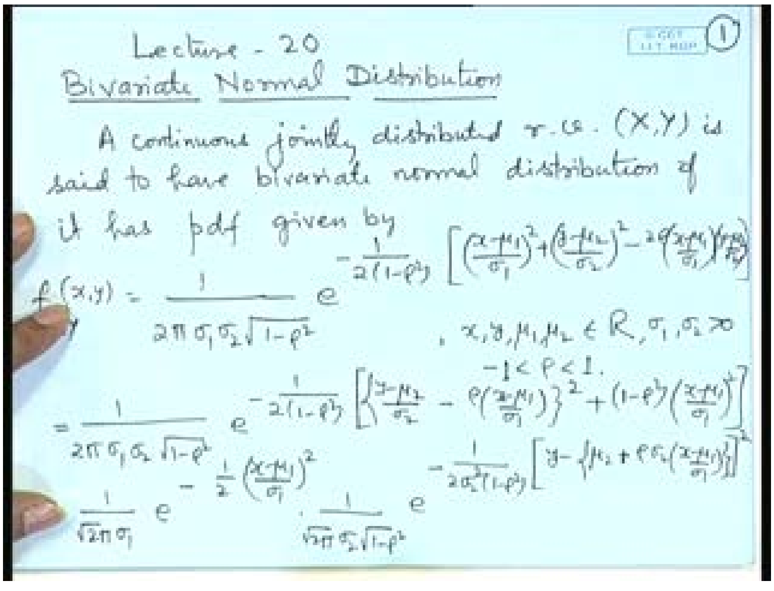
Probability and Statistics

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

Module No. #01
Lecture No. #20
Joint Distributions-III

In the last two lectures we havediscussed the distributions of bivariate random variables. So, we looked at how to derive the marginal distributions and the conditional distributions, we also discussed various characteristics of the joint distribution such as the moments, product moments, covariance and the coefficient of correlation, and we alsolooked at some of the features of thesecharacteristics. Today, I will introduce a particular joint distribution it is known as bivariate normal distribution.

(Refer Slide Time: 00:56)

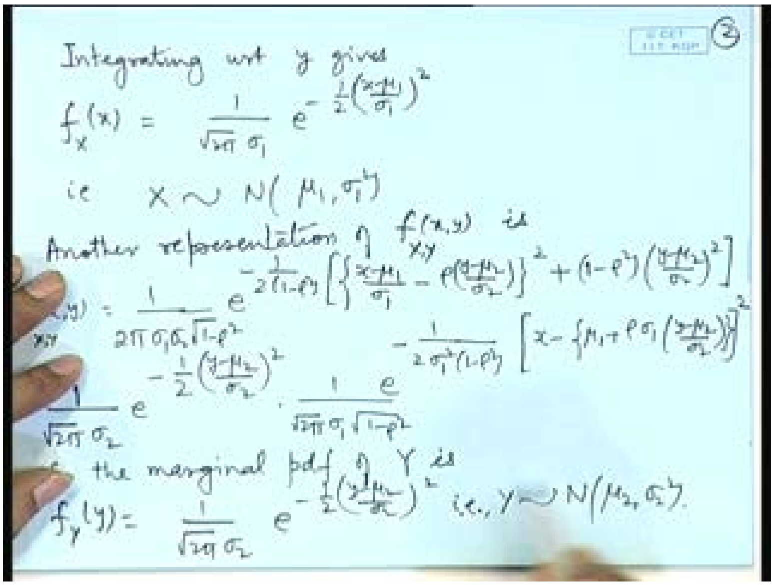


So, a continuous jointly distributed random variable XY is said to have bivariate normal distribution if it has the probability density function given by fxy equal to 1 by 2 pi sigma1 sigma2 root 1 minus rho square e to the power minus 1 by 2 1 minus rho square x minus mu1 by sigma1 square plus y minus mu2 by sigma2 square minus 2 rho x minus mu1 by sigma1 y minus mu2 by sigma2. Here the range of xy, mu1, mu2 is the whole real line and sigma1 sigma2 are positive and rho is between minus 1 and plus 1. So, first of all we look at that what are the marginal distributions and the conditional distributions and overall structure of this bivariate normal distribution, we will like to study.

Suppose we want to find out the marginal distribution of x, in that case we need to integrate this joint distribution with respect to y.A closer examination of the density function reveals that in the exponent we have a term which is a term like which appears in theexponent of the normal distribution. So, if we want to integrate with respect to y, we can convert it into a density with respect to y, so, that suggests that we make a perfect square in y; so, we can factorize it as 1 by 2 pi sigma1 sigma2 root 1 minus rho square e to the power minus 1 by 2 1 minus rho square, so, here if we make a square in y, then we have y minus mu2 by sigma2 minus rho x minus mu1 by sigma1 whole square.Now, this square corresponds to this and the cross product term corresponds to this, so that means, I have added rho square into x minus mu1 by sigma1 whole square. So, if we subtract this, I will get the term as 1 minus rho square x minus mu1 by sigma1 square, which we can write as 1 by root 2pi sigma1 e to the power minus 1 by 2 x minus mu1 by sigma1 square and 1 by root 2pi sigma2 root 1 minus rho square e to the power minus 1 by 2 sigma2 square 1 minus rho square y minus mu2 minus rho sigma2 x minus mu1 by sigma1, so, this I can write as plus and arrange it in the bracket, square.

You can see here that the first term is a normal density for x andthe second term is a normal density for y. So, if we want to find out the marginal distribution of x, we can integrate this with respect to y, and we notice here that this entire term denotes a distribution, which is normal, with mean mu2 plus rho sigma2 x minus mu1 by sigma1 and variance sigma2 square into 1 minus rho square, so if we integrate with respect to y, this term will give us unity and we will get only this term.

