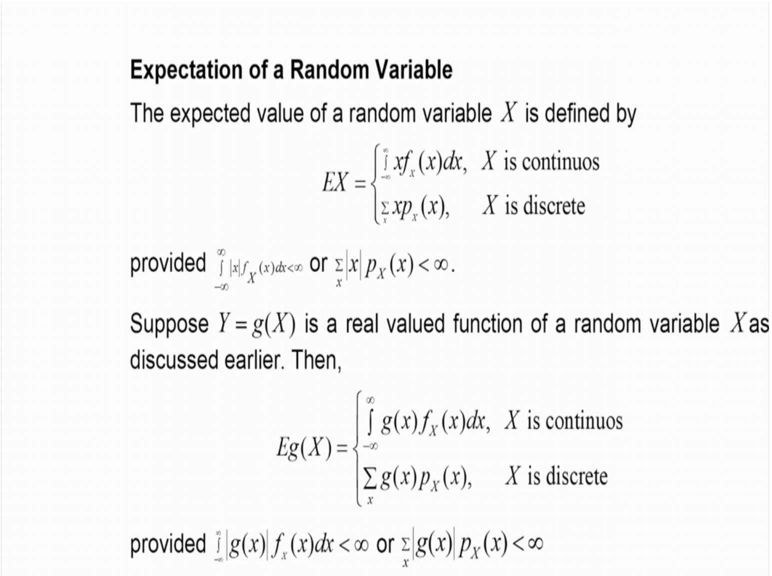
Advanced Topics in Probability and Random Process
Prof. P.K. Bora
Indian Institute of Technology Guwahati

Lecture -03
Random Variable- II

(Refer Slide Time: 00:25)



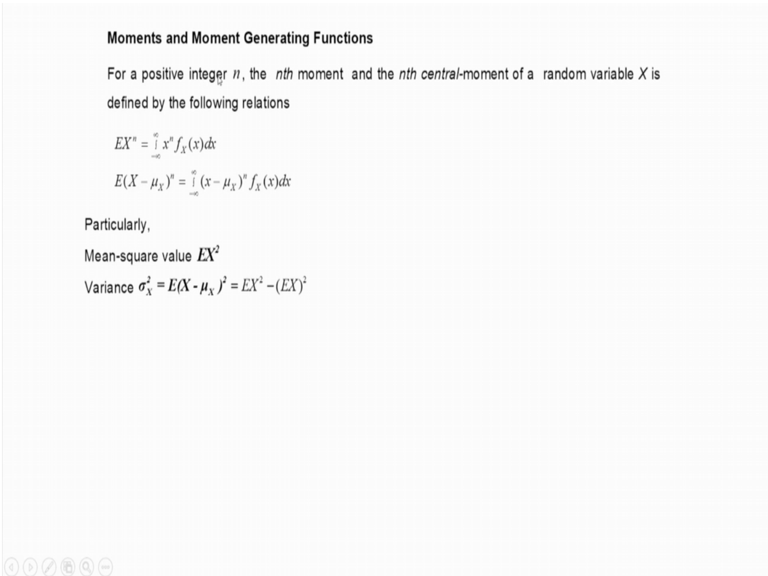
So, we discussed random Variables and how to characterize them with the help of CDF, PDF, PMF etcetera, the CDF, PDF and PMF they are complete representation of a random variable and now we will see how we can represent a random variable, partially we can characterize a variable partially by moments.

So, we will start with the definition of expected value of a random variable. The expected value of a random variable x is defined by this 2 relationship, if x is continuous and then this is the integration of x into f x with the x d x and similarly if X is discrete it is the summation of the product of x into p x of x and the limit of integration or the summation we have to properly define, when we define expectation we wonder this value is unique.

So, product one condition is imposed that is x mod of x f X of x that is absolutely. So, we want mod of x into f X of x should be integrable over the limit minus infinity to infinity or in the case of discrete mode of x into P X of x should be summable for the range of the random variable x. So, we have defined the expected value or average value this is the average value of the single random variable x by this relationship.

Now, we will define the expected value of a function of random variable, suppose y is equal to g X is a real valued function of a random variable X. Then E of g X this is even by again it is integration from minus infinity to plus infinity g x into f X of x dx so, in the case of continuous x. Now if I consider discrete X then it will be summation over x of g x into P X of x, so that way we can find out the expected value of any function any real valued function of g X. Of course, we have to be careful that this expected value is also a unique and for this will acquire these 2 conditions. So, we have defined the expected value of a function of a random variable, particularly we will discuss moments. So, moments first of all we will define moment.

