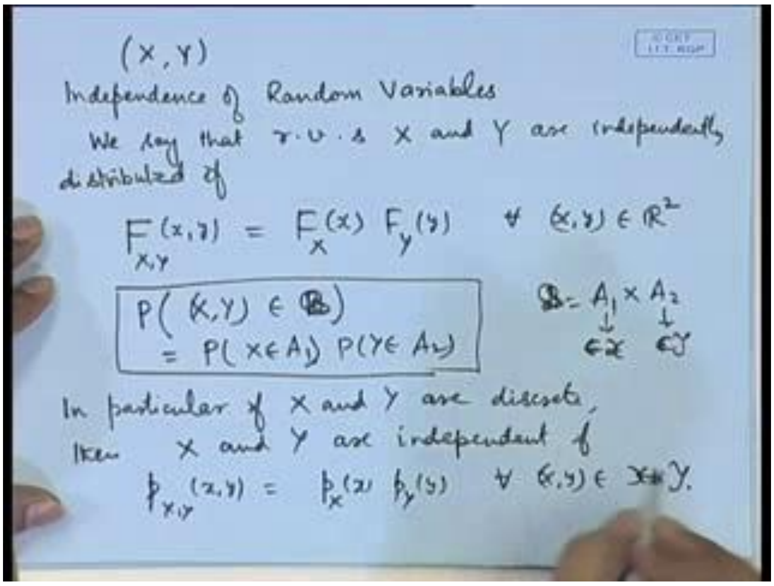
Probability and Statistics
Prof. Somesh Kumar
Department of Mathematics
Indian Institute of Technology, Kharagpur

Lecture – 37
Independence, Product Moment

Yesterday we are introduced the concept of jointly distributed random variables, because many times we will be interested in recording the numerical characteristics of several phenomena in one random experiment. For example for a newborn 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 in a semester during a course is measured in terms of his marks in say homework assignments, his performance 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 may be discrete, all may be continuous, some may be mixed etcetera.

(Refer Slide Time: 01:36)



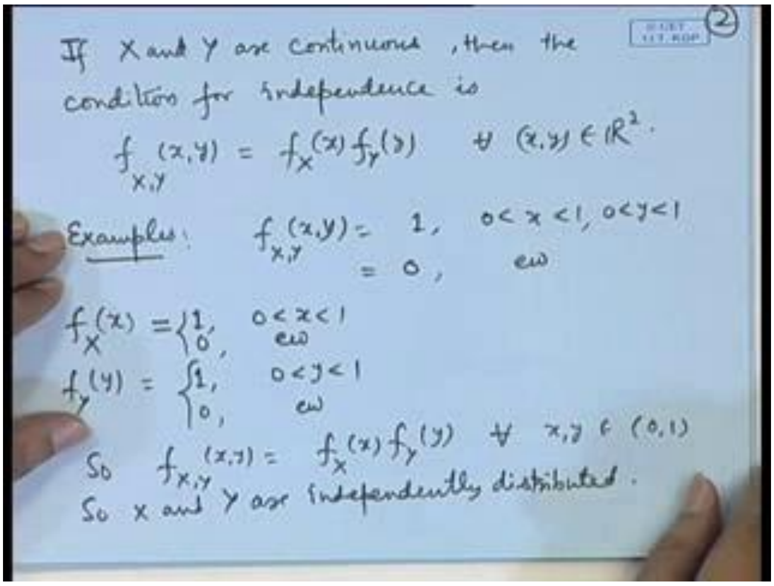
In particular when we are considering a jointly distributed random variable X, Y; we have considered two specific cases where both X, Y may be discrete such as the marks of the students or both may be continuous for example, height and weight of a child at birth.

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 and using this we can calculate any probability statement regarding the joint distribution of X and Y or marginal distribution of X and Y or conditional distribution of X and Y.

Next we discuss the concept of independence of random variables. So, we say that random variables X and Y are independently distributed, if the joint CDF is equal to the product of marginal CDFs at all points. So, basically if we do not make use of the CDF; we should be 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 two sets A 1 and A 2, where A 1 is in a 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 1 and into 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 p, x, y is equal to p X x; p Y y for all x y belonging to x cross y.