(Refer Slide Time: 05:39)



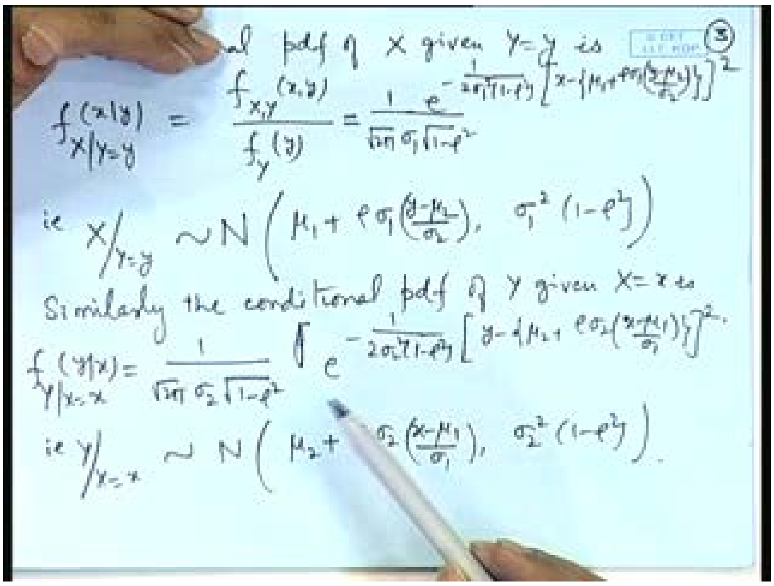
So, integration with respect to y givesfxx equal to 1 by root 2pi sigma1 e to the power minus 1 by 2 x minus mu1 by sigma1 squarethat is, the marginal distribution of x is normal mu1, sigma1 square. In a similar way, we can split this term when we want to integrate with respect to x then I make it as a perfect square in x, so, we will write x minus mu1 by sigma1 minus rho y minus mu2 by sigma2. So, another way of writing is, another representation offxy can be 1 by 2pi sigma1 sigma2 root 1 minus rho square e to the power minus 1 by 2 1 minus rho square, and now I make a square with respect to x, s, x minus mu1 by sigmaminus rho y minus mu2 by sigma2 whole square- so, comparing with the joint density, we can see here that x minus mu1 by sigma1 whole square, that is coming here, and the cross product term is minus 2rho x minus mu1 by sigma1 into y minus mu2 by sigma2, which is a term appearing here, so we have added the term rho square y minus mu2 by sigma2 square- so, subtracting this we get, 1 minus rho square y minus mu2 by sigma2 whole square. So, we can write it as 1 by root 2pi sigma2 e to the power minus 1 by 2 y minus mu2 by sigma2 whole square 1 byroot 2pi sigma1 root 1 minus rho square e to the power minus 1 by 2 sigma1 square 1 minus rho square and x minus mu1 plus rho sigma1 y minus mu 2 by sigma2 whole square.

So, notice here that the second term is a density of normal random variable with mean mu1 plus rho sigma1 into y minus mu2 by sigma2 and variance sigma1 square into 1 minus rho square. So, if we integrate this joint density with respect to x, the term, this term integrates to 1 and we are left with a normal density. So, the marginalpdf of y is fyy, that is equal to 1 by root 2pi sigma2 e to the power minus 1 by 2 y minus mu2 by sigma2 square that is, the marginal distribution of y is normal mu2 sigma2 square.

So, we come across this interesting phenomenon that if xy follow a joint bivariate normal distribution, then the marginal distributions of x is normal mu1 sigma square and then marginal distribution of y is normal mu2 sigma2 square that means, given a joint bivariate distribution the marginal distributions are univariate normal.

Now, we also calculate the conditional distributions of x given y and y given x. Now, if we look at the conditional distributions of X given Y, then we have to divide the joint distribution of x, y by the marginal distribution of y. Now, from this break up, we can see that this joint distribution if we divide by the marginal of y, this term gets cancelled out and we are left with this particular term, which is nothing but the normal distribution.This proves that the conditional distribution of X given Y is normal and the mean and the variance are specified here.

(Refer Slide Time: 10:38)

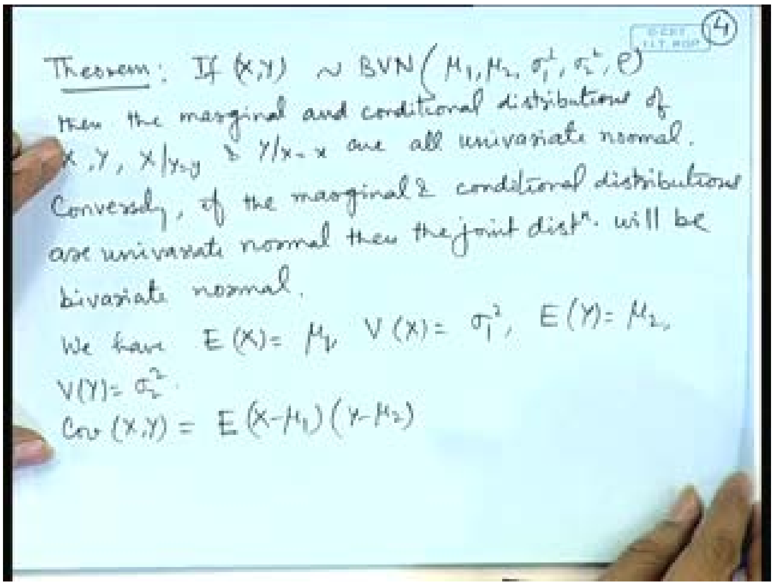


So, we have the conditional probability density function of X given Y is equal to y, that is obtained as the joint distribution divided by themarginal distribution of y. So, after simplification it is equal to 1 by root 2pi sigma1 root 1 minus rho square t to the power minus 1 by 2 sigma1 square 1 minus rho square x minus mu1 plus rho sigma1 y minus mu2 by sigma2 whole square.That is, we can say that X given Y is equal to y follows normal with mean mu1 plus rho sigma1 y minus mu2 by sigma2 and variance sigma1 square into 1 minus rho square.

In a similar way, notice here that the joint distribution of xy was earlier factorized like this and if we divide by the marginal distribution of x, then this term gets cancelled out and we are left with this term, which is again a normal distribution with a certain mean and a certain variance. So, this proves that the conditional distribution of, similarly, the conditional pdf of Y given X, that is obtained as 1 by root 2pi sigma2 root 1 minus rho square e to the power minus 1 by 2 sigma2 square 1 minus rho square y minus mu2 plus rho sigma2 x minus mu1 by sigma1 whole square.That is Y given X is equal to x follows normal with mean mu2 plus rho sigma2 x minus mu1 by sigma1 sigma2 square 1 minus rho square. So, we conclude that if the joint distribution is bivariate normal the marginal's as well as the conditional distributions are univariate normal.

Now, the converse of this is also true.If the conditionals and the marginal are univariate normal, the joint distribution will be bivariate normal. So, this is also a characterizing property of the bivariate normal distribution.

