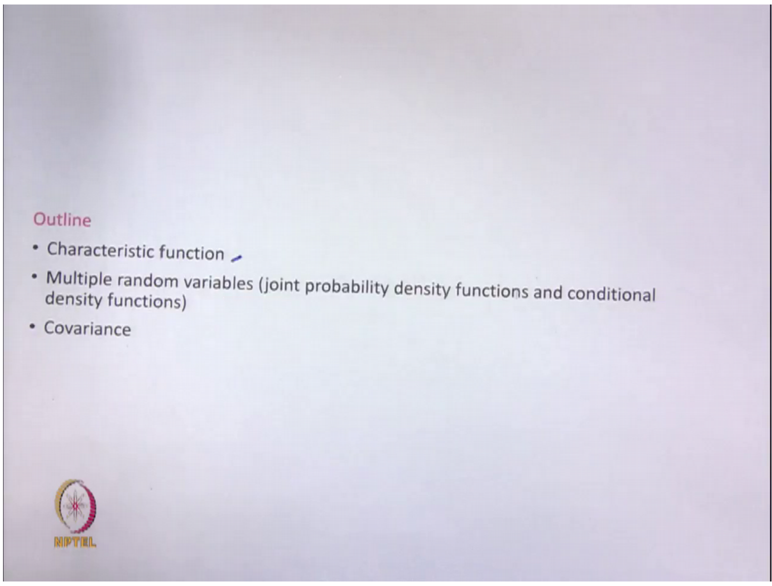
Principles of Digital Communication
Prof. Abhishek Dixit
Department of Electrical Engineering
Indian Institute of Technology, Delhi

Lecture - 09
Random Variables & Random Processes: Multiple Random Variable

Good morning. Welcome to lecture number 3rd on Random Processes. In the last lecture, what we did was we defined what are known as continuous random variables. And we have looked into the probability density function, which is used to characterize the probability of continuous random variables. And then we looked into uniform distributed random variables, and Gaussian distributed random variables.

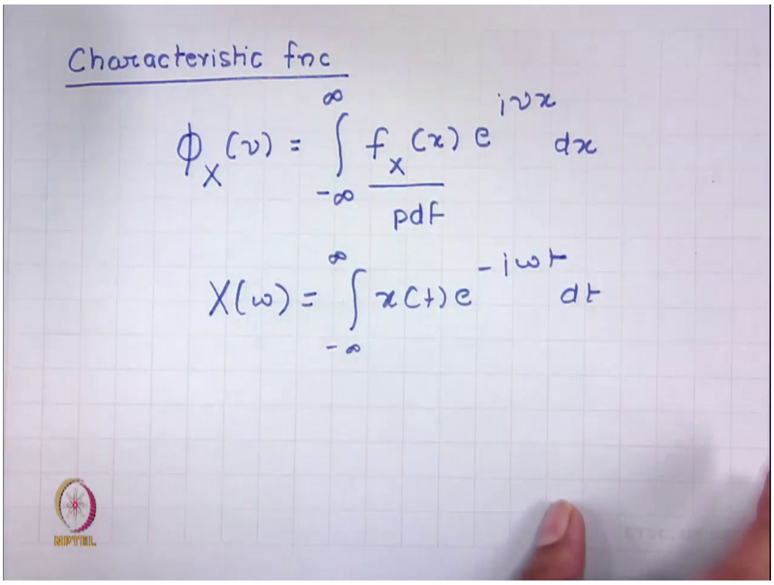
We said Gaussian distributed random variables are really important random variables, and we will use a lot of them in this course in digital communication. Then we derived an important property related to Gaussian random variables, and that property is that any linear function of a Gaussian random variable is a Gaussian random variable.

(Refer Slide Time: 01:09)



In this lecture, what we will like to do is to talk about another important concept of characteristic functions. Then we will talk about multiple random variables. So, far what I have been doing is I have been just taking a one single random variable. Now, we will see what happens if we have more than one random variable in picture, and then we will talk about the covariance that is the new concept.

(Refer Slide Time: 01:31)

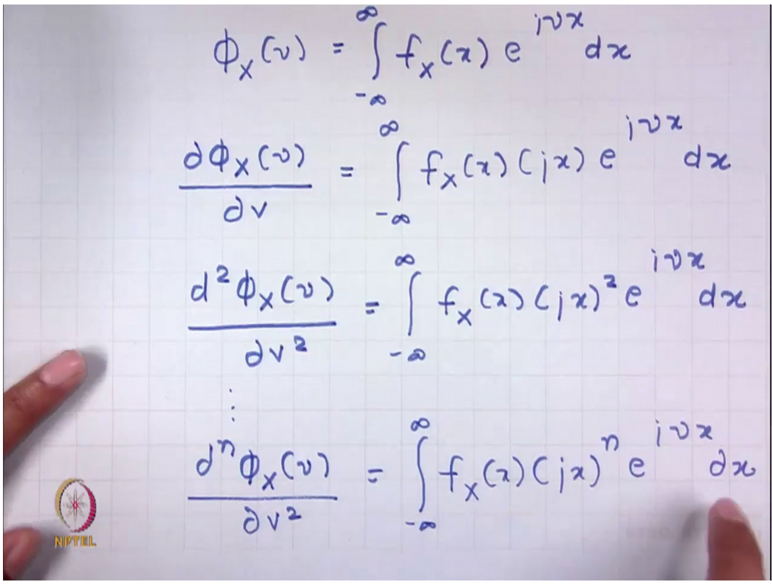


So, starting with characteristic function. First let me define what is the characteristic function, and then I will explain how this function is so useful. So, if we have to define a characteristic function for a random variable X, the characteristic function is expressed like this ok. So, what I have here is the probability density function of this random variable X, and I have multiplied this with a complex exponential. And I am integrating it with respect to X, so that is how the characteristic function is defined.

Now, let us see two important things, this is very similar to Fourier transform of a signal. So, if I have, let me define Fourier transform of a signal, so this is the expression for the Fourier transform of the signal just reminding you about this. And this expression looks very similar to Fourier transform of the signal, other than that there is a negative here, and there is a positive here. So, other than that this expression looks exactly a very similar to the Fourier transform of the signal. So, I can think characteristic function as the Fourier transform or inverse Fourier transform of the probability density function ok.

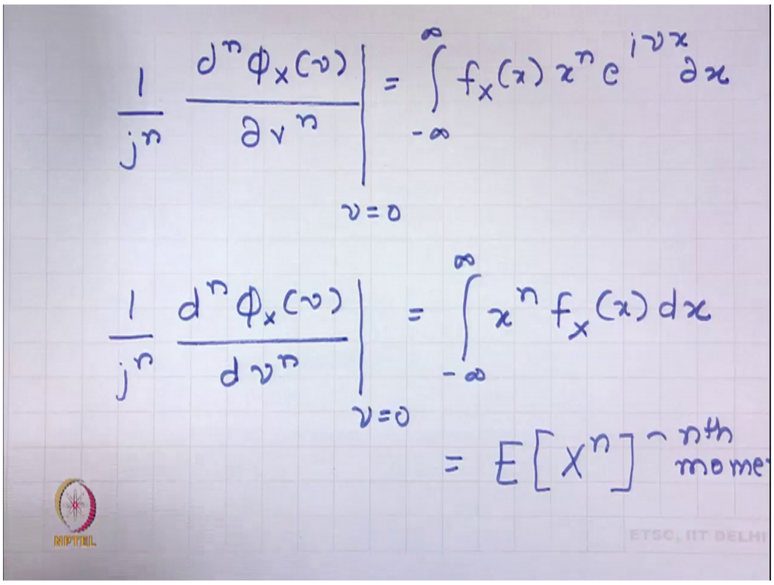
So, then the second thing that you have to you can think in this, and it is important to notice that this integration is with respect to running variable X right. And this is a definite integral that means, you are putting a limit from minus infinity to plus infinity. So, when you will do this integration, the resultant function will not be a function of X right, because the running variable is x. If you are running, if you are integrating something with respect to x, and if it is the definite integration, then you in the output you would not have any function of X. The output would if would be a function of this nu right. So, just to appreciate that this characteristic function is a function of nu, where nu is this variable, it will not be a function of x, because the integration has been done with respect to x.