(Refer Slide Time: 04:42)

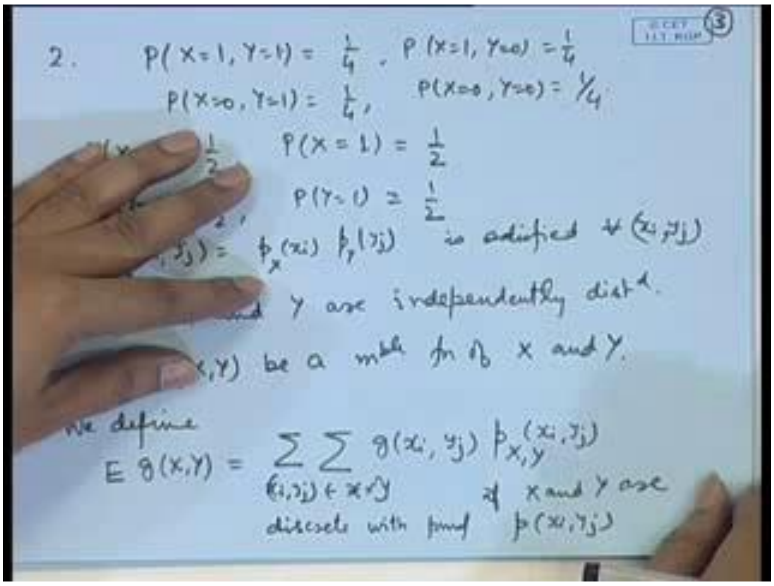


In case of continuous random variable, if X and Y are continuous, then the condition for independence is the joint pdf is equal to the product of marginal pdf’s.

Let us take some examples let f x 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 f x. So, we will be integrating with respect to y from 0 to 1, so we get 1. And similarly if we integrated this with 2 x from 0 to 1, we will get 1 and of course, 0 elsewhere.

So, you can easily see that the product of f X x and f Y y is equal to f X 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 a passenger to the railway station between anytime from say 7 am to 8 am and y may denote the arrival time of the train between 7 am to 8 am. So, both may be independently distributed and this could be the one of the possible distributions.

(Refer Slide Time: 07:13)

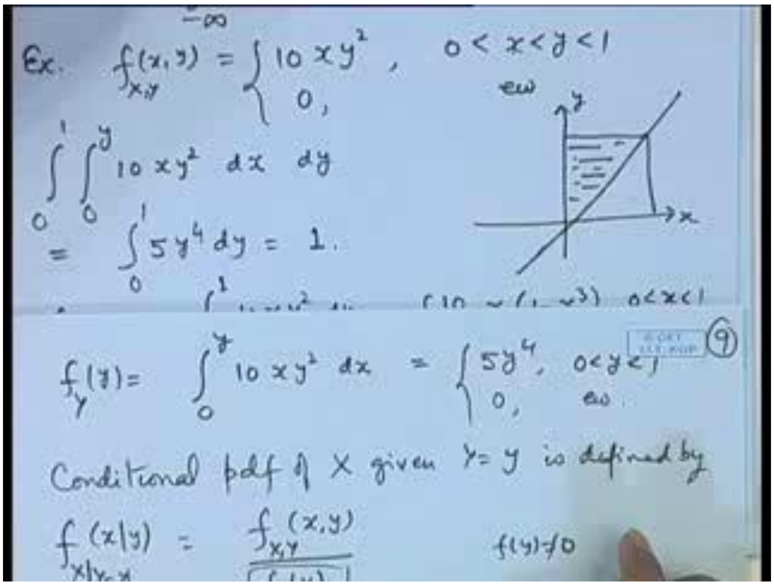


Let us take another example, suppose probability X is equal to 1; Y is equal to 1 is 1 by 4; probability X is equal to 1; Y is equal to 0 is 1 by 4, probability X equal to 0, Y is equal to 1 say 1 by 4, probability x equal to 0 and y is equal to 0 is say one by 4. So, the joint distribution of X and 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, Y is equal to 1; that is equal to half. Probability X equal to 1 in a similar way, it is equal to the sum of these two probabilities again it is half. If you look at probability of Y is equal to 0 that is the sum of these two probabilities that is half and probability of Y is equal to 1 that is the sum of these two numbers; that is again half.

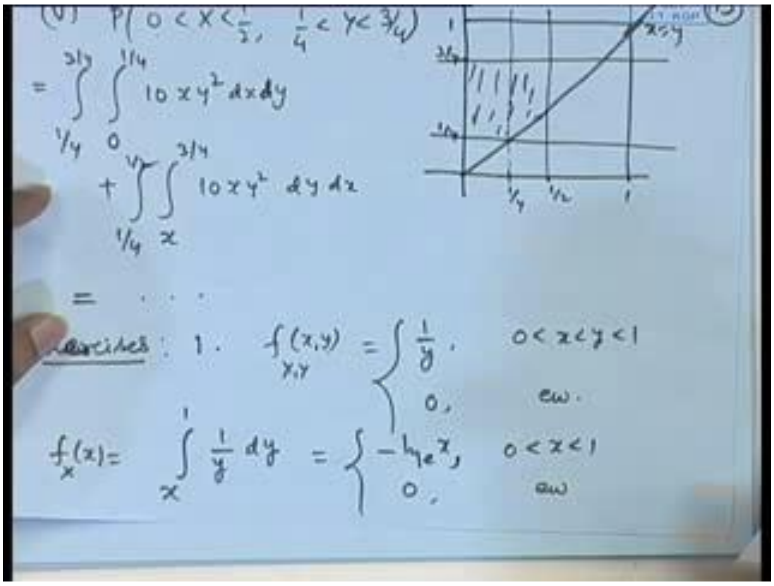
So, we can see here that the condition that p x y, x i; y j is equal to p X; x i, p Y; y j is satisfied for all x i, y j in the range of x and y random variables.

