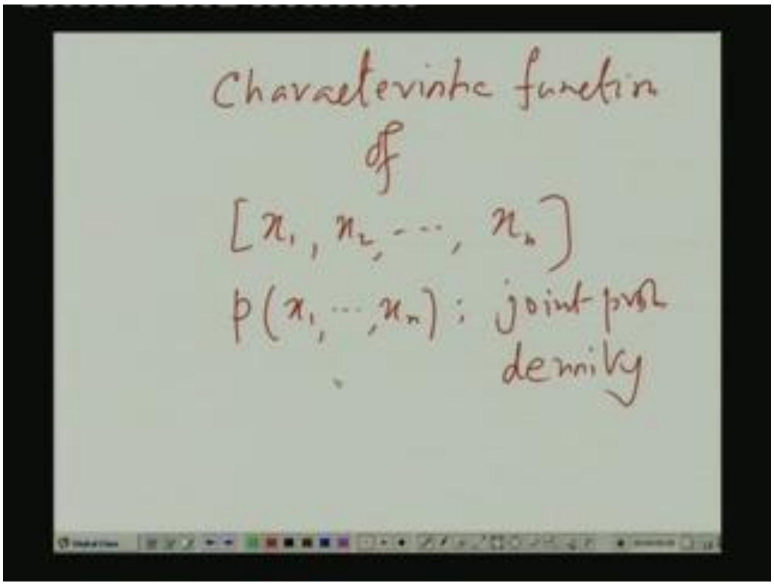
Probability & Random Variables
Prof. M. Chakraborty
Department of Electronics and Electrical Communication Engineering
Indian Institute of Technology, Kharagpur

Lecture - 26
Characteristic Functions and Normality of a Random Vector

So far we have been in the last class we have been considering random vectors. For last few lectures only you have been on this topic. So, today we will be you know considering characteristic functions for random vectors. So, you remember I mean earlier we considered only a single random variable and with respect to a single random variable, we consider a characteristic function. So, that time it has a function of just one frequency variable omega 1, then we extended that to the domain up to random variables. That time also we had a characteristic function. It was a function of two variables, two frequency variables omega 1 and omega 2. So, now that whole approach will be generalized to a random vector that has got say n number of random variables, right.

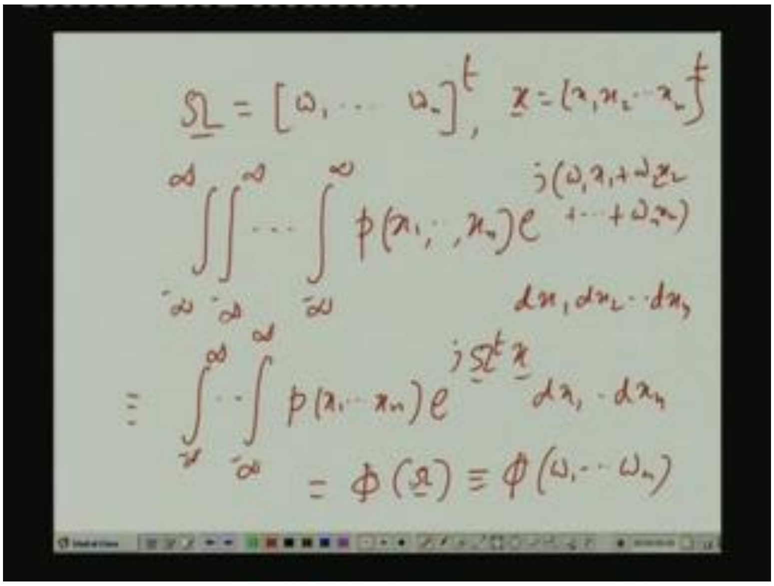
(Refer Slide Time: 01:48)



So, we will be considering random x 1, x 2, dot, dot, say x n. There are n random variables and they are jointly random. They have got a joint density that is you can say joint density, joint probability density. It is for probability. Obviously you can understand that since there are n random variables, we will have n frequency variables.

Now, omega 1 associated with x 1 omega 2 associated with x a omega 2 associate with x 2 dot dot dot omega n associated xn.

(Refer Slide Time: 02:45)



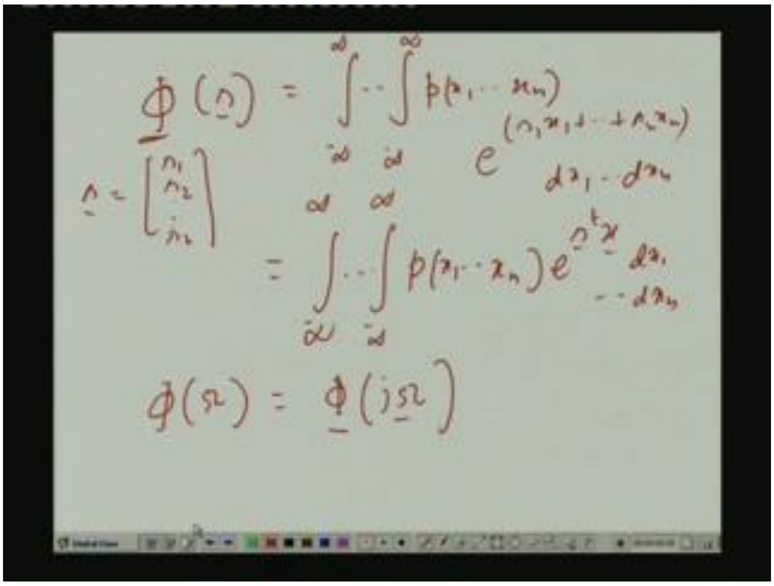
So, characteristic function I mean if I define a vector, in my case all vectors are actually column vectors whereas, in the book by Papoulis, he normally takes vectors as row vectors. There is a difference you may find. This is row. If you put a transpose, it becomes a column vector. So, the characteristics function actually becomes a function of n random n variables n frequency variables omega 1 to omega n. In fact, it should be this. Take the joint density and multiply it by j omega 1 x 1 plus omega 2 x 2 plus dot dot dot omega n x n integrate with respect to x 1 x 2 up to x n.

If I also define a vector x as x 1 x 2 dot dot dot x n transpose column vector, you can also equivalently write this as P of x 1 to x n e to the power j. You can see I can always write it as omega transpose x, that is omega transpose gives you this row vector omega 1 to x omega n multiplied by a column vector x 1 to x n. If you do this row into column, you will get this term omega 1 x 1 plus plus omega 2 plus dot dot omega n x. This is compact way of writing and this is as before.

So, this is the characteristic function. It is a function of omega 1 to omega n or since we have put them in a vector form within m capital omega, you can also call it phi omega.

There is equivalent to saying this is phi omega 1 to omega n, in short phi omega capital omega.