(Refer Slide Time: 13:46)

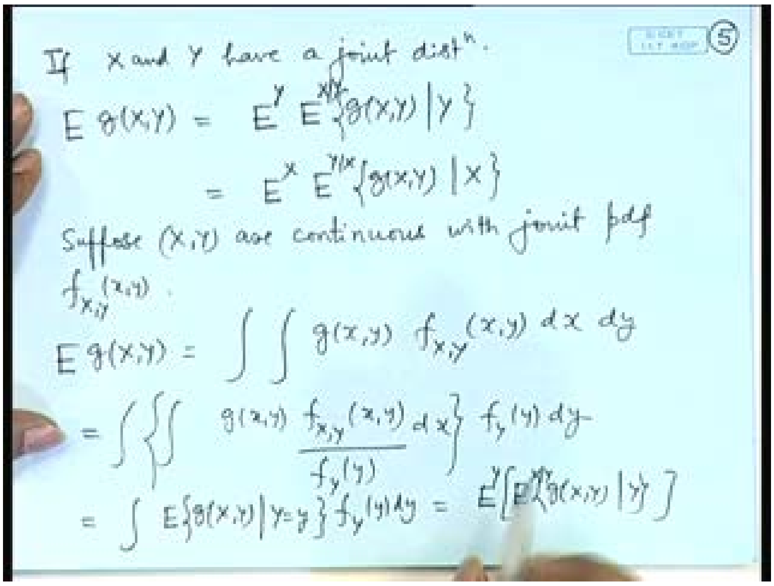


We can state it as a theorem-if xy follows bivariate normal with some parameter say mu1 mu2 sigma1 square sigma2 square and rho, then the marginal and conditional distributions of XY, X given Y and Y given X are all univariate normal.Conversely, if the marginal and conditional distributions are univariate normal, then the joint distribution will be bivariate normal.

So, this is quiteuseful inobtaining any probability related to marginal, or the conditional distributions of the X and Y because we can make use of the standard normal distribution by making a suitable transformation. Any joint probability statement about bivariate normal distribution will need the tables of a standard bivariate normal distribution; by a standard bivariate normal distribution we mean mu1 is equal to 0 mu2 is equal to 0 sigma1 square and sigma2 square is equal to 1, but rho will still be there and therefore, several tables will berequired with respect, which will berelated to the joint probabilities of the bivariate normal distribution.

So, since the marginal distributions are identified we have expectation of X is equal to mu1, variance of X is equal to sigma1 square, expectation of Y is equal to mu2 and variance of Y is equal to sigma2 square.Now, we also consider the covariance term between X and Y. So, the covariance term between X and Y is expectation of X minus mu1 into Y minus mu2.Now, this product, central product moment can be calculated bythe joint integration of the density function multiplied by this function however, at this stage we introduce some formula for evaluation of the joint moments.

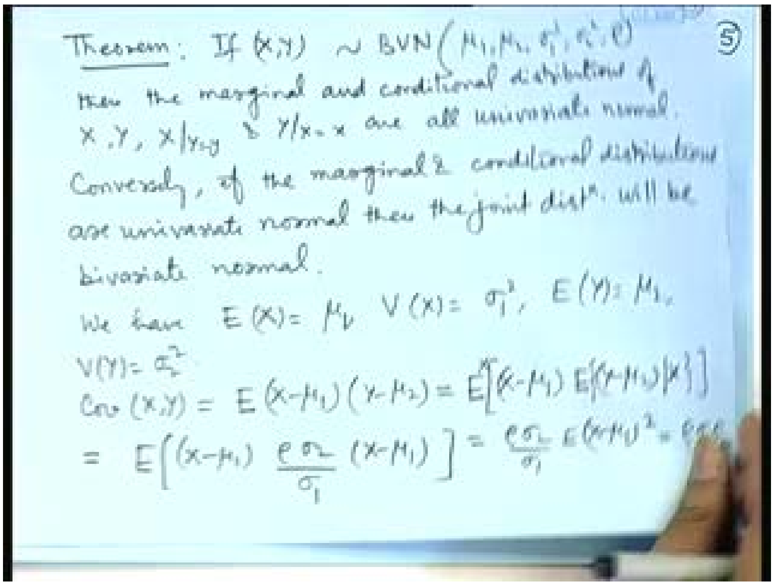
(Refer Slide Time: 17:17)



IfX and Y have a joint distribution, then in general, expectation of a function can be calculated in stages.We may calculate firstly, the conditional and then with respect to marginal, or alternatively we may consider it as expectation of gXY given X, Y given X, provided of course, the expectations do exist. Let me give a rough sketch of the proof.

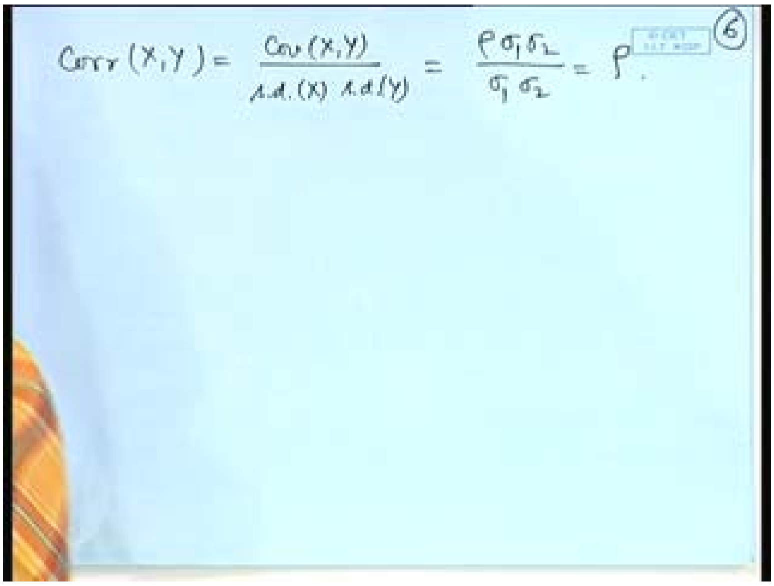
Suppose X and Y are continuous with joint pdf sayfxy. So, expectation of g X, Y we can express as integral gxyfxyxy; suppose we keep the order of integration as d x d y then this we can express as gxyfxy divided by fy multiplied by fydy; so, this quantity, inner quantity is nothing but the expectation of gXY given Y is equal to y multiplied by the density of y, which is nothing but expectation of expectation gXYgiven Y. That means, the joint expectations can be calculated in stages firstly, with respect to a conditional distribution and then with respect to a marginal distribution in either order.

