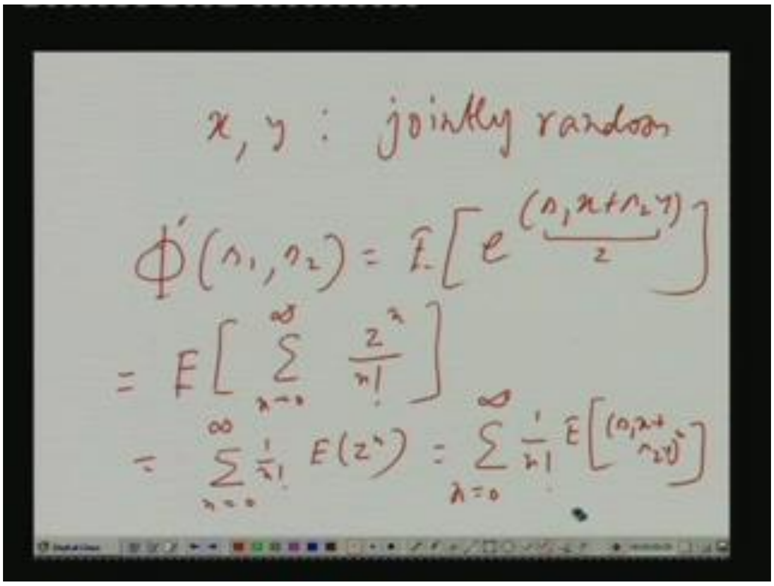
Probability and Random Variables
Prof. M. Chakraborty
Department of Electronics and Electrical Communication Engineering
Indian Institute of Technology, Kharagpur

Lecture - 19
Joint Conditional Densities

So, in the previous class, we were discussing this joint moments, rather this moment generating function. We start from there again. There will be little overlap, but there is nothing wrong if there is a little overlap with the last phase of previous day’s lecture.

(Refer Slide Time: 01:16)

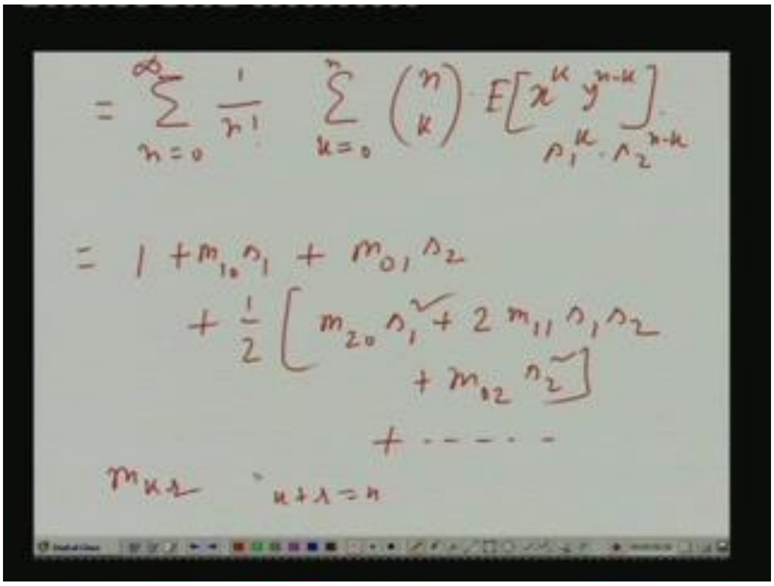


So, we are given two random variables as before: x, y – jointly random. Then, we define this function phi; maybe you can put a phi prime. I will tell you why I am putting phi prime. s 1 s 2 as the expected value of e to the power s 1 x plus s 2 y. Now, listen; earlier, we had a situation, where s 1 was equal to j omega 1 s 2 was j omega 2; and the entire thing was called the joint characteristic function. It was written as phi of omega 1 omega 2. Since j omega 1 is replaced by s 1 here and j omega 2 is replaced by s 2, and j is missing on this side; I am giving it a new name – phi prime. That is the only difference. This is called the moment generating function. We have already seen what is the moment – joint moment; that is, x to the power r, y to the power k; its excepted value of this product is called the joint moment of order k plus r equal to say n. Now, this function will help us in getting those moments of various orders.

To understand that, let us first do this. This is an exponential. So, s 1 x plus s 2 y – we can even call it Z. So, e to the power Z. You expand e to the power Z into a power series. We all know what that power series is; there will be a summation of terms – an infinite summation actually. And then expectation is a linear operator. You can apply expectation on each of the terms in the summation separately. If you do that; that is, first, you have expected value n equal 0 to infinity – Z to the power n factorial n. This is the exponential series. And then I will apply this expectation operator on each term in this summation. Factorial n is a constant.

So, it remains outside. Then, what happens? That is E… Now, we know what is Z. Z is this factor. So, replace Z by its actual form s 1 x plus s 2 y whole to the power n. Now, s 1 x plus s 2 y whole to the power n is actually a binomial series of n plus 1 terms. Again, that can be… So, I can expand this term into this binomial series. So, I get another summation; but, unlike the outer summation, that will be a finite summation now, because there are total terms equal to n plus 1. And again, expectation can be brought inside that summation and can be applied on each term present in the binomial expansion.

(Refer Slide Time: 04:50)



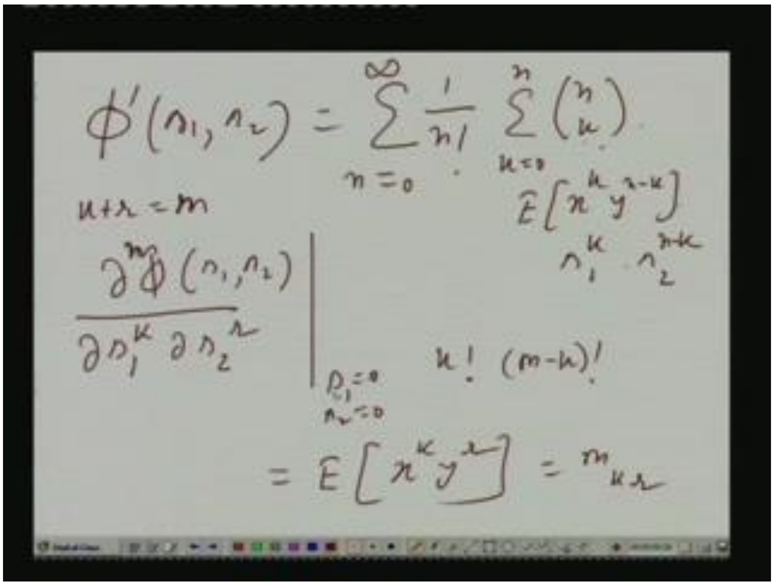
If I do that, what we get? Is this... If there will be k. So, k should be from 0 to n. Total number of terms is n plus 1. We know that binomial series n k here – factorial n divided by factorial k into factorial n minus k. Expectation of a pointer does not work in it, because this is not random. Then, there will be this product – s 1 x whole to the power say k s 2 y whole to the power n minus k; out of which, s 1 and s 2 can be separated out. So, we get E working on x to the power k y to the power n minus k; and again, s 1 to the power k – s 2 to the power n minus k. You see this moment is coming here. x to the power k, y to the power n minus k, its expected value. Summation of the two powers is equal to n. So, it is a moment of n-th order. Also, note one thing; this is a summation over n in 0, n equal to 1 and n equal 2 up to n equal to infinity.