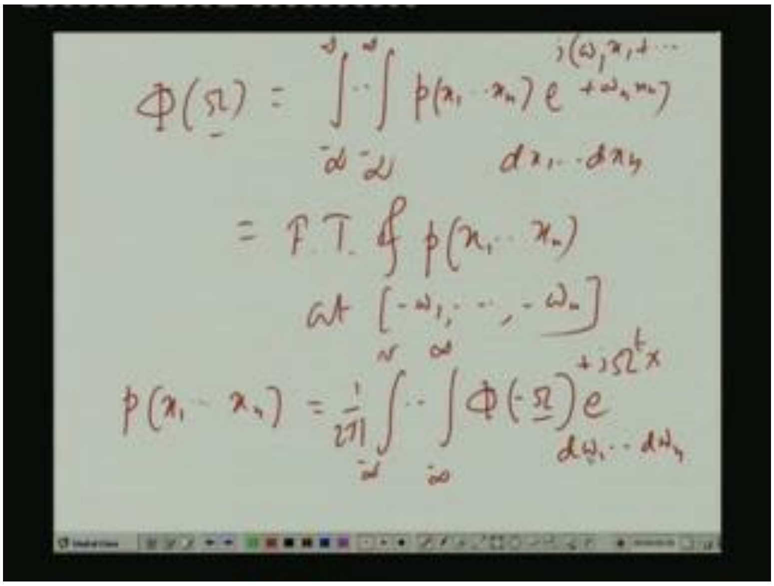
(Refer Slide Time: 05:34)



This is a characteristic function and like before we had like earlier we had something called the moment generating function which was phi s. Here also you can have moment generating function phi s, where s is a vector. It consists of these elements s 1, s 2 dot dot dot s n. S 1 associated with x 1, s 2 associated with x 2 so on and so forth, and you have got these integrals. Now, E to the power s 1 x 1 plus dot dot plus s n x n d x 1 dot dot dot d x n which also you can write as p x 1 x a E to the power transpose x d x 1 dot dot dot d x n.

Obviously, then the characteristic function phi and moment generating function though I am using the same symbol, I am putting at underscore here, so that it is called I say that is phi underscore. It is a defined function. It is a moment generating function whose expression is given this, where I just write phi. If it is a characteristic function, you can easily say that phi omega, it is nothing but moment generating function phi bar where s is replaced by j omega. Here if s is replaced by j omega vector, you get back your characteristic function. That is very simple, right. Also, as before you can get back the joint density from the characteristic function because after all this is a Fourier transforms relationship.

(Refer Slide Time: 07:36)



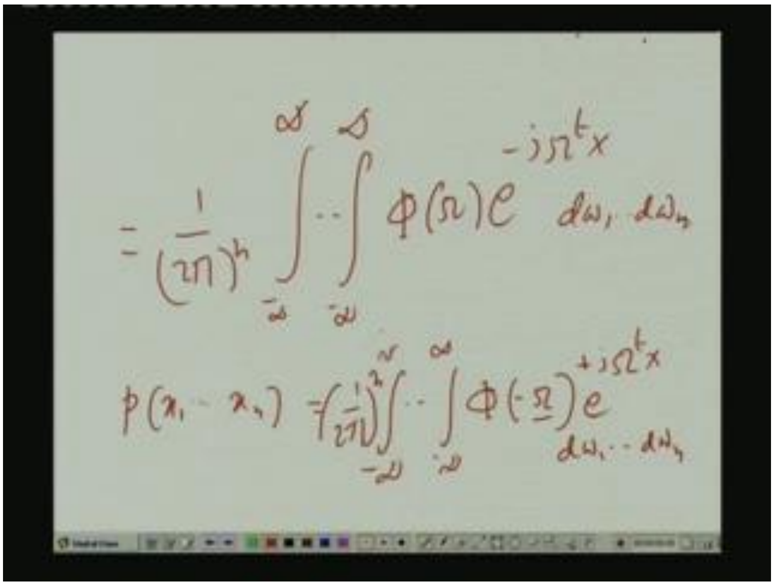
As we have seen characteristic function phi omega vector is nothing but I am rewriting the expression again. Joint density multiplied by E to the power j omega 1 x 1 plus dot dot dot plus omega n x n d x 1 dot dot dot d x n. You can easily see that this is nothing but Fourier transform of F distance. For Fourier transform, this function p x1 to p x n at these frequencies minus omega 1 dot dot dot minus omega n, that is if you really evaluate the Fourier transform of this, but put a negative sign with the frequencies, then you will get the expression and there is no negative sign that comes in this exponential because negative and negative cancels and becomes positive.

How to get back p x 1? How to get this back? From this by the inverse Fourier transform relation 1 by 2 pi integrals minus infinity to infinity dot dot minus infinity to infinity. There are Fourier transform at these frequencies, right. So, that means I have to put back phi, but this phi minus omega is actually the Fourier transform of the joint density at omega 1 to omega n. So, joint density is nothing but inverse Fourier transform of phi minus omega. So, phi minus omega if you put omega is the vector here, capital omega that consists of all elements omega 1 to omega n, small omega 1 to small omega n, and then you have got j minus. This is inverse relation, right. Omega transpose x, sorry it will be plus inverse Fourier transform of phi minus omega.

So, phi minus omega you can also write it as phi of minus omega 1, small omega 1, small 2 minus omega 2 dot dot dot minus small omega n, but I am writing in a compact from using this capital omega notation and this j capital omega transpose x actually is nothing but omega 1 x 1 omega 2 x to dot dot dot, this expression omega 1 x 1 dot dot dot omega n x n and nothing else. This is inverse Fourier transform of phi minus omega, right and you are integrating with respect to these components d omega 1 2 d omega n.

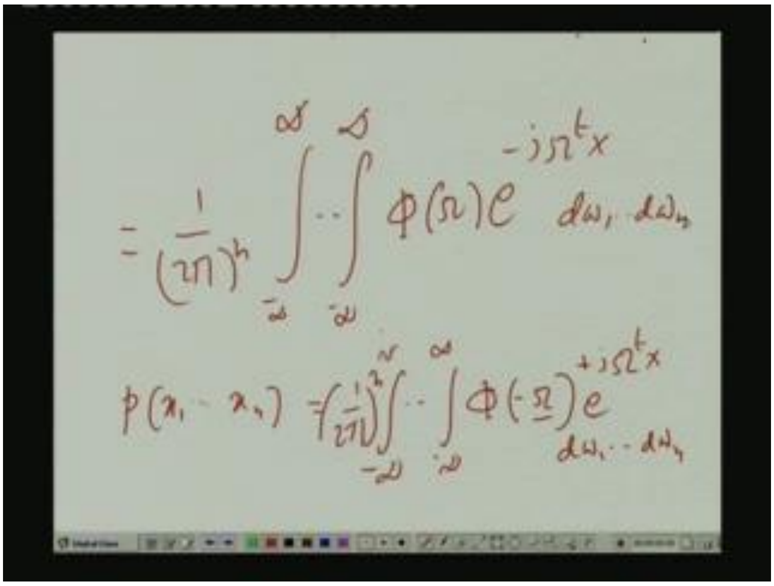
Now, as before you can say that if you can make a substitution say minus capital omega, you can call it omega prime. So, the integrals limits get interchanged infinity. I mean infinity to minus infinity, but the differentials d omega 1 now becomes minus of previous d omega 1 and like that. Those minus signs can be observed it reversing the limits of the integral again. So, again from plus infinity to minus infinity, you get back minus infinity to plus infinity. So, you get back just so omega is replaced by minus omega. So, you just get this thing, just a minute.