(Refer Slide Time: 04:21)



So, let us see how this characteristic function is useful, and let me write this again. The characteristic function we are writing for a random variable X is nothing but in some form it is the Fourier transform of the probability density function. Let me differentiate this everything with respect to nu. So, if I differentiate with respect to nu, I will get this right, it is a simple differentiation. Let me differentiate it again, let me differentiate it 2 times, what I would end up with is instead of j x, I will have j x square ok. I can keep on doing this differentiation let me do it for n times, I will end up with this.

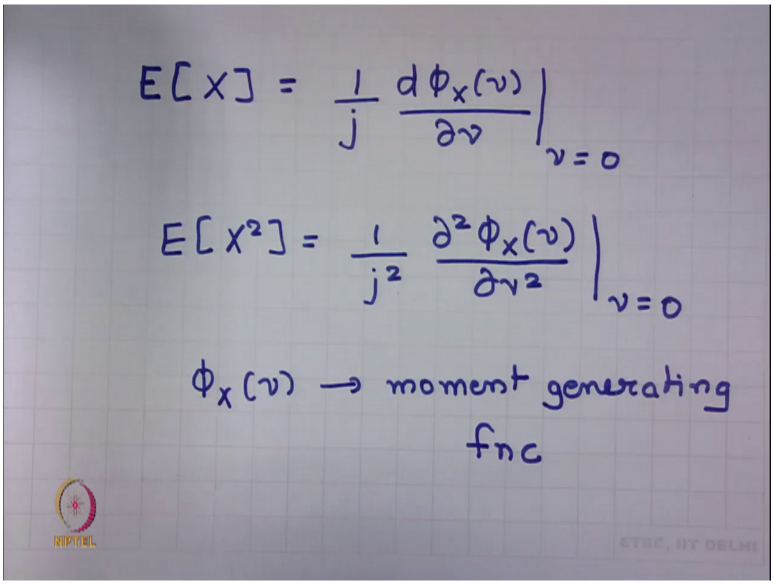
(Refer Slide Time: 06:15)



So, what I am saying is from this expression. So, this should be n sorry so what I can do is I can take this j to the power n to this side ok. So, let me now get rid of this and concentrate on this expression. So, do you get some sense of this why can this be important function, let us just evaluate this function at nu equals to 0. So, if I evaluate this function at nu equals to 0, so this here I have to substitute nu equals to 0, and this becomes at nu equals to 0 yes. And what is this? This is nothing but it is the expected value of random variable X to the power n right.

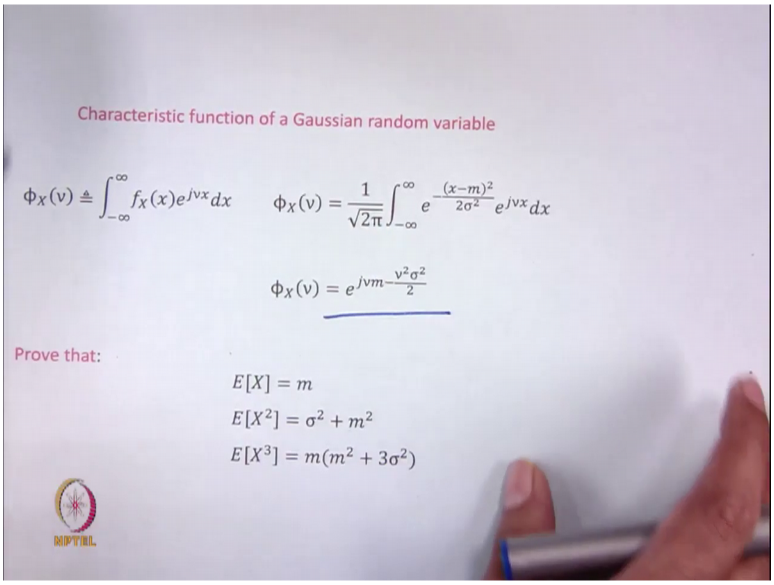
So, what I am saying is the expected value of a random variable, and this expected value of a random variable which is raised to power n is nothing but it is the nth movement of random variable X ok. So, I can find nth moment of random variable X in terms of characteristic function right. So, if I am given a characteristic function of any random variable, I can find it is all moments.

(Refer Slide Time: 08:43)



For example, if I have to find first moment of a random variable X, then this would be nothing but this. So, first movement of your random variable X can be obtained from that expression like this. You can find similarly the second moment like this snd you can generate all moments of random variables; all moments of your random variable in term in terms of this characteristic function. So, if I am giving the characteristic function of a random variable, I can generate all its moments, and that is why this characteristic function is also known as moment generating function. And this is very useful, because by using this I can derive all moments of a given random variable.

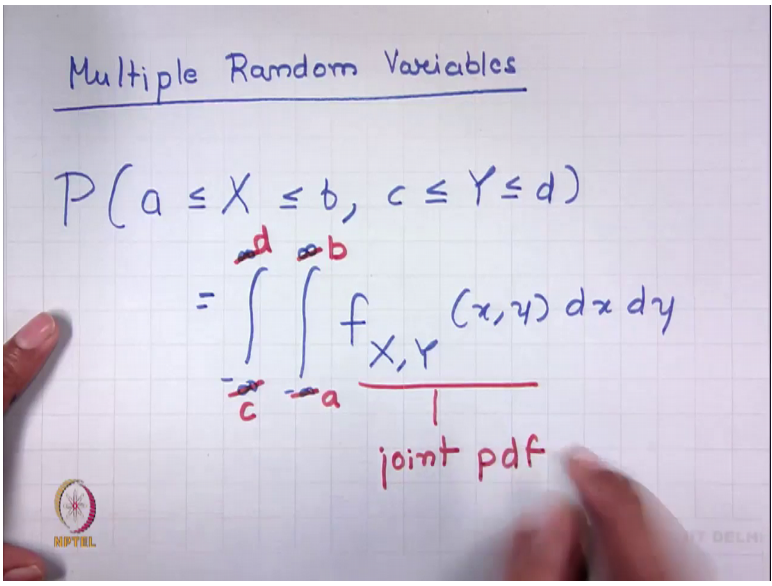
(Refer Slide Time: 10:07)



So, let us now see one example, where I will use this characteristic function. So, if you look at the characteristic function of a Gaussian random variable, now I am particularly looking into a Gaussian random variable. As I have said this characteristic function is defined like this. If you want to find the characteristic function of a Gaussian random variable, the probability density function has to be substituted for a Gaussian random variable which is this probability density function, and then you solve this integration.

After solving this integration, what you get is characteristic function of a Gaussian random variable is this ok. I request you to solve this integration, and derive that a characteristic function of a Gaussian random variable is this. And using this characteristic function try to find all moments of a Gaussian random variable. And prove that the expected value of X, where X is the Gaussian random variable is m, expected value of X square is sigma square plus m square, expected value of X cube is m times m square plus 3 sigma square. So, using this single characteristic function, you can derive all moments of the random variable of interest. And hence this characteristic function is very useful in the probability theory.

(Refer Slide Time: 11:29)



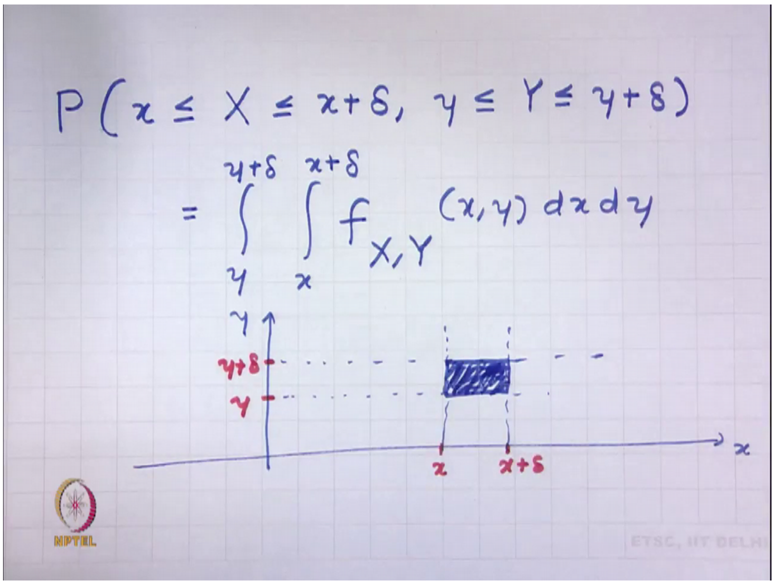
So, let us now introduce the next concept of the day and as I talked about, we will be thinking about multiple random variables. So, what we study here in this under this topic is up to now we were just concentrating on a single random variable. Now, the question is what if I have more than one random variable? So, let us say I am interested in finding the probability that a random variable X lies between a and b, and the random variable Y lies between c and d ok.