(Refer Slide Time: 03:35)

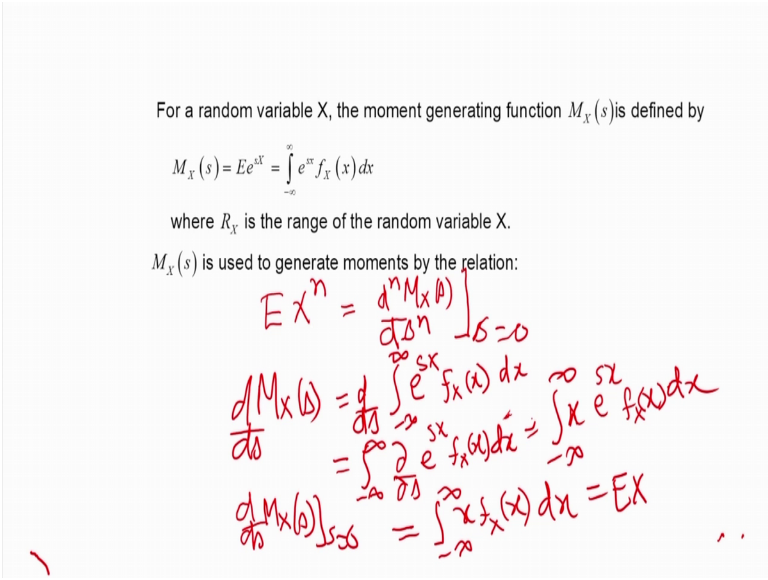


For a positive integer nth, the nth order moment is defined as E of X to the power n that is the nth order moment, this is given by integration from minus infinity to plus infinity x to the power n f X of x d x this is the nth order moment.

Similarly, nth order central moment we can define we have to consider the derivation of X from the mean X minus mu x that is, so E of X minus mu x to the power n we have to consider that is called nth order central moment and it is given by from that is integration from minus infinity to plus infinity x minus mu x to the power n effects f x of x dx, so that way we define the nth order central moment. Particularly we are more interested with E of X square that is when n is equal to 2 that is the mean square value of a random variable.

Similarly variance that is sigma x square so it is a measure of variability of a random variable, so it is given by E of X minus mu x whole square and this we can also write in terms of that is mean square value E of X square means square minus square the mean this expression we can simplify. So, that way we can define mean and variance which are 2 important parameters any of any random variable X. So, we have defined moment and we will define our moment generating function M X of s, as the name implies it is used to generate the moments. So, moment generating function is a function of a variable s, so the s is usually a real number.

(Refer Slide Time: 05:40)



So, this is the finest E of to the power s x that is expected value of e to the power s x and that is equal to integration from minus infinity to infinity it will require e s x f X of x dx, so this is the moment generating function.

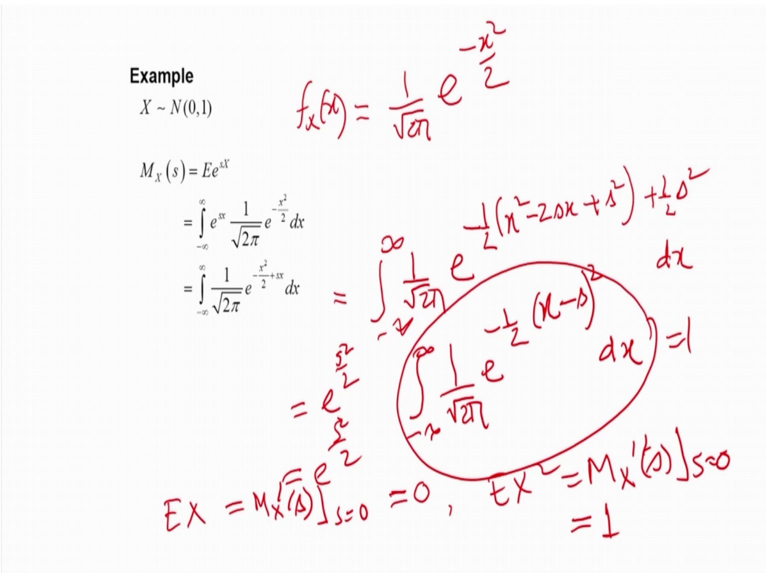
So, it is defined over this minus infinity to infinity that is the range of the random variable X. So, that we assume that this random variable X is defined over the range R X, which is from which is the real line from minus infinity to plus infinity. So, that way we have defined the moment generating function, so it is essentially a function of single parameter s and M x can be used to generate the moments by the relation. We will see how we can have the relationship that is E of x to the power n is equal to nth derivative.