(Refer Slide Time: 12:01)



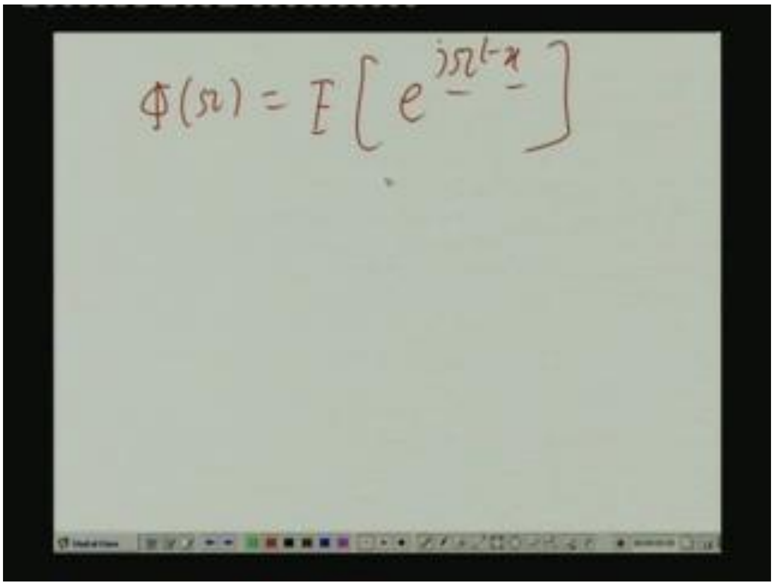
I hope you understand what I said because I have done this exercise many times in past. So, I am not doing it again. I am just telling it orally. That replaces each minus omega 1 replace each omega 1 by say minus omega 1e. You call it a minus omega prime like that. So, this integral limit, they get interchange plus infinity to minus infinity and all differential d omega 1 becomes minus d omega 1 like that. Each minus sign can be observed in the integral limits again, so that you get back the original position of the limits from minus infinity to infinity, and no negative sign there phi minus 1 minus omega becomes phi plus omega and in the cardinal in this exponential, you get a minus sign which then gives you this thing that this joint density is nothing but one thing this should be 1 by 2 pi to the power n, because there are n variables. I am sorry I made this mistake. So, 1 by 2 pi to the power n minus infinity to infinity dot dot dot minus infinity to infinity, and phi omega E to the power may be I erase this here because it is getting bit crowded just here. So, better I write it somewhere else.

(Refer Slide Time: 13:53)



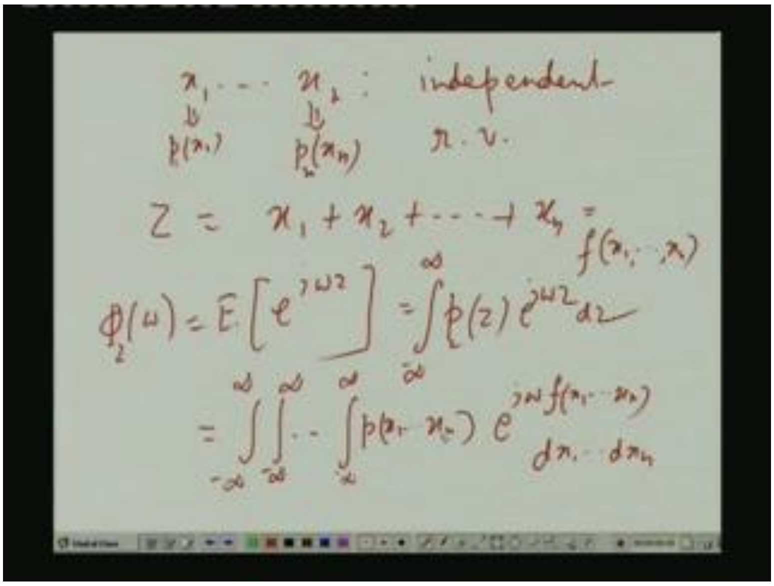
So, this becomes nothing but now a minus sign comes omega transpose x d omega 1 dot dot dot d omega n.

(Refer Slide Time: 14:23)



As before you can also say that the characteristic function phi omega is nothing but after all you are multiplying the function E to the power j omega transpose x by the joint density. That means, you are taking expectation of that. So, it is nothing but expected value of j omega transpose x. Now, a very interesting special case of this again and we have considered that case also in a past.

(Refer Slide Time: 14:50)



In a context of single random variable and two random variables here, we generalize x 1 and x n, they are independent random variables. R v for random variable and you form a random variable z single random single random variable as just summation. Each random variable x 1, it has own density probability density p x 1, it has its probability density p x n, but remember I am writing p x 1 p x n. They are not the same function.
It is not that I have got the same function and x 1 is replaced by x n. Actually, I should put p 1 p n and like that, but I suppose I can keep this subscript. You can understand px 1 is just a density function for x 1.

P x n is a density function of xn and they are not the same function. This is just loosely I can I mean omit this subscript because you know just things become bit congested if I write equations, so many subscripts and all that, but ideally the subscript should be there.
I may be for the time being I can still put it my question is what given the probability densities of x 1 to x n. What is probability density of z? It is first characteristic function of z. What is a characteristic function of z phi z? Some omega, sorry it is not a vector

omega. It is scalar omega. Now, phi z omega is nothing but expected value of e to the power j omega z, z is actually a function we can view it like this is a function of x 1 to x n.

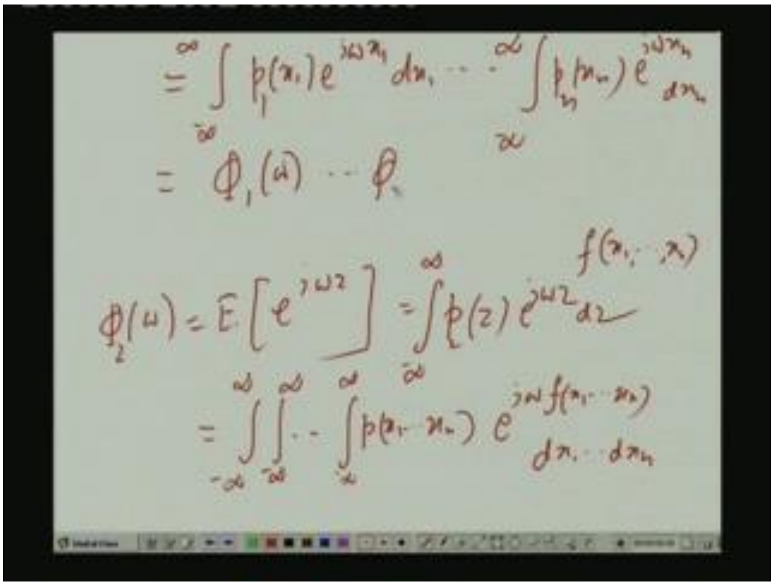
Now, when you write this expected value of e to the power j omega z, actually what you do? You multiply this by probability density of z, take it different function. It is not the same p I wrote. May be you can put p z subscript here, e to the d power j omega z dz, right, but we have also seen earlier that e to the power j omega z which is a function of x 1 to x n, there are random variables n and independent random variables given your evaluated function and is taking its expectation. That expectation is same as taking the expectation of j omega z replaced by the function x 1 dot dot dot x n and multiplied by. We have done this in the case of single variables, and then two random variables. That is expected value of the function of some random variable, may be one or two random variables.

It is same as replacing that function. Those I mean like let me repeat this that is suppose you are given a function z of say two random variables x 1 and x 2. Then, you are evaluating the expected value of E to the power j omega z. So, multiplying by the probability density and integrating, we have seen at least in the contests single and two random variables, that is same as evaluating the function first not as z, but directly x 1 to x n. I am taking it as expression with respect to x 1 to x n there multiplying by the joint density and integrating both will be same that we have seen. If you do that and that treatment come be extended to n variable case that I am not going into because intuitively that should be clear.

