So, $E_{X,Y}(X,Y)$ this is one term you get, then minus from this minus E x y of this product now here this is a constant this is not random. So, it will go out alright it will go out ok. Anyway, I need a small result here.

So, once I derive that then I should come back. So, let me for the time being not pursue this up to this is fine. I have to go to something called conditional density and then I will come back to this, this you know using linearity, breaking it up in 4 terms and applying E on this and then trying to see what immerges and all that is later one more result I need. So, let me go up to this and then go to the other one. Next topic which is joint sorry conditional probability density let me erase this part.

Suppose I am doing the same experiment. So, suppose I am doing the same experiment. So, in this experiment we are given 2 joint jointly random variables as before. So, this much is my $x$, this is $dx$, this is $y$ and this is $dy$ as before plus this point is $x\ y$. So, I have to find out the chance of $X, Y$ pair falling here that we have already done now I will do the same thing in another way ok.

We have already seen probability of $X$ lying from $x$ to $x + dx$ simultaneously $Y$ equivalently meaning $X, Y$ lying here that we have already seen in terms of the joint probability density this much, but we can arrive at the same thing in another way. Suppose we first restrict or whether we first 1 minute restrict the choice of the joint case of capital $X$ to be here only. So, in my experiment's capital $X$ will take any value from here to $-\infty$ to $\infty$, but I ignore all others I will consider only those cases where $X$ is from $x$ to $x + dx$.

$$Prob(x \leq X \leq x + dx, y \leq Y \leq y + dx) = P_{X,Y}(x,y)dx\,dy$$

So, we first restrict the choice of $X$, but that will not always happen that will have a chance factor that I will bring in later ok. Because $X$ as such will take value from $-\infty$ to $\infty$.

Ok there is nothing that, says that, you know, that ensures that $X$ will be here only, but I am considering only those experiments, in which $X$ produces a result from either $x$ to $x + dx$. So, that will have a chance factor because that will not always happen that chance factor will bring later, but suppose I am considering only those cases, where $X$ takes values from either $x$ to $x + dx$ only those cases, that is we first restrict the choice of $X$ to alright. Subject to these, now in these cases where $X$ fixed here, subject to these, subject to the above, that is $X$ is lying here only, then the probability that $Y$ will be from here to here, that is from here to here, which will mean $X, Y$ that will lie here. Because $X$ is lying from $x$ to $x + dx$ and now, I am in those cases when it is already you know, guaranteed to lie here only because I am only considering those cases where $X$ lies from $x$ to $x + dx$, only in those cases ok. Again, that is keeping that fixed if I measure $Y$ there is no guarantee capital $Y$ will lie from here to here or then the pair will lie from you know will lie on this box is it not.

So, subject to the above probability of $Y$ lying between $y + dy$, which is equivalent to $x\ y$ lying in this box, lying in the box, in the shaded box. There is subject to this because capital $X$ is already lying from $x$ to $x + dx$ only those cases are considered. Then again capital $Y$ could have been anywhere from minus infinity to infinity, then I am taking the only those cases, here while capital $X$ is fixed here, then subject to that only those cases where capital $Y$ is lying from here to here ok that will again be proportional to $dy$, where $dy$ is small. So, if $dy$ is large ok, if $dy$ is large that probability will go up transits will go up or if $dy$ is less transits will be less, since $dy$ is infinitely small I will take this probability to be linearly proportional to $dy$ and therefore, it is equal to a proportionally constant time $dy$. Proportionally constant it is not just a function of $Y$ because capital $X$ also is playing a role, capital $X$ is taking a value up to small $x$ ok.

$$Prob(y \leq Y \leq y + dy)\ \alpha\ dy$$

So, it is basically a function of again capital $X$, I mean where this slot is located which is small $x\ y$. So, we write like this it is just a notation capital $Y$ instead of writing $Y, X$, we write $^{Y}/_{X}$, this is a notation, but it actually means it is a function of both the random variables $X$ first and $Y$ and in this order you write $^{y}/_{x}$ it means $X$ is constrained to take value from $x$ to $x + dx$. $x$ means $x + dx$, $x$ to $x + dx$ subject to that $Y$ taking value from $y$ to $y + dy$ ok.