So, nth derivative of M x of s at point S is equal to 0. So, this is the relationship we can find out the nth order moment by from the moment generating function by taking the nth order derivative at S is equal to 0, how we can get this relationship? Let us see E of how we can find out this relationship let us examine, we will start with suppose d dx of d ds of M x of s ok. If we see this what it will be it is the d ds of that integration from minus infinity to infinity e to the power s x f X of x d x ok.

Now this is a differentiation with respect to s, we can take differentiation inside and we will get minus infinity to infinity, so since it is a function of x and s we will take the partial derivative del del s of e to the power s x f X of x dx ok. So, this we will get as minus infinity to infinity and now this will be if I take the partial derivative with respect to s we will get x into x into e to the power s x f X of x dx.

So, therefore if I put d d x of M x of s at s is equal to 0 what we will get this one will become 1. So, we will get that is equal to minus infinity to infinity x f x of x d x, which is nothing but E of X. Continue in a similar manner we can arrive at this relationship that E of X to the power n is the nth derivative of M x of s with respect to s. Of course nth derivative with respect to s of M x of s at point s is equal to 0, S is equal to 0. So, that way this is the utility of the moment generating function from, so it and kept sorry it encapsulates all the moments. So, therefore from the moment generating function we can generate all the moments.

(Refer Slide Time: 09:58)



We can consider one example X is suppose standard normal variables standard normal random variable or standard Gaussian X, this is the notation X is normal 0 1. What does it mean? It is the normal distribution with mean 0 and variance 1, so that means in this case f X of x this is the PDF that will be equal to 1 by root over 2 pi e to the power minus x square by 2 this is the PDF, here we have to find out M x of s and that is equal to E of e to the power s x. So, if I put the expression for the expected value we will get this as integration from minus infinity to infinity of E to the power s x, we will derive the moment generating function this standard normal random variable.

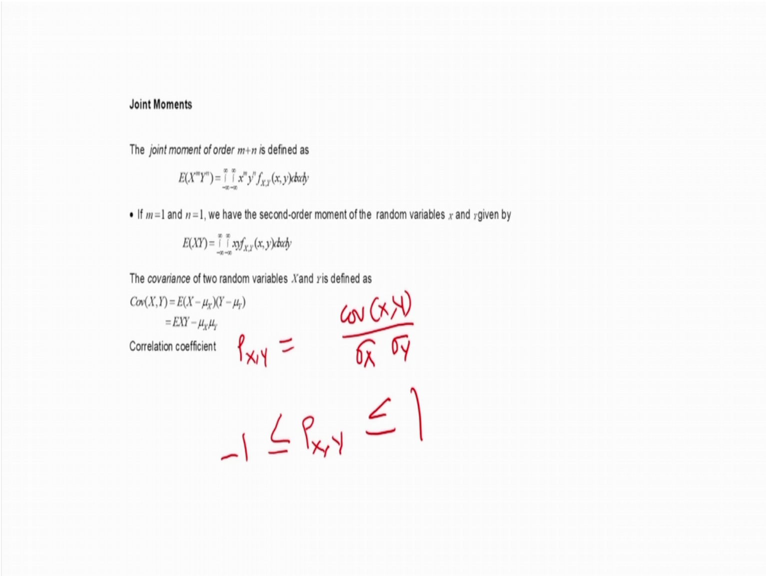
So, M x of s by definition it is E of e to the power s x. So, it is given by integration from minus infinity to infinity it e to power s x then PDF is 1 by root over 2 pi into e to the power minus x square by 2 d x, so this we can simplify like this so plus this same exponential so it will be plus s x.

Now, this we can write further s 1 by root over 2 pi e to the power if I take half common minus half common then s square minus 2 s x. So, now we can complete this to be a perfect square if we add a square here then we have to subtract, so because of that we will have another term half of s square into dx ok. So, this is the expression now let us see we can take e to the power half s square common, so e to the power s square by 2 then integration minus infinity to infinity 1 by root over 2 pi e to the power minus half of s, so what we will have x minus s whole square d x.