So, if we now replace f of x 1 to x n by just this summation because that is how this function is what you get e to the power j omega within bracket x 1 plus x 2 plus dot dot dot dot x n. Remember these random variables x 1 to x n, they are statistically independent. So, this joint density is nothing but product of p x 1 into p x 2 dot dot dot into p x n product of the individual densities, and then multiplied by e to the power j omega. Then, within bracket this function is nothing but x 1 to plus x 2 plus dot dot dot x n which can be then separated.

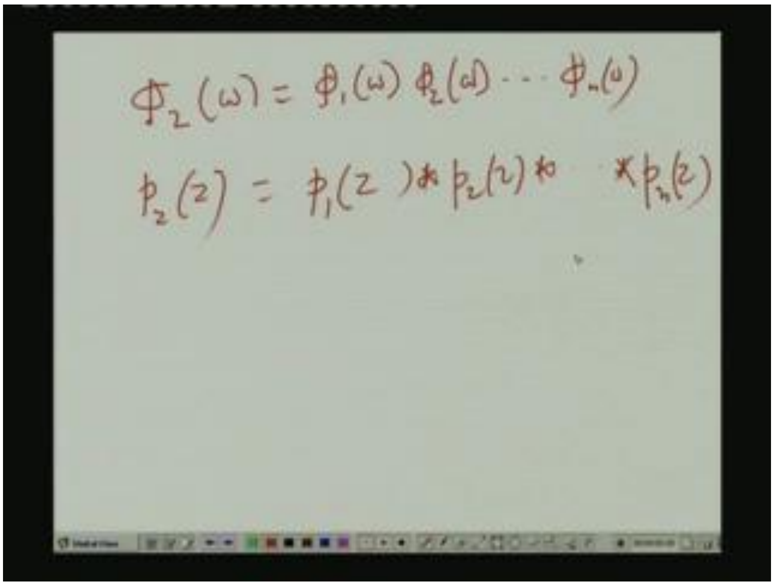
So, I can have want to one term p of x 1 into e to the power j o omega x 1 under one integral with respect to x 1. Then term p of x 2 into e to the power j omega x 2 under another integral that the integration is with respect to x 2 so on and so forth.

(Refer Slide Time: 20:09)



That means there is nothing but omega is common. There is only one omega 1. Now, that is the difference because omega came from the characteristic function of z. Z was the single variable. So, it has only one single frequency omega. We will continue on all the integrals multiplied by dot dot dot dot p x n. I would call it p 1 p n e to the power j omega x n d x n. This is nothing but the characteristic function of x 1 with frequency omega 1 dot dot dot characteristic function of x n with again same frequency omega. So, that means phi z omega is nothing but phi 1 omega dot dot dot multiplied by then phi n omega, right. So, product of several Fourier transforms gives rise to the characteristic function of z.

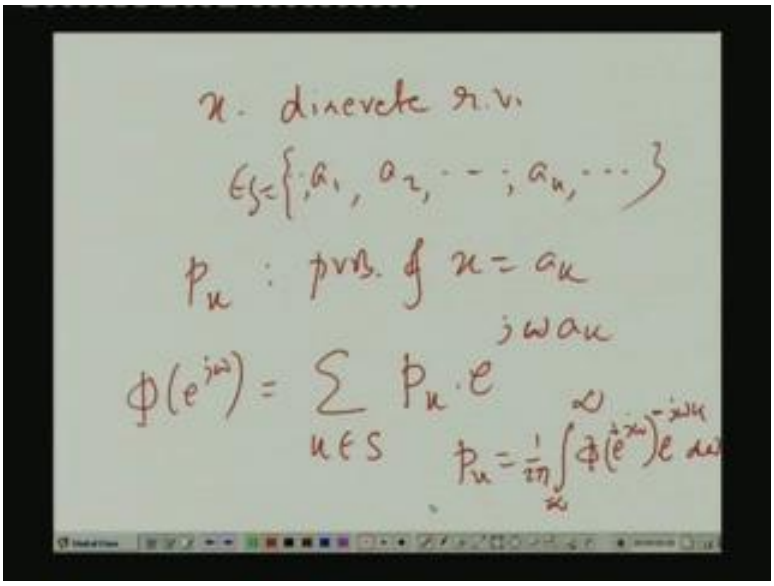
(Refer Slide Time: 21:47)



So, we know the convolution theorem this means may be equal to, so you got this phi z omega. So, from convolution theorem, I know if I take the inverse Fourier transform, I get here probability density of z. That will be nothing but convolution of the inverse transforms of these functions. Inverse transform here is p 1, but some of the variable is z mind you resulting convolute being, but the independent variable that will come out of this will be z. So, p 1 z star p 2 z star stands for convolution star p z p n, sorry p n z.

So, if a random variable z is formed as a summation of n number of independent random variables, then the resulting probability density for z is nothing but convolution of the individual densities of those random variables. Same result can be extended in the discrete case also. In the discrete also you can define characteristic functions. May be we can quickly do that. Suppose x is we did not do this in the case of a single variable or two variables also.

(Refer Slide Time: 23:26)



So, quickly suppose x is a discrete random variable. It takes values from this set say x 1, x 2. These are some constant values now or maybe I replace x because x 1 x 2, I have already used. It takes values like you know a 1 a 2, they are all numbers may be 10, may be 15, may be 17, may be minus 10 like that. It can be finite, it can be infinite. Dot dot dot I am putting a k dot dot dot like that. For each value, there is a probability. It is not density anymore because it is not a continuous function. It is the discrete function for which p k is a probability of x equal to a k then, and of course you can imagine that suppose a 1 is 10 and a 2 is 15 and there is no number between them.

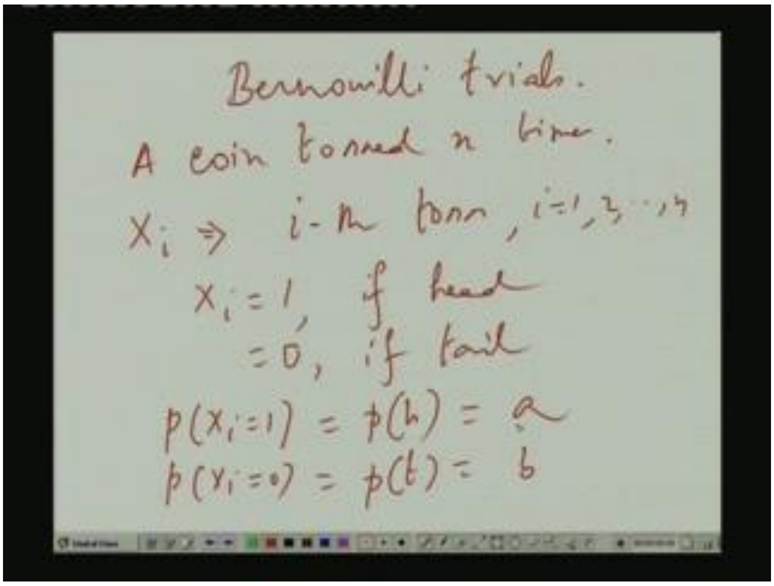
You can assume that there are still numbers 11, 12, 13, 14 like that, but x has probability 0. I mean the probability that x can take the value 11 or 12 or 13 and 14 is 0. So, that way you can say that x is a discrete random variable which takes any value, any integer values from minus infinity, but for certain integers, it has got non-zero probability and for certain integers, it will be zero probability. That way you can define the characteristic function as a k. If you call this set s k or s multiplied by p k, sorry, probability multiplied by j. Now, this frequency omega, it is a discrete frequency. Its digital frequency unit is radian.

