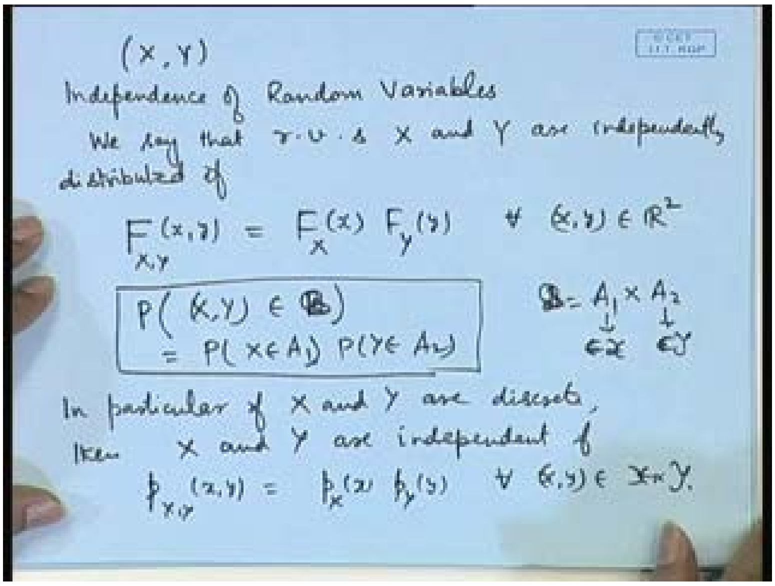
Probability and Statistics

Prof. Dr. Somesh Kumar
Department of Mathematics
Indian Institute of Technology, Kharagpur

Lecture No. #19
Joint Distributions-II

So, yesterday we had introduced the concept of jointly distributed random variables. So, because many times we will be interested in recording the numerical characteristics of several phenomenain one randomexperiment;for example, for a new born child, we will be interested in recording its height, weight and pulse rate at birth.A doctor records several characteristics of the patient who visits him for certain disease;the performance of a student is in a semester during a course ismeasured in terms of it his marks in say homework assignments, hisperformance in mid semester examination, his performance in the end semester examination, etcetera. We have seen that the types of jointly distributed random variables;some of them may be discrete, some of them may be continuous, all maybe discrete, all maybe continuous, some maybe mixed, etcetera.

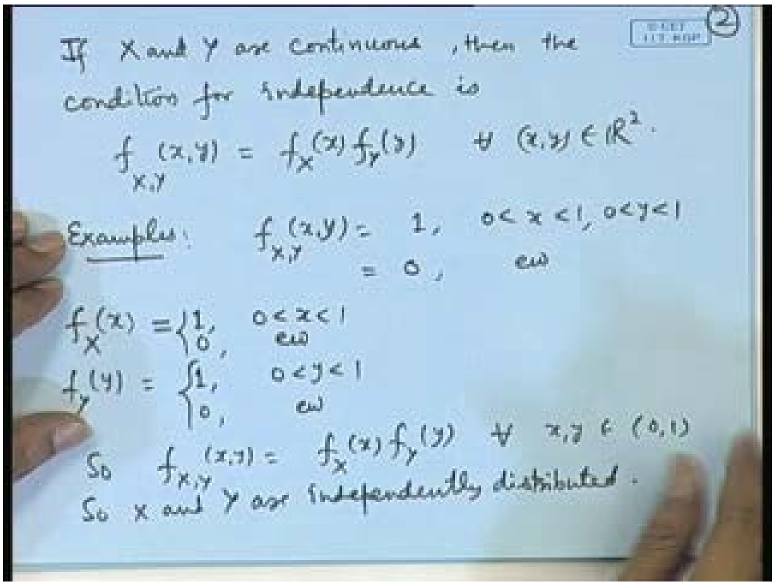
(Refer Slide Time: 01:34)



In particular, when we are consideringa jointly distributed random variable X, Y,we have considered 2 specific cases where both X, Y maybe discrete such as marks of the students or both maybe continuous. For example, height and weight of a child at birth. So, in the case, when both of them are discrete - we have a probability mass function and when both are continuous - we have a probability density function.We have seen the description of this andusing this we can calculate anyprobability statement regarding the joint distribution of Y and Y or marginal distribution of X and Y or conditional distribution of Xand Y.

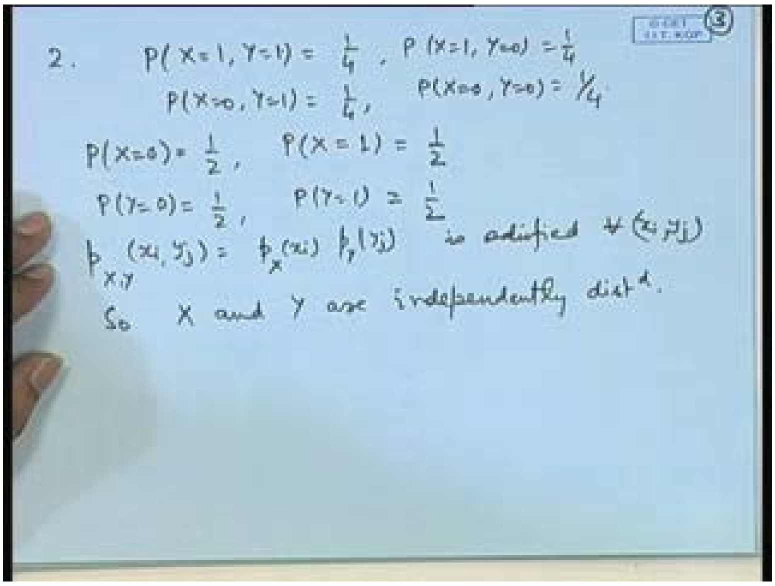
Next we discuss the concept of independence of random variables. (No audio from 02:20 to 02:29) So, we say that random variables X and Y are independently distributed. If the joint cdf is equal to the product of marginal cdfsat all points. So, basically if we do not make use of the cdf, we shouldbe able to say like that. That if we are considering a set in the two-dimensional plane where B can be expressed as product of 2 sets A 1 and A 2 where A 1 is in the space of x values and A 2 is in the space of y values. Then this should be equal to probability of X belonging to A 1and in to probability of Y belonging to A 2. If that happens for all combinations of this type of sets then they are independent. However, this condition is equivalent, if we state in terms of the cumulative distribution function.In particular,if X and Y are discrete then X and Y are independent; if pXY is equal to pX (x) p Y(y)for allxy belonging to x cross y.

(Refer Slide Time: 04:42)



In case of continuous random variable, if X and Y are continuous, then the condition for independence isthe joint pdf is equal to the product of marginal pdfs. Let us take someexamples; let fX,Y be equal to say 1, 0 less than x less than 1, 0 less than y less than 1, and 0 elsewhere. Now, this is a jointly distributed continuous random variable. Let us consider say fX.So, we will be integrating with respect to y from 0 to 1, so, we get 1. And similarly, if we integrate with respect to x from 0 to 1, we will get 1;and of course, 0 elsewhere.So, you can easily see that the product of fX (x) and fY (y)is equal to fX, Y(x,y). This condition is satisfied for all x and y. So, X and Y are independent. So, this may be considered as say arrival timings of aa passenger to the railway station between any time from say7 am to 8 am, and y may denote the arrival time of the train between 7 am to 8 am. So, both maybe independentlydistributed and this could be the 1of the possible distributions.