Now, for each n – for a particular n, we have got this summation, where total terms – total number of terms is n plus 1 first. And then remember – for that given n, powers of x and powers of y in this summation – they should be such that, the summation of those two powers is always equal to n. So, if n is 4, you can have 0 here, 4 here, 1 here, 3 here, 2 here, 2 here, 3 here, 1 here and 4 here, 0 here. You got everything else. So, this is one thing you please remember, because subsequent derivations will need this property. So, for a particular n as chosen from outside, we will have only those powers of k and n minus k, those powers of x and those powers of y. So, that summation of those two powers is always equal to n. So, this is what we get as the moment generating function for x and y.

Now, I erase this. As I told you, this n-th order moments are present. So, maybe if we expand Z little bit, you will see the various moments are coming up. For instance, we start with n equal to 0 first. So, 1 by factorial n is 1; n equal to 0 means only one term in this summation, that is, k equal to 0. So, that will give rise to what? Basically, that will give rise to 1 – s 1 – x to the power 0, y to the power 0. And then s 1 to the power 0, s 2 to the power 0, and this is 1 here, so that 1; and 1 by factorial 0, that is, 1. So, first term for n equal to 0 is 1. Then, take n equal to 1. n equal to 1 means in this inner summation, I have got two terms: one is for k equal to 0; another is for k equal to 1. So, for k equal to 0 first, we have got this term – s 1 to the power 0, that is, 1; but, s 2 to the power 1. So, s 2. Similarly, here E of y. s to the power 0 is 1. So, E of y. And 1, 0; that will give rise to 1 only. So, you will get one term as E y s 2. And E y is… Similarly, now, put k equal to 1. So, x to the power 1, but y to the power 0, so E of x only. Similarly, s 1 to the power 1, but s 2 to the power 0, so only s 1. And again here you get 1; and outside also 1 of course. So, this time you get…

Now, take n equal to 2. For n equal to 2 here, we have – in this summation, we have got three terms; k equal to 0 first; or, maybe we start with k equal to 2; then 1; then 0. If k equal to 2; 2 here and 2 here means actually it will give rise to 1. And x square y to the power 0. So, E x square. Similarly, s 1 square, but s 2 to the power 0. So, this will give rise to… And of course, 1 by factorial 2, that is, 1 by 2. So, that will give rise to what? 1 by 2; maybe we can put in a bracket – E of x square, that is, m 2 0 s 1 square. Then, k equal to 1; k equal to 1 means 2 here, 1 here. So, factorial 2 divided by factorial 1 into factorial 1. So, that will give rise to 2. So, we will have a term – 2. Then, E of x to the power 1 y to the power 1. So, E x y, which is m 1 1. And again s 1 to the power 1 s 2 to the power 1. Then, k equal to 0 will give rise to what? 2 0 here will be give rise to again 1; factorial 2 divided by factorial 2 into factorial 0, that is, 1. And E of y square s 2 to the power 2. So, that will give rise to m 0 2 s 2 square plus so on and so forth. So, you see all the moments are present here. Suppose given this moment generating function, we have to find out m k r; that is, if k plus r is equal to say some n; then this is n-th order moment m k r how do you about it?

(Refer Slide Time: 11:20)



So, for that, I rewrite that term again that, phi prime s 1, s 2. Now, you see one thing; if I differentiate this guy; if k plus r is equal to some n and if I differentiate this n times; out of which, with respect to s 1 k times and s 2 r times; and then after differentiation, put s 1 equal to 0, s 2 equal to 0. What do you get? Now, let us see how to proceed. First, since k plus r equal to n, consider in this summation, n equal to… See if you are confused; if you want rather; if you want, I can make it m. So, if it is m… In this case, first, consider n equal to m; after all, n is 0, n is 1, n is 2 dot dot dot up to infinity. So, you it can m also. So, take the case n equal to m. So, when n is m from outside; here I have got a binomial series; where, k starts from 0 goes up to m; total number of terms is m. Here this summation of the two powers: power of x, power of y – that is always equal to m then – m. That is always is equal to m, because I am talking from outside; I am taking the particular case, where n equal to m. So, n equal to m here and I have got a binomial series. So, E of x to the power k y to the power m minus k.

Now, similarly, s 1 to the power k s 2 to the power m minus k. Now, suppose I first differentiate it with respect to s 2; how many times? r times; r means m minus k times. So, if you differentiate this factor m minus k times, what happens? You do not get any power of s 2. Each differentiation leads to reduction of power by 1. First time we differentiate, power becomes m minus k minus 1; next time, m minus k minus 2; so on and so forth. So, if you differentiate it m minus k times; then you finally, get s 2 to the power 0, that is, 1. But, a factorial of m minus k comes. So, that will give rise to m minus k factorial; everything else remains same in that particular term.

