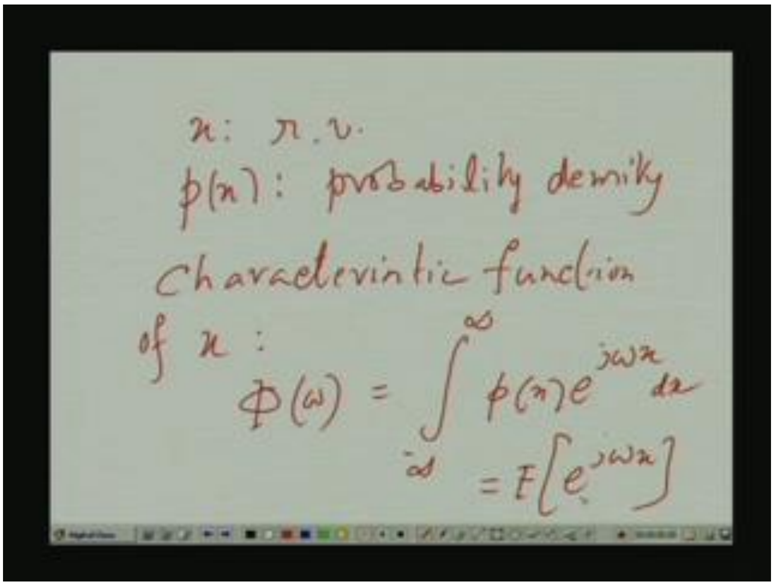
Probability and Random Variables
Prof. M. Chakraborty
Department of Electronics and Electrical Communication Engineering
Indian Institute of Technology, Kharagpur

Lecture - 11
Characteristic Function

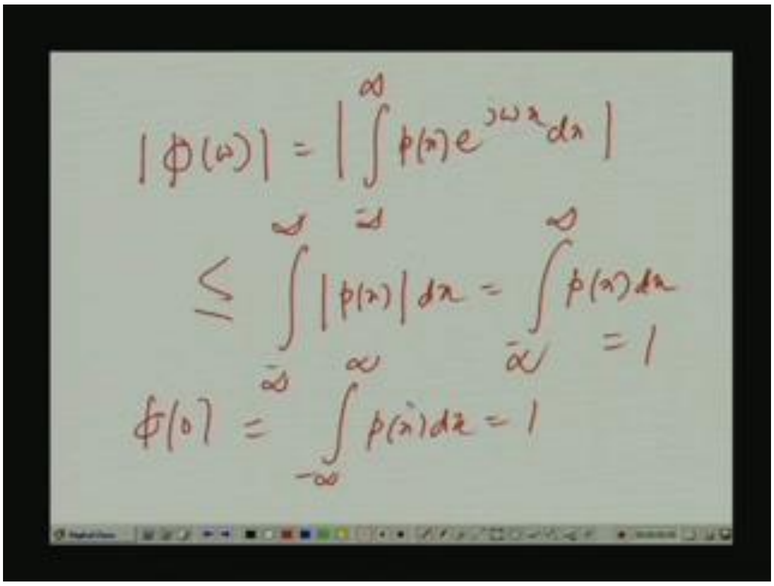
So, today, we consider what we discussed what we stated last time. We consider a very important topic, which is called the characteristic functions. Giving a random variable x with probability density p x, what is its characteristic function first.

(Refer Slide Time: 01:20)



So, given x – random variable; p x – probability density; then, the characteristic function of x is denoted as phi omega; this integral. It is something very similar to Fourier transform; only it is that, in Fourier transform, you have a minus sign here; here you do not have that. So, phi omega as such is a complex function of omega. Phi omega is a complex function of omega. It is called the characteristic function or sometimes the first characteristic functions of the random variable x. Obviously you can see that, we can also write this integral as the expected value of e to the power j omega x, because we have seen, given a function jx, its expected value is jx times p x integral. So, that is what is happening here. So, characteristic function is also characteristic function of a random variable x with probability density p x is actually the expected value of the function e to the power j omega x.