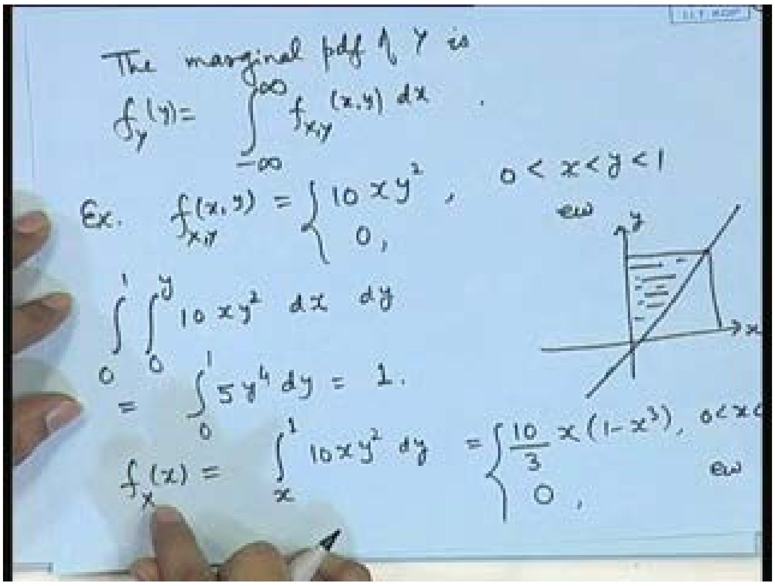
(Refer Slide Time: 07:12)



Let us take another example, suppose probability X equal to 1, Y is equal to 1 is 1 by 4; probability X is equal to 1, Y is equal to 0 is 1 by 4; probability X equal to 0, Y is equal to 1 is say1 by 4; probability X equal to 0 and Y is equal to 0 is say1 by 4.So, the joint distribution of X and Y is given by this. Let us look at the marginal distribution of X. So, what is probability of X equal to 0?It is obtained from summing the probability of X equal to 0, Y equal to 0, and probability X equal to 0 and Y is equal to 1that is equal to half.Probability X equal to 1 in a similar way, it is equal to the sum of these 2 probabilities again it is half,if we look at probability of Y is equal to 0;that is the sum of these 2 probabilities; that is half, and probability of Y is equal to 1; that is the sum of these 2 numbers; that is again half.So, we can see here that the condition that p X, Y(x i, y j) is equal to p X(xi), p Y(y j) is satisfied for all x i, y j in the range of x and y random variables.So, x and y are…

(No audio from 08:40 to 08:49)

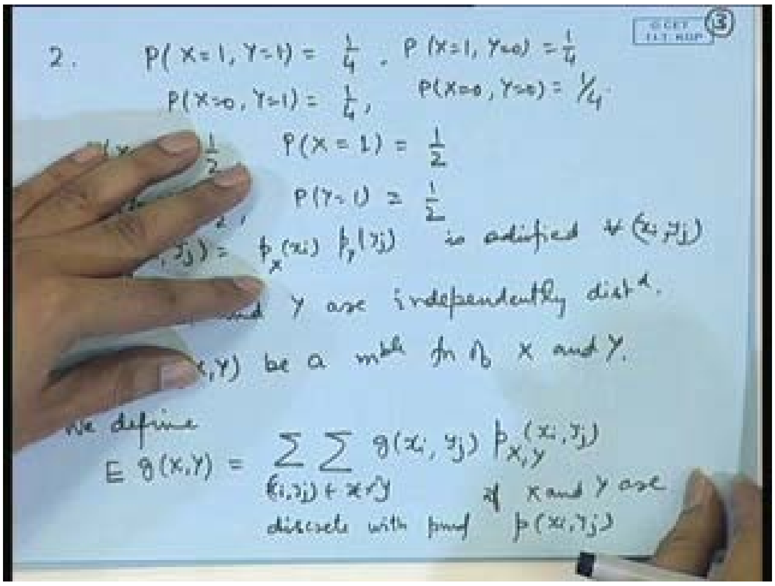
(Refer Slide Time: 09:00)



On the other hand, if we consider a distribution such asf x y is equal to 10 x y square where x and y are satisfying the condition; that 0 is less than x less than y less than 1.Then the marginal distributions of x was 10 by 3 x into 1 minus x cube and the marginal distribution of y is 5 y to the power 4, 0 less than y less than 1. So, you can see that the product of f (x) into f(y) is not equal to f (x, y), because the product here will give you 50 by 3 x into 1 minus x cube into y to the power 4 whereas f (x, y) is 10 x y square. So, the condition is never satisfied. So, x and y are obviously not independent. If we consider another problem,sayf (x,y) is equal to 1 by y for satisfying similar condition that 0 less than x less than y less than 1 then the marginal distribution of x is minus log(x) and the marginal distribution of y is uniform. So, once again you can see that the product of these 2 does not give 1 by y. So, the distributions are not independent.

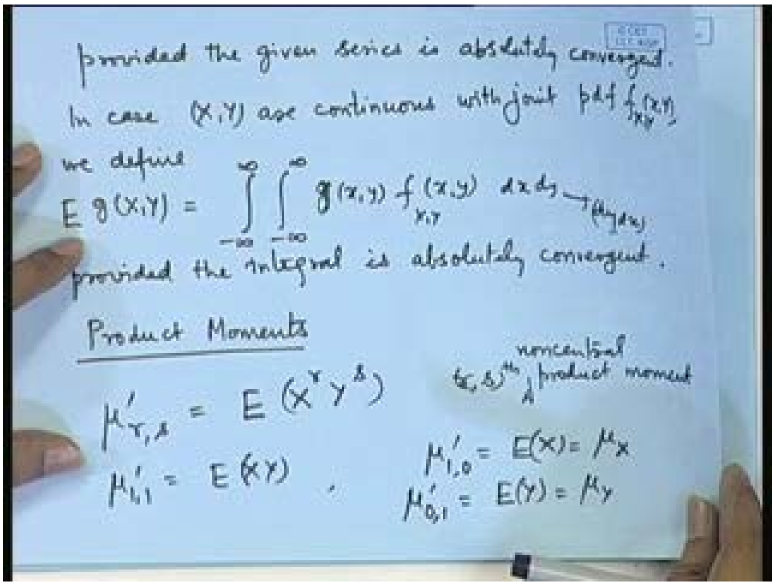
So, the main important role of the condition of independence is that - it helps us to obtain the joint distributions for independently distributed random variables. Many times what happens that; we know these phenomena are independent and we know the individual distributions.Now, in order to study certain characteristics of the joint distribution, we can obtain the joint distribution by simply multiplying the 2 distributions. Where in general, marginal distributions do not give the joint distribution; the joint distributions give the marginal distributions, but theconverse is not true.However if the random variables are independent, then simply by multiplying the marginal distributions, we can get the joint distribution.Therefore, this is a quite useful concept.Now, let us look at the concept of expectation in case of joint distributions.