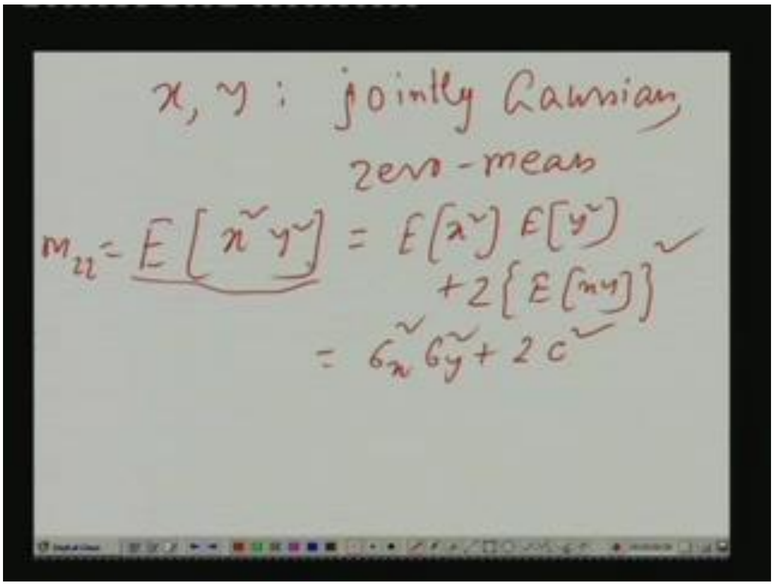
And now, if you differentiate this with respect to s 1 k times; again, you get s 1 to the power 0 in the end, that is, 1; but, a factorial k comes. There are other terms also in this series first. You are differentiating them. But, remember – in this series, either you have got this situation, where this is k and this is m minus k. If you take some other term, may be the power here is higher; but, power here is less. See if you differentiate this s 2 to the power m minus… Suppose you consider for example, other term; s 1 to the power k plus 1 s 2 to the power m minus k minus 1. If you consider that term and again differentiate it with respect to s 2 m minus k times; obviously, you get 0 here, because power is less. Power then… because now you have got s 1 to the power k plus 1; but, s 2 to the power m minus k minus 1; whereas, we are differentiating it m minus k times. So, obviously, this will give rise to 0. So, product will be 0.

On the other hand, if you take a term s 1 to the power k minus 1 and here s 2 to the power m minus k plus 1. So, this power here is higher by 1; for s 1, power is less by 1; that is, power is k minus 1. And now, differentiating with respect to s 1 k times, again, this factor will become 0, because power is k minus 1, but you are differentiating k times. So, you understand that, other terms after differentiation becomes 0. You even do not have to put s 1 equal to 0. S 2 equal to 0. Then, further, I took n equal to m. If it is n equal to m minus 1 or m minus 2 or m minus 3 and likewise; wherever it is obvious, that powers are less, and differentiation will give rise to 0.

I will consider the case, where n equal to either m plus 1, m plus 2, m plus 3; that is, when n takes higher values. There you will see; there you will get some – such terms like this. Maybe consider the case, where n equal to m plus 1. So, this becomes m plus 1 minus k and this is s 1 to the power k. So, I agree… I mean here if you differentiate this product with respect to s 1 k times and s 2 m minus k times as here, this product will not become 0. There will some term either s 1 or s 2 or may be some power of that. But, then we substitute s 1 equal to the 0, s 2 to equal 0. That will eliminate them. So, you understand that, even if I take n equal to m plus 1 or m plus 2 or m plus 3 whatever terms and consider the binomial series, do this differentiation; upon differentiation, if I substitute s 1 equal to 0, s 2 to 0, all terms varies. Only when I am considering this n equal to m and the corresponding binomial series here and then differentiating every term with respect to s 2 r times and s 1 k times; where k plus r equal to m. So, then that is, s 2 to the power m minus k, which is r. So, s 2 to the power r, s 1 to the power k; this is differentiated; s 2 to the power k differentiated; s 2 to the power r differentiated r times; s 1 to the power k differentiated k times. So, both give rise to 1 and 1 and factorial k, factorial m minus k; but, no s 1, no s 2. So, it does not matter whether I have to put s 1 equal to the 0 or s 2 equal to 0. So, only that term remains. And these two factorials have come up: factorial k, factorial m minus k

Now, consider n k here. After all, what is this? Factorial n. That will be canceled with factorial n here from outside. And in the denominator, you have got factorial k; which will cancel with this factorial k. Another term is factorial m minus k; which will cancel with… I am considering n equal to m; let me make the correction. So, it is m k here. So, factorial m; and n equal to m is considered. So, factorial m here. So, factorial m, factorial m cancels. In the denominator here, we have got factorial m minus k and factorial k; this cancels with this factorial k, factorial m minus k. So, what I am left with is basically this term: E of x to the power k, y to the power r; where, k plus r equal to m. So, this will give rise to… That is, m k r; where, k plus r equal to some m. This m stands for moment; it is not the m that I was talking of earlier. So, this is… This shows that, giving this function – phi s 1 comma s 2, if you are interested in finding out a moment m k r, then you differentiate this with respect to s 2 r times and with respect to s 1 k times; then upon differentiation, whatever you get, substitute s 1 equal to 0 there and s 2 equal 0 there. So, we have proved this result.

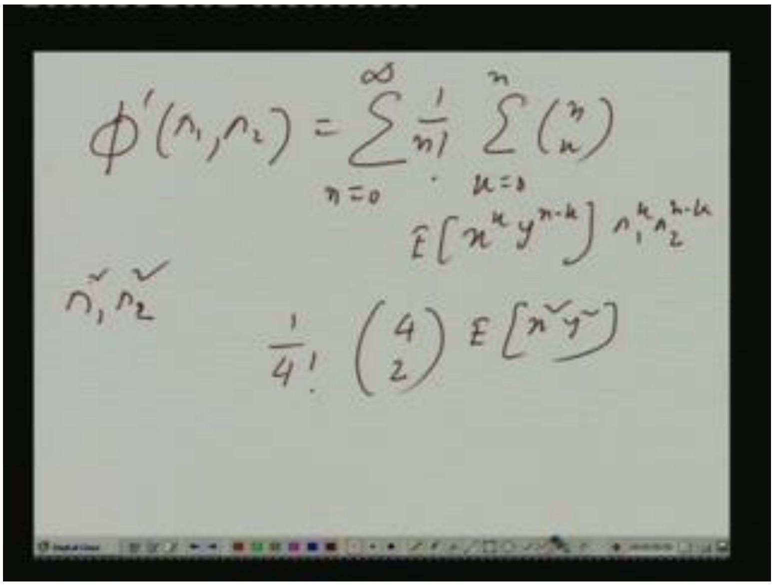
(Refer Slide Time: 20:50)



Now, let us see and let us consider an example. Suppose it is given that, x and y – they are jointly Gaussian, zero-mean. You have to find out… The question is show that, E x square y square, which is nothing but m 2 2. It is nothing but E x square E y square plus twice E x y whole square; where, you can also say equivalently sigma x square. After all, mean of x is 0. So, E of x square means E of x minus 0 whole square; that is, variance of x – sigma x square sigma y square here. And obviously, this is the covariant C, because what is covariance? Expected value of x minus mu x times y minus mu y. But, both the means are 0. So, it turns out to be expected value of just x, y. We can call it C. So, this is twice C square. We have to prove this.

Now, one thing is this that, since it is m 2 2, what you can do; you can simply apply the previous formula; that is, first, find out phi prime – the function phi prime as a function of s 1 and s 2. And then 2 2. So, k equal to 2; r equal to 2. So, k plus r, that is, 4. So, differentiate that function, which is s 1 two times, s 2 two times. Whatever you get; now, put s 1 equal to 0 in that expression, s 2 equal to 0 in that expression. That will give you this. But, that will be little cumbersome. So, we follow just… It will more… An approach, which is slightly more different – slightly more direct. Let us see what it is.