(Refer Slide Time: 19:37)



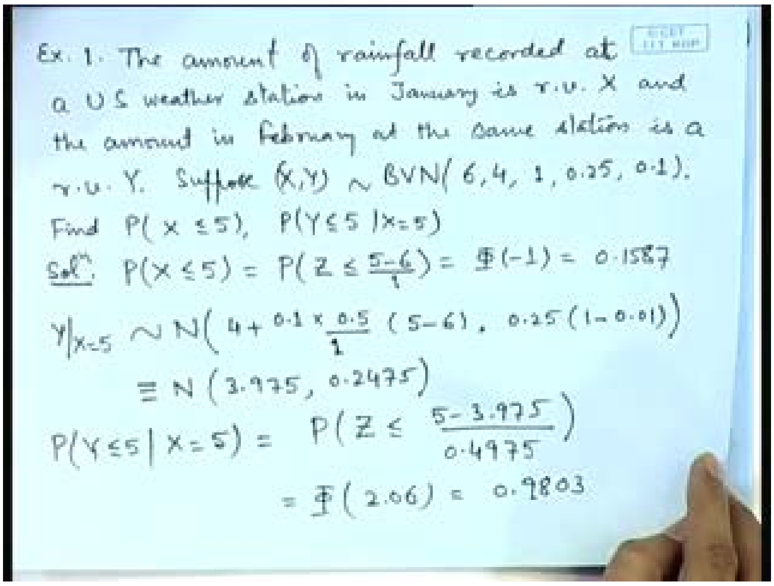
So, if we make use of this, then expectation of X minus mu1 into Y minus mu2, we can write it as X minus mu1 into expectation of Y minus mu2 given X. So, inner expectation is the conditional expectationwith respect to the distribution of Y given X and the outer is with respect to X-the conditional distribution of Y given X was calculated to be a univariate normal distribution andthe mean was mu2 plus certain term, so expectation of Y given X will be mu2 plus rho sigma2 x minus mu1 by sigma1-therefore, expectation of Y minus mu2 given X will be equal to X minus mu 1rho sigma2 by sigma1 X minus mu1, which is nothing but rho sigma2 by sigma1 expectation of X minus mu 1 square, which is sigma 1 square, so, it is rho sigma1 sigma2.

(Refer Slide Time: 21:06)



So, we conclude that the covariance of the X Yin a bivariate normal distribution is given by rho sigma1 sigma2. Therefore, we can calculate the coefficient of correlation between X Y that is equal to rho sigma1 sigma2 by sigma1 sigma2 that is equal to rho. So, the parameter rho of a bivariate normal distribution denotes the correlation coefficient between the random variables X and Y.

(Refer Slide Time: 21:49)



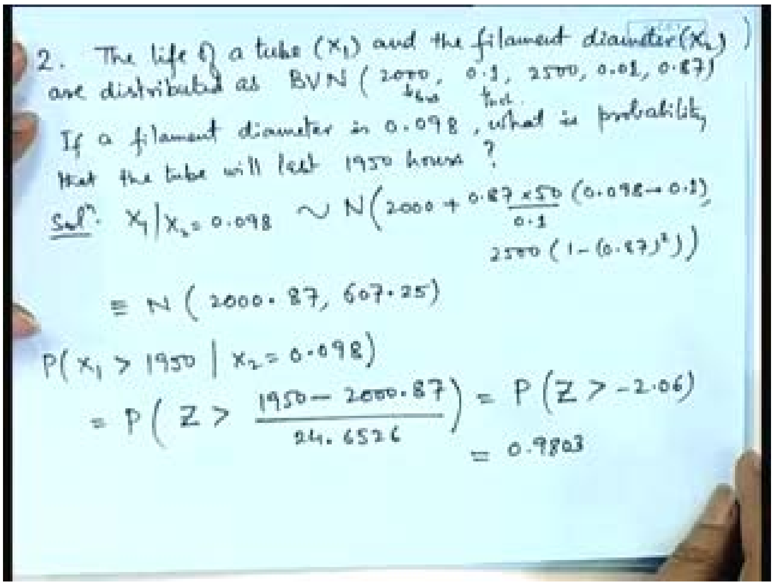
Let us look at a problem here.The amount of rainfall recorded at a US weather station in January is a random variable X and the amount of rainfall recorded inFebruary at the same station is a random variable Y.Suppose the distribution of X and Y is observed to be a bivariate normal distribution with mean 6. So, the mean of random variable X is 6 the mean of random variable Y is 4. So, suppose it is in measured in inches because it is amount of rainfall or centimeters, the variances are 1 and 0.25 and rho is equal to 0.1.We are interested to calculate what is the probability that X is less than or equal to 5, or what is the probability of Y being less than or equal to 5 giving that X is equal to 5.

So, notice here probability of X less than or equal to 5 can be calculated from the marginal distribution of X, which is having mean 6 and variance unity; so, it is simply transform to the standard normal probability as z less than or equal to 5 minus 6 by 1, here z denotes the standard normal random variable; so, from 5we have subtracted the mean of X and divided by the standard deviation, which is equivalent to the cdf value of standard normal variate at minus 1, which we see from the tables of normal distribution as 0.1587.

Suppose, we are interested in the probability of Y less than or minus to 5 given that in January the rainfall is 5. So, we need the conditional probability of Y less than or equal to 5 given X equal to 5.For this we firstly, calculate the conditional distribution of Y given X equal to 5.Now, making use of the conditional distribution of Y given X, which is given by normal with mean mu2 plus rho sigma2 x minus mu1 by sigma1, so, here mu 2 is 4,rho is 0.1, sigma2 is 0.5, sigma1 is 1 and the point x is, smallx is 5, so x is 5 and mu1 is 6, so, this is the mean of the conditional distribution of Y given X, so, after simplification this turns out to be 3.975, the variance of the conditional distribution is sigma2 square into 1 minus rho square, which is 0.25 into 1 minus 0.01, so, it is evaluated to be 0. 2475.

So, the conditional probability of Y X less than or equal to 5 given x equal to 5 can be calculated from this distribution. So, we transform it to the standard normal distribution. So, it isz less than or equal to 5 minus 3.975 divided by square root of this that is,0.4975. So, after simplification it turns out to be Phi of 2.06, which is 0.9803, which isquite high probability, but that is understandable because in January there is more rain, so, since the variables are correlated it is affecting the probability of Y also.Let us take up another example of a similar nature.