(Refer Slide Time: 11:20)



So, let g (X, Y) be a measurable function of X and Y. So, we define expectation of g (X, Y) that is equal to if X and Y are discrete, it is equal to g (x i, y j), p X, Y(x i, y j) for all x i and y j in the range of x cross y.This is if x and y are discrete with pmf p (x i, y j). Now, this is valid provided the series is absolutely convergent.

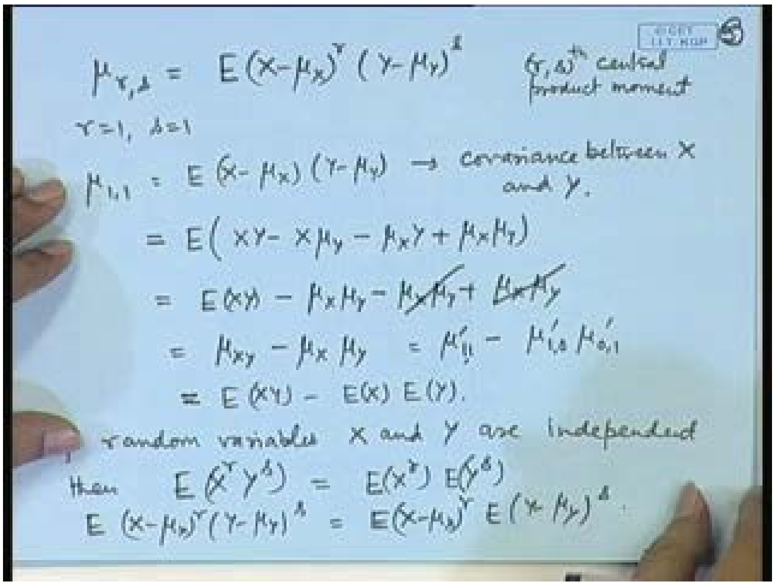
(Refer Slide Time: 12:30)



Provided the given series is absolutely convergent. In case X and Y are continuous with joint pdfsayf X,Y, then we define expectation of g X,Y as double integral g (x, y) f (x, y) dx dy or this could be dy dx also. Provided the integral is absolutely convergent. Here the range of integration is over the range of the density. Whateverbethe appropriate range of the density here. So, in particular, we can consider given random variables we can define an expectation of x plus y,expectation of x minus y, expectation of log of x square y, any type of function of the random variable x and y, we can find out its expected value.

Now, in particular,we will be concerned about product moments. So, we define mu r, s prime as expectation of X to the power r Y to the power s. This is r,s thproduct moment. Now, here this is non-central moment as we had seen in the univariate case also, we can define central and non-central moments. So, in particular, we can consider mu 1,1prime that is equal to expectation of X into Y; if we consider mu 1, 0 prime that is expectation of X, that is the mean of X or the expected value of the random variable X. Similarly, if we consider mu 0,1 prime that is equal to expectation of Y, that is equal to the expectation of random variable Y. Now, using this mu X and mu Y, we can define product moments which are central.

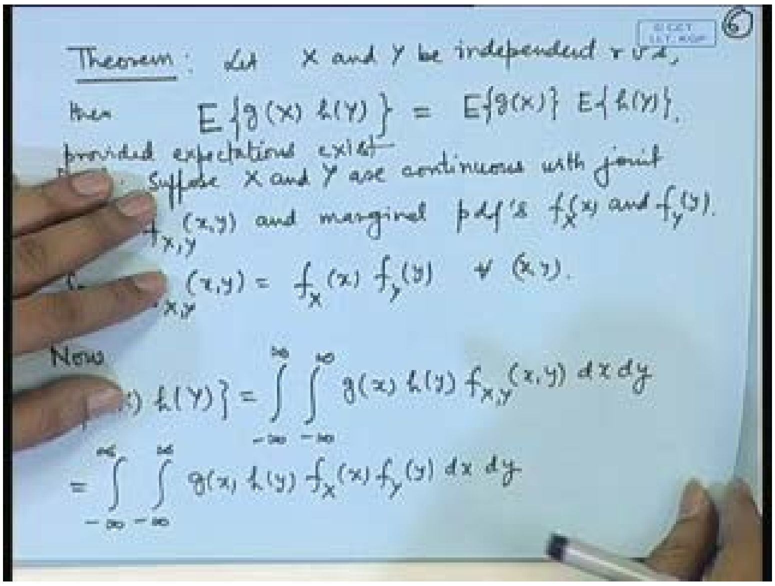
(Refer Slide Time: 15:39)



So, we can talk about mu r, s that is equal expectation of X minus mu X to the power r Y minus mu Y to the power s. Now, if we consider a special case r is equal to 1, s is equal to 1. So, this is in generalr,sth central product moment.So, if we consider both r and s to be 1, that is expectation of X minus mu X into Yminus mu Y; this is called the covariance between X and Y.We can further simplify this, we can write it as expectation of XY minus X into mu Y minus mu XY plus mu X mu Y; that is equal to expectation of XYminus… So, nowhere if wetake expectation, this is mu X mu Y minus mu X mu Y plus mu Xmu Y. So, this term cancels off.And we are getting, we can use this notation mu XY here or mu 1,1 prime minus mu1, 0 prime mu 0,1 prime. We can further see the effect of independence on the moments.

If random variables X and Y are independent, then expectation of X to the power r,Y to the power s, this will become expectation of X to the power r into expectation of Y to the power s.Similarly, expectation of X minus mu X to the power r Y minus mu Y to the power s, that will become expectation of X minus mu X to the power r expectation of Y minus mu Y to the power s. To see this let us prove a general result regarding the independent random variables.

(Refer Slide Time: 18:36)

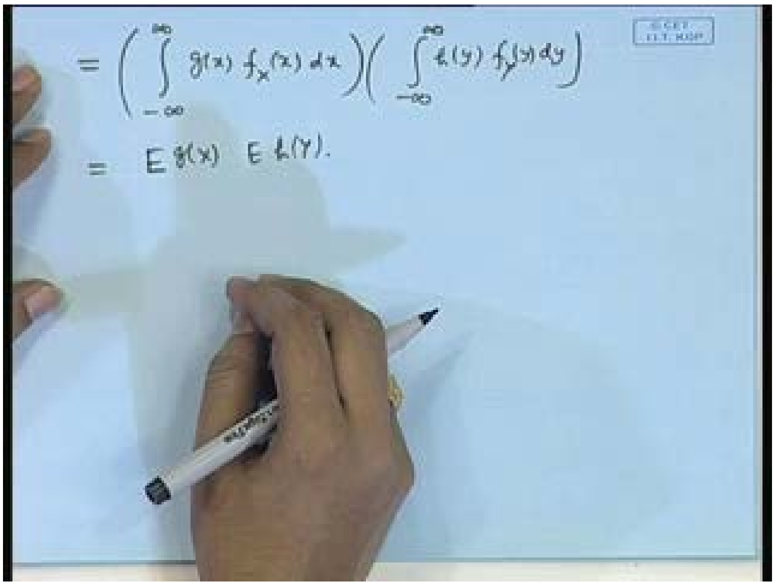


Let X and Ybe independent random variables. Then if we are considering expectation of a product function,then this is equal to product of the expectations. Let us assuming that X and Y are continuous with joint pdff X, Y and marginal pdfsayf X and f Y.So, since X and Y are given to be independent, we have f X, Y is equal to f X into f Y for all x, y. So, nowconsider expectation of g X into h of Y. So, of course, these statements are valid provided these expectations exist.