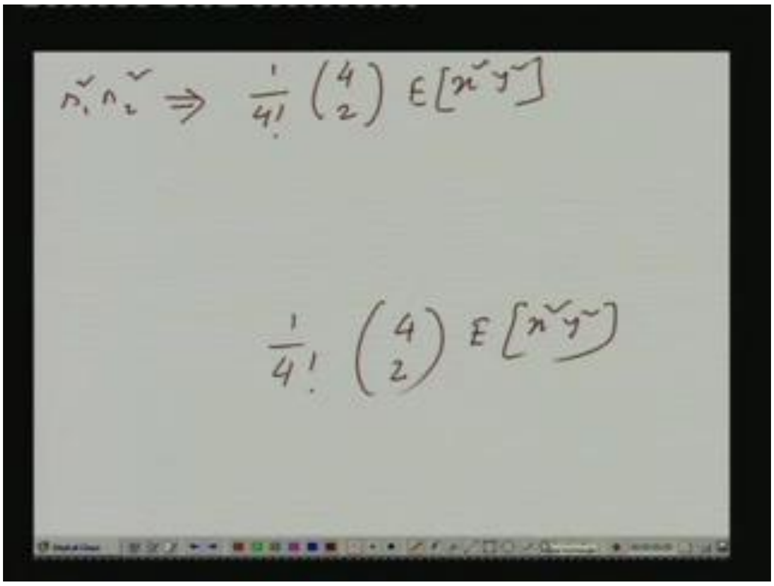
(Refer Slide Time: 23:11)



As you know, we have already seen that, phi prime… This we have already seen. And I am interested in finding out E x square y square; that means k should be 2. And n minus k should be 2; that is, n minus 2 should be 2. So, n should be equal to 4. So, in this summation, from the outer summation, I pick up the particular case for n equal to 4; where, n equal to 4. And the corresponding term… For n equal to 4, you consider this binomial series; they are for one term, I get E of x square y square. For what k? k equal to 2. So, first, n equal to 4 I fix. So, k equal to 0 to 4. I have got a binomial series of five terms. Out of the five terms, I consider k equal to 2 case; if k equal to 2, then only you have got x square y square. And therefore, we have got it s 1 square s 2 square. So, you can say… Remember this s 1 square, s 2 square – this combination – this comes up only in this case; where, n equal to 4 and then k equal to 2.

If n equal to 5, you would not have… If suppose n equal to 5, you have got another binomial series from k equal to 5. But, there are six terms. But, in no term, you will have a situation like s 1 to the power 2 s 2 to the power 2. Simply because the sum of the two powers must be then equal to what? Since n equal to 5, it should be 5. But, when it is s 1 square s 2 square; that means total power should be equal to 4; n should be equal to 4. So, this combination occurs; only this combination: s 1 square s 2 square – this occurs only when n equal to 4 first; that is, 2 plus 2 is 4. Then, there you take k equal to 2. So, one is 2; other one is 2. So, I can say that, in this infinite series, I have to find out the term, find out the coefficient associated with this term – s 1 square s 2 square. What is the coefficient? At first, we have said n should be 4. So, that will be 4, k equal to 2, and from outside, 1 by factorial 4 and then E x square y square. So, I erase this.

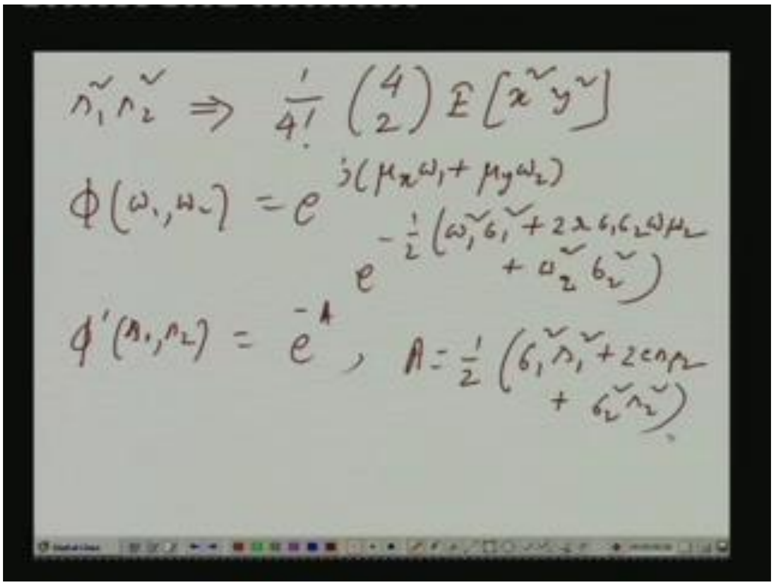
(Refer Slide Time: 27:03)



I rewrite it here. So, s 1 square s 2 square gives rise to this term. ((Refer Slide Time: 27:23)) our interest is in this thing in this expected value; that is, E of x square y square. We are interested in this. On the other hand… So far, I have not used the fact; they are zero mean – jointly Gaussian with mean zero. I have only taken the general expression for phi prime s 1 s 2. This is general expression whether x, y or jointly Gaussian or not; whether they are zero mean or not, this is always true that, in that moment generating function, that particular term that was s 1 square s 2 square – that will have this coefficient. So, I have just written down the coefficient. But, now, I come to this business that, x and y are jointly Gaussian and mean zero.

Now, we know; we have already seen earlier that, in such case, what is phi s 1 s 2 phi prime. In fact, we have found out the characteristic function in such case. And phi prime s 1 s 2 square is quiet close to that; j omega 1 and j omega 2 are to be ((Refer Time: 28:46)) by s 1 s 2. That will give rise to this. If I can remember, in case you have forgotten, just remember… As we have to rewrite it here, because that page has gone from here. So, just bear with me; there is some problem here. So, I just have to write down once again.