(Refer Slide Time: 09:01)



So, x and y are; on the other hand if we consider a distribution such as f 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 of 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.

(Refer Slide Time: 10:05)

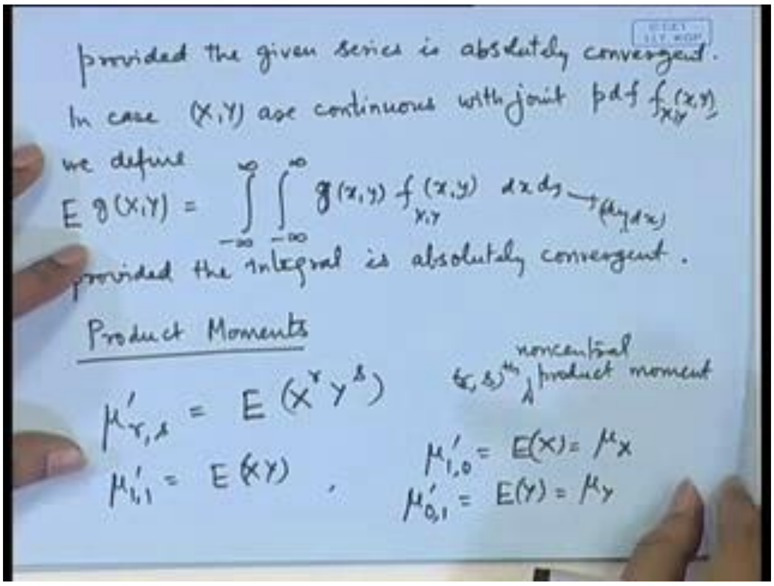


If we consider another problem; say f x y is equal to 1 by y for satisfying similar condition that 0 less than x less than y less than 1 then the marginal distribution of x is minus log of x and the marginal distribution of y is uniform. So, once again you can see that the product of these two does not 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 that 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 two distributions, where in general marginal distributions do not give the joint distribution. The joint distributions give the marginal distributions, but the converse 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 quite useful concept.

Now let us look at the concept of expectation in case of joint distributions. 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: 12:28)

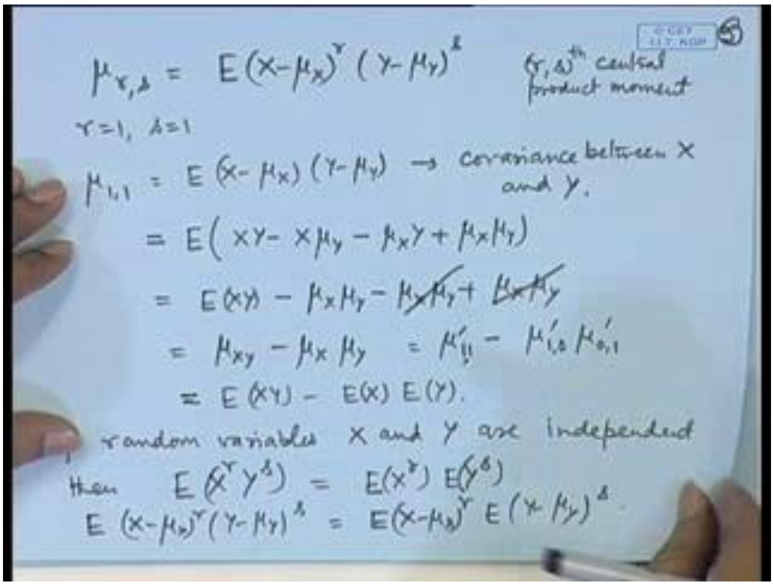


Provided the given series is absolutely convergent; in case X and Y are continuous with joint pdf say f x y, then we define expectation of g; X Y as double integral g; x y, f x y; d x d y or this could be d y, d x also; provided the integral is absolutely convergent; here the range of integration is over the range of the density whatever be the appropriate range of the density here. So, in particular we can consider given random variables x y we can define 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 sth product moment. Now here this is non centre 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, 1 prime 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 products moments which are central. So, we can talk about mu r s that is equal to expectation of X minus mu X to the power r; Y minus mu Y to the power s.