$$= P_{Y/_X}(^{y}/_{x})dy$$

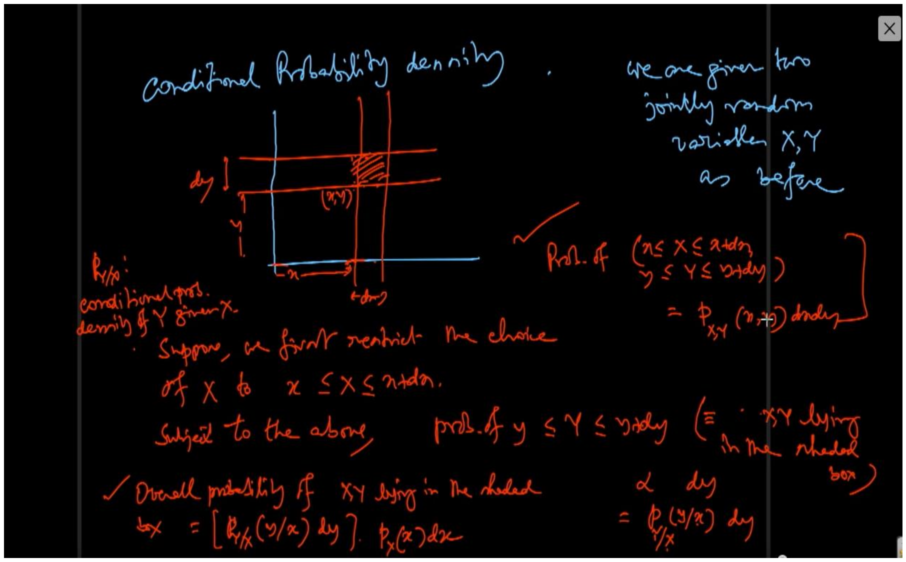
There is a meaning of this. This is the proportionally constant that times the width ok. So, therefore, or equivalently because $X$ is already lying from here to here and then under that $Y$ lying from here to here, which means under $X$ lying from here to here, what is the probability of that pair lying in the box, that is this much ok, but as I told you $X$ will not always occur here, only it will occur here only with a chance factor. So, overall chance of this overall chance of $X, Y$ falling here will be what this probability times the chance of this happening, $X$ times between $x$ to $x + dx$. So, overall probability of $X, Y$ lying in the shaded box will be what this chance this actually means is a function of $X$ and $Y$ by we write like this.

$$[P_{Y/_X}(^{y}/_{x})dy]P_X(x)dx$$

We write like this. it means basically $X$ takes a value from $x$ to $x + dx$ subject to that $Y$ takes a value from $y$ to $y + dy$, then the probability is this which would mean then $X, Y$ will lie here, but under the condition that $X$ is only prefixed to take values from here only those cases are considered. So, only then the then the chance is so much, but $X$ will not always take value from here. So, overall chance of $X, Y$ falling here will be these chance times the chance of $X$ occurring here. So, $X$ occurring here chance we have already seen in the first lecture it is. So, now, this and this they are same this was written in direct way, this is obtained in another way.

This is called conditional probability density of $Y$ given $X$. So, this two are same. If you now equate, if you equate the two, $dy\,dx$, $dy\,dx$, they are common.

**(Refer Slide Time: 27:43)**

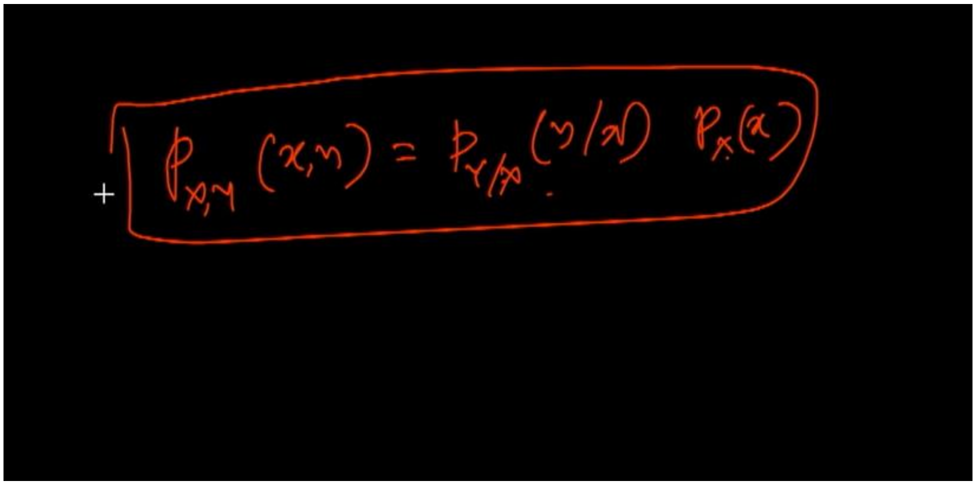



So, take the bound. so that means, joint density joint density let me write in the next page.