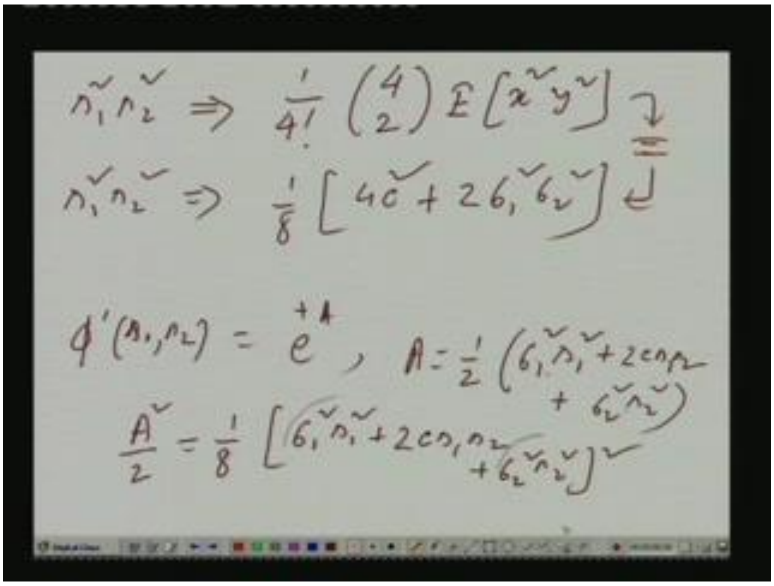
(Refer Slide Time: 29:19)



s 1 square s 2 square; that was given rise to this coefficient. Now, we are considering this fact that, x and y are jointly Gaussian with mean zero. Now, earlier, we have seen the characteristic function – joint characteristic function for such random variables. What was that? Phi omega 1, omega 2; we have seen it earlier. This is equal to e to the power j mu x omega 1 plus mu y omega 2 times – again e to the power minus half; and. no j here – omega 1 square sigma 1 square plus twice r sigma 1 sigma 2 omega 1 omega 2 plus omega 2 square sigma 2 square. So, you can write down what is phi prime s 1 s 2. Simply… If you remember this moment generating function, the formula for moment generating function is nothing but characteristic function, where j omega 1 was to be replaced by s 1, j omega 2 to be s 2. So, what does it give rise to? And in any case, mu x and mu y are given to be 0 here. So, e to the power 0; this term goes; this is 1. And here we get e to the power… So, j omega 1 is s 1. So, s 1 square is minus omega 1 square. So, minus omega 1 square. s 1 square is minus omega 1 square, just a minute. So, I directly write down in fact.

As the book says, this is e to the power minus A; where, A is… But, twice r sigma 1 sigma 2 is nothing but… This r into sigma 1 sigma 2 is nothing but the covariance, because r is correlation coefficient. So, you can say this is covariance C. So, you can say twice C s 1 s 2 plus – plus what? s 2 square. There is sigma 2 square s 2 square. Now, what we do is very simple. This e to the power minus A – that is an exponential series. So, it will be having terms like 1 plus A plus A square by factorial 2 plus A cube by factorial 3 plus dot dot dot dot dot. In this series, I will find out where we get this term like s 1 square s 2 square. Take out the coefficient; equate that with this coefficient. That will give as a result. So, let me erase this.

(Refer Slide Time: 33:38)



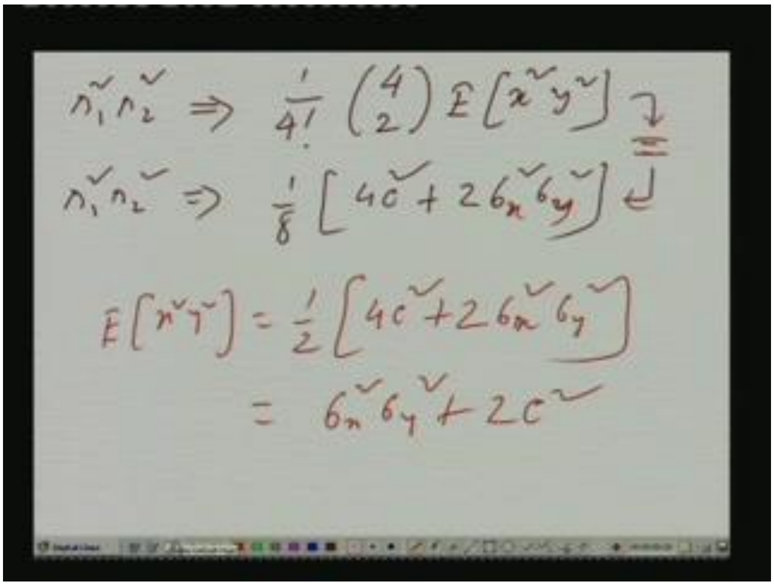
You see one thing; first term is 1; forget that because that has nothing to do with s 1 square s 2 square. Second term is A; A means this term. But, it has either s 1 square or s 2 square, but s 1 s 2; no s 1 square s 2 square together. So, forget that term too. Then, s square by 2. Now, in s square by 2, you have the square of this. The square of this… Square of this term comes. So, there is one case, where you have got s 1 square s 2 square and a product of these two. That also has s 1 s 2 square. So, in A square by 2, I get this s 1 square s 2 square term.

How about A to the power 3? In A to the power 3, of course, you would not get it, because obviously, you see either this is s to the power 3, it is s to the power 3; or, square of this times this or square of this times this. You would not get it; you can easily see you would not get s 1 square s 2 square. So, only A square by 2 is where you get this term – s 1 square s 2 square. What is A square by 2? Remember how are we getting A square by 2? That is, I expanded e to the power minus A into its power series. In fact, it should be to the power plus A; you remember, because in the characteristic function, we had terms like minus omega 1 square minus omega 2 square. And if s is j omega 1; then s square is minus omega 1 square. This raised to s 2 is minus omega 2 square. And s 1 s 2 is minus omega 1 omega 2. So, it should be actually plus.

Now, s square by 2 comes as the third term of the exponential series. First is one; forget it; it has no such term – s 1 square s 2 square. Second is A. Again, that has got no term like s 1 square s 2 square; forget it. Only the second term… This is the third term, which is A square by 2, that we have expressions like s 1 square s 2 square. A square by 2 means 1 by 8, 1 by 2 and whole square of this 1 by 4; 1 by 4, 1 by 2, 1 by 8 times this whole square.

