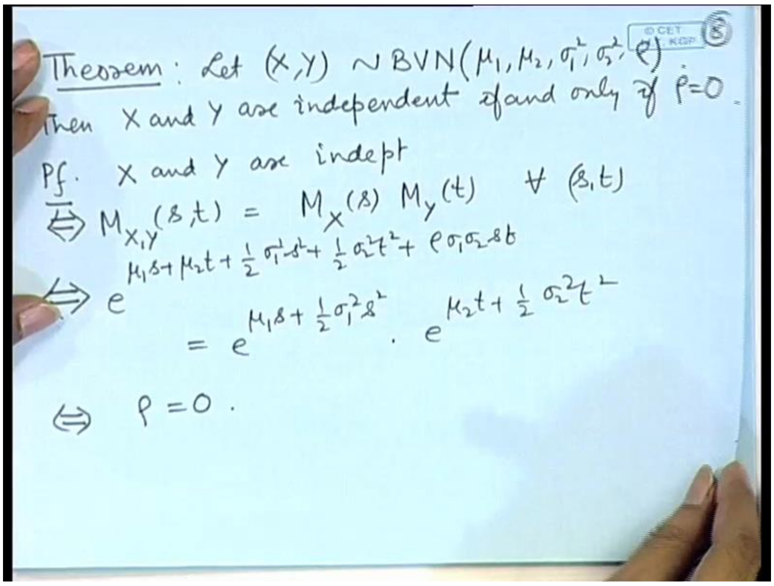
Probability and Statistics

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

Lecture - 40
Bivariate Normal Distribution – II

(Refer Slide Time: 00:22)

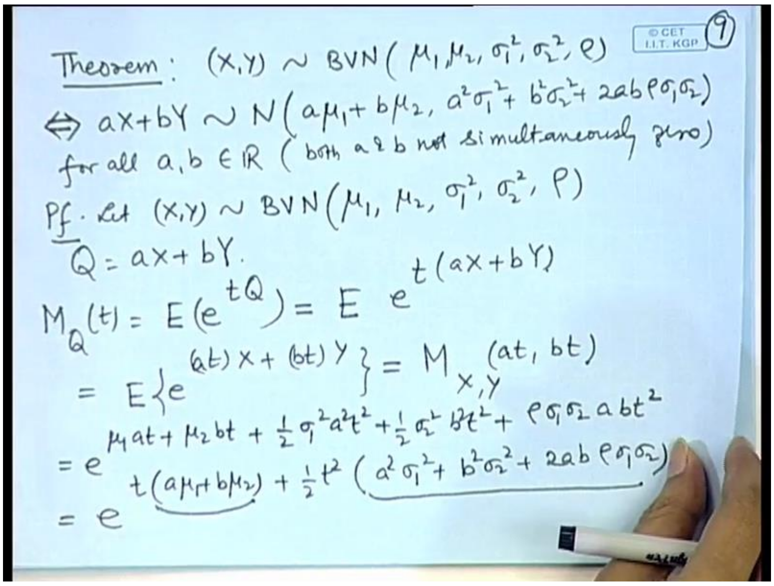


Let X Y follow a bivariate Normal Distribution with parameters mu 1, mu 2, sigma 1 square, sigma 2 square and rho. Then X and Y are independent if and only if rho is equal to 0.

Now we already know that if X and Y are independent, then correlation is 0, so rho will be equal to 0 will be true. To prove the reverse, we make use of the joint M g f, so X and Y are independent; this is equivalent to the statement M X Y; s, t is equal to M X; s, M Y; t for all s, t.

Now, this is equivalent to e to the power mu 1 s plus mu 2 t plus half sigma 1 square s square plus half, sigma 2 square t square plus rho sigma 1, sigma 2 s t equal to e to the power mu 1 s plus half sigma 1 square s square; e to the power mu 2 t plus half; sigma 2 square t square, so this is equivalent to the statement that rho is equal to 0. So, although in general correlation 0 does not imply independence, but in the case of bivariate normal distribution; independence and correlation is equal to 0 is equivalent. We prove another property of bivariate normal distribution using the moment generating function.