(Refer Slide Time: 25:21)

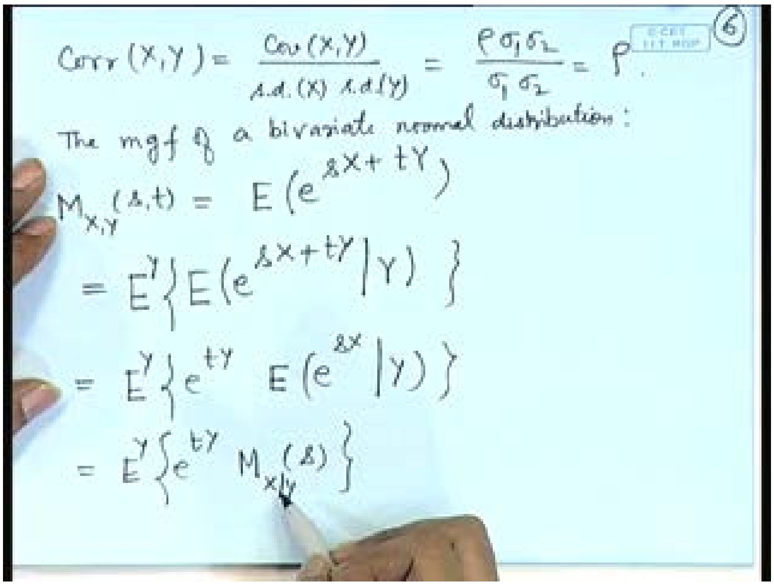


The life of a tube, which is measured as random variable X1 and the filament diameter, which is measured as a random variable X2, the life is measured in say, hours and the diameter is measured in inches, they are distributed as a bivariate normal distribution with mu1 is equal to 2000 hours, mu2 is 0.1 inches, the sigma1 square is 2500, sigma2 square is 0.01 and the coefficient of correlation is 0. 87. So, themanufacturer may use thefilament diameter (()), which can be measured to estimate the life of the tube. So, if a filament diameter is 0.098, what is the probability that the tube will last 1950 hours? So, we are interested to calculate what is the probability of surviving till 1950 hours, given that the diameter is 0.098 inches.

For this we need the conditional distribution of X1 given X2 is equal to 0.098. So, we make use of the formula for the conditional distribution of X given Y here. So, that is mu1 that is, 2000 plus rho,0. 87, sigma1 is 50 divided by sigma2 is 0.1, y minus mu2, so y is the point at which we are conditioning, that is 0.098 minus mu2 that is,0. 1. So, after simplification this turns out to be 2000.87 and the variance here is sigma1 square into 1 minus rho square, which is equal to 607.25 So, the conditional probability of X1 greater than 1950, given that X2 is equal to 0.098 can be calculated using this univariate normal distribution. So, after transformation to standard normal we get it as probability of z greater than minus 2.06, which is evaluated as 0. 9803.

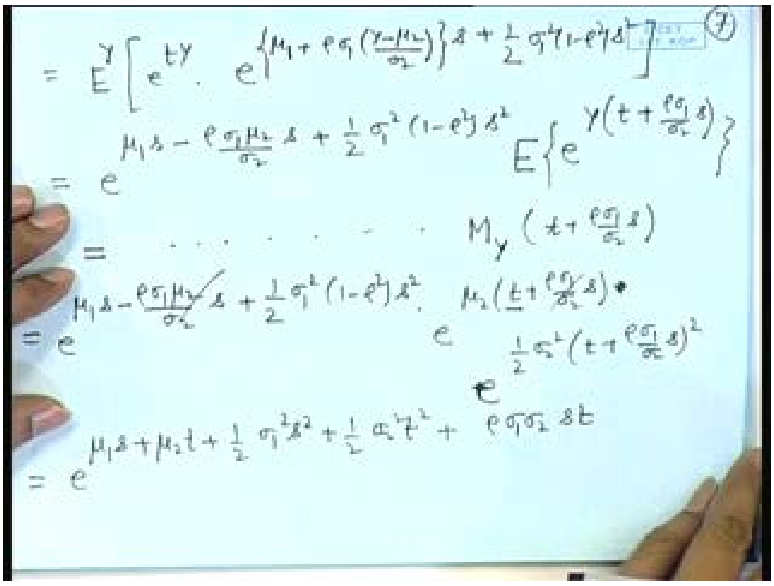
So, likewise any probability statement related to the marginal distributions, or the conditional distributions of X or Y, or X given Y or Y given X can be calculated using the univariate normal properties.We also look at the moment generating function ofa bivariate normal distribution.

(Refer Slide Time: 28:08)



So, the moment generating function of a bivariate normal distribution. So, it is defined as MXYst, that is equal to expectation of E to the power sXplus tY. Now, again, you can see that this is some function g of xy. So, the joint expectation we can calculate easily in terms of conditional and the marginal expectations. So, we will use that, we can write their expectation of expectation E to the powersX plus tY, given say- in the previous one we have done the calculation usingconditional distribution of X, so we can use the conditional distribution of Y here- now, given Y this e to the powertY terms is fixed, so we can separate it out and we are left with expectation of e to the powersX given Y; now, notice here that this inner expectation is nothing but the moment generating function of the conditional distribution of X given Y; so, this is equal to expectation of e to the power tY into the moment generating function of the conditional distribution of X given Y at the point s; now, here the conditional distribution of X given Y is univariate normal, we already know the form of the moment generating function of a univariate normal distribution, suppose the normal mu sigma square distribution is there then we have seen that the mgf is represented as e to the power mu t plus half sigma square t square; so, here the point is s in place of t and X given Y, the distribution has the parameters mu1 plus rho sigma1 y minus mu2 by sigma2 and sigma1 square 1 minus rho square, so, we make use of this.

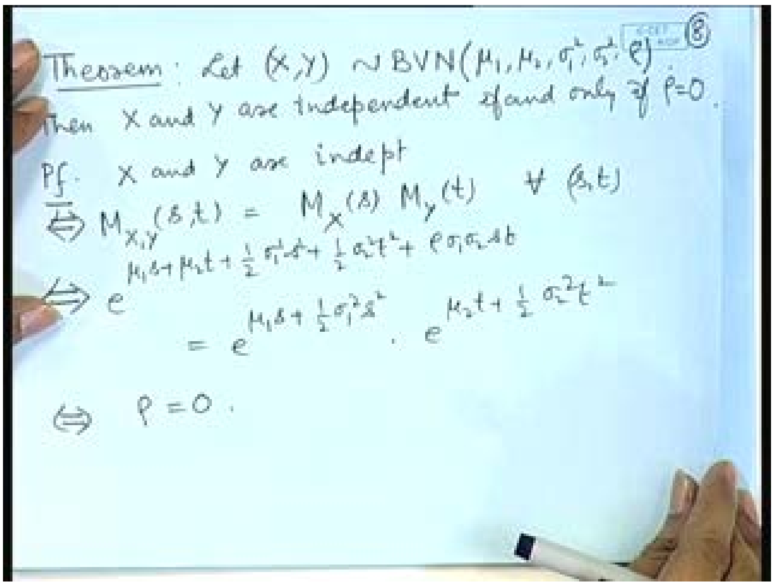
(Refer Slide Time: 30:47)