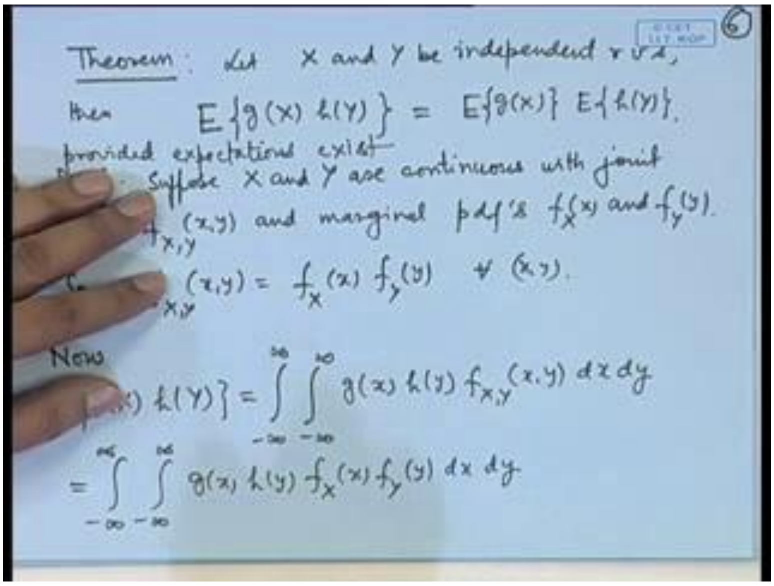
(Refer Slide Time: 15:40)



Now, if we consider the special case r is equal to 1, s is equal to 1. So, this is in general r sth central product moment. So, if we consider both r and s to be 1, that is expectation of X minus mu X into Y minus mu Y; this is called the covariance between X and Y. We can further simplify this we can write it as expectation of X; Y minus X into mu Y minus mu X Y plus mu X, mu Y; that is equal to expectation of X Y minus, so now here if we take expectation this is mu X, mu Y minus mu X, mu Y plus mu X, mu Y. So, this term cancels out and we are getting we can use this notation mu X Y; here or mu 1 1 prime minus mu 1 0 prime mu 0 1 prime.

We can further see the effect of independence on the movements; 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.

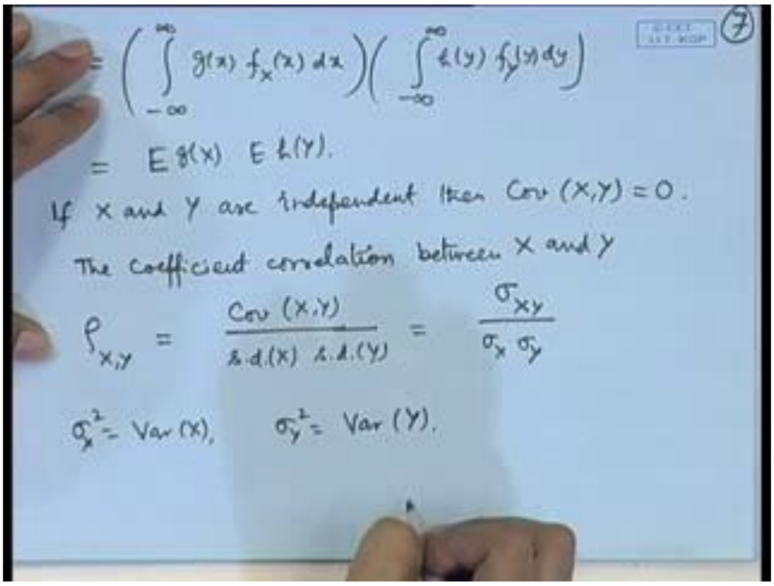
(Refer Slide Time: 18:36)



Let X and Y be independent random variables then if we are considering expectation of a product function then this is equal to product of the expectations. Let us assume that X and Y are continuous with joint pdf; f x y and marginal pdf say f 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, now consider expectation of g X into h of Y. So, of course, these statements are valid provided these expectations exist, 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, d x; d y or d y d x. Since the random variables are independent, we can make use of the condition of independence; that means, this is equal to g x into h y; f x into f y; d x; d y.

(Refer Slide Time: 21:19)



Now, this is nothing, but the product of the integrals g x; f x; d x and h y; f y; d y 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 in case the random variables are discrete because here the integrals can be replaced by the summations and the density function can be replaced by the mass function. The proof is little bit involved in the case of mixed random variables and when one of them may be 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 a product moment is equal to, product of the individual moments and this statement is valid for non central as well as central moments.

Now, if we make use of this condition on the first central product moment which is we call as the covariance then expectation of X into Y; will be equal to expectation X into expectation Y and therefore, covariance term will become 0. So, we have the consequence of this above theorem that if X and Y are independent then covariance of X and Y is 0.

Using the covariance one defines, the coefficient of correlation between X and Y; we denote rho X Y; that is defined as covariance between X, Y divided by standard deviation of X into standard deviation of Y; we can use the notation sigma X Y divided by sigma X; sigma Y where sigma X square was the 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.