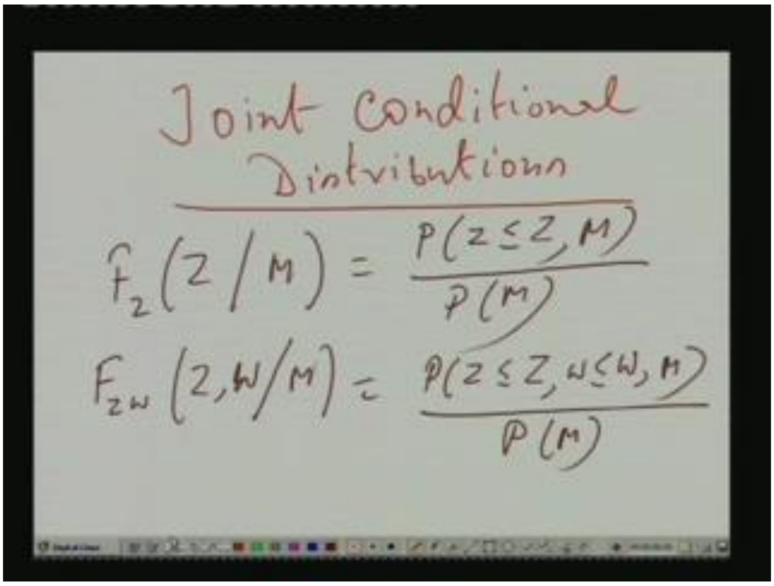
In this expression now, just find out when, on what occasions you get s 1 square s 2 square? First whole square of this – that will give rise to s 1 to the power 4; forget it. Whole square of this; that will give rise to s 2 to the power 4; forget it. But, whole square of this – that will give rise to 4 C square s 1 square s 2 square. So, I will take it. 4 C square s 1 square s 2 square. So, this time I am getting 4 C square by 8; 1 by 8 is for 1. So, one term is 4 C square. And there will be a term like twice this and this. That will gives rise to s 1 square s 2 square, but twice sigma 1 square sigma 2 square. Twice sigma 1 square sigma 2 square. So, this is what I will have with s 1 square s 2 square in this term A square by 2. In A square by 2, there will be several terms; but, s 1 square s 2 square will have this coefficient. So, these two are to be equated. These two are to be equated – this one and this one.

(Refer Slide Time: 38:16)



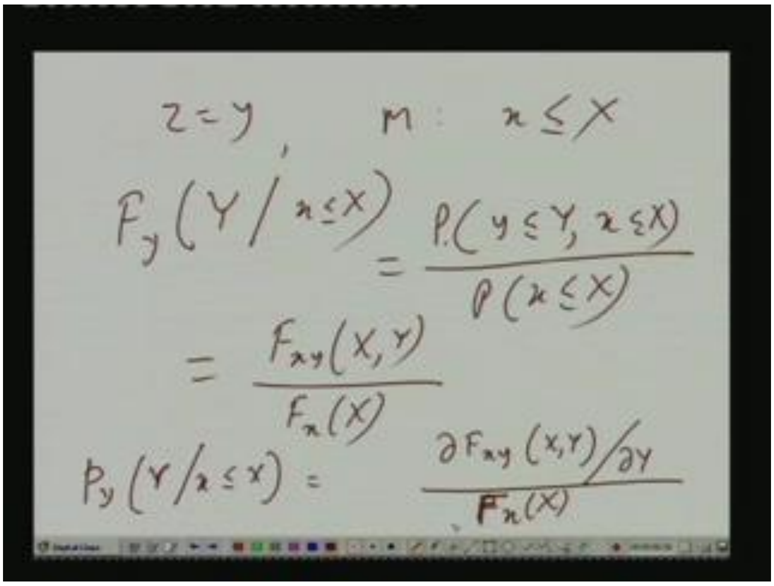
If you equate this, you directly get the result. After all, you get factorial 4 here; that cancels with factorial 4. In the denominator, you have got factorial 2; factorial 2, which is basically 4. So, 1 by 4 and here 1 by 8. So, it gets cancelled. So, you get only 1 by 2; which means E x square y square is 1 by 2 four C square plus twice sigma 1 square. Sigma 1 or rather sigma x sigma y – sigma x square sigma y square; we can write sigma x square sigma y square plus twice C square. This is what we wanted to prove. So, we have proved it. So, that is all about this moment generating function.

(Refer Slide Time: 39:06)



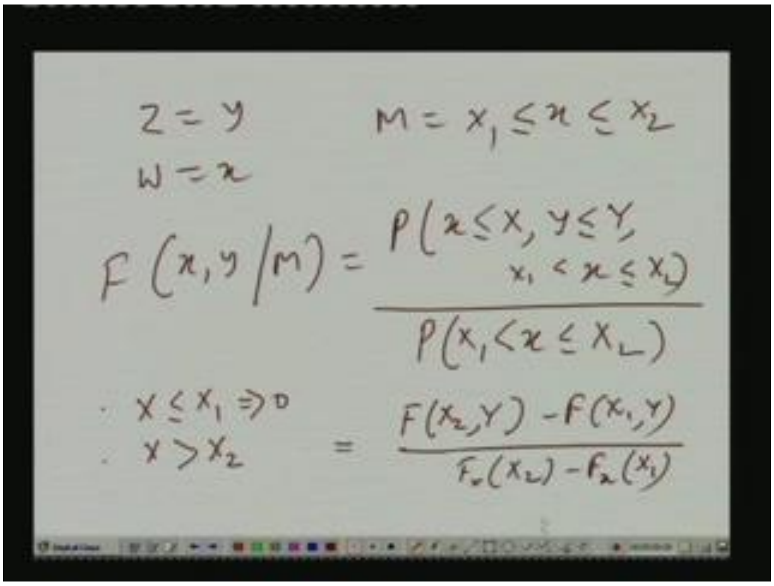
Another thing I want to do, that is, joint conditional distribution; that is, either Z is a function of x, y. There is a function of – single function of two random variables. We have got maybe two functions z and w – two different functions of two random variables. If both case, you can say that, F z – some capital Z subject to the event M is same as probability of z taking value capital Z less than equal to capital Z along with M. So, this is joint event; both M taking place and this event taking place. There is z taking value less than equal to capital Z divided by probability of the event M. So, this is the case for a single function of two random variables. And you have got… Where you have got two such functions: z and w, capital Z, capital W by m with the same token will be this joint density.

(Refer Slide Time: 41:05)



Let us take some example. Suppose z is y. z is a function of x and y, but z is supposed to be y. And that event M is nothing but x taking values less than equal to some capital X; that means we have to find out P y… Here you have to find out this conditional thing. This is nothing but probability of this joint probability. This is y taking values less than equal to capital Y, x taking values less than equal to capital X divided by the probability of x. The probability… This is the joint density we know. This is the joint distribution F x y; you can say X, Y; and this is F x X. So, how about the density? That is more important. Density is not simply derivative of this with respect y. So; that means subject to this is nothing but… We modify this example; make it slightly more general now.