So, I have two random variables here X and Y. And I am interested in this probability that random variable X lies between a and b, and y lies between c and d. And I am interested that both these events happened at the same time. So, when X takes the value between a and b, Y takes the value between c and d, so what is this probability?

So, if we have to evaluate this probability, and X and Y are continuous random variables, then what I would end up with is; so what I will end up with is, so because I am interested in the random variable x taking a numerical value between a and b, this limit has to be from a and b. And as I am interested in random variable Y taking the value between numerical value between c and d, this limit has to be between c and d.

So, what I am saying is the probability that a random variable X takes a value numerical value between a and b, and Y takes a numerical value between c and d, I can evaluate this probability in terms of a new density function. So, this is a new density function that I have and this is known as joint probability density function. So, this is a joint probability density function. So, using a joint probability density function, I can calculate or compute the probability of two random variables lying in a region ok.

(Refer Slide Time: 14:25)

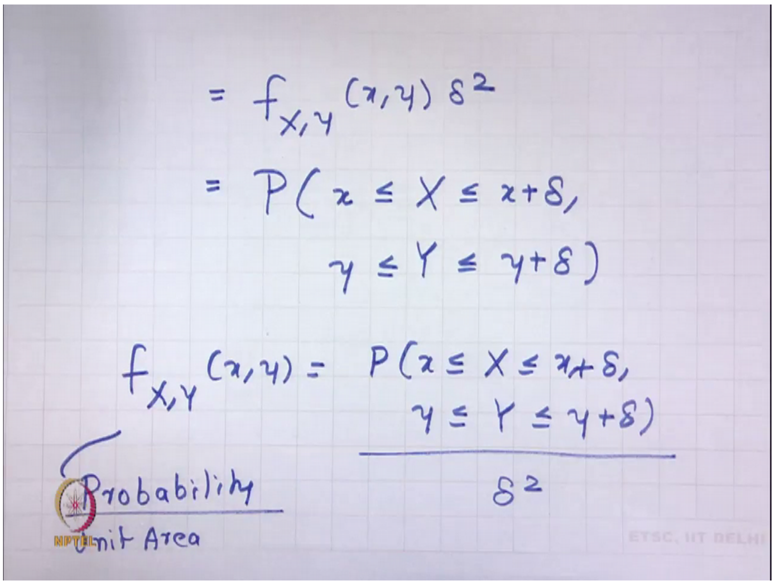


So, let us try to interpret this joint probability density function little bit more, and try to see what does this mean. And as we have done for a single random variable case the approach is essentially similar. So, what I am saying is I would be interested in the probability that this random variable X lies between x and x plus delta. And the second random variable Y takes a value between y and y plus delta.

And if I use the expression that I just have written before, what I would get is y going from y plus y plus delta x going from x to x plus delta ok, then I have f X, Y. So, these are the random variables, and this is the numerical value of the random variable dx dy. So, graphically what it means is so I have now two random variables say x and y. And I am interested in what is the probability that random variable X lies between x and x plus delta, and random variable Y lies between y and y plus delta.

So, if you look at this, basically what we are interested in is what is the probability that the random variables lie in this region ok, so what is the probability that random variables lie in this region. So, if the random variables X and Y lies in this region, X lies between x and x plus delta, and Y lies between y and y plus delta ok. So, thinking that this probability density function is constant, because this delta is really small. So, what I am assuming that my probability density function is constant in this limit.

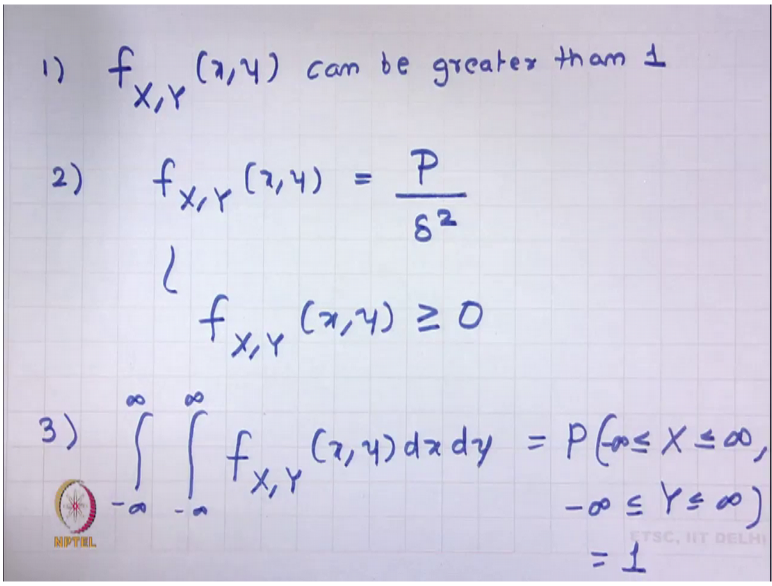
(Refer Slide Time: 16:43)



So, assuming that this is constant in this limit, what I would get is ok. So, with constant I can pull this out of integration, and dx into this thing will become delta, and dy into this thing will become delta. So, I will have f of X, y x, y that means, joint probability density function of x, y into delta square. And this quantity is nothing but; what is the interpretation? The interpretation is that this quantity is nothing but the probability that X lies between X and x plus delta. And Y lies between Y and y plus delta ok.

So, I can write f X, Y x, y joint probability density of random variables x and y. So, what is the unit of this quantity? This is probability that we know. And this is delta square that means, this area right. So, this quantity unit is probability per unit area ok. So, remember as I have said for a single random variable case, this probability density function is not probability. In single random variable case, the probability density function was probability per unit length. And in case of two random variables, the joint probability density function is probability per unit area, it is not probability. So, it can be more than 1 ok.

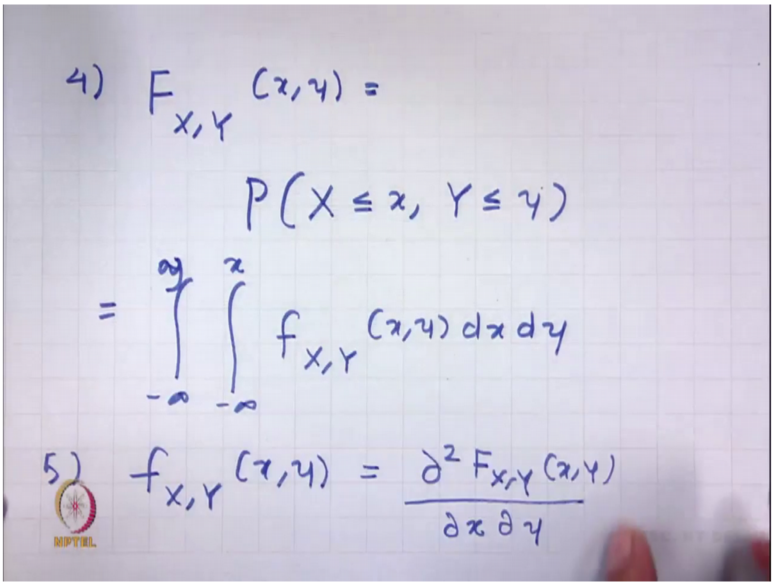
(Refer Slide Time: 18:43)



So, it is f of X and Y x, y that is the joint probability density function of random variables X and Y can be greater than 1, because this is not probability ok. And as this quantity is probability per unit area, and probabilities can never be negative. The joint probability density function is also always non-negative that means, it is either greater than 0 or it is equals to 0, but it can never be negative.

The third point, if I integrate this joint probability density function by taking the limits of x from minus infinity to plus infinity, and limits of y infinity from minus infinity to plus infinity, what I am saying is this quantity will be equal to the probability that random variable X takes a value between minus infinity to plus infinity, and random variable Y takes a value between minus infinity and plus infinity ok. And you know that this probability will be 1, because X will definitely lie between minus infinity and plus infinity, and Y will also definitely lie between minus infinity and plus infinity. So, this probability will be going to be 1, this probability will be 1 ok.