(No audio from 20:19 to 20:28)

So, let us consider expectation of g X into h Y. So, by definition of the expectation it is equal to g (x) into h (y)f X, Y (x,y)dx dy or dydx. Since the random variables are independent we can make use of the condition of independence. That means, this is equal tog x into hy f x into f ydxdy.

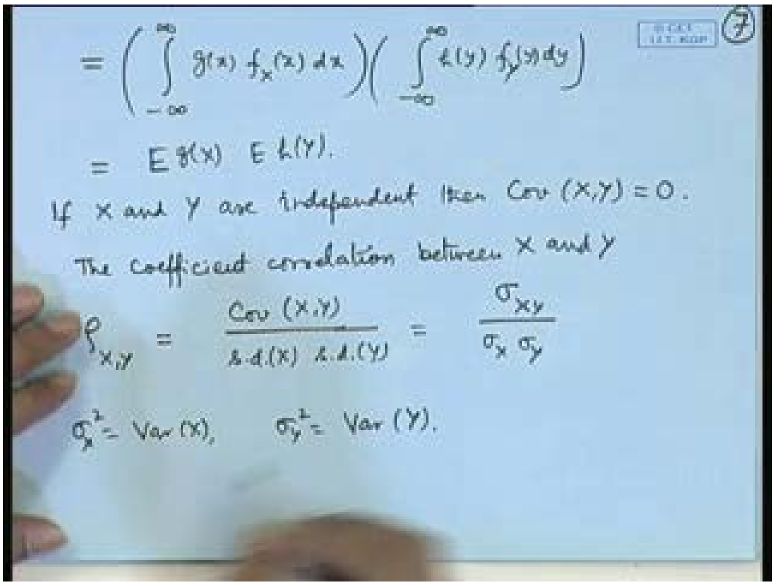
(Refer Slide Time: 21:16)



Now, this is nothing but the product of the integrals g (x)f (x)dx and h (y)f (y)dy which is nothing but expectation of g x into expectation of h y. So, the expectation of a product is equal to the product of expectations in case of independent random variables.

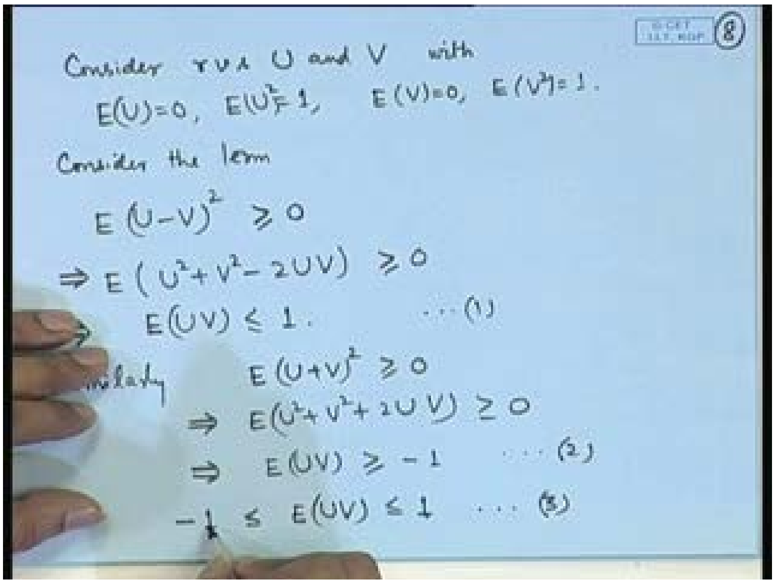
A similar proof can be given incase the random variables are discrete, because here the integrals can be replaced by thesummations and thedensity function can be replaced by the mass function.The proof is little bit involved in the case of mixed random variables and whenoneof them maybe discrete or continuous etcetera.So, in particular if g (X) is X to the power r and h (Y) is Y to the power s. So, expectation of the product moment is equal to the product of the individual moments, and this statement is valid for non-central as well as central momentsNow, if we make use of this condition on the first central moment - first central product moment which we call as a covariance, then expectation of X into Y will be equal to expectation X into expectationY, and therefore, covariance termwill become 0.

(Refer Slide Time: 23:17)



So, we have the consequence of this above theorem. That if X and Y are independent, then covariance of X and Yis 0. Using the covariance,onedefines the coefficient of correlation between X and Y.We denote rhox, y that is defined as covariance between X, Y divided by standard deviation of X into standard deviation of Y. We can use a notation sigma XYdivided by sigma X sigma Y where sigma X squarewas a variance of X and sigma Y square denotes the variance of Y. Now, the question arises that what this correlation coefficient between X and Y represent.So, we claim that this gives a measure of linear relationship between the random variables X and Y.

(Refer Slide Time: 24:47)

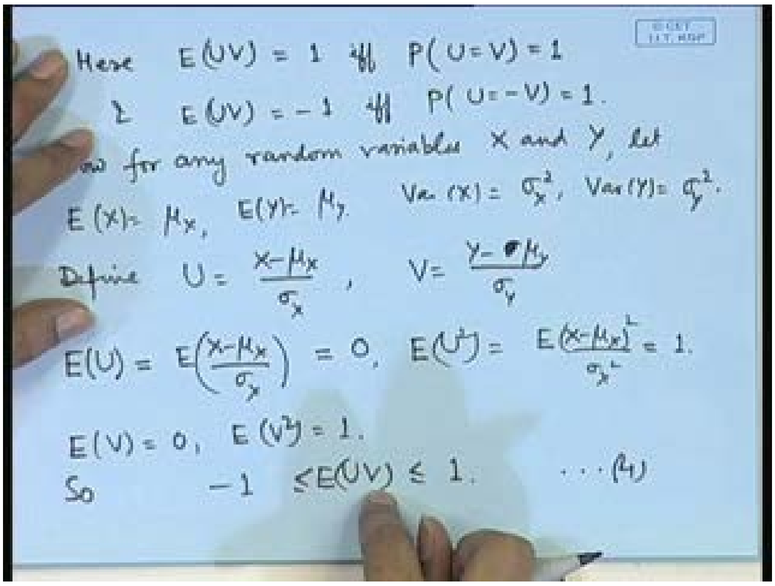


So, in particular, let us consider say consider say random variables U and Vwith say expectation of U is 0, expectation of U square is equal to 1, expectation of V is equal to 0, expectation of V square is equal to 1. If we consider these 2 random variables,consider the term expectation of U minus V whole square. Now, naturally this is greater than or equal to 0 being the average value of a non-negative term.Now, this will imply expectation of U square plus V square minus 2UV is greater than or equal to 0. Substituting the value of expectation U squareand expectation of V square as 1, this relationship is reducing to expectation of UV is less than or equal to 1.