This joint density is this into this that is joint density is $Y/X$, $x + dx$. If $x$ is constrained to lie between $x$ to $x + dx$ subject to that $Y$ lying from $y$ to $y + dy$ probability density of that, the next term will be $X$ here. If this is the constraint here then it will be free here this variable this is how we get all right.

$$P_{X,Y}(x,y) = P_{Y/X}(y/x)P_X(x)$$

**(Refer Slide Time: 28:38)**

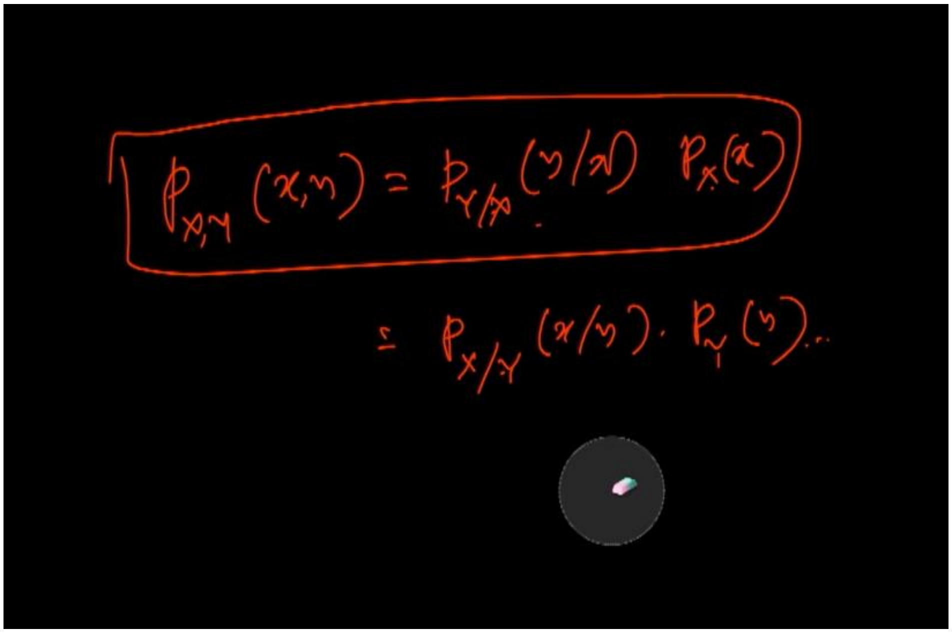


Now, in the previous page I constrained $X$ to lie from here to here subject to that I took the cases when $Y$ lying from here to here.

I can reverse that we interchange that. So, I can first constrain $Y$ to lie from $y$ to $y + dy$ and then see how often $X$ lies from this much to this much. So, just same procedure interchanging $X$ and $Y$. So, then it will give me equivalently $P$ capital $X$ by $Y$ that is capital $Y$ is constrained to take value from $y$ to $y$ to $y + dy$ subject to that $X$ lying from $x$ to $x + dx$ there is a probability that times now this will be $P_Y(y)$ all right either way ok. Now if $X$ and $Y$ are called statistically independent or maybe I will do this later this one not now. I rather stop here now because another topic if I begin in 2 5 minutes you know it will be I will not be able to complete that we will just begin.

$$= P_{X/Y}(x/y) P_Y(y)$$

**(Refer Slide Time: 30:14)**




So, it will remain this continuous. So, to stop here in the next class I will build up on this conditional probability you will see me further some properties and all that then I will pick up that case of covariance correlation and then instead of taking just two random variables I will then take 3 $X, Y, Z$ that will show up further properties and then I will generalize it to n number of random variables together I will call them $x_1$ $x_2$ capital $X_1$ capital $X_2$ dot dot dot capital $X_n$, then will come their correlation matrix covariance matrix and all those ok and then their properties we will go by this ok. So, thank you very much if you have any question you can raise in the forum ok and please follow my lecture here also the NPTEL video lecture clearly it will be clear ok. Thank you very much see you then bye bye.