Now, we observe that this is a Gaussian this is a Gaussian with mean s and of course variance 1. So, if I integrate this Gaussian from minus infinity to plus infinity this integral will be equal to 1 so this integral will be 1, therefore M x of s is equal to e to the power s square by 2. So, moment generating function for the standard Gaussian is e to the power s square by 2. So, now we can use this moment generating function for example, if I take the first derivative E of x will be equal to first derivative that is M x dash of x at s is equal to 0 ok, so this will get s.

So, this will be equal to 0 similarly E of x square is equal to M x double s is equal to 0 and this will be equal to again 1. So, that way we can find out all the moments of this random variable namely this standard Gaussian. So, that way we see the utility of the moment generating function and I will be using the moment generating function in our subsequent proofs.

(Refer Slide Time: 14:22)



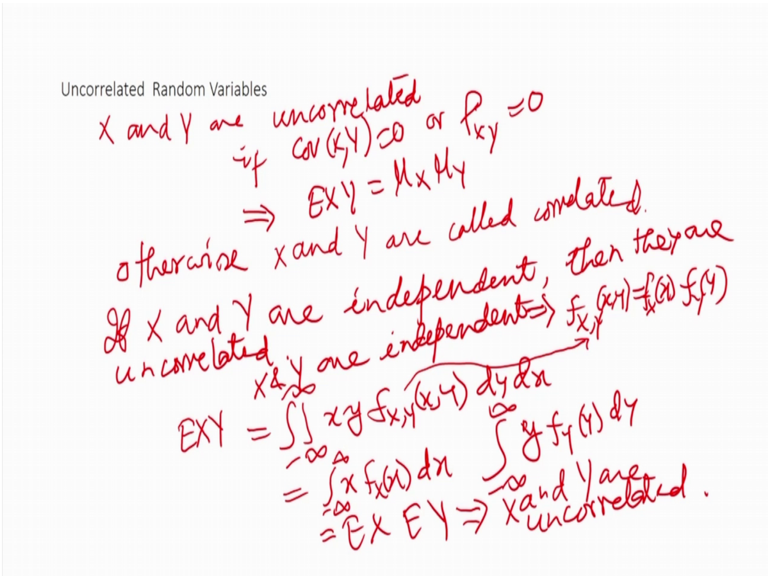
So, we define the moment for a single random variable, we can define the moments for multiple random variable, so for 2 random variable x and y the joint moment m plus n order moment, so joint moment of order m plus n. So, this is finest E of x to the power m y to the power m that is equal to again it will be double integral, then you are multiplying s to the power m and y to the power n with the joint PDF f xy at point xy and you will get with respect to dx and dy. So, this is the expression for joint moment of order m plus n.

So, now if I take m is equal to 1 n is equal to 1, we have the second order moment which is also called correlation between x and y and it is given by E of xy is equal to now here m is 1 and is n is 1, therefore it is xy into joint PDF into dx into dy integration from minus infinity to plus infinity so, that where we can define the joint expectation or second order moment of the random variables x and y. We can define the covariance of the random variable x and y, so the covariance between x and y is given by this curve of x y that is equal to u of x minus mu x into y minus mu y.

So, we consider the derivation from the each random variable x minus mu x and similarly deviation from y mu y; y minus mu y and this quantity is known as the covariance. And we can also define the correlation coefficient rho x y that is equal to covariance of x y divided by sigma x into sigma y. Where sigma x where sigma x is the standard deviation of x which is the square root of variance, sigma y is the standard deviation of y this coefficient rho which usually lies between minus 1 and plus 1. So, therefore the correlation coefficient rho x y is defined as covariance of x y divided by sigma x into sigma y, where sigma x is the standard deviation of x which is the square root of the variance of x similarly sigma y is the standard deviation of y.

Now, this correlation coefficient is an important parameter to characterize these 2 random variables x and y and we can show that rho x y lies between minus 1 and plus 1. Now once we define the covariance or correlation coefficient we can define uncorrelated random variable uncorrelated random variable x and y x and y are uncorrelated.