I hope I mean you have done little bit of DSP. You know the difference between discrete time Fourier transform analog and Fourier transform. Here j omega a k and if I go by that notation of Oppenheim and Schafer, then the discrete time I mean order the notation used in DSP. This is actually written as a function of not omega, but to the power j omega just to emphasize that we get power series in terms of it by j omega to both e and j are constant omega is variable, but this is the style of writing and this omega unit is different from earlier. It is now radian.

Earlier it was radian for the unit of the variable. If the x is time radian per second, if x is some space radian per that you need a space and like that, but here it is just radian omega. This comes from DSP, and then again p k found out by the inverse d t f t place inverse discrete time Fourier transform relation, that is p k. You can easily see as we have all studied in DSP that this function is actually periodic in omega. If you replace omega by omega plus 2 pi, you will get the same number because a k are integers and using that you can derive this, the inverse d t f t relation. Only thing is this exactly not d t f t because there is the plus sign. So, I will call it d t f t of p k d t f t of the sequence p k, but at frequency minus omega, so inversed d t f t. That means, phi e to the power minus j omega is the d t f t of p k.

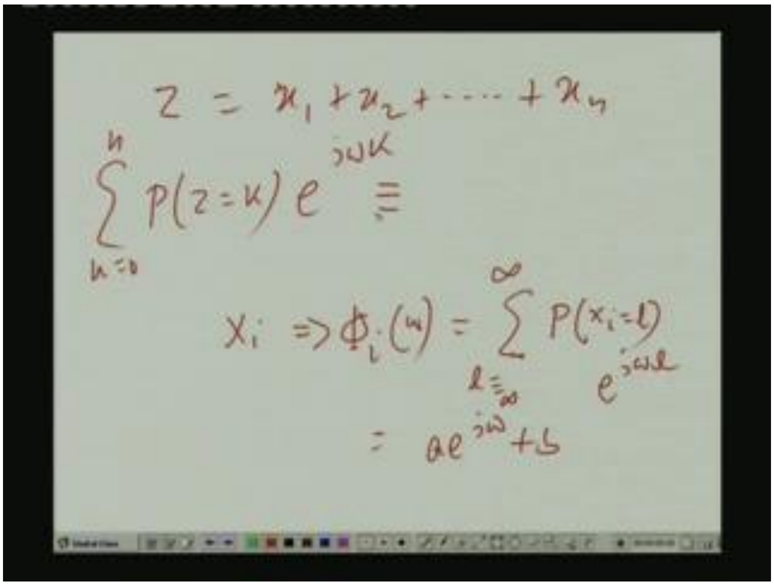
So, p k is nothing but inverse of d t f t phi e to the power minus j omega. So, you can put minus j omega to start and then, it to the power j omega n. This is the formula j omega, sorry k d omega. There is a minus sign to start with, but then you can replace omega by minus omega integrals and then, it becomes infinity to minus infinity d omega becomes minus d omega. That minus sign can be observed here. So, you get the same limits again minus infinity to infinity phi. I mean you get phi e to the power j omega back because earlier it was minus sign and now it has become plus, but in this here it becomes minus. I do not think I have to repeat this because similar things we have done just a while back. You can generalize this in the two variable cases, and n variable case, there will be just multi-dimensional d f t and multi-dimensional inverse d t f t. I am not going to that.

(Refer Slide Time: 28:55)



As an example, let us consider Bernoulli trials. Suppose a coin is tossed n times x i. It is related to i th toss or i can be 1, 2 up to n. Xi is a random variable. It takes value 1 if head comes and it takes value as 0 if tail. So, that means probability of x i equal to 1 is nothing but probability of head which could be say a. I am not saying that coin is a good coin. I may be biased coin. So, it is not the tail and head have the same probability half and half. It is more general a, and p x i is equal to 0, that is x i takes the value 0. It is nothing but the probability of tail occurring which is a b of course a plus is b is 1. You can also see that the random variables x i, that is x 1 x 2 up to x n, they are mutually independent because one toss has got no influence on the subsequent toss or the toss that was executed before. So, they are independent, right.

(Refer Slide Time: 30:49)



So, we can use I mean we know now suppose next we form a random variable z as x 1 plus x 2 plus dot dot dot x n. So, out of n trails, we find out z. So, z can take various values. It may so happen that both x 1 x 2, I mean up to x n, they all took value 0 that is all were tail. So, z minimum, minimum of z is 0. So, it can happen that out of n 1 is head, rest at tails then z can be 1, and z can be z 2 and maximum value of z that is permissible is n when all are head. So, that way you can find out what is p probability of z equal to k. That means, k element k variables take k of this n variables. They take value 1, that is corresponds to the head and rest correspond to tail.

This we can find out directly by finding out the probability directly, but we can show and we know the expression also, but you can show that by using the characteristic function theory also, we get the same probability for p equal p of z n equal to k. We have already seen considered the characteristic function of z. What is the characteristic function of z? It is e to the power j omega z and z is from 0 to n. I would not call it z. It is actually k, sorry. So, z taking value 1, z taking value 2, z taking values each of the probability and that k comes here. K is from 0 to n. That we have just a while back we have seen the characteristic function for a variable random variables which take discrete values, right.

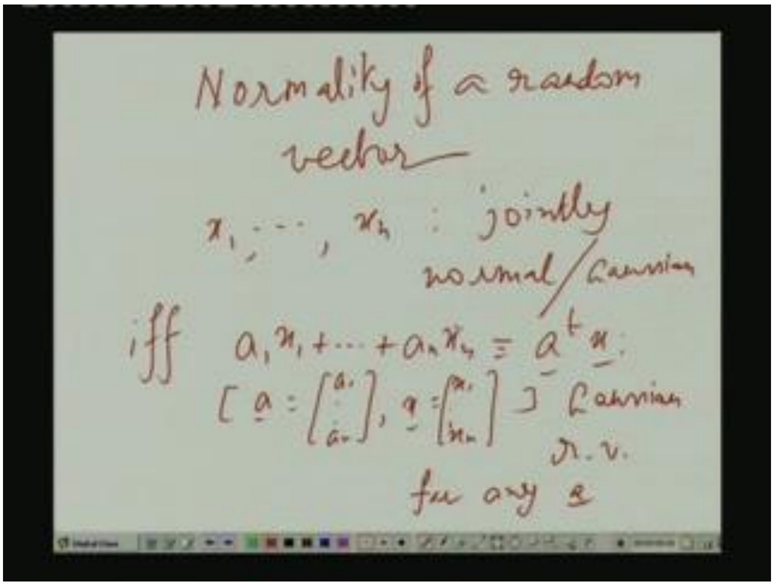
So, this is the characteristic function, but we have also seen that if n random variable which are mutually independent or added, then the resulting characteristic function is nothing but product of the individual characteristic functions. Now, what is the characteristic function for say the random variable xi random variable xi? It can take only two values 0 and 1 and all other. All others are 0. I mean for 0 1 1, for 1 it has got a value. I mean the probability of a for taking value 0, it has a probability b for taking any other value. Probability is 0. So, for this what is the characteristic function phi i z omega. Remember characteristic function of z at a particular frequency omega is a product of characteristic function of x 1 to x n, but at the same frequency omega that omega I choose here.