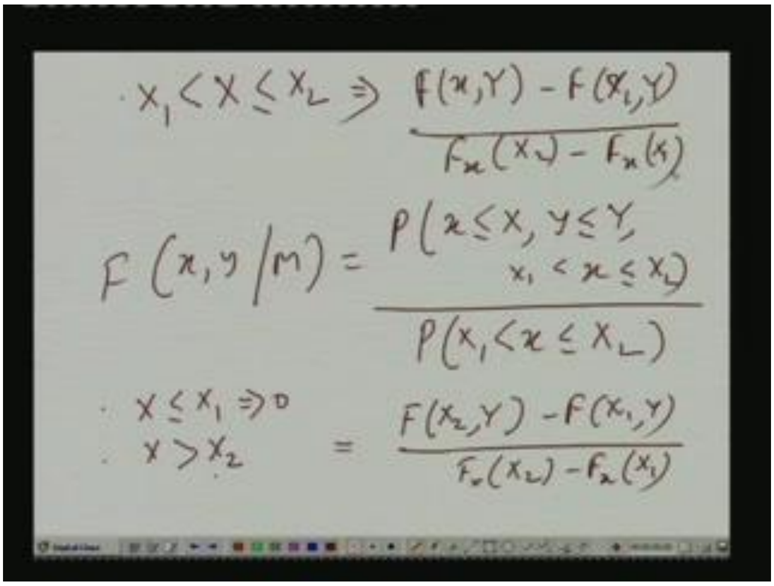
(Refer Slide Time: 43:52)



Now, two functions… z is y, w is x and M is this thing; x less than equal to some x 2, x 1. So, here what is subject to this M? This is nothing but probability… It is joint probability of x… Now, you see one thing… Suppose we choose x; we have got two limits; we have got two limits: one is x 2, another is x 1. Then, this capital X is such that it is less than x 1. This is x 1 limit. So, the two things are jointly occurring; x has to be less than equal to some given capital X. At the same time, x should be greater than x 1 less than equal to some x 2. Maybe this capital X is less than or equal to x 1; then two things cannot take place together; that x is simultaneously above capital X 1 and less than equal to X 1; that is not possible.

So, if x is less than equal to x 1, this is 0. Similarly, if x is greater than x 2, then what happens? If x is greater than x 2, then x is less than equal to x 2; x is less than equal to a capital X; and x is less than equal to capital X 2. Out of these two, I will consider this condition, because capital X 2 is less than x. So, in that case, it is simply because joint probability of y falling less than equal to capital Y and x lying within this range. That will be nothing but what? The distribution X 2 Y… And divided by this is of course, F x X 2 minus F x X 1. I am nearly through; just one more expression; and that will be over for today.

(Refer Slide Time: 47:51)



And, the other case – the other case is if x is lying within this range. If x is… Then, obviously, I will take small x to be greater than X 1, but less than equal to capital X. On the lower side, leave it as x 1; but, on the upper side, it will go up to capital X as coming from here. In that case, it will be similar to the previous one. Now, to find out density, you will see one thing that, in one case, we have got 0. So, density will be 0 because derivative after all. In the other case also, if you find out the… This is a constant thing. After all, joint density means what? You have to take partial derivative – del square F del x del y; there is no x here; it is constant. So, derivative will be 0. Only here there is x; only here there is x. So, only here you will get this density. And what will that density to be? This is independent of x. And here del square F del x del y will give rise to p of x, y. So, here p of x, y divided by whatever we have here – this constant. In the other two cases, it will be 0. So, I stop here today. That is all for today. And in the next class, we start from here.

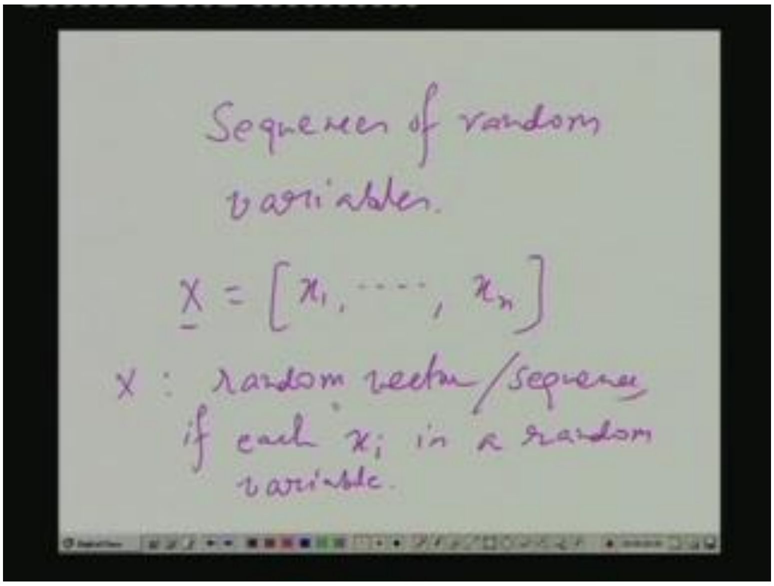
Thank you very much.

Preview of Next Lecture

Lecture - 21

Sequences of Random Variables

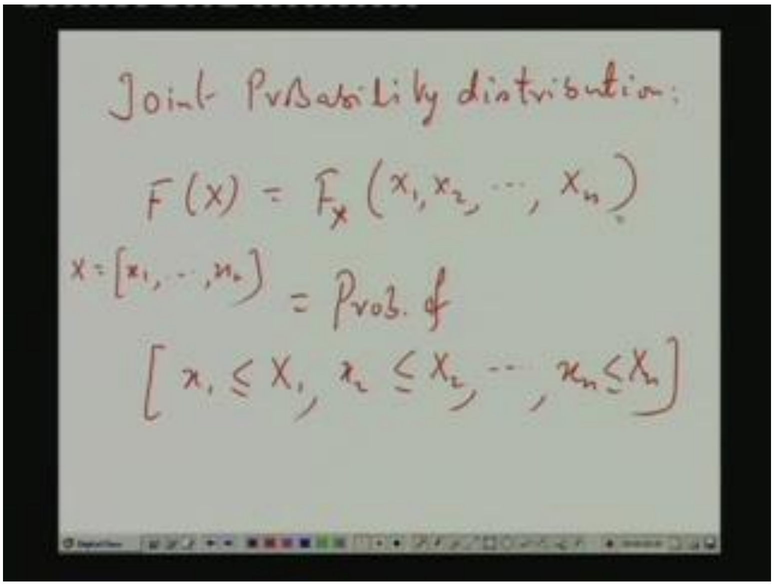
(Refer Slide Time: 50:06)



So, today, we discuss an important topic, that is… In fact, we start this topic – sequences of random variables. You see we started with first – a single random variable. I will consider a function of a single random variable. Then, we generalize it to two random variables. So, you can even say that, it is a sequence just of two elements – two random variables – two-element sequence. And we consider a function of two random variables;