So, this can be expressed as expectation of Y e to the power tY e to the power mu1 plus rho sigma1 Y minus mu2 by sigma2 into s plus half sigma1 square 1 minus rho square s square- so, this is the value coming after substituting the value of themoment generating function of the conditional distribution of X given Y, which is univariate normal and therefore, the form is known to us; so, now, here, there are certain constant terms and we can separate it out, e to the power mu1 s minus rho sigma1 mu2 by sigma2 s plus half sigma1 square 1 minus rho square s square, we have expectation of e to the power Y t plus rho sigma1 by sigma2 s; so, we notice here that this is nothing but the moment generating function of Y at the point t plus rho sigma1 by sigma2 s. So, this term is there. So, notice here that the distribution of Y is again univariate normal with parameters mu2 and sigma2 square therefore, the moment generating function has an own form, in place of the point t we substitute t plus rho sigma1 by sigma2 s; so, we get it as e to the power mu1 s minus rho sigma1 mu2 by sigma2 s plus halfsigma1 square 1 minus rho square s square e to the power mu2 t plus rho sigma1 by sigma2 s plus, so, I will write it as e to the power half sigma2 square t plus rho sigma1 by sigma2 s whole square; so, we have e to the power mu1 s plus mu2 t, that is this term- now, we note here minus rho by rho sigma1 sigma by sigma2 mu 2 s this term is coming here also as a plus sign plus rho sigma1 by sigma2 mu 2 s, so this term gets cancelled with this term- then, we have half sigma1 square s square andhalf sigma2 square t square- now, when we take square here, it is becoming twice rho sigma1 by sigma2 s t, so, sigma2 and sigma2 square- so you will get it as plus rho sigma1 sigma2 s t- and the square term here,rho square sigma1 square s square with a half here will get cancelled with minus half sigma1 square rho square s square- so, we are left with this term as the mgf of the bivariate normal distribution.

So, notice here that e to the power mu1 s plus half sigma1 s square denotes the mgf of the normal distribution with parameter mu1 andsigma1square, that is the mgf of X. Similarly, e to the power mu2 t plus half sigma square t square denotes themgf of Y. So, we have these terms and an additional term coming here. So, using this we can prove certain more properties regarding the bivariate normal distribution.

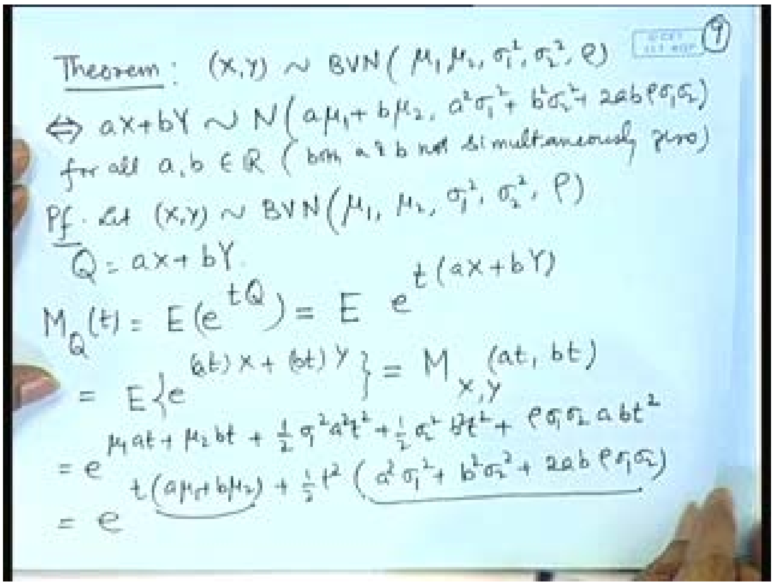
(Refer Slide Time: 35:20)



Let X Yfollow a bivariate normal distribution with parameters mu1, mu2, sigma1 square, sigma2 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 mgf. So, X and Y are independent, this is equivalent to the statementMXYst is equal to MXs MYt for all s t; now, this is equivalent to e to the power mu1 s plus mu2 t plus half sigma1 square s square plus half sigm2 square t square plus rho sigma1 sigma2 s t equal to e to the mu1 s plus half sigma1 squares square e to the power mu2 t plus half sigma2 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 ofBivariate normal distribution independence and correlation is equal to 0 is equivalent. We prove another property of bivariate normal distribution using the moment generating function.

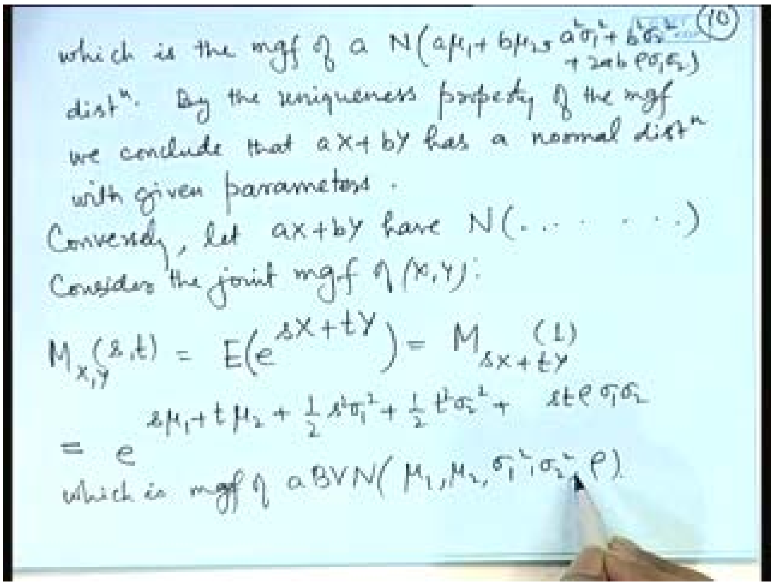
(Refer Slide Time: 37:34)