(Refer Slide Time: 18:04)



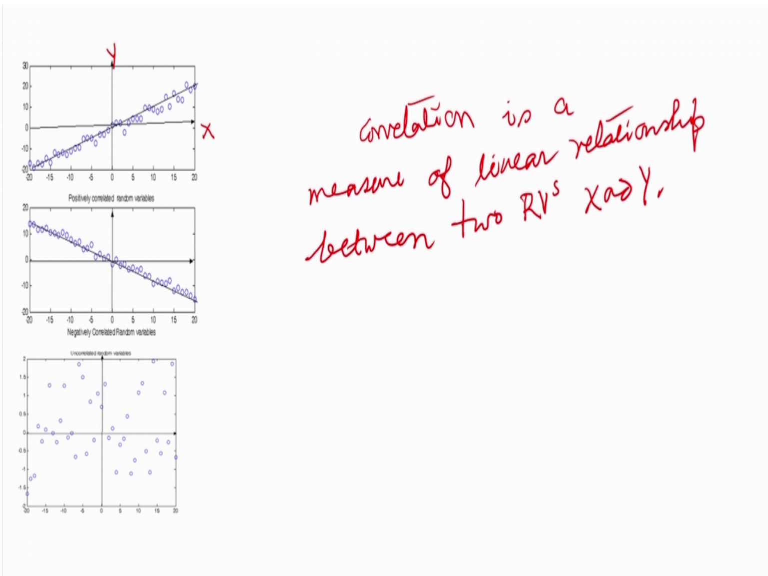
Let uncorrelated if covariance of x y is equal to 0 or rho x y is equal to 0, correlation coefficient is 0. Implies that E of x y is equal to mu x into mu y this is the relationship for x and y to be uncorrelated.

Similarly so x and y are uncorrelated if a covariance of x y is equal to 0, equivalently E of x y is equal to mu x into mu y, otherwise x and y are called correlated. So, what is correlated random variable? So, covariance is not equal to 0 we will discuss this concept of correlated random variable, but do we want to establish one relationship, we know when x and y are independent.

Suppose if x and y are independent y are independent then they are uncorrelated. But the converse is not generally to true, so independent random variables are always uncorrelated, but uncorrelated random variables need not be independent. We can consider this result we know what is x and y are independent x and y are independent imply that joint density f xy at point x y is equal to f X of x into f Y of y, joint density is product of marginal density. Therefore, this is the definition of independence and this is for all values of x and y, so therefore, we can find out E of x y for that for 2 independent random variable x and y E of xy will be equal to integration double integration from minus infinity to infinity xy into f xy at point xy dy dx ok.

Now, we can write this joint density as the product of marginal density and then this x is there y is there. So, we can write this integral in the product of 2 integral that is integration from x from minus infinity to infinity of x into f X of x dx into integration from minus infinity to infinity y into f Y of y dy. So, this is E of x and this is E of y therefore E x into E y. Therefore, the in the case of independent random variable E of x y will be always equal to E x into E y implies that x and y are uncorrelated. So, therefore if x and y are independent they will be always uncorrelated converse is not generally true.

(Refer Slide Time: 22:28)

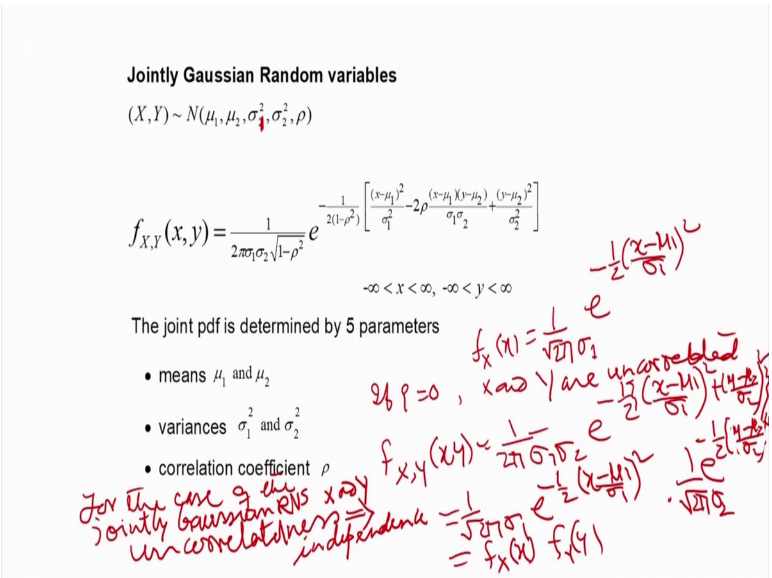


So, what does correlation implies, correlation is a type of dependence what type of dependence. If we can show that if x and y are correlated suppose this is x this is y. Now we can approximate y in terms of x by a line, so correlation means that if x and y are correlated we can have an approximate straight line relationship between x and y. So, this is these are the points suppose for a particular value of x and y these are the particulars of x and y in x y plane and this can be approximated by this line. So, this one this is an example of positively correlated random variable similarly and this slope will be determined by the correlation coefficient.