Similarly, expectation of U plus V whole square is greater than or equal to 0. This yields expectation of U square plusV square plus 2UV is greater than or equal to 0. Once again substituting the values of expectation U square and expectationV square as 1, we get expectation of UV greater than or equal to minus 1.So, we have got that expectation of U intoV lies between plus 1 and minus 1 providedexpectation of U and expectation ofVis 0, and expectation of U square and expectation ofV square is 1.

Now, we can consider when the equality will be attained. So, here equality at 1 will be attained when expectation of U minusV whole square is equal to 0. Now expectation of a non-negative random variable is 0 if and only if the random variable itself is 0. That means U must be equal toVwith probability 1.In a similar way, the equality at minus 1 will be attained, if U is equal to minusV with probability 1.

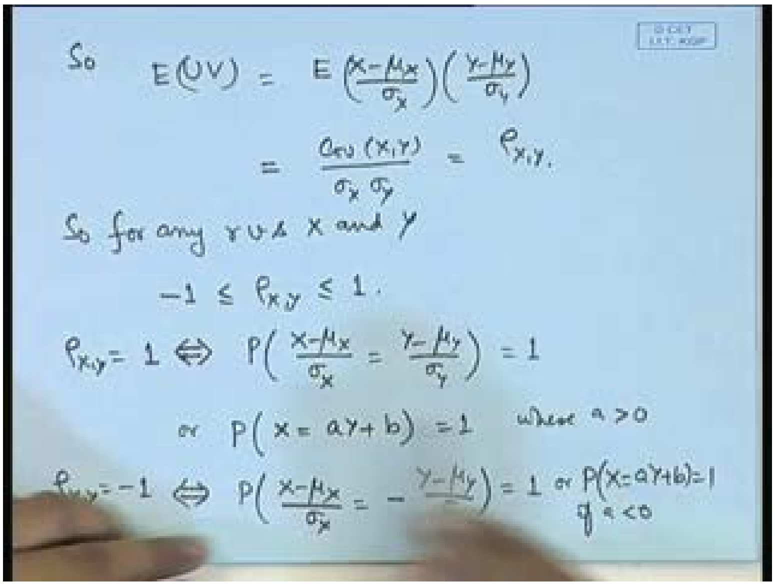
(Refer Slide Time: 27:58)



Let us consider the expressions for this. Here expectation ofUV is equal to 1 if and only if probability of U is equal toV is equal to 1, and expectation of U intoV is equal to minus 1.So, now let us take for any random variables X and Y, let us use a notation that expectation of X is equal to mu X, expectation of Y is equal to mu Y, variance of X is equal to say sigma X square and variance of Y is equal to sigma Ysquare;define U is equal to X minus mu X by sigma X, V is equal to Y minus mu Y by sigma Y. So, if we takes expectation of U,this is equal to expectation of X minus mu X by sigma X, and by the linearity property of expectation, it is expectation X minus mu X by sigma X, that is simply 0.

If we consider expectation of U squarethat is equal to expectation of X minus mu X square by sigma X square. Now, the numerator is simply variance of X that is sigma X square. So, it is 1. So, in a similar way, you cansee that expectation ofV is 0 and expectation ofV squareis 1. So, if we make use of the equality that we have proved here for U, V random variables with the property that expectations are 0 and the expectation of the squares are 1. So, we get expectation of UV between minus 1 to 1.Now, for any random variables X and Y, when U andV are defined like this; what is expectation of UV representing?

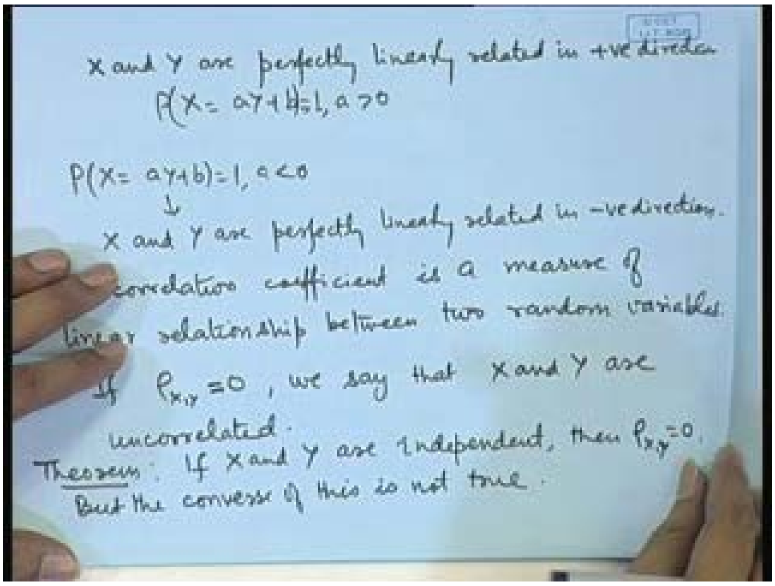
(Refer Slide Time: 30:43)



So, expectation of UV is equal to expectation of X minus mu X by sigma X into Y minus mu Y by sigma Y. So, the numerator here is simply the covariance termbetween X, Y divided by the standard deviations of the X and Y. That is the correlation coefficient between. So, for any random variables X and Y - the correlation coefficient lies between plus 1 and minus 1; the least value is minus 1 and themaximum value is plus 1.

Now, rhoX, Y is equal to 1. So, we look at the conditions for attaining the equality, expectation of UV was 1, if and only if probability of U is equal toV is 1. So, this will be satisfied if and only if probability that X minus mu X by sigma X is equal toY minus mu Y by sigma Y is equal to 1 or you can say probability that X is a linear function of Y.That is x is equal to sum a times Y plus b where a is a positive number, because sigma X by sigma Y.Similarly, rhoX, Y is equal to minus 1 if and only if X minus mu Xby sigma X is equal to minus Y minus mu Y by sigma Y is equal to 1 or probability that X is equal to aY plus b is equal to 1 if a is negative. So, this condition that X is equal to aY plus b where a is apositive; this is known as that X and Y.