What is phi omega? Phi omega means I have to just run this summation, but x i can take 2 only values for which probability is non-zero. The x i is equal to you can say some integer l e to the power g omega l. L can be this is the general formula, where l could be from minus infinity to infinity, but we have seen that x i can take only two values 0 and 1. For x i equal to 0, I have got a non-zero value of this probability which is b, and for x i equal to 1. We have a non-zero value of this probability p that is a or any other value of l, this probability is 0. That means, i in this summation I just take this as a general formula for the characteristic function, but in this case we take only two values. l equal 0, l equal to 1 for l equal is 0. This is b and b e to the power **0** that is 1.

So, you get term b and then, for x i equal to 1 p of that probability is nothing but probability of head occurring which is a. So, a into e to the power j omega l equal to 1, so e to the power j omega. So, this gives rise to a e to the power j omega plus b, and I have to just multiply this probability, this characteristic function. So, this is the characteristic function for the i th random variable, but it is so general that is independent of i. That means the characteristic function will be same for all the random variables. So, when I multiply, it is nothing but raising this to the power of n that is this equal to this is equivalent to now you can use the binomial theorem. You can use the binomial theorem and you will see this is nothing but n c k a to the power k a to the power j omega k b to the power n minus k.

So, you can just equate the terms from both sides. This will give rise to probability of z equal to k is nothing but n c k e to the power k e to the power n minus k. This is what in the formula for Bernoulli distribution, we earlier derived it directly, but you see for the characteristic function approach also you get the same result. We next move to normal vectors.

(Refer Slide Time: 37:06)

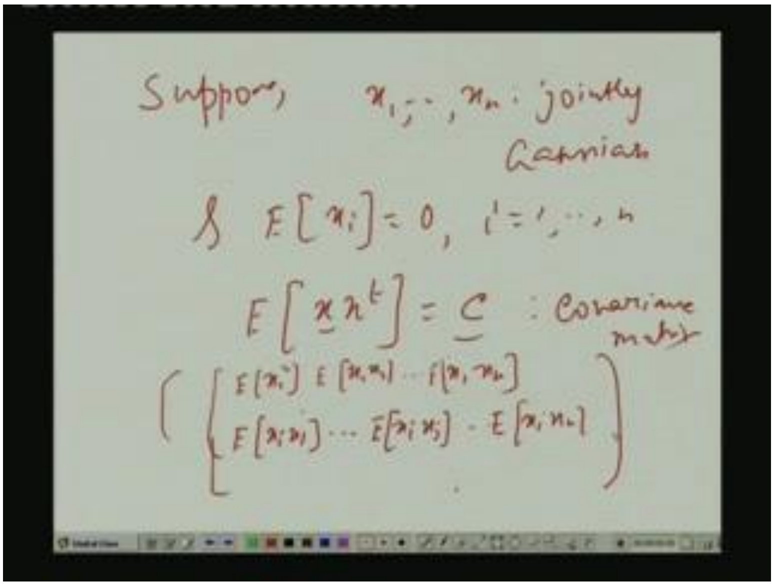


The notion of normality, a set of random variables are given x 1 to x n. We will say that jointly normal and Gaussian, they are same equivalently. If and only if a 1 x 1 plus a n x n. By the way I am considering only real value variables and therefore, the coefficients that I am bringing you want to be there also real values, but this can be generalized to the complex case which I am not doing here. You can also write it equivalently as a vector, sorry a vector transpose x vector where a is nothing but this a 1 to a n x nothing but as before if this is which is a scalar variable is a Gaussian r v random variable for any a whereas, if you choose any coefficient say 1 to n from this summation, you get a random variable that must be Gaussian.

See if that is true for any choice of a that we choose, a vector for the resulting variable is Gaussian, then only it will be called x 1 to x a. They are jointly Gaussian. Obviously you can see that if they are jointly Gaussian n is subset also is Gaussian jointly. First suppose you take a 1 to be 1 and a to 0 a 3 0 a n 0. So, you get back only x 1. So, if this is jointly Gaussian, that means, x 1 is Gaussian and same for x 2, same for x n. Then take a 1 equal to 1 a 2 equal to 1 raise 0. So, that means or may be a 1 and a 2 non-zero and rest zero.

So, that means a 1 plus x 1 plus a 2 x 2 that is Gaussian for any choice of a 1 a 2. That means, x 1 and x 2, they are jointly Gaussian so on and so forth. Now we will see. So, this is general definition.

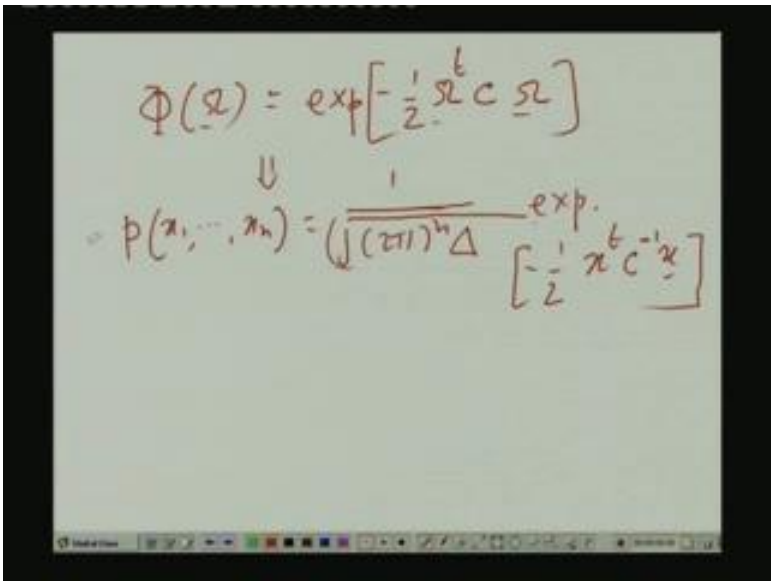
(Refer Slide Time: 40:21)



Now, we will see that suppose jointly Gaussian and e x I, they are zero mean to make life simple. They are zero mean for i equal to 1 to n, that is each has zero mean and E of this vector x x transpose which is in this case covariance matrix, you see that is covariance matrix. Normally in the case, covariance computation, we first deduct subtract the mean from each random variable, but mean here is 0, right. So, we simply take x vectors x transpose. You remember we have done this earlier. This is a symmetric matrix Hermitian matrix. The structure actually is like this to expand it e of x 1 square e of x 1 x 2 dot dot dot dot e of x 1 x n. In general, e of x i x 1 dot dot dot in general term i th row and j th column x i x j dot dot dot e of x i x n so on and so forth. This is called the covariance matrix.