(Refer Slide Time: 02:33)

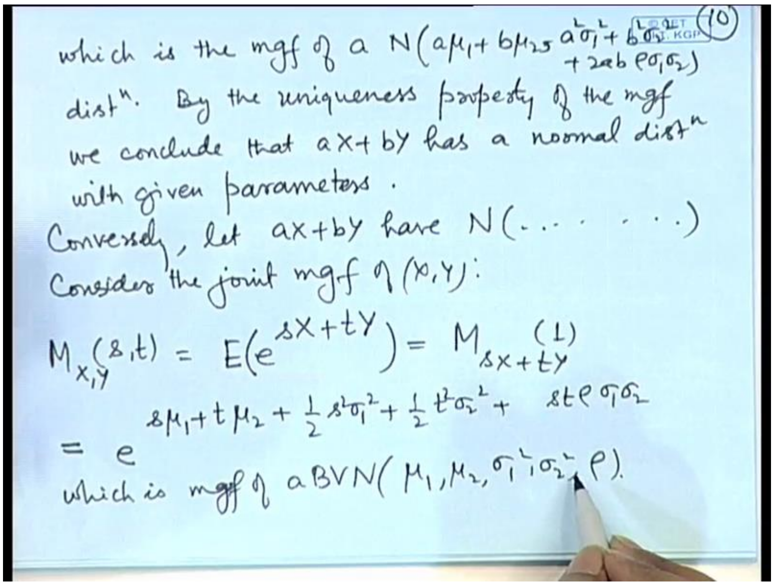


X Y follow a bivariate normal distribution with parameters mu 1, mu 2, sigma 1 square, sigma 2 square, rho; if and only if a x plus b y follows a univariate normal distribution with parameters a mu 1, plus b mu 2, a square sigma 1 square plus b square sigma 2 square plus twice a b rho sigma 1, sigma 2 for all a b real. Of course, both a and b not simultaneously 0; they were very strong property because it says that giving that joint distribution is bivariate normal any linear combination will be univariate normal conversely given every linear combination is a univariate normal, the joint distribution will be bivariate normal.

So in order to prove this statement let X Y have bivariate normal distribution with the given parameters mu 1, mu 2, sigma 1 square, sigma 2 square and rho. Let us write the random variable say Q as a x plus b y, then the moment generating function of Q that is equal to expectation of e to the power t Q that is equal to expectation of e to the power t a x plus b y; that is equal to expectation of e to the power a t; X plus b t Y, this is the joint mgf of X Y at, a; t, b; t. Since X Y has a joint bivariate normal distribution, the form of the joint mgf of X Y at a; t, b; t can be obtained by substituting s is equal to a t and t is equal to b t; in the expiration given just now. So, this becomes e to the power mu 1 a t plus mu 2 b t plus half sigma 1 square, a square, t square plus half sigma 2 square; b square, t square plus rho sigma 1, sigma 2 a b t square.

So after combining the coefficient, we get it as t a mu 1 plus b mu 2 plus half t square a square, sigma 1 square plus b square sigma 2 square plus twice a b rho sigma 1; sigma 2. Now this is nothing but the mgf of a normal distribution with the mean this term and variance this term. So, because of the uniqueness of the M g f, we prove that a x plus b y is having this particular normal distribution.

(Refer Slide Time: 06:25)