(Refer Slide Time: 33:04)

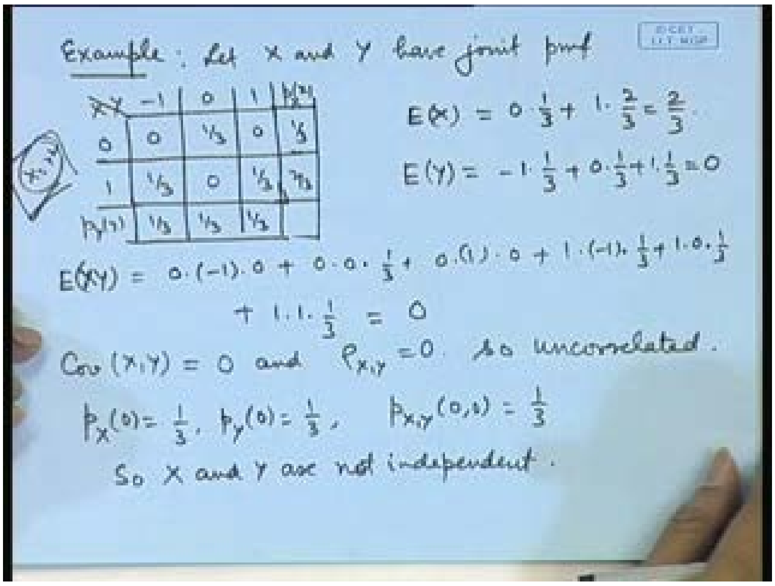


So, we can write it that X and Y are perfectly linearly related in positive direction. That is X is equal to aY plus b for a positive and probability of this statement is 1.And if we say probability of X equal to aY plus b is equal to 1 where a is negative, then we say that X and Y are perfectly linearly related in negative direction.Now, this gives an interpretation for the coefficientof correlation. So, we can see that - in general coefficient of correlation lies between minus 1 to plus 1. The bounds minus 1 and 1 are attained. So, minus 1 is attained when there is a perfect linear relationship in a negative direction or you can say perfectly negatively linearly related, and the equality at 1 is attained when it is perfectly linearly related in the positive direction or perfectly positively linearly related.

So, in general any value between minus 1 to 1 denotes the degree of the linear relationship between random variables X and Y. Suppose I say the correlation coefficient is equal to 0.7 that shows that there is a good positive correlation between X and Y, and here the relationship is of the linear type. If we say rho X,Y is equal tominus 0.3; it shows that there is a lower degree of negative linear relationship between the random variables X and Y.When rhoX, Y is equal to 0, we say that the random variables X and Y are uncorrelated; now, here uncorrelated means that the linear relationship is not existent between random variables X and Y. So, we can interpret these statementsas... So, correlation coefficient is a measure of linear relationship between 2 random variables.If rhoX, Y is 0, we say that X and Y are uncorrelated. At this point, it is important to understand the difference between uncorrelatedness and independence. If we say that the random variables are uncorrelated, it does not mean that they are independent.Of course if X and Y are independent, it will imply uncorrelatedness, becauseif X and Y are independent then covariance term is 0,thereforecorrelation term will also be 0. So, we have the following result, if X and Y are independentthen rhoX,Y is 0, but the converse of this is not true. So, proof is of course obvious that – if X and Y are independent then covariance term is 0and therefore, the correlation term is 0.Let us look at the converse of it.

(No audio from 37:34 to 37:42)

(Refer Slide Time: 37:35)

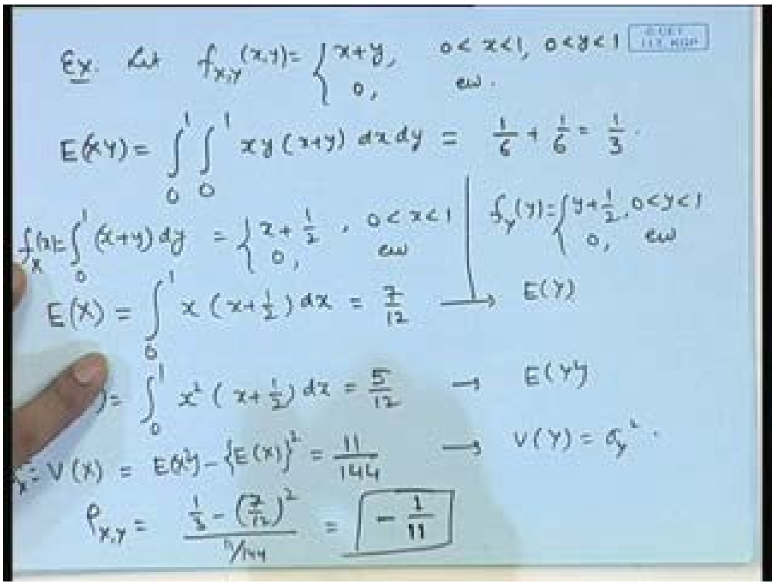


Let X and Y have joint pmf asX is taking values 0,1,Y is taking values minus 1, 0and 1.The probabilities are 0, 1 by 3, 0, 1 by 3, 0, 1 by 3. So, here we consider say expectation of X. So, here you can calculate the marginal distributions by adding the rows and columns. So, the marginal distribution ofX is 1 by 3, 2 by 3, the marginal distribution of Y is obtained as 1 by 3,1 by 3,1 by 3. So, expectation of X is equal to 0 into 1 by 3 plus 1 into 2 by 3;that is equal to 2 by 3. If we look at expectation of Y that is equal to minus 1 into 1 by 3 plus 0 into 1 by 3 plus 1 into 1 by 3; that is equal to 0. If we look at expectation of X into Y; so, we look at all the possibilities of theX and Y values here. So, X is 0,Y is minus 1 with probability 0 plus X is 0,Y is 0 with probability 1 by 3,X is 0,Y is 1 with probability 0,X is 1,Y is minus 1 with probability 1 by 3,X is 1,Y is

0 with probability 1 by 3 plus X is 1,Y is 1 with probability1 by 3.You can see these terms vanish, we are getting minus 1 by 3 and plus 1 by 3. So, it is 0. Therefore, covariance term is 0 and consequently rhoX, Y is 0.But we can see here that the product of the marginal distributions,for example, pX, 0 and p Y, 0 both are 1 by 3, but pX, Y. So, let us write here, pX, 0 is 1 by 3, pY, 0 is 1 by 3, but if we consider pX,Y 0, 0; that is also 1 by 3. So, X and Y are not independent. So, they are uncorrelated, but not independent.

Now, this further brings out the contrast between the concept of independence and correlatedness.So, independence simply means then that the random variables or we can say the observance of the 2 phenomena has nothing to do with each other; that isonephenomena which yieldsthe random variable X and the phenomena which yields the random variable Y, they are totally independent; whereas the correlation gives a degree of linear relationship between the random variables. So, if they do not have a linear relationship, the random variables may become uncorrelated. But that does not mean that they are independent.For example, even in this problem, it may happen that X is actually Y square,because the probability that X equal to 0 is same as probability Y is equal to 0, and probability X equal to 1 is samesum of probabilityY is equal to minus 1 and probability Y is equal to 1.So, this could be a non-linear relationship.So, let us take up a few examples for calculation of the covariance and correlation term.