Suppose this is given as symmetric. Obviously, you can take the transpose of this, you get the same thing. E of x 1 x 2 and e of x 2 x 1, they are same and likewise. E of x i e x j and e of x j x i, they are same because x i into x j or x j into x i, they amount to the same quantity. This is we have dealt with these cases earlier. So, I am not getting into the properties of correlation matrix or covariance. In this case correlation covariance matrix, they are same because mean is 0. So, we will now show that in this case that is when they are jointly Gaussian and each has zero mean and that covariance matrix is c, then the characteristic function that the characteristic function phi omega capital omega as before is a vectors small omega one to small omega n.

(Refer Slide Time: 42:39)

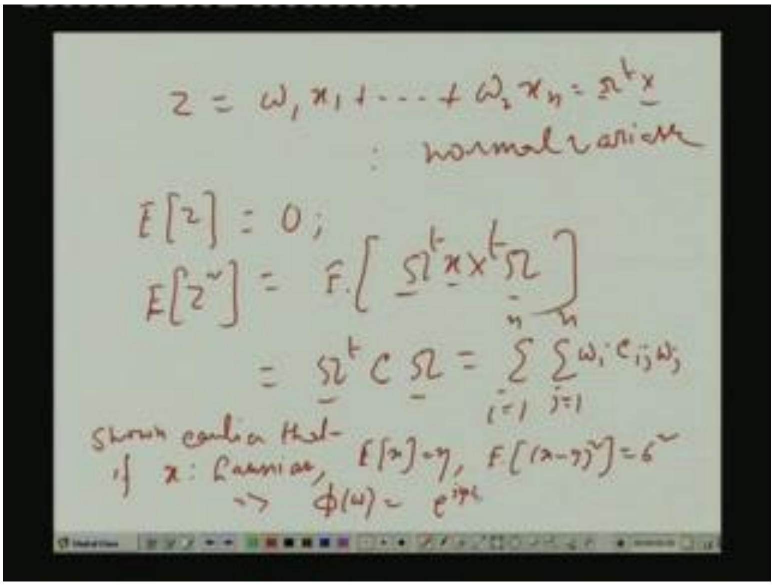


So, I am not rewriting it any more. This is nothing but this exponential minus half omega transpose row vector because my omega is capital and my omega is column vector always unlike Papoulis, where the standard definition where he takes the vector to be always in row form my vectors are in column form. So, I put transpose phi transpose c phi. So, matrix into column vector gets a column vector row vector into column vector, it is a scalar number. So, exponential of a scalar number, but it is a function of all those omegas 1 to omega 2 because each omega vector is nothing but I mean consists of all those elements, right.

We will show these and from this also, we will show that then the joint density p of x 1 to x n joint density which will be obtained from the inverse relation, actually we will be given by sorry exponential always remember this. This is a general formula for multivariate or multivariable Gaussian distribution half x transpose inverse of this matrix c inverse x. How we do that? We just concentrate on proving this phi omega. Actually this part we will not prove because this is nothing but the inverse transform Fourier transform of phi omega at minus omega.

Now, if you want to carry out the inverse Fourier transform, then some mathematical exercises will be there. So, that part if required I will take up later, but that will give you this formula. There is nothing probabilistic or statistically that it is just inverse Fourier transform calculation from phi omega at minus omega. So, you concentrate on proving this fact that characteristic function is of this form. How do we show that?

(Refer Slide Time: 45: 54)



Now, you see suppose you form a variable z as omega 1 x 1, then omega n x n since x 1 to x n, they are jointly normal, z is then a Gaussian variable or normal variable. You can also write it as omega transpose x. What is e z mean of z will be 0 because mean of z is nothing but mean of omega 1 x 1, that is omega 1 into e of x 1 which is 0 here. So, E of z is 0. Obviously, what is E of z square? That is covariant variance in this case variance because single variable or that is nothing but E of omega transpose x and since, it is a scalar number, you can take its transpose because scalar number and its transpose, they are same, so z into z.

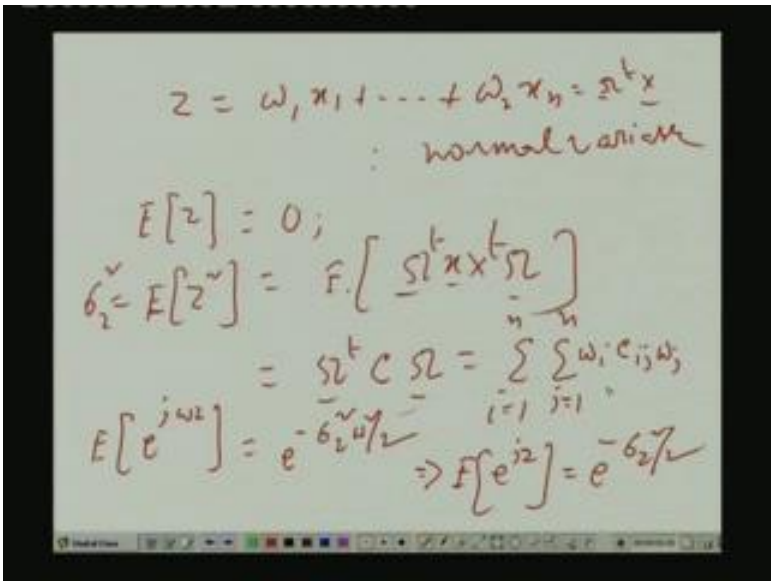
So, this is z and then, I write z transpose because z and z transpose, they will be same as long as z is a scalar, but if you take the transpose z, you get x transpose omega, and you can push this E operation directly on x x transpose because there is only random part. So, omega transpose c because E of x x transpose is c, then you input correlation matrix. This you can see that this you can also write as omega i. So far i for particular omega i c i j omega j, so forget about omega i. Keep it in the outer summation first. This inner summation c i j omega j will give you what i th row of c matrix you are moving over columns, that is you are varying j and multiplying the w's.

So, i th row times i th row of c times the column vector omega, you get a scalar number that you get out of this. You multiply with the i th element of this omega transpose omega transpose which is omega transpose which is omega i, and then now do it for all i's. That is how you get this. Now, you remember one thing. Earlier we had considered one thing long back that suppose there is a Gaussian random variable z and its characteristic function Gaussian random variable z.

Suppose we had shown earlier that if x a single Gaussian variable or normal e x, it is not 0, but some eta and e x minus eta square is sigma squared, then the corresponding characteristic function for this Gaussian case was shown to be e to the power j eta omega times e to the power minus omega square sigma square by 2. This you have seen earlier.

Here we also have got a Gaussian random variable z, but that has a mean eta equal to 0, so that we put 0 here. The moment you put 0, you get only this function e to the power minus omega square sigma square by 2. Sigma square is nothing but the variance of z that is e of z square which is this, that means using this I erase this part.