(Refer Slide Time: 20:53)

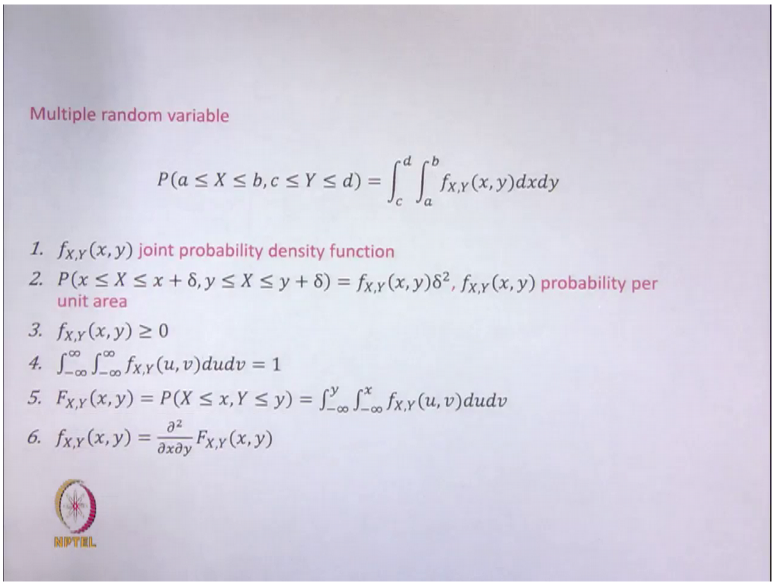


Now, the next important property that we can understand is deriving the cumulative distribution function or cdf. The cdf in case of two random variables is nothing but it is the definition is very similar to the case of a single random variable, it is nothing but what is the probability that random variable X takes a value less than x, and random variable Y takes a value less than y ok.

And as you know this is nothing but; this will be nothing but this quantity. So, x limit has to go from minus infinity to x, because random variable should fall in this limit. And for y the limit should go from minus infinity to y, because random variable y should fall in this limit. So, what I have done is I have been able to express a cdf or cumulative distribution function in terms of joint probability density function.

And I can also go back, if I wish I can find joint probability density function in terms of cdf, by taking the partial derivatives of this quantity. So, by doing the partial differentiation of cdf, I can find the joint probability density function right. So, this is essentially how can you think about two random variables. This is the summary of all the properties that I have done.

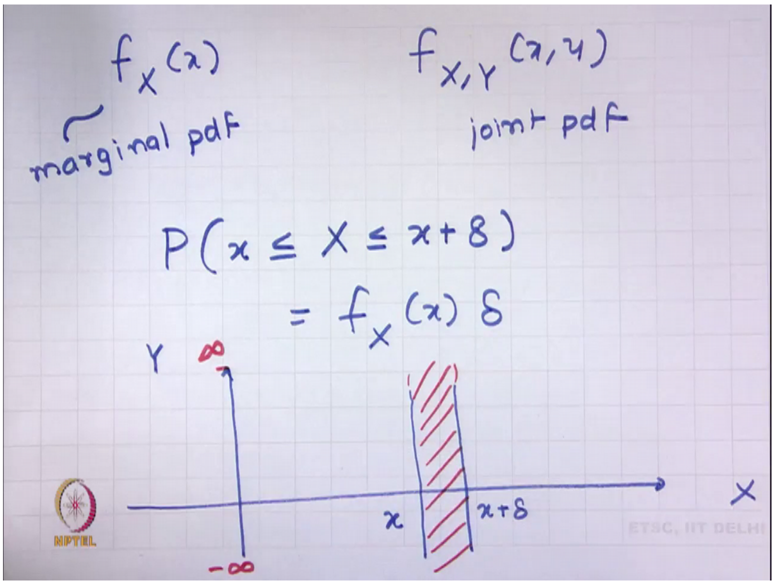
(Refer Slide Time: 22:47)



What I have said is if we want to find the probability that a random variable X takes a value between a and b, and random variable Y takes a value between c and d, then this probability can be computed in terms of joint probability density function of x and y. I have said that the joint probability density function is nothing but it is a probability per unit area, it is not probability enhance, it can be more than 1.

But, because it is used to calculate probability, this quantity has to be non-negative. If you integrate this quantity from minus infinity to plus infinity for u, and from minus infinity to plus infinity for v, then this will give you this will be 1, because then we are saying what is the probability that x lies from minus infinity to plus infinity, and y lies from minus infinity to plus infinity, and that probability will be 1. Then we have said that you can think about the cdf in terms of pdf by doing the double integration of this quantity. And you can get pdf from cdf by doing the partial differentiation of cdf.

(Refer Slide Time: 24:03)

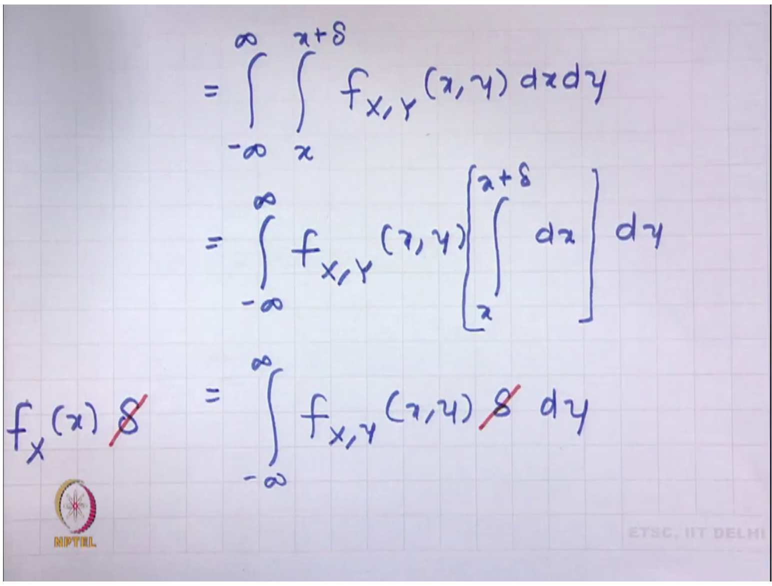


So, let us now think about the probability density function of a single random variable, in terms of the probability density function of two random variables. So, we are trying to establish a relationship between this pdf in terms of this pdf. And we also called this as marginal probability density function, and this is nothing but joint probability density function.

And essentially what we are trying to do is to find out an expression for marginal pdf in terms of joint pdf. And to think about this, let me write what is this quantity. So, if I am interested in finding the probability that x random variable X lies between x and x plus delta, what is this? This could be written as by definition in terms of marginal pdf, that means here I am just thinking about a random variable X a single random variable. And the probability that a single random variable lies or takes a value between x and x plus delta is nothing but it is the marginal pdf of that random variable times delta ok.

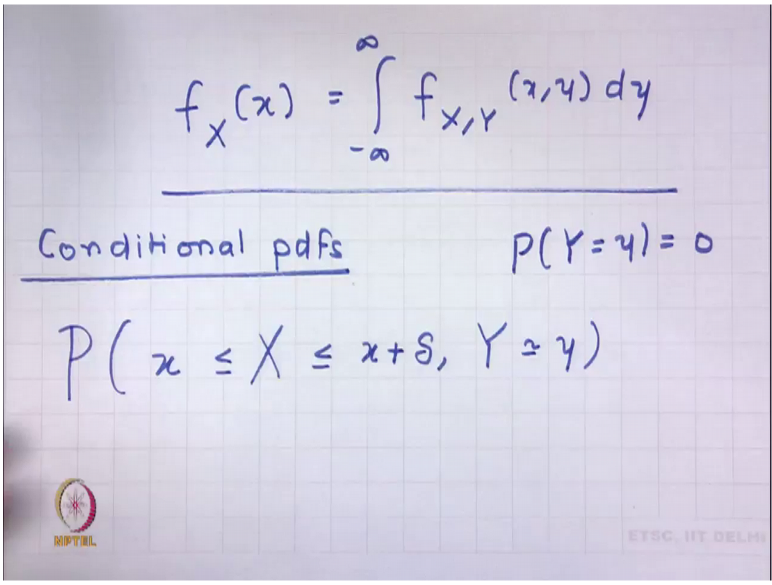
And let me try to interpret this quantity in terms of joint pdf to establish a relationship between a marginal pdf and joint pdf. So, if I am thinking about this probabilities in terms of joint pdf, so if I have two random variables, let us say X and Y, and I am trying to think what is the probability that X lies between x and x plus delta. So, as you can see this probability will be this region. So, I am interested in finding what is the probability that X lies between x and x plus delta, and what are the limits for the Y, Y can take any value between minus infinity to plus infinity right. So, Y can take any value from minus infinity to plus infinity. So, we have to take all values of Y ok. So, what you will get is a thin strip like this.