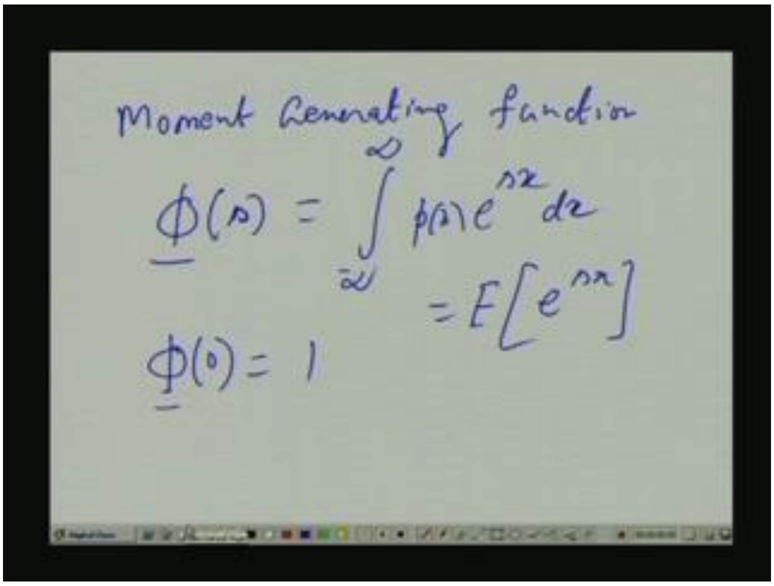
(Refer Slide Time: 03:10)



Now, as you know, using the triangle inequality, we all know that… But this is less than equal to… You know this triangle inequality, right? Sum of two sides is greater than equal to sum of third side. That we generalized in the case of… Here it is not discrete sum, but it is a continuous sum. So, mod of a summation – mod is equivalent to the length – length of the overall summation is less than equal to the individual lengths. So, if we apply that; so summation individual magnitudes p x – and magnitude of this is 1. So, it is transferred to be this. And mod of p x and p x – they are same, because p x is a real value – real and non negative value quantity. So, this is actually minus infinity to infinity equal to 1.

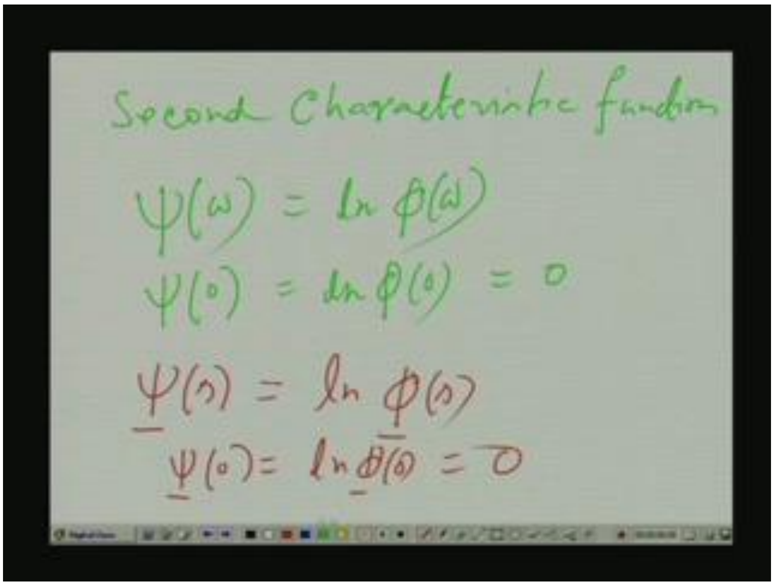
So, you see this characteristic function magnitude is upper bounded by 1. And when is it equal to 1? You can see that, what is phi 0? Phi 0 is nothing but p x dx, because e to the power j omega x at x equal to 0, at omega equal to 0, is 1. So, you have p x dx, which is 1; which means phi 0 is real and its value is 1. And the function phi omega – the magnitude of the function phi omega attains its maximum value of 1 at origin at omega equal to 0 and the magnitude falls after that. So, this is the property of the characteristic function. I repeat – phi omega – its magnitude rather attains its maximum value at omega equal to 0, that is the origin and the maximum value is 1; maximum value is 1. And then, at all other omega other than 1, other than 0, its value – its magnitude is less.

(Refer Slide Time: 05:50)



Now, you can generalize this concept and get some other definitions also, which is also equally useful. One of them is called the moment function – rather moment generating function. Here instead of phi, I put underscore. So, phi underscore. I restore omega – it is a general complex variable s – something what you are familiar with in the case of Laplace transform, where s is a general complex variable. This is this; which means we can also write it as expected value of this function – e to the power sx. And as before, you can see phi underscore 0; you can put 0 here. This is 1. And integral of p x dx, that is, 1. So, phi underscore 0 is 1.