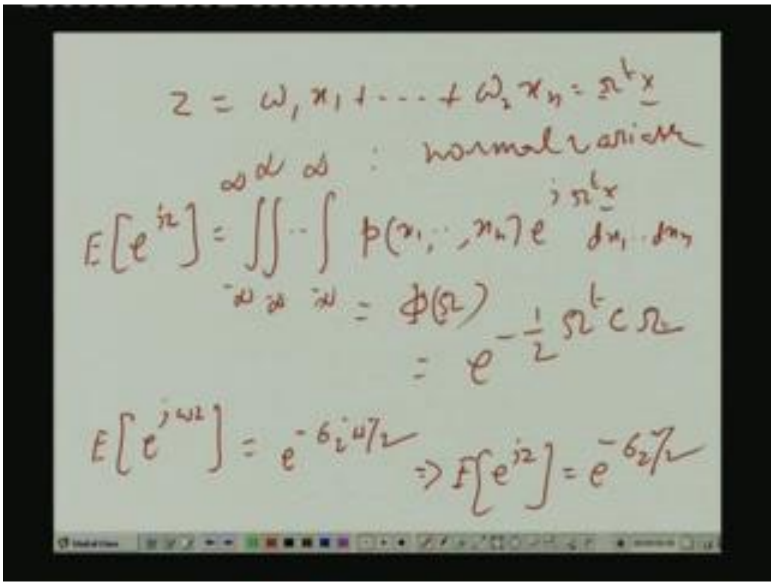
(Refer Slide Time: 50:46)



Using this we have the characteristic function of z, that is E of e to the power j omega z. There is the characteristic function of z. That is nothing but e to the power minus sigma square sigma square is this expression omega transpose c omega. You can call it sigma z square. So, minus e to the power sigma z square omega square by 2 is the characteristic function at particular omega. Suppose I took you know omega equal to 1 now. So, this gives raise to e to the power expected value of e to the power j z. What is that? E to the power minus just sigma z square y 2, but what is e to the power.

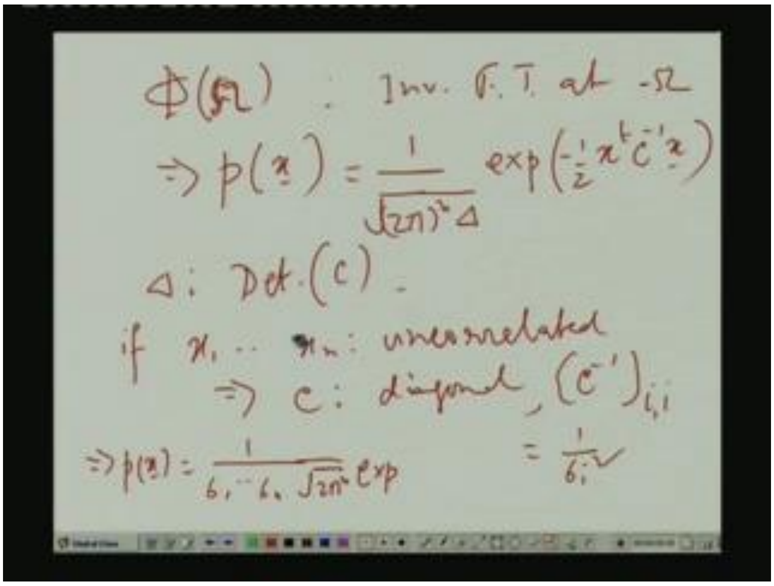
What is expected value of e to the power j z? That is multiply e to the power j z by probability density of j z integrate, but z is a function of x 1 2 x n. So, expected value of e to the power z can also be written as multiplication of e to the power j, then this quantity omega 1 x 1 plus dot dot dot plus omega n x n, that is capital omega transpose x multiplied by the joint density of x 1 2 x n multiplied by the joint density of x 1 to x n. What is it? That is nothing but the characteristic function of x 1 to x n at frequencies omega 1 to omega n. Let me erase some part and explain this further, but e to the power expected value of e to the power j z, it is also same as e to the power j because z is a function of x 1 to x n.

(Refer Slide Time: 53:33)



So, e to the power j you replace z by the function which is nothing but omega transpose x and multiplied by the corresponding densities and integrate, you will get the same expected value. I have just stated, a little while back we have done similar thing for the single variable case, two variable case and that can gets extended to the n variable case, that is either you can take the expected value of e to the power j z by multiplying it by its probability density that is p of z integrate or replace z by its function in terms of x 1 to x n, which is this omega transpose x multiplied by the joint density of x 1 to x n and integrate, but omega transpose x if you put back here, you get to see this is nothing but the characteristic function of x 1 to x n that is nothing but phi omega. So, phi omega is nothing but e to the power minus and sigma z square. We have already seen is nothing but this omega transpose c omega. By taking inverse Fourier transform relation at minus omega, you can get the corresponding joint density function that we do not have to do exercise. You take that formula for the joint density function.

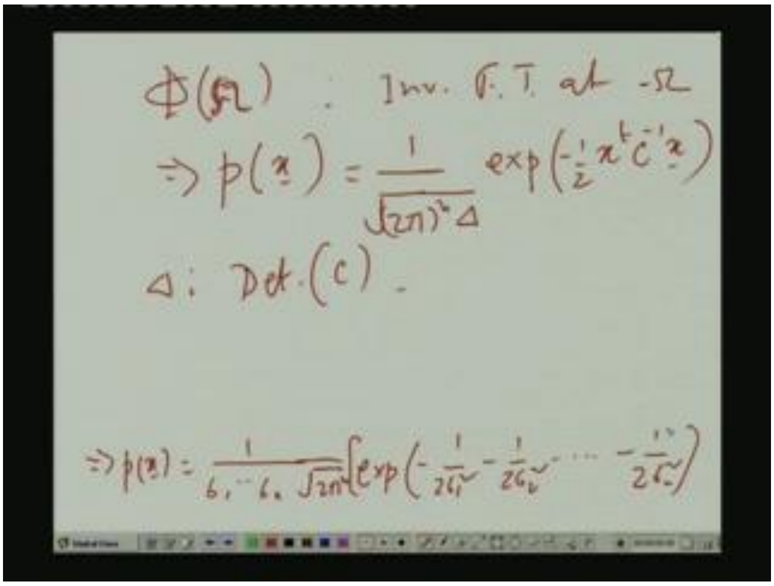
(Refer Slide Time: 55:24)



Finally, that is you just take the inverse Fourier transform of phi omega, take the inverse Fourier transform of this at minus omega that is you take inverse f t at minus omega, that will give you probability density of this x vector, that is x 1 x 2 up to x n. So, that inverse Fourier transform integration is just a calculus exercise. That I am not doing here. Time does not permit and it is not required. You can just trust me and you can use this expression always that is I am rewriting only. Delta is the determinant of the matrix c. I forgot to mention where delta is determinant of c.

Suppose x 1 to x n, they are statically independent, then obviously c is a diagonal matrix because the correlation terms are 0. C inverse will give I mean will be a diagonal matrix also. I mean ith term will be nothing but 1 by sigma i square 1 by because of inversion and sigma i square sigma i square corresponds to the variance of x i. So, in that case, you know that is if x 1 to x n are uncorrelated, then diagonal c is a diagonal matrix. C inverse matrix its ith element, that is i th diagonal element will be 1 by sigma i square, where sigma i square is the variance of x i. Obviously, determinant of c will be you can see product of this various terms sigma 1 square into sigma 2 square into dot dot dot sigma n square because we have got only one diagonal entry.

(Refer Slide Time: 58:21)



So, in that case and that is under square root, in that case p x becomes just 1 by sigma 1 after square rooting dot dot dot sigma n 2 pi to the power n square root into exponential minus 1 by 2 sigma 1 square minus 1 by 2 sigma 2 square dot dot dot minus 1 by 2 sigma n square. It amounts to just multiplying n individual Gaussian density functions for n random Gaussian random variables. That has zero mean and variances sigma 1 square 2 square dot dot dot dot sigma n square. So, we stop here today. In the next class, we talk about stochastic convergence and we go towards central limit theorem.

Thank you very much.