(Refer Slide Time: 26:37)



So, if I want to compute this probability in terms of joint pdf, what I am saying is the Y should go from minus infinity to plus infinity, and X should go from x to x plus delta. And this quantity will give me the probability that X falls on this strip ok. And assuming that this joint pdf is constant, because delta is very small in the limit x and x plus delta, I can pull this out of this integration assuming that this is constant. And what I get is ok. And if you look, then that quantity in terms of marginal pdf was this by definition. So, what we get is we can cancel this delta with this delta.

(Refer Slide Time: 28:15)



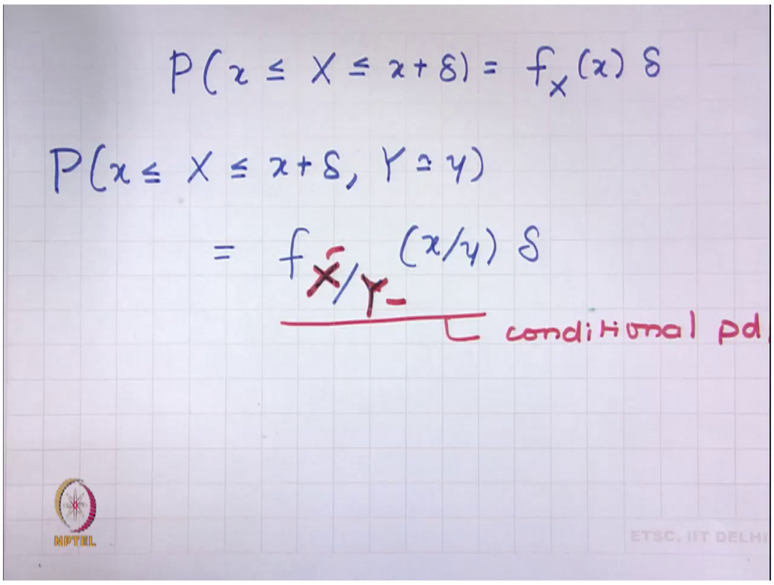
And what you get is a very neat expression that marginal pdf is nothing but this expression. So, this expression allows us to compute this marginal pdfs in terms of joint pdf, and this relationship is quite useful. And this is telling me if you want to think about the probability density function of x, you can think of it in terms of joint pdf you integrate this joint pdf with respect to y and you take the limits from minus infinity to plus infinity, because there is no restriction on what values Y can take. Y can take any value between minus infinity to plus infinity, and this integration would give you the marginal probability density function of x. So, this is quite useful.

And the another relationship that will also be quite useful in this course is thinking about conditional pdfs. So, let us start with this conditional pdfs. So, what does this conditional pdfs helps me to calculate is if I am interested in finding the probability that random variable X takes in the value between x and x plus delta, but now I have been given an additional information, and that information is that random variable Y takes in a value close to y ok.

So, under this information, new information or a new universe that I have been given that random variable Y takes a value around y. I am interested in finding, what is the probability that X takes a value between x and x plus delta. There is one question that we are saying random variable Y takes in a value around y, we are not saying that random variable Y takes in a value y, because that probability is 0 right is a continuous random variable.

So, the probability that probability that Y takes in a value y, this probability is 0. So, we cannot say that Y is a random variable y takes a numerical value small y, we have to say that random variable Y takes in a value around y, it is very close to y, but it cannot be y ok. So, just remember this question in mind. So, you have to say that random variable Y takes in a value around y. So, under this universe under this condition what I have to do is I have to investigate this probability, and this is in we are going into the regime of conditional pdfs.

(Refer Slide Time: 31:19)

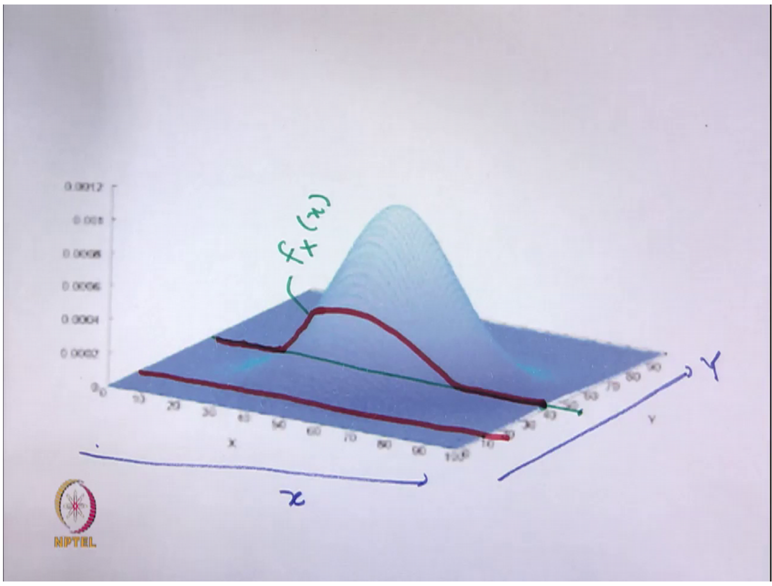


So, let us rewrite a definition. So, if I if you are interested in the probability that a random variable X takes in a value between x and x plus delta by definition you write that this is nothing but it is the marginal pdf of x times delta right. So, if I have to define what is it conditional pdf, in conditional pdf situation what we are given is that I am still interested in what is the probability that X takes a value between x and x plus delta, but under the assumption that random variable Y takes in a value around y.

By definition I can write by the same analogy I can write this as see we are introducing a new notation. So, new notation here is so this is the conditional pdf. So, what this conditional pdf represents is if you take a conditional pdf, and you multiply this with delta what you would get is the probability that a random variable takes in a value between in this case, for example it takes a value between x and x plus delta, and the second random variable value is given to you. So, under this condition, we use the conditional pdf right.

And let us now try to think what this conditional pdf is so just let me introduce the notation here. We are saying that we are interested in finding the probability of random variable X, what we are given is a numerical value of Y ok. So, I am interested in finding the probability of X given that I have some information about Y, and in the parentheses I also denote this by dividing the numerical value of x by numerical value of y. So, this is just notational, you could have used any other notation to know this effect. So, this is the notation that we would we use for conditional pdf, it is represented. So, what you have is I am interested in this random variable given the information about Y.

(Refer Slide Time: 33:45)

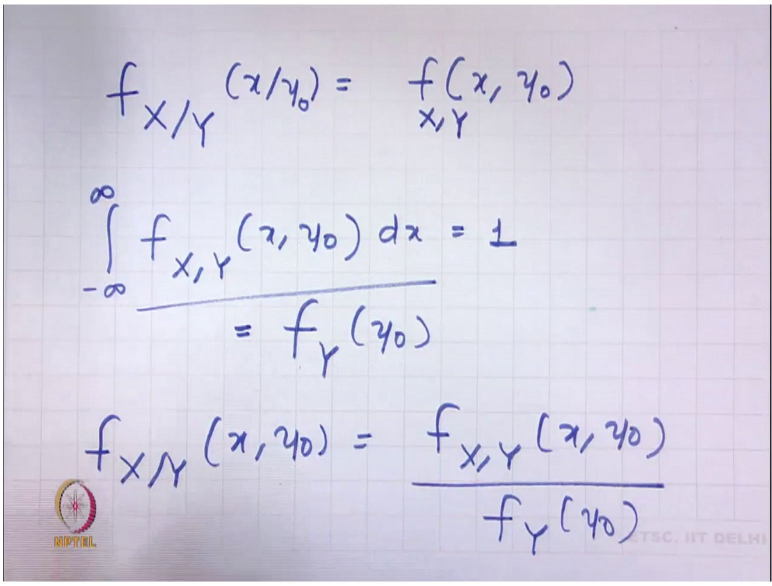


So, let us try to interpret this little more. So, let me first plot a joint probability density function, so this is a joint probability density function. So, you have on this axis you have Y, and on this axis you have X ok. So, now what I am saying is I am giving some information about Y. So, let me consider this strips. So, I am given that my Y lies here ok. So, suppose my Y takes in this numerical value. So, I am given that my Y lies or takes this value, and I am interested in finding what is the probability density function of X right.