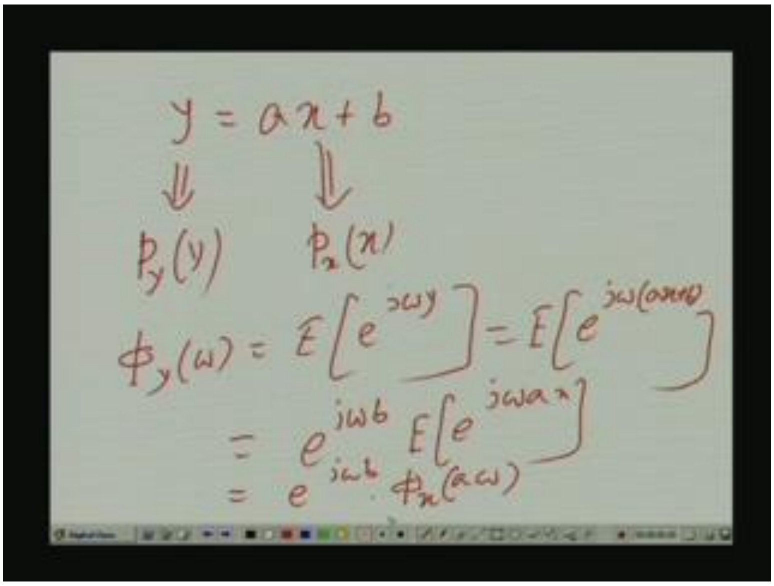
(Refer Slide Time: 07:05)



And then there is another function; that is called second characteristic function. Here instead of taking the phi omega or phi underscore omega or phi underscore s, we rather concentrate on their logarithms. So, instead of phi omega, we take ln of phi omega; or, instead of phi underscore s, we take ln of phi underscore s; both are second characteristic functions. So, the second characteristic function psi omega is nothing but ln phi omega. And you can see what is psi 0; ln phi 0; but phi 0 equal to 1; we have seen. So, this is equal to 0.

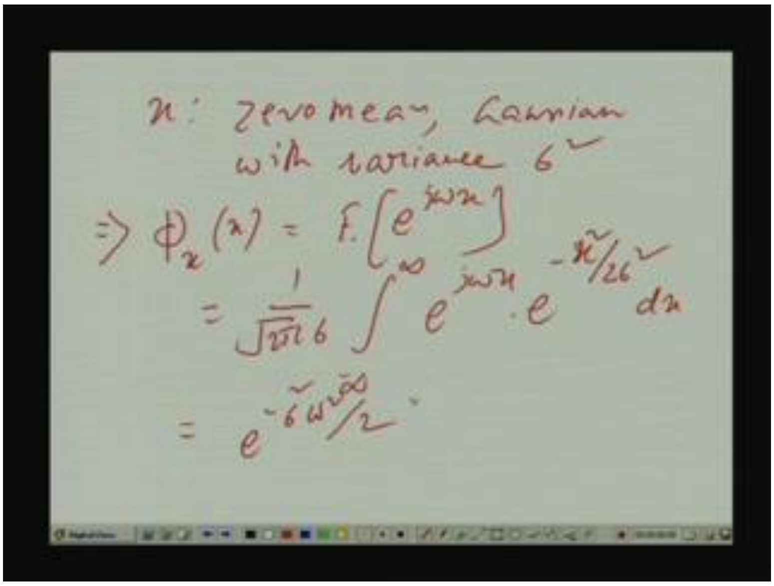
So, when you are dealing with phi omega, you take ln of that; you get psi omega. And psi omega is the second characteristic function. But suppose you are dealing with the moment function or moment generative function – phi underscore s; I am putting underscore because phi and phi underscore – they are two different functions. And also, it is a function of the variable omega; it is a function of the generalized complex variable s. If you take ln of that, instead of doing here – instead of doing this; if you take ln of phi underscore s, then again you get a different function; I call it psi underscore s. This is also equally useful. Either you live with this definition psi omega or psi underscore s. And as before, psi underscore 0 ln phi underscore zero. And phi underscore 0 was 1 we have seen. So, this is equal to 0 as before.