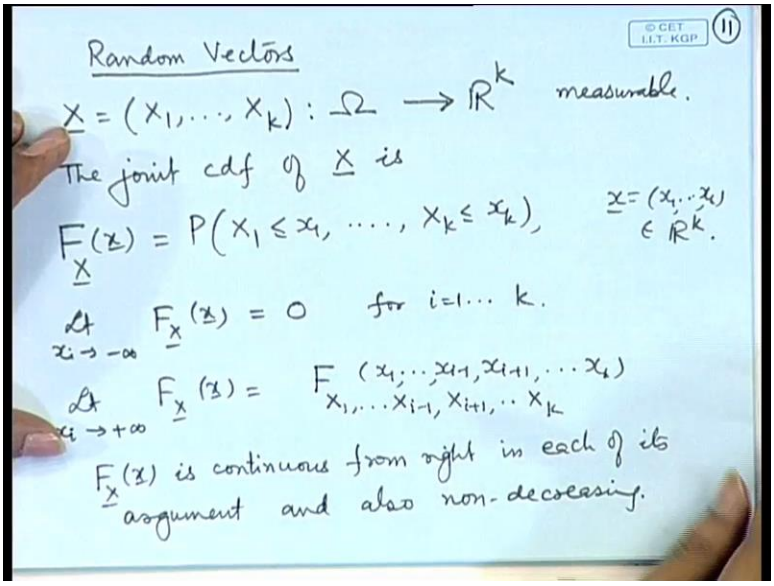


Which is the mgf of a normal a mu 1 plus b mu 2 and a square sigma 1 square plus b square sigma 2 square plus twice a b rho sigma 1 sigma 2 distribution, so by the uniqueness property of the mgf, we conclude that a X plus b Y has a normal distribution with given parameters. Now conversely assume that let a X plus b Y have normal distribution with the desired sector. Now consider the joint mgf of X Y that is M; X Y, s t that is expectation of e to the power s X plus t Y.

Now, notice here that this is nothing, but a linear combination of X and Y. We are assuming that every linear combination has a univariate normal distribution with desired parameters. So, this becomes nothing, but the moment generating function of s X plus t Y, at the point 1 which is known to us because the distribution of s X plus t Y is assumed to be normal with mean s mu 1 plus t mu 2 and s square sigma 1 square plus t square sigma 2 square plus twice s t rho sigma 1 sigma 2. So, since the mgf of the normal distribution is known, we substitute this here and it becomes equal to e to the power s mu 1 plus t mu 2 plus half s square sigma 1 square plus half t square sigma 2 square plus s t rho sigma 1; sigma 2, which is the mgf of a bivariate normal distribution with the parameters mu 1 mu 2 sigma 1 square sigma 2 square and rho.

So, once again the uniqueness of the mgf proves that X Y must have a bivariate normal distribution. So, notice here that this joint mgf is extremely useful improving certain characterization properties of the bivariate normal distribution. We also looked at the generalization of the concept of joint distributions to more than 2, so in general we may consider a k dimensional random variable.

(Refer Slide Time: 10:13)

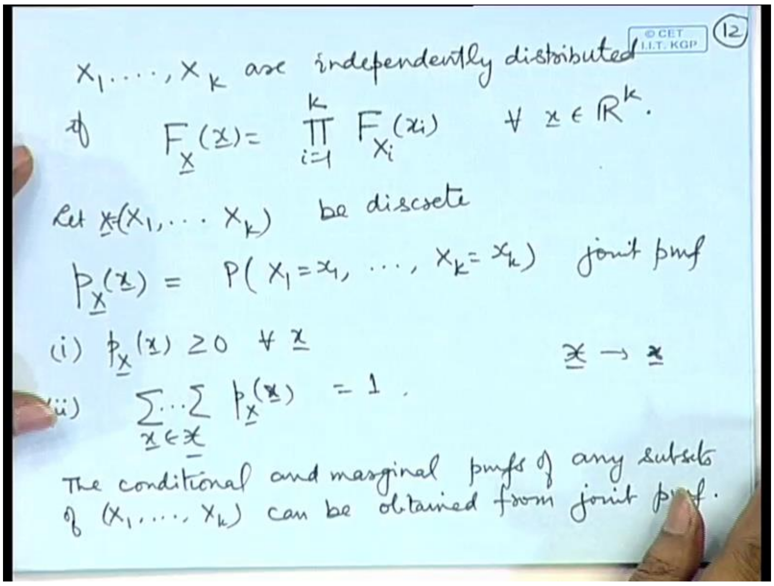


So, we call it random vectors in general, so X is equal to X 1, X 2, X k, so this is a k dimensional random vector; it is defined to be a measurable function from omega into R k and of course, the function should be measurable. Now you may have the random variables; some of the random variables X i is as discrete, some of them as continuous, we may have some of them as mixtures. So, all types of possibilities of the type of the random variables are there, we may make use of the joint c d f; joint c d f of X is defined as F X; x as probability of X 1 less than or equal to X 1; X k less than or equal to X k where this point X is equal to X 1, X 2, X k belongs to R k.