So, if you look at this when Y takes in this value, if I am finding the probability density function it might be something like this ok. So, this quantity we are using that this is the probability density function of x. Now, it is not a function of Y, because Y has taken a value. So, this is this will gives me this probability density function the red probability density function would give me what is the probability density function of x.

Now, if Y would have taken just to appreciate, if Y would have taken this numerical value, so I draw a strip. And if you look at the probability density function of x is more or less a straight line right that means, all x are virtually taking the same probabilities ok. So, there is no preference for any X. But, if you look at this strip, what you would see is that some numerical values of x are more likely. So, what I am trying to say is depending upon what numerical value Y takes, you can have different distributions for x. So, now when I am giving that this Y has happened, I need to choose the appropriate probability density function of x ok.

(Refer Slide Time: 36:13)



So, let me say what I am trying to state is the probability density function a conditional probability density function of X, when Y's information is given is let me say that Y is y naught. So, I am being given that Y takes a value around y naught. So, this so what you have to do then is you have to take that join probability function, and you have to substitute that numerical value of y naught. So, depending upon y naught, this probability density function would take a specific value right. So, this is the idea behind, how can you define the conditional probability density function again in terms of joint pdf.

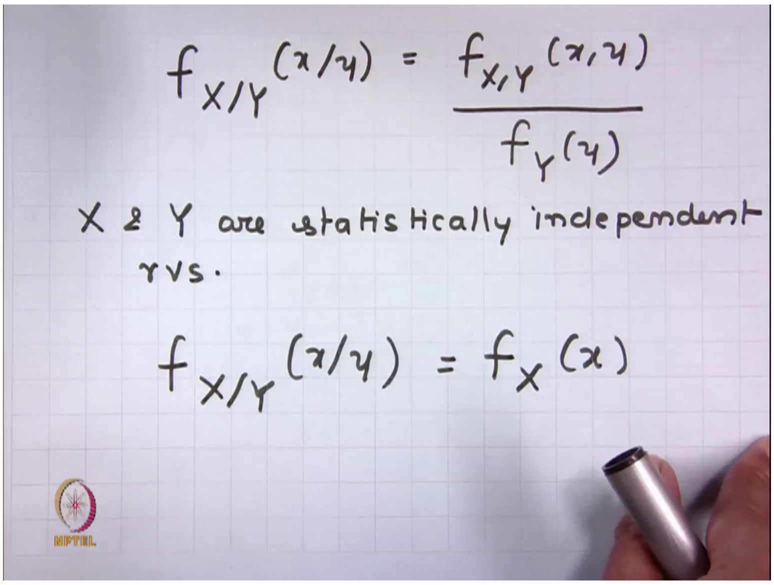
So, let me ask you a question, do you agree with this that this definition is correct or do I have to do something more than this. For example, does this satisfy a property that this is 1. So, if it is a correct probability density function, when you integrate it from minus infinity to plus infinity, it is not a function of y anymore, because y has chosen a numerical value. So, we just have to integrate it with respect to x. So, when you are integrating with respect to x, x takes a value between minus infinity to plus infinity. If it is a valid pdf, then they should integrate to 1. If it is not a valid, if it is not a valid pdf or if it is not integrating to 1, then it is not a valid pdf right.

So, and this quantity as we have defined before is nothing but so this integration by definition is this quantity right. So, we have already said how can you interpret the marginal pdf in terms of joint pdf. So, in terms of marginal pdf, this quantity is nothing but this. And this in general will not be 1, this will in general it is not 1. So, what we are saying that if you integrate this pdf from minus infinity to plus infinity, it is not going to 1.

So, we have to do proper scaling in this pdf to make sure that this resultant pdf always gives us 1, when it is integrated from minus infinity to plus infinity. So, the correct definition is this, so this is the correct way in which you can interpret this conditional pdf, so I am repeating. So, for conditional pdf or to interpret a conditional pdf, what you have to do is you have to first look at the joint pdf in terms of marginal pdf. And this marginal pdf is evaluated, when the Y has taken a specific numerical value.

So, now we have to look at the probability density function of x at the place, where Y has taken a specific numerical value, and because this value will not integrate to 1. We need to have the proper scaling factor, and the proper scaling factor that we need to use is this, so that when this quantity it is integrated from minus infinity to plus infinity, this turns out to be 1, to satisfy definition of a probability density function. So, this is how we can think about conditional pdf.

(Refer Slide Time: 04:07)

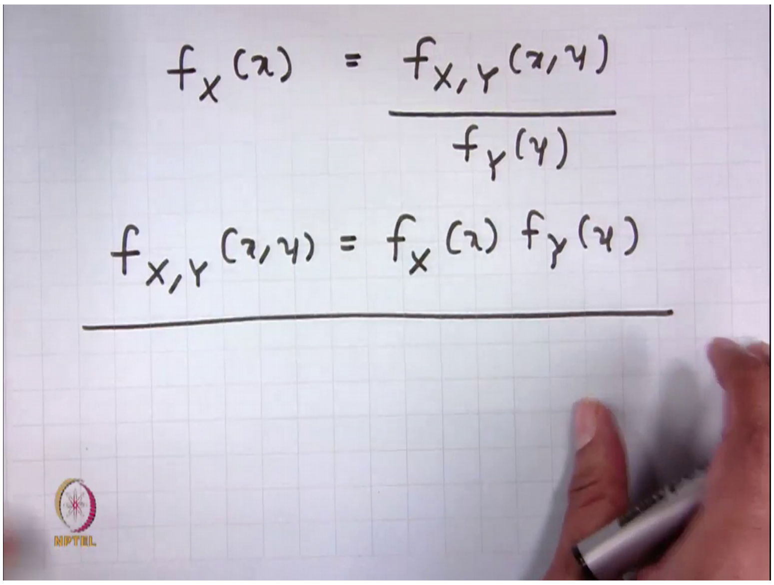


Now, let us do something interesting in this conditional pdf. So, let me rewrite the definition of conditional pdf ok. Now, what we are saying is let us say that X and Y are statistically independent independent random variables, what do you think, what will be the meaning of this statistically independent random variables means that if you have some information about Y, you do not get any extra information about X right that means, given if I tell you some information about Y, your knowledge about X does not increase, you have the same information about the X.

So, in terms of probability density functions what this (Refer Time: 41:17) to is the conditional pdf. So, now the probability density function of X given that Y has happened is nothing but it is the same as probability density function of x that means, you do not get any extra information about the pdf of X, when you are given some information about Y right.

So, whether you are given Y or not given Y as in this case the pdf's are same. So, this is the meaning of statistically independent random variables. If two random variables are statistically independent, the conditional probability density function of x given y is same as the probability density function of x ok. So, this is the basic interpretation about a statistically independent random variables.

(Refer Slide Time: 42:27)

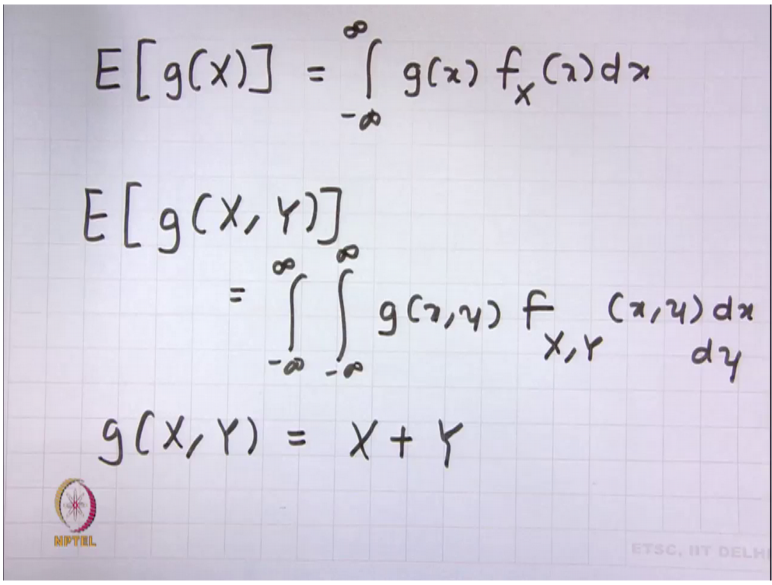


So, let me try to substitute this quantity in this expression. So, if I am given that two random variables were statistically independent, then this quantity (Refer Time: 42:31) to just the pdf of x ok. Under the assumption that my two random variables are statistically independent, I can replace the conditional density of X in terms of just the probability density function of x.