Similarly, if the correlation coefficient is negative, this is a sorry this is an negatively correlated random variable as x increases y decreases and here also we can approximate the relationship between x and y by a straight line. But when action y are uncorrelated we cannot have a straight line relationship between x and y, that is the meaning of uncorrelated nests.

So, if so that way correlation is a measure of of linear relationship between 2 random variables x and y. Now if x and y are uncorrelated we cannot approximate one by the other by plotting or by plotting a straight line approximate straight line between them. So, that way correlatedness is a special case of dependence. So, that dependence is linear dependence that x and y have linear dependence between them that can be represented by the correlation or correlation coefficient.

(Refer Slide Time: 25:13)



We will give one example of 2 dimensional random variable, this is the notation x y it is distributed as normal mu 1 mu 2 sigma 1 square this is sigma 1 square sigma 2 square and rho, so there are 5 parameters and the expression for the joint PDF is this. So, we see that these parameters this is the mean of this random variable x mu 2 is the mean of the random variable y and then sigma 1 square is the variance of X, sigma 2 square is the variance of Y and rho is the correlation coefficient, how x and y are correlated.

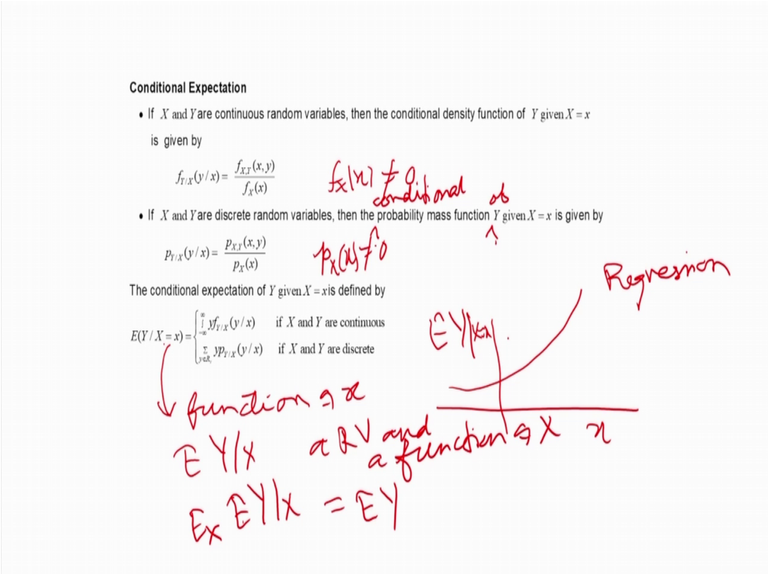
So, this is the 2 dimensional Gaussian and we see that here these 5 parameters are controlling the characteristics of these 2 dimensional Gaussian and we also observed that we can show that if suppose x and y are jointly Gaussian, then x will be individually Gaussian. So, what will the effects of X effects of x in this case it will be individually Gaussian 1 by root over 2 pi into sigma 1 into e to the power minus half of x minus mu 1 whole square. So, this is the marginal PDF of x similarly marginal PDF of y will be also Gaussian with mean mu 2 and variance sigma 2.

So, that way if x and y are jointly Gaussian, then x will be Gaussian and y will be Gaussian. But if converted not again true if x is Gaussian y is Gaussian x and y need not be jointly Gaussian. Now we see that this correlation coefficient which is m measure of correlation between x and y and we see what happens if rho is equal to 0, if rho is equal to 0 that means x and y are uncorrelated x and y are uncorrelated. So, in that case what will happen that f x y at point x y is equal to now this rho will be equal to 0, so everywhere if we put rho is equal to 0. Then we will get one by 2 pi sigma 1 sigma 2 into E to the power minus half of x minus mu 1 by sigma 1 whole square plus x minus y minus mu 2 divided by sigma 2 whole square.

So, this now we can write as 1 value 2 pi sigma 1 into E to the power minus half of x minus mu 1 by sigma 1 whole square into 1 way root 2 pi sigma 2 into e to the power minus half of y minus mu 2 whole square and this is equal to that is marginal density of x into marginal density of y. So, what we observe that if x and y are uncorrelated and jointly Gaussian, then x and y will be independent also. Therefore, what is our conclusion if x and y are uncorrelated and jointly Gaussian then x and y are independent also. Usually uncorrelatedness does not imply independence, but in the case of jointly Gaussian random variables x and y that uncorrelatedness imply independence. So, for the case of Gaussian jointly Gaussian random reverse x y uncorrelatedness implies independence ok.