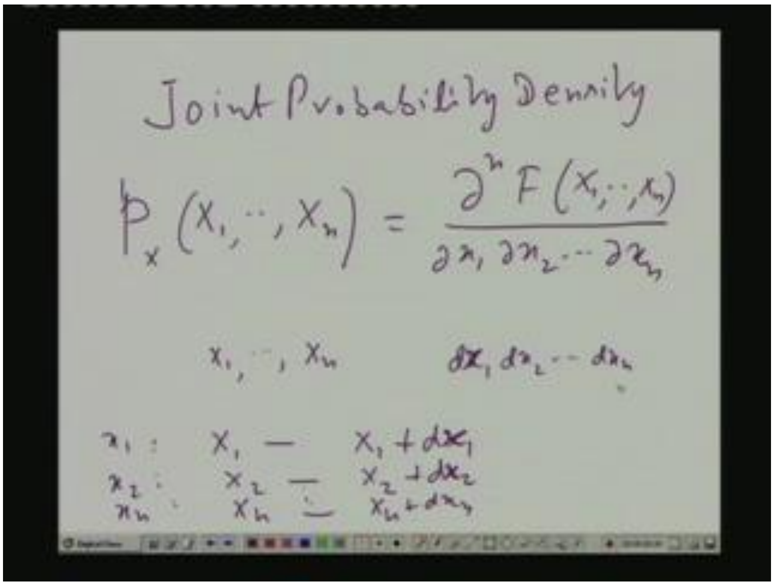
then two functions of two random variables and all that. That whole treatment will now be generalized, where we will be considering a vector say X as x 1 dot dot dot dot x n. This X will be called a random vector if each element x 1 to x n is a random variable; that is, X will be called a random vector or say sequence if each x i is a random variable. So, there you see earlier we considered two random variables. So, we had vector of two elements: x 1 and x 2. We call it x, y. This is a different notation. But, now, it is more general in such random variables. So, for the hind side, you can see that, if you talk of the joint density or joint probability distribution of X, it will be basically a function n variables – x 1 to x n.

(Refer Slide Time: 52:05)



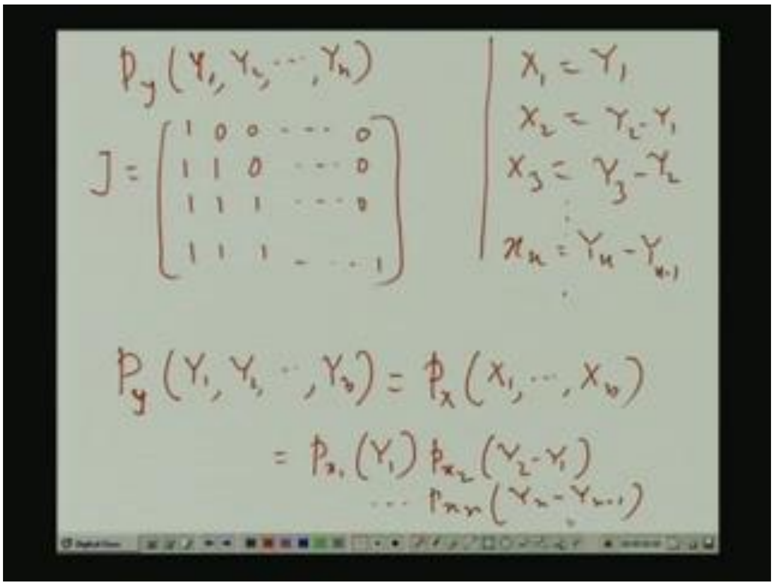
So, we can define joint probability distribution; actually I should write like F x for x; say takes values x 1; takes specific values x 2 dot dot dot x n. This capital X actually is the vector considering all the variables – small x 1 dot dot dot small x n. Here capital X 1 is a particular value for small x 1; x 1 is the variable. Capital X 2 is a particular for the variable small x 2 and likewise. What does this mean? It means the probability – probability of this event that, x 1 is less than equal to capital X 1; x 2 less than equal capital X 2 dot dot dot x n less than equal capital X n. This should occur jointly. There are n joint events; that is, one is small x 1 less than equal capital X 1, another is small x 2 less than equal capital X 2, so on and so forth. This should occur jointly – simultaneously. That is why it is called joint distribution. This is denoted by Fx x 1 up to x n. Sometimes when I do not need, I may skip this subscript x and directly put x 1 to x n. I think you can easily understand sometimes. But, when there is confusion when suppose I am dealing with two or more than two such distribution functions; then to differentiate between the functions, I may put the subscript. Now, this is for the joint probability distribution. By the same token, you can then define… And you can see easily see that, this is now a function of n variables x 1 to x n. By the same way, I can next define the joint probability density function.

(Refer Slide Time: 54:26)



Joint probability density – recently, we know that, we have to differentiate the joint probability distribution function. So, you we define like this – P X X 1 up to X n is nothing but del n – n times – now, skipping this subscript here – X 1 dot dot dot X n with respect to del x 1 del x 2 dot dot dot del x n. This is the joint probability density. What does it mean actually? It means… But, we should now consider an n-dimensional another space, whose X's are X 1, X 2 up to X n. And there if you take an infinitesimally volume, I mean, mass… In fact, it was the origin, whose volume is say d x 1, d x 2 up to d x n. That is, along x 1 axis… that is, suppose in an n-dimensional space, you are located at a point up to x n. At this point, you are considering an infinitely small or infinitesimal small region or cell, whose sides are… We call it actually hyper cube, because actually real life cube means it has got only three sides; but, in the n-dimensional space, I will call it n-dimensional hyper cube, whose sides are d x 1, d x 2 up to d x n. So, volume is this. So, basically, defines an area, where x goes x… – the variable x 1 goes from capital X 1 to capital X 1 plus d x 1; capital X 2 goes for, that is, this zone. This for variable x 1. Therefore, variable x 2 again. This is the range – dot dot dot.

(Refer Slide Time: 57:06)



You can easily see; this will be nothing but P x at this choice X 1 up to X n divided by the determinant, that is, 1. And we are giving the fact that, the random variables x 1 up to x n – they are independent. So, this join density is nothing but the product of the individual density. So, P x 1 and capital X 1 is Y 1; P x 2, which is Y 2 minus Y 1 dot dot dot P x n Y n minus Y n minus 1. So, you can easily see that, the concepts involved here – they are nothing new; they are simple generalizations of the concepts that were valid or that were introduced rather in the case of two random variables. That is why I am not giving you any proof and all that; you can argue about this on your own.

So, I stop here today. In the next class, I consider this issue further and go into things like mean, covariance, correlation; in fact, we will have correlation matrix now, covariance matrix and things. Again the characteristic function issue as relevant here. And that takes us to a very important theorem called central limit theorem. So, that is all for today.

Thank you very much.