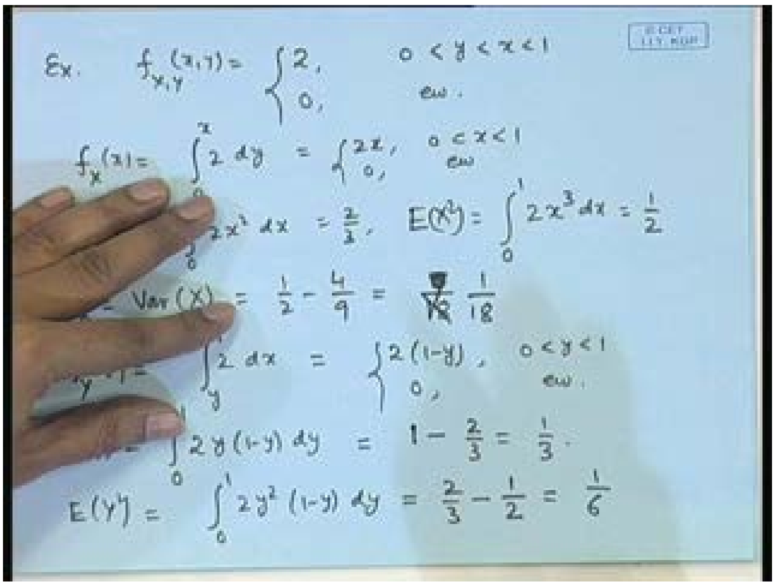
(Refer Slide Time: 42:28)



So, let fX, Y be equal to x plus y for 0 less than x less than 1, 0 less than y less than 1, and 0 elsewhere.In order to calculate the coefficient of correlation, we need the first and second moments of x and y and also the first product moment of the joint distribution. So, expectation of X into Y that is equal to double integral xy into x plus y dx dy, 0 to 1.So, this is simply integral of x square y plus x y square which we can easily evaluate and it is equal to 1 by 6 plus 1 by 6 that is equal to 1 by 3. In order to calculate the moments of x and y separately, we can make use of the marginal distributions. So, you can see here this will be equal to x plus half for 0 less than x less than 1,and similarly the marginal distributions of Y will become y plushalf for 0 less than y less than 1, and 0 elsewhere.

The distributions of X and Y are same. So, it is enough if we calculate the moments for 1of them. So, expectation of X becomes the integral of x into x plus half from 0 to 1 which is obviously7 by 12.So, the same value will be expectation of Y; and expectation of X square likewise can be calculated; that is equal to 5 by 12 same as expectation Y square, and therefore, variance of X that is equal to expectation of X square minus expectation of X whole square that is equal to 11 by 144;that is sigma X square and sigma Y square.So, correlation coefficient is then equal to 1 by 3 minus 7 by 12 square divided by 11 by 144 which after simplification is minus 1 by 11. So, this means that there is a low degree of negative linear relationship between the random variables X and Y. We can take up some examples which we read yesterday.

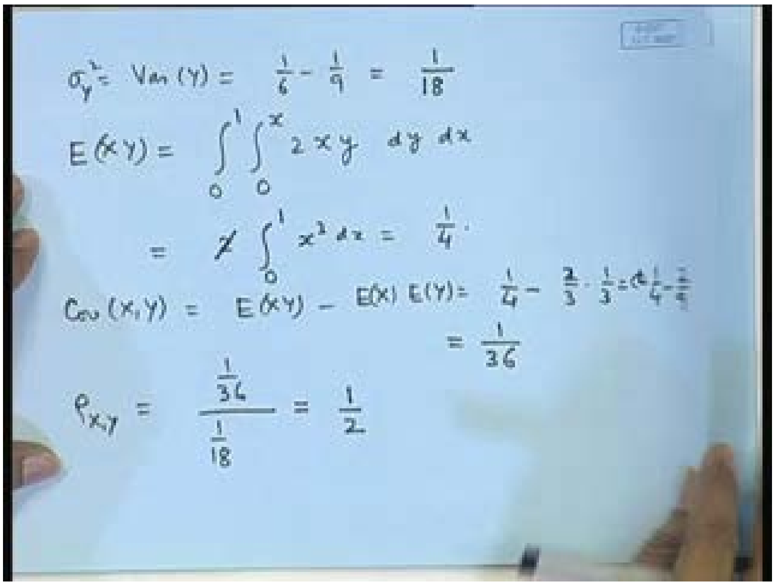
(Refer Slide Time: 45:44)



So, let us consider say fX, Y is equal to say2, 0 less than y less than x less than 1, 0 elsewhere.Here, suppose we want the distribution of X then it is equal to integral with respect to y,2dy from 0 to x that gives 2x.So, expectation of X is equal to 2x square dx from 0 to 1; that is 2 by 3. Expectation of X square is equal to integral 2x cube dx, 0 to 1; that is equal to half.Therefore,sigma X square that is variance of X is equal to half minus 4 by 9; that is equal to 5 by 18.Iam sorry this is not 5 by 18, it is 1 by 18. So, we are able to calculate the mean and variance of the distributions of X.

Similarly, let us calculate the marginal distribution of Y that is integral 2 dx from y to 1; that is equal to 2 times 1 minus y for 0 less than y less than 1, 0 elsewhere.So, expectation of Y is equal to integral 2y into 1 minus y dy from 0 to 1; that is equal to twice y. So, that is y square that is 1 minus 2 y square that is 2 by 3 that is equal to 1 by 3. Expectation of Y square will be equal to 2y square 1 minus y dy integral from 0 to 1. So, once again 2y square;so, y square is integrated toy cube by 3. So, this 2by 3 minus 2y cube, y cube is integrated to y to the power 4 by 4. So, that is 2 by 4; that is half.So, it is equal to1 minus4 minus 3 by 6; that is 1 by 6.So, wecan calculate the variance of Y that is equal to 1 by 6 minus 1 by 9. So, that is equal to 1 by 18.

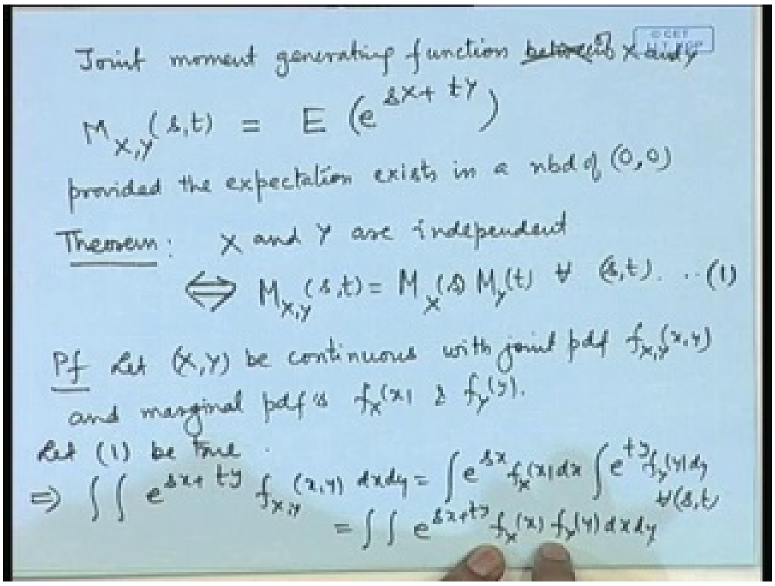
(Refer Slide Time: 48:44)



Now, we have calculated the means and variances of the random variables X and Y. In order to get the correlation coefficient, we need the product moment also. So, expectation of X into Y that is equal to 2xy. Now, here we can choose the order of integration,we may do dy dx. The range of y is from 0 to x, and the range of x is from 0 to 1. So, this is equal to 2 times integral 0 to 1.