So, this relationship uncorrelated generally does not imply independence, but in the case of jointly Gaussian random variable uncorrelatedness imply independence. So, that way we define the moments of the random variable particularly joint moment or second order moment covariance correlation coefficient. Also we discussed about the jointly Gaussian random variables how it is characterized by 5 parameters and what happens when the correlation coefficient is equal to 0.

(Refer Slide Time: 31:10)



Next we will discuss about conditional expectation, we know the definition of conditional PDF and PMF suppose X and Y are continuous random variable. If X and Y are continuous then the conditional PDF of Y given X is equal to x, so given X is equal to x the PDF of Y is given by this that is joint PDF divided by marginal PDF. Of course, will equal that f X of x need not be equal to 0; f X of x not equal to 0. Similarly if x and y are discrete random variables then the conditional probability mass function of Y given X is equal to x is given by this. This is the conditional PMF at point Y given that X is equal to small x that is equal to joint PMF divided by marginal PMF, here also we require that P X of x is not equal to 0.

So, once we have the conditional PDF of the random variable Y given that X is equal to small x, similarly conditional PMF of Y given X is equal to small x then we can define the conditional expectation of Y given that X is equal to x E of Y given that X is equal to x.

So, this is for the jointly continuous random variable x and y it is given by this integral relationship and for the discrete random variable x and y, this is given by this some relationship. So, that way we have defined the conditional expectation of Y given X is equal to x, we observe that this conditional expectation. Suppose for a given X what is the conditional expectation or average value of y, if we fix X then what are the likely value of Y and what are the likely values of Y and what is the average of Y that is the conditional expectation.

Given that we have an observation X is equal to small x this is a very important parameter and this is suppose if I have X what is the expected value of Y. So, we can plot suppose x versus E of y given X is equal to small x, so we can get some curve which is known as a Regression; regression of Y on x and this parameter is also. Suppose if we have to predict Y from s this is the best prediction that x is given then what is the best prediction under the means questions.

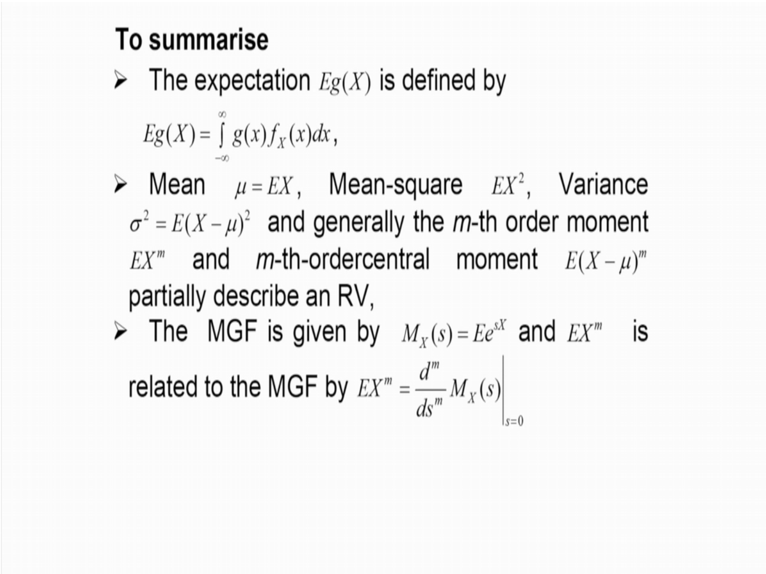
Suppose you want to minimize the mean square, here then the solution will give us that the resulting relationship is E of y given X is equal to x, so that is the best prediction of Y given that X is equal to small x. We also observe that E of Y given X is equal to x this is a function of x. So, now if I consider x to be variable because, this is for a particular X and if X if we consider X to be variable then we want this conditional expectation E of Y given X.

Because, X will be taking different values and this quantity is a random variable, of course this is a random variable and this can be used to for example to find out the expected value of Y. So, that way this is a random variable a random variable and it is distribution will be with respect to suppose if X is fixed then it will be a PDF with respect to Y.