Now this function has in the case of two variables, this is giving complete information about the types of random variables X i's are and also the probability distributions of individual X is are conditionals for example, if I take limit as X i tending to minus infinity in any i, then this will be 0.

If we take limit as say X i tending to plus infinity then that will yield the CDF of all the variables accept the ith one. We may also obtain the marginal distributions of only X 1 or only X 2 by taking the limits of all other variables tending to infinity. The function F x is continuous from right, in each of its argument and also non-decreasing.

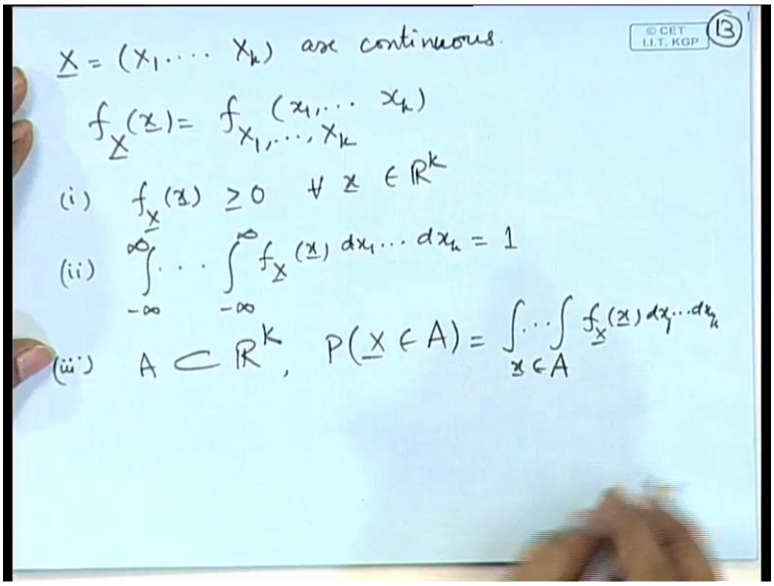
(Refer Slide Time: 13:28)



Making use of this joint CDF we can define the concept of independence X 1, X 2, X k are independently distributed if the joint CDF can be written as the product of individual CDF for all x belonging to R k. Now, we can take the particular cases that is when all of the X is are discrete or all of the X is are continuous because in that case we can define a joint probability mass function and joint probability density function respectively. So, let us take up these two cases; let X 1, X 2, X k be discrete; that means, all of the components are discrete. So, we have a probability mass function that is probability of X 1 is equal to X 1 and so on X k is equal to X k. It will satisfy the usual properties that is it should be non negative function and if we sum over all possibilities of X 1, X 2, X k; it should add up to 1. So, this is the joint probability mass function; it will satisfy the properties that P X; x is greater than or equal to 0 and the sum over all the components must be 1; where X is the set of values of x.

The marginal distribution of any subset of X 1, X 2, X k can be obtained by summing over the remaining variables. For example, if we want the marginal distribution of X 1 then we will sum over the joint pmf over X 2, X 3 up to X k. Suppose we want the marginal pmf of say X k minus 1 and X k, then we will sum over the variables X 1, X 2, X k minus 2. Likewise we can define the conditional probability mass functions of any subset of X 1, X 2, X n given any other subset of X 1, X 2, X n. So, the conditional and marginal pmfs of any subsets of X 1, X 2, X k can be obtained from the joint pmf.