And then what I can do is I will get this expression. And this gives us an important insight, it is telling me if two random variables are statistically independent that means, information about one does not add any extra formation about another random variable. If two random variables are statistically independent, then their joint probability density function is nothing but it is the multiplication of their marginal probabilities right.

So, the joint probability density function of x and y is nothing but the marginal probability of x times marginal probability density function of y. The marginal probabilities multiplies two give you joint probability density function, if the random variables involved are statistically independent. This is an important property, and this is an important insight which helps us to simplify few expressions.

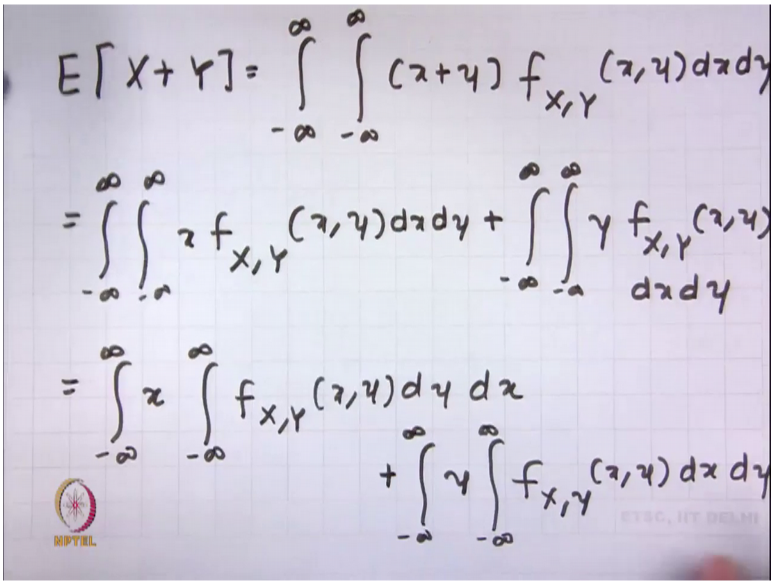
(Refer Slide Time: 44:17)



So, let us try to do, again we will like to come back to this expectation, we said is really useful. And let me remind you, what is the expectation of an function of X, this is what we did in previous lectures. I am just recalling you that we said that this is nothing but, so this was in the case of a single random variable.

In case of two random variables, for example we can calculate the expectation operator essentially in the same way. So, this by definition, because now two random variables are involved, we need to have double integration. And two random variables are involved, we need to have joint pdf’s that is it nothing else changes in this expression. So, this is how you can interpret the expectation of function of two random variables. And this is exactly similar to how we define expectation for a function of a single random variable ok. Let me now do certain cases, for example let us assume that this function of X and Y is nothing but it is a sum. So, it is the sum of a random variable X plus Y. So, what does, what expectation we get in that case?

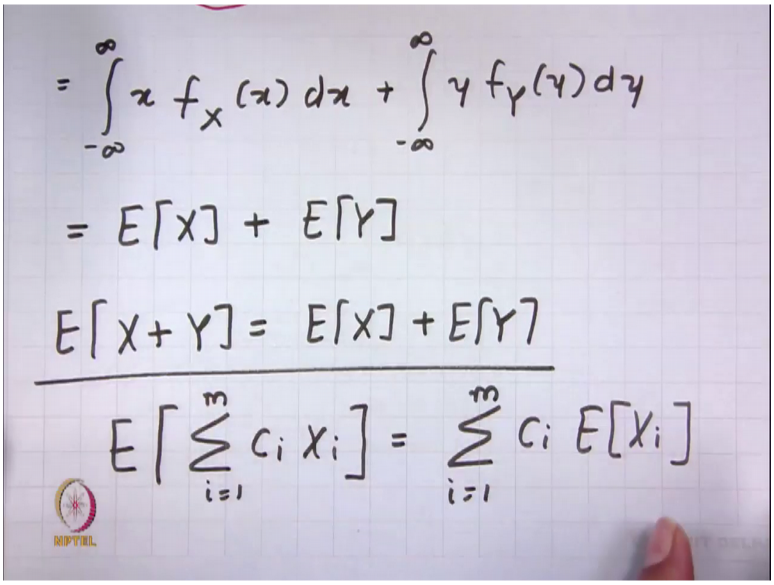
(Refer Slide Time: 45:51)



So, we can say that expectation of X plus Y will be nothing but we have now to substitute the numeric values. So, X takes a numeric value x, Y takes a numeric value y ok. Now, let us try to simplify this ok. Now, if you look at this very carefully, I am just going to split this integration double integration.

And in this course, we will not worry about whether we can change the order of the integration or change the integration or summation or whatever we assume always we can do that, because that is more generic case and we do not worry about nit-pickings. So, you have to believe me that I can change the order of integration without any ambiguity always in this course. I can rewrite this changing the order of integration, and similarly for that part I can write this ok.

(Refer Slide Time: 48:11)

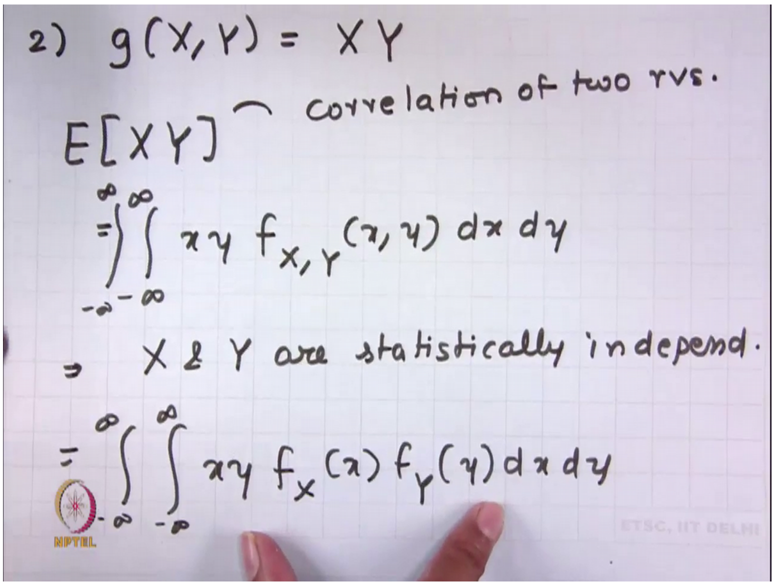


So, using this expression, so let me just use the top one, because these two are exactly similar what we can do is; so what does this quantity integrate to? So, if you look at this quantity, what does this will give to me is it is nothing but it is the marginal probability density function of x times dx. And similarly for the other integration, we would have and this is nothing but it is the expectation of X by definition, and this is nothing but this is the expectation of Y. And this is an important result.

So, what we are saying is the expectation of random variable X plus Y is nothing but it is the expectation of X plus expectation of Y that means, expectation is a linear operation. So, whether we have not said anything about whether these random variables are statistically independent or any other thing, what we have just said this is for any random variable howsoever they are related, we have not made any assumption while evaluating this. So, we have said that expectation of X plus Y, where X and Y can have any distribution they can be statistically independent, they can be statistically dependent whatsoever is nothing but it is the expectation of X plus expectation of random variable Y.

So, in general I can say that expectation of summation of C i X i, where x random variable, and C are some constant is nothing but it is the linear combination of expected value of random variables, this is where we are trying to generalise this result. So, we are saying that the expectation of a linear combination of a random variable is nothing but it is the linear combination of expectation of random variables.