Now, firstlywe are integrating with respect to y. So, we get y square by 2. So, this termcancels out and that gives us x square. So, we are left with x cube dx that is equal to 1 by 4. So, covariance term between X and Y that is equal to expectation of XY minus expectation X into expectation of Y; that is equal to 1by 4 minus 2 by 3 into 1 by 3; that is equal to 1 by 4 minus 2 by 9.So, once again it is equal to 1by... So, 9 minus 8 by 36. So, 1 by 36. So, the coefficient of correlation is equal to 1 by 36 divided by 1 by 18 both variances of X and Y are 1 by 18. So, sigma X into sigma Y will be 1 by 18. So, it is equal to half.So, this shows that there is a moderate degree of relationship, moderate degree of positive linear relationship between the random variables X and Y. We also introduce the concept of the joint moment generating function between the random variables X and Y as follows.

(Refer Slide Time: 51:35)

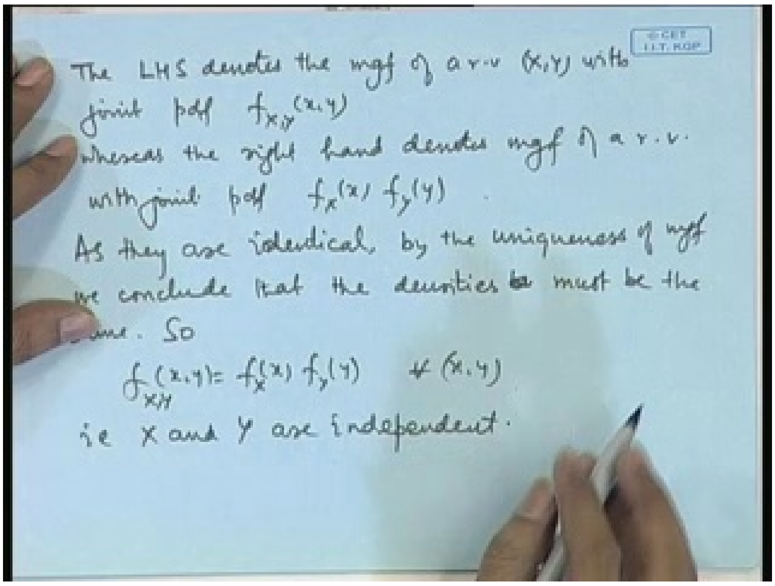


Joint moment generating function between X and ofX and Y; so, we use a notation M X,Yat the point s,t. It is expectation of e to the power sx plus ty, provided the expectation exists in a neighborhood of(0, 0). At(0, 0), this always exist and is equal to 1. So, in a neighborhood of(0, 0) this should exist. Now, the nature of this term is nice, it suggests certain properties.For example, it is equal to expectation of e to the power sx into e to the power ty. So, the first thing that we observe that if the random variables are independent, it will become product of the expectations of e to the power sx and e to the power ty which are nothing but the individual moment generating functions of x and y. However, because of the uniqueness property of themgf’s, we have a stronger result here.

X and Y are independent; this implies and implied by that M X, Y s,t is equal to M Xs M Yt for all st.In order to prove this, let us consider the case of continuous random variable with jointpdfsay fX, Y andmarginal pdf’ssay fX and fY. Now, here let us notice that - if the random variables are independent, then this expectation of a product will be equal to the product of the expectations. So, the jointmgf will be equal to the product of the individual or marginalmgf’s.

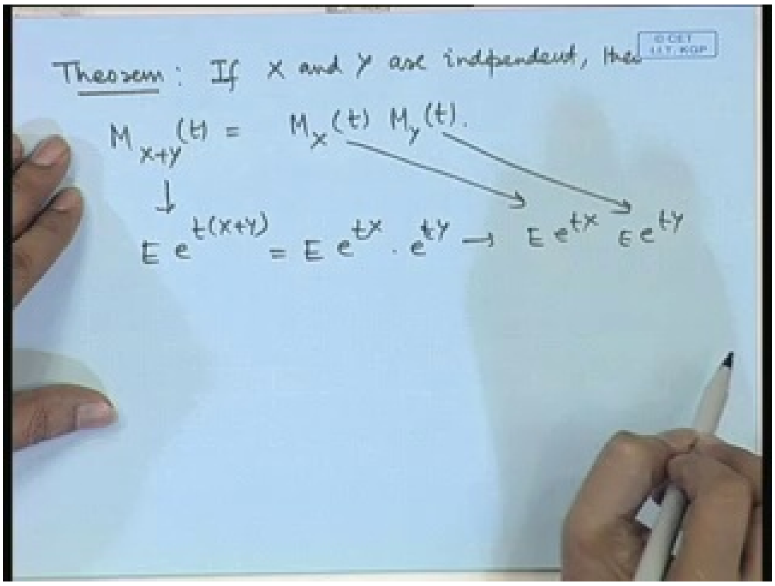
Let us look at the converse. Let this(1) relation be true, then this implies that integral e to the power sx plus ty fX x,y dx dy is equal to integral e to the powersx fx dx into integral e to the power ty fy dy for all st.Now, the right hand side we can write as product of thetwointegrals can be written as a combine integral e to the power sx plus ty fx into fy dx dy. Now, note here the left hand side denotes the jointmgf of random variables X and Y then the pdf is fX,Y. The right hand side denotes the joint distribution - the jointmgf of the random variables X and Y then their joint distribution is given by fX into fY. And this statement is true for all st. So, by the uniqueness of themgf, the twodensities must be same.

(Refer Slide Time: 56:08)



So, the left hand side denotes the moment generating function of a random variable X,Y with joint pdf fX,Y(x,y) whereas the right hand side denotes mgf of a random variable with joint pdffX into fY.As they are identical by the uniqueness of mgfwe conclude that the distributions must be the same. So, we should have fX,Y(x,y) is equal to fX into fY for all xy. That is X and Yare independent.So, this is quite a strong property and that is true, because of the uniqueness property of the moment generating functions. So, independence of random variables can also be proved through the consideration of the joint mgf.

(Refer Slide Time: 58:05)



We end this onewith another result, if X and Y are independent, then M X plus Y that is the moment generating function of random variable X plus Y; it is equal to the product of MX and M Y.The proof of this is almost trivial. The left hand side is expectation of e to the power t x plus y and so if independence is there, this can be written as tx into e to the power ty which will become expectation of e to the power tx into expectation of e to the power ty which will be this term and this term respectively.Thisresult is extremely useful in obtaining the distributions of sums of various random variables, because if we know the mgf's of those random variables, then we can identify the distribution of the sum by identification of the mgf of the sum as the mgf of certain random variable. So, we will consider the bivariate normal distribution in the next lecture.