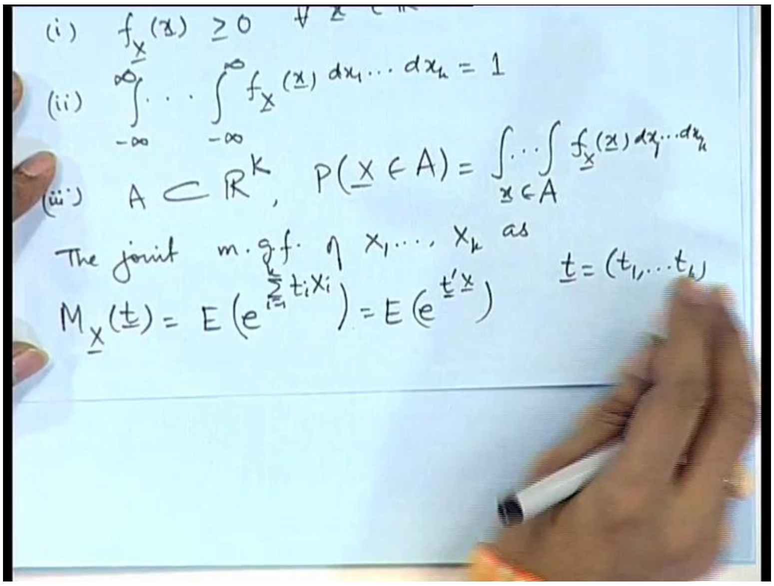
(Refer Slide Time: 16:51)



In a similar way, we may talk about the case when all of the X i's are continuous. In this case we will have a joint probability density function and it will have the properties that the function is nonnegative, the integral over the entire space must give 1 and if I take a to be any subset of the k dimensional Euclidean space, then probability of X belonging to A is where the integrand is integrated over the range A. Once again the marginals or conditionals of any subset of X 1, X 2, X k can be obtained by integrating over the remaining variables.

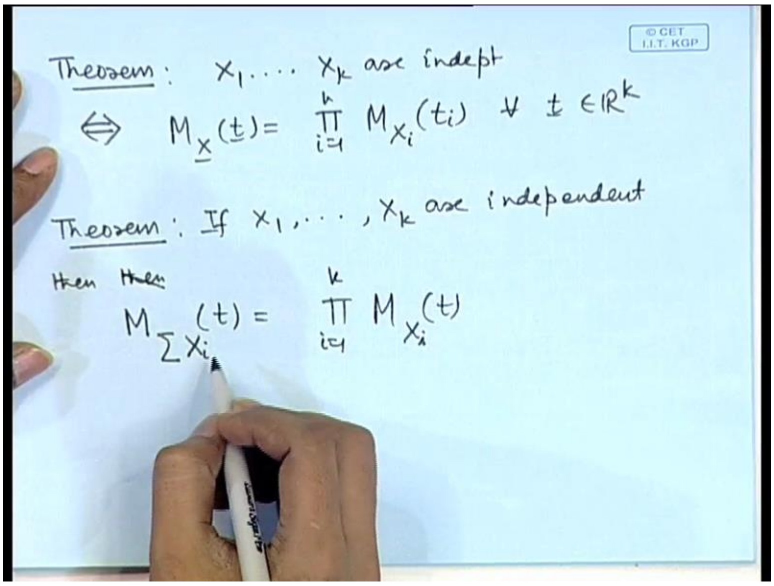
For example if I want the marginal distribution of X 1 and X 3; then living only X 1 and X 3; we will integrate the joint distribution with respect to X 2, X 4, X 5 and so on. Similarly we may talk about say conditional distribution of X 3, X 5 given X 2, so that will require the joint distribution of X 2, X 3 and X 5 and the marginal distribution of X 2.

(Refer Slide Time: 18:46)



We can define the joint moment generating function of X 1, X 2 and X k as M X; t, where t is the point t 1, t 2, t k as expectation of e to the power sigma t i; X i, i is equal to 1, 2 k; that is expectation of e to the power t prime X; where t prime denotes the transpose of the vector t.

(Refer Slide Time: 19:21)



Using this we can prove the theorems as in the case of bivariate that X 1, X 2, X k are independent; if and only if the joint mgf is the product of the individual mgf's for all t. Similarly, if the random variables X 1, X 2, X k are independent then the mgf of the sum is the product of the mgf's. Now this is a very useful tool in evaluating the distributions of the sums of random variables, given that certain random variables are independently distributed; if you are interested in the distribution of the sum, then we simply multiply the mgf's of the individuals and notice that what is the form of that, if it is identify with certain distribution, then we know the distribution of the sum without going through the usual procedure of transformations, from mgf itself we can derive the joint g f.

Using this we will show the additive properties of certain distributions in the next lecture and we will also see some a special joint distributions. So, we will stop today's class here.