(Refer Slide Time: 50:41)

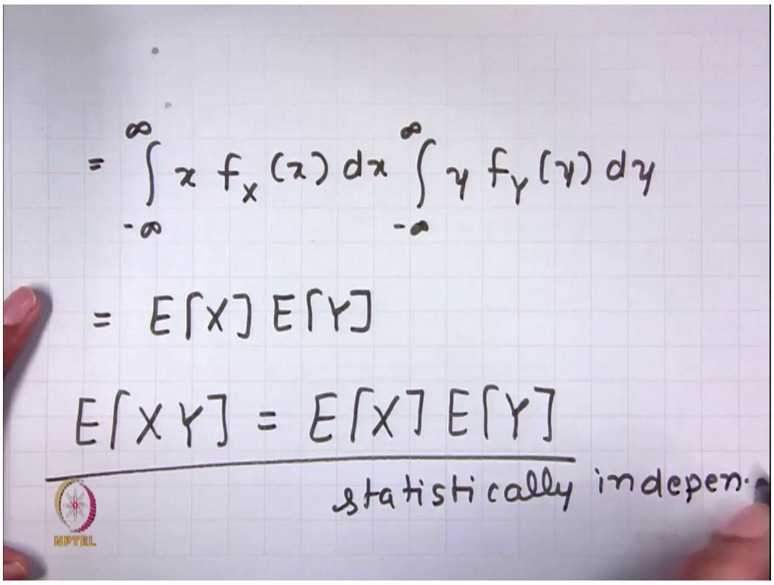


So, let us try to find for another function. So, let us now see what happens if a linear function of X, Y or if it is a function of X, Y is nothing but X Y ok. So, now we are trying to investigate what happens if g X, Y is nothing but it is the product of two random variables. So, essentially what we are evaluating is expected value of X into Y, and this operation this value is known as correlation of two random variables ok. So, this is the correlation operation of two random variables. The correlation of two random variables is nothing but it is given by the expected value of the product of those two random variables.

Now, so by definition this is nothing but and again changing the order of integration, but before doing that I let me assume that these two random variables are statistically independent. So, let me make an assumption that X and Y are statistically independent.

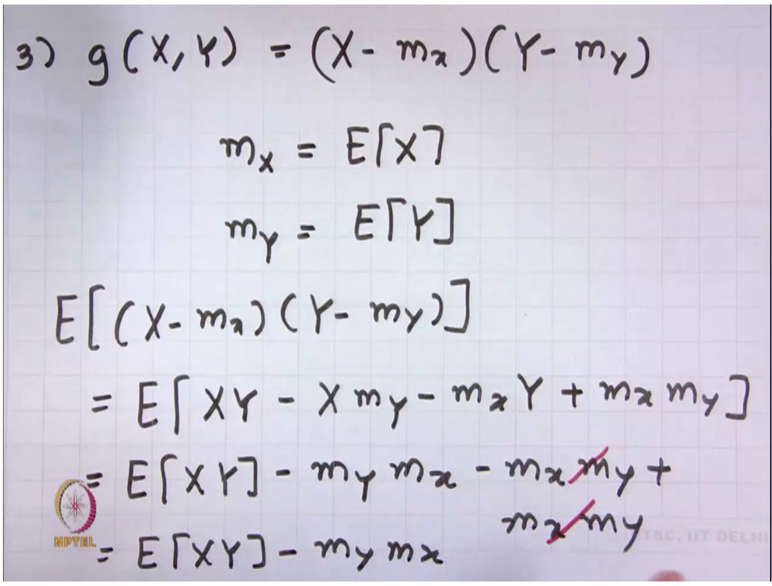
And let us see does something interesting happens if the two random variables are statistically independent. If two random variables are statistically independent, then their join pdf can be replaced by the multiplication of their marginal pdfs. So, in the case when these two random variables are statistically independent, I can simplify this expression. So, I am just replacing the joint pdf in terms of marginal probability density functions.

(Refer Slide Time: 53:03)



I can start with this expression, change the order of integration as I have done before without caring about anything, and I get this. So, what is this expression now? So, as you can see that this is nothing but it is the expected value of X, and this is expected value of Y. So, we have derived an important property that if two random variables are statistically independent, then the expected value of X Y is nothing but expected value of X times expected value of Y. So, this will happen if the two random variables involved are statistically independent.

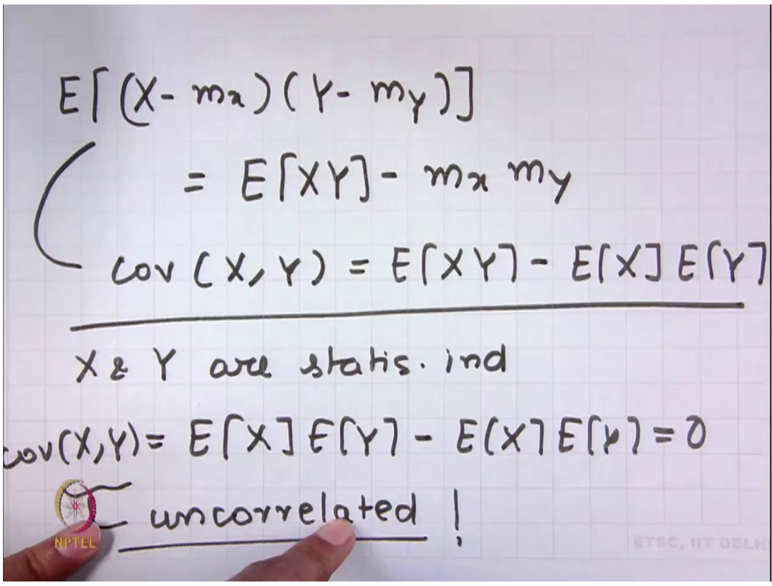
(Refer Slide Time: 54:13)



So, let us do the last concept for today. Now, we have assume the third case in which this function of X and Y, I have assumed this to be X minus m x into Y minus m Y, and let me define my notation here. So, m X is nothing but it is the expected value of X, so it is the average value of random variable X or it is also known as the mean of the random variable X. And m Y is nothing but expected value of Y ok.

So, now we have to evaluate this expression, and because my expectation is a linear operation as I have already said I can write this as ok. And we can pull this expectation operation inside m Y is a constant, so m Y would be pulled out. What we will have is expectation expected value of X, which is m x similarly here you will have m x times m y ok.

(Refer Slide Time: 56:07)



And so this one m x m y cancels with each other. What you would end up is this expression. So, what we are saying is expected value of X minus m x into Y minus m y is nothing but it is the expected value of X Y minus m x times m y ok. This quantity is known as covariance of random variable X and Y ok.

Covariance is very central concept to this course, we will use it a lot more. So, it is very it will be very helpful for you, if you understand this clearly. So, what is the covariance of two random variables, it is nothing but it is the expected value of X minus m x where m x is the mean of that random variable into Y minus m y where, m y is the mean of the random variable Y. And this quantity covariance of the two random variables is nothing but this ok.

Let us just do a specific case in which X and Y are statistically independent, what happens to covariance? If X and Y are statistically independent, then expected value of X Y, we can write this is expected value of X into expected value of Y, so covariance of X and Y is expected value of X into expected value Y minus expected value of X into expected value of Y. So, the covariance becomes 0 right.

If the covariance of the two random variables is 0, we say that these two random variables are uncorrelated that is very important. So, what we are saying is if the two random variables are statistically independent, the covariance becomes 0. And if the covariance becomes 0, then the random variables are referred to as uncorrelated random variables that means, the statistically independent random variables are always uncorrelated.

In the next lecture, we will ask you a question is the reverse also true that means, if the random variables are uncorrelated can we conclude, they are statistically independent. And the answer strangely will be no, if the two random variables are uncorrelated they need not be statistically independent. But, if they are statistically independent, then they are guaranteed to be uncorrelated. So, here we will like to conclude our lecture.

And today we have covered a lot of central concepts that will be very useful, we will continue talking about the covariance, we have just started with is, but there is lot more to say in this covariance, because it will be very useful. But, other than covariance we have also established what is a joint probability density function, we have also established what is a conditional probability density function, we have also looked into how can you take in a joint probability density function, and find the marginal probability density functions.

We have also defined a statistically independent random variables, and we have said if the two random variables are statistically independent, then their conditional pdf of those two random variables is nothing but it is the marginal pdf. So, we conclude here.