So, E of Y given X is a random variable and a function of X is a random variable it is a random variable and a function of X a function of X random variable X. Therefore, we can take the expected value of this quantity with respect to X E of X Ex of E of y given X. So, when we take the expected value here it is a random variable is Y given X suppose is equal to suppose small x, then we can take the expected value with respect to Y. Now we will take the expected value with respect to X this will be equal to E of Y, so this is one important relationship.

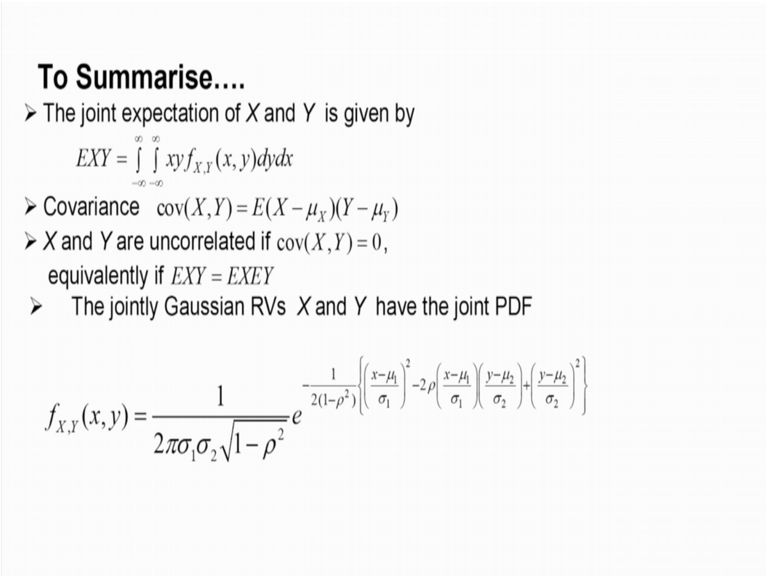
So, that way we will using the concept we will be using the concept of conditional expectation; conditional expectation it is a measure of based linear prediction given that X is equal to small x. So, we discussed about conditional expectation E of y given that X is equal to small x and also we discussed that considering s to be variable it is a function of X and from this expected value if we take the expectation again with respect to X, we will find E of y and also the plot of E of y given X is equal to x versus x. So, that is the regression of Y on X and this value suppose for a given X what is the base prediction under some sense, under the mean minimum mean square regression that is given by the regression curve. So, that way this condition just like conditional PDF and PMF the conditional expectation is also important quantity and that we will be using in our subsequent studies.

(Refer Slide Time: 38:43)



Let us summarize the lecture first we define the expectation for a function g x E of g X is defined by this integration, so basically g x into f X of x integration from minus infinity to infinity. Provided of course this g x f x should be absolutely integrable then we define different moments. For example, mean mu is equal to E of x mean square E of X square variance sigma square is equal to E of X minus mu square and then generally mth order moment E of x to the power m mth order central moment E of X minus mu to the power m. So, these moments partially describe the random variable unlike CDF which completely described a random variable the moments partially describe the random variable. Next we discussed about Moment Generating Function the MGF is given by M x of s that is the expected value of E to the power s x and the utility of the moment generating function is that we can find out the moments from the MGF. So, that way E of x to the power m is related to MGF by this relationship that is g of x to the power m is the mth derivative of M x of s at point s is equal to 0.

(Refer Slide Time: 40:22)



We also talked about joint expectation, suppose joint expectation of X and Y E of XY and that is the integration from minus infinity to infinity from minus infinity to infinity x y into joint PDF dy dx. So, this is the integration double integration of the product of XY and joint PDF. So, this E of xy is an important quantity, so it partially describes the random variables X and Y. We also define the covariance, covariance of XY it is given by the expected value of X minus mu X into Y minus mu Y and we are we discussed about the importance of covariance. How it expressed the linear relationship between 2 random variables X and Y, also X and Y are called uncorrelated if covariance of XY is equal to 0 and that also equivalently we can say that if E of XY is equal to EX into EY.

We define jointly Gaussian random variables X and Y, this randomly joint Gaussian random variables X and Y have the joint PDF given by this. So, it is determined by it is an exponential quantity and it is determined by 5 parameters namely mu 1 that is the mean of x, mu 2 that is the mean of y, then sigma one variance of x sigma 2 variance of y and rho rho is the correlation coefficient. That also we define defined that correlation coefficient is the covariance normalized by sigma X into sigma Y.

Thank you.