X Y follow a bivariate normal distribution with parameters mu1, mu2, sigma1 square, sigma2 square,rho if and only if aX plus bY follows a univariate normal distribution with parameters amu1 plus bmu2 a square sigma1 square plus b square sigma2 square plus twice abrho sigma1 sigma2 for all a b real of course, both a and b not simultaneously 0.This is a very strong property because it says that given 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 mu1, mu2, sigma1 square sigma2 square and rho.Let us write the random variable say Q as aX plus b Y then, the moment generating function of Q, that is equal to expectation of e to the power tQ, that is equal to expectation of e to the power taX plus bY, that is equal to expectation of e to the power at X plus bt Y, this is the joint mgf of X Y at at,bt-since X Y has a joint bivariate normal distribution the form of the joint mgf of X Y at at,bt can be obtained by substituting s is equal to at and t is equal to bt in the expression given just now- so, this becomes e to the power mu1 t plus mu2 bt plus half sigma1 square a square t square plus half sigma2 square b square t square plus rho sigma1 sigma2 abt square; so, after combining the coefficients we get it as t amu1 plus bmu2 plus half t square a square sigma1 square plus b square sigma2 square plus twice abrho sigma1 sigma2. Now, this is nothing, but the mgf of a normal distribution with the mean this term and variance this term

(Refer Slide Time: 41:23)

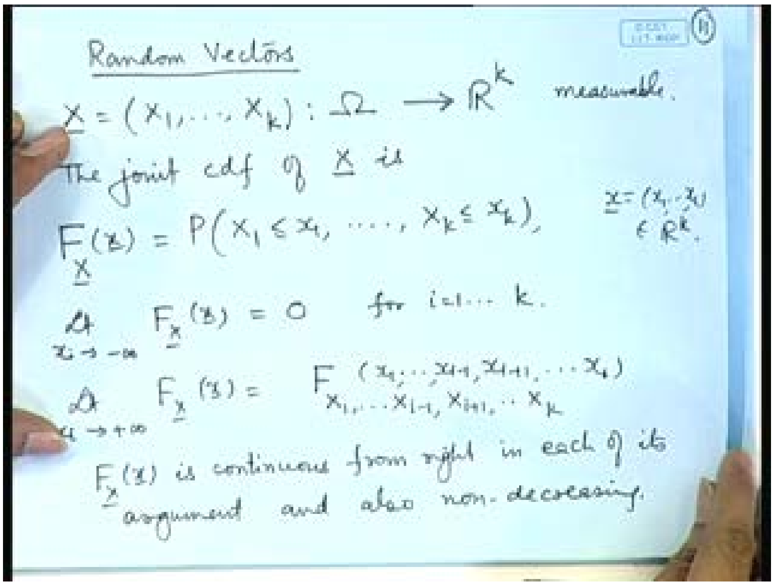


So, because of the uniqueness of the mgf we prove that aX plus bY is having this particular normal distribution,which is the mgf of a normal amu1 plus bmu2 and a square sigma1 square plus b square sigma2 square plus twice abrho sigma1 sigma2 distribution. So, by the uniqueness property of the mgf, we conclude that aX plus bYhas a normal distribution with given parameters.

Now, conversely, assume that let aX plus bY have normal distribution with the desired setup.Now, consider the joint mgf of X Y that is, MXYst that is, expectation of e to the power sXplustY, 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 sXplus tYat 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 mu1 plus t mu2 and s square sigma1 square plus t square sigma2 square plus twice strho sigma1 sigma2; so, since the m g f of the normal distribution is known we substitute this here and it becomes equal to e to the power s mu1 plus t mu2 plus half s square sigma1 square plus halft square sigma2 square plus strho sigma1 sigma2, which is the mgf of a bivariate normal distribution with the parameters mu1, mu2, sigma1 square, sigma2 square and rho.

So, once again the uniqueness of the mgf proves that XY must have a bivariate normal distribution. So, notice here that this joint mgf is extremely useful in proving certain characterization properties of the bivariate normal distribution.

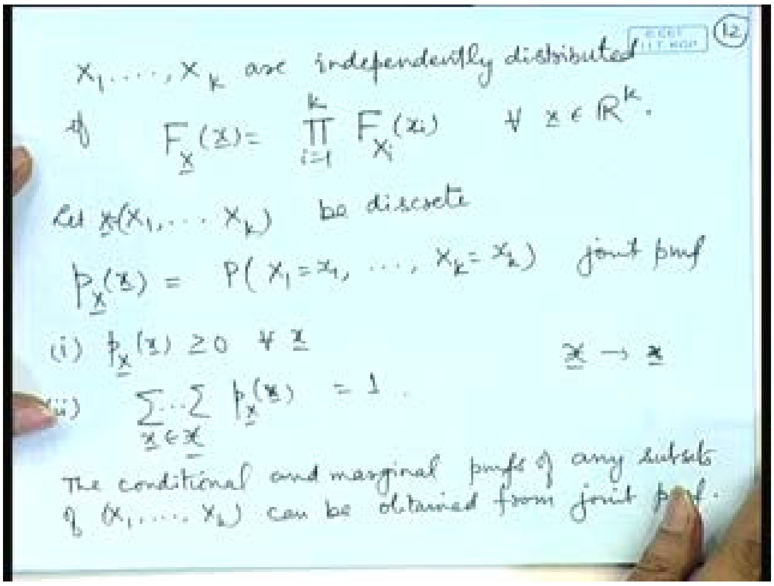
(Refer Slide Time: 45:14)



We also look at the generalization ofthe concept of joint distributions to more than 2. So, in general we may consider a K dimensional random variable, so, we call it random vectors in general. So, X is equal to X1, X2, Xk, so, this is a k dimensional random vector, it is defined to be a measurable function from omega intoRk and of course, the function should be measurable.Now, you may have the random variables,some of the Random variables xis as discrete, some of them is 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 cdf, joint c f of X is defined asFXx as probability of X1 less than or equal to x1, Xk less than or equal toxk, where this point x is equal to x1, x2, xk belongs toRk.

Now, this function, as in the case of two variables, this is giving complete information about the types of random variables xis are and also the probability distributions of individual xisor conditionals.For example, if I take limit as xi tending to minus infinity in any I, thenthis will be 0; if we take limit as say xi tending to plus infinity, then that will yield the cdf of all the variables except the ith one; we may also obtain the marginal distributions of only x1 or only x2 by taking the limits of all other variables tending to infinity. The function FX is continuous from rightin each of its argument and also non decreasing.