(Refer Slide Time: 09:34)



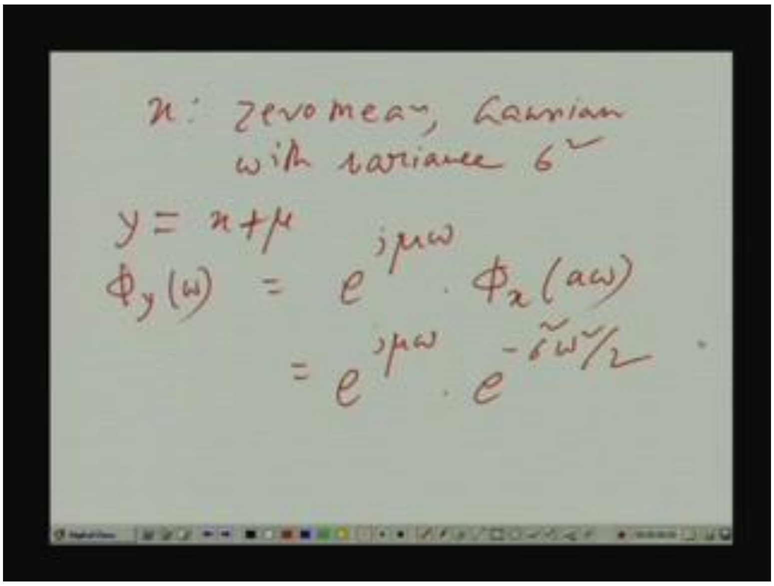
Now, we see one thing; suppose x is the random variable and you know its probability density function p x. At y, is a function of x; it is a linear function of x given by this equation y equal to ax plus b. Then, what is… At y, has its own probability density function p y y; but this is just the symbol; where this is the variable; the y within parenthesis is the variable actually. And this subscript y is just a symbol showing that – meaning that, this is the probability density of the variable y; and x is the random variable that has got probability density p x x. Once again this x within parenthesis is the actual variable and this subscript is x denotes, that is, the function for the random variable x. Then, what is phi y omega? This is nothing but expected value of… But y is ax plus b. So, this is again… And you have this quantity. This is outside expected value of ax; which means e to the power j omega b phi x – not omega, but a omega. This is a very useful derivation and used in the practice.

(Refer Slide Time: 11:30)



As an example see this. Suppose x – zero mean Gaussian with variance... In that case, what happens to phi x x? This is expected value of e to the power j omega x. This is nothing but e to the power j omega x multiplied by the probability density e to the power minus x square by twice sigma square dx, because it is Gaussian. This integral is well-known; I am not deriving it; I mean this is standard result for this integral that is equal to e to the power minus sigma square omega square by 2. This is for the zero mean random variable. But now, suppose… But suppose now x – it is not zero mean, but it has got a finite mean mu. Then, what happens? That is… First, let me erase this.

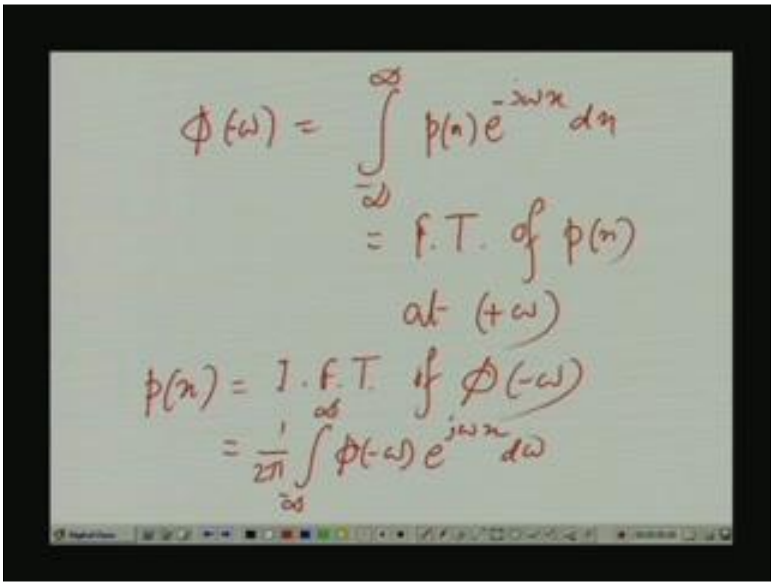
(Refer Slide Time: 13:21)



Suppose y is a random variable; once again Gaussian with the same variance sigma square, but a mean mu; which means y is actually x plus mu. What will be the characteristic function now? What is phi y omega? Phi y omega as you have seen is nothing but… This equation you can interpret as ax plus b. So, a equal to 1 and b is mu. So, from our previous result, it will be e to the power j mu omega and then phi x a omega; but a equal to 1. So, you will get this quantity only; which means e to the power j mu omega e to the power minus sigma omega square by 2.