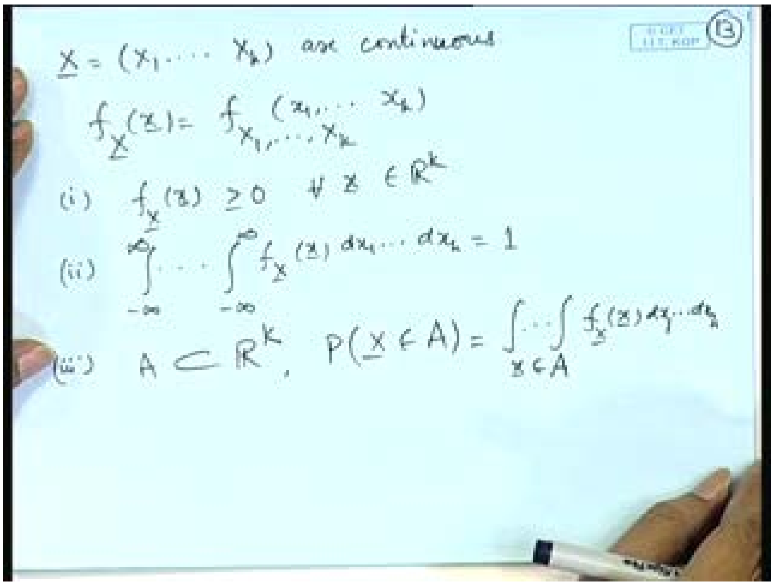
(Refer Slide Time: 48:27)



Making use of this a joint cdf, we can define the concept of independence, X1, X2, Xk are independently distributed if the joint cdf can be written as the product of individual cdfs for all x belonging toRk. Now, we can take the particular cases that are when all of the xis are discrete, or all of the xis 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 X1, X2, Xk be discrete that means, all of the components are discrete. So, we have a probability mass function that is, probability of X1 is equal to X1 and so on, Xk is equal toXk- it will satisfy the usual properties that is, it should be non negative function, and if we sum over all possibilities of X1, X2, Xk, it should add up to 1. So, this is the joint probability mass function, it will satisfy the propertiesthatpXx 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 X1, X2, Xk can be obtained by summing over the remaining variables. For example if we want the marginal distribution of X1, then we will sum over the joint pmf over X2,X3 up toXk;suppose we want the marginal pmf of say, Xk minus 1 andXk, then we will sum over the variables X1, X2, Xk minus 2.Likewise we can define the conditional probability mass functionsof any subset of X1, X2,Xn given any other subset of X1, X2, Xn. So, the conditional and marginal pmfs of any subsets of X1, X2, Xk can be obtained from the joint pmf.

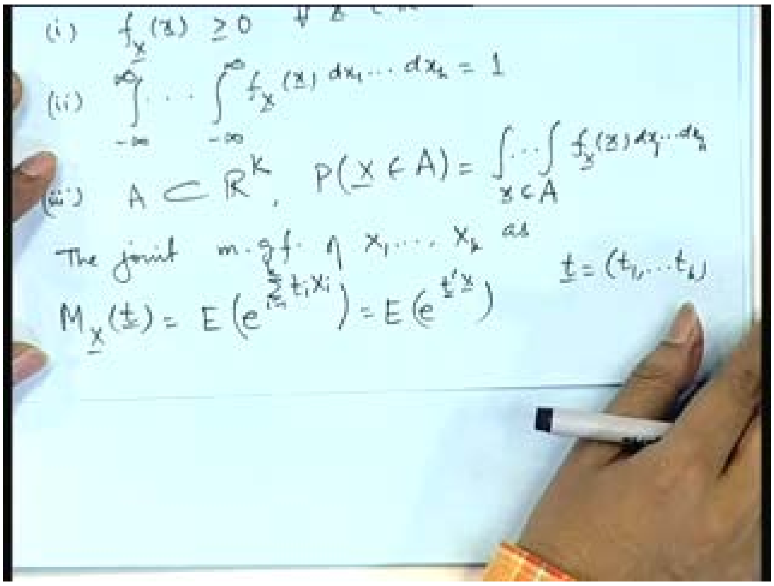
(Refer Slide Time: 51:49)



In a similar way, we may talk about the case when both, when all of the xis are continuous.In this case, we will have a joint probability density function and it will have the properties that the function is non-negative, 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 integrant is integrated over the range A. Once again, themarginal or conditionals of any subset of X1, X2, Xk can be obtained by integrating over the remaining variables.For example, if I want the marginal distribution of x1 and x3, then leaving only x1 and x3, we will integrate the joint distribution with respect to x2,x4,x5 and so on.

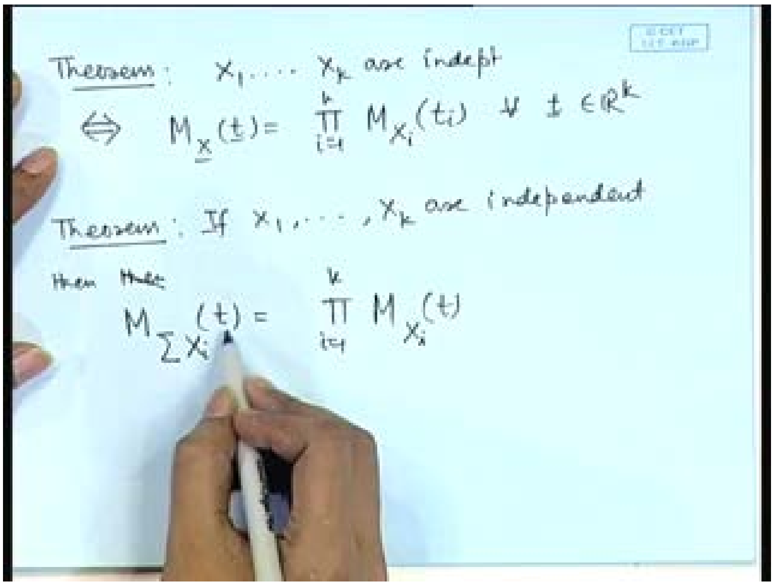
Similarly, we may talk about say conditional distribution of X3,X 5 given X2, so, that will require the joint distribution of x2,x3 and x5 and the marginal distribution of X 2.

(Refer Slide Time: 51:49)



We can define the joint moment generating function of X1, X2, Xk as MXt, where t is the point t1, t2,tk, asexpectation of e to the power sigma tiXi, i is equal to 1 to k, that is expectation of e to the power tprime x, where tprime denotes the transpose of the vector t.

(Refer Slide Time: 54:19)



Using this we can prove the theorems, as in the case of bivariate that X1, X2, Xk are independent if and only if the joint mgf is the product of the individual mgfs for all t. Similarly, if the random variables X1, X2, X k are independent, thenthe mgf of the sum is the product of the mgfs.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 we are interested in the distribution of the sum, then we simply multiply the mgfs of the individuals and notice that what is the form of that, if it is identifiable with certain distribution, then we know the distribution of the sum going through the, without going through the usual procedure of transformations, from mgfs itself we can derive the joint mgf .Using this we will show the additive properties of certain distributions in thenext lecture and we will also see some for special joint distributions. So, today we will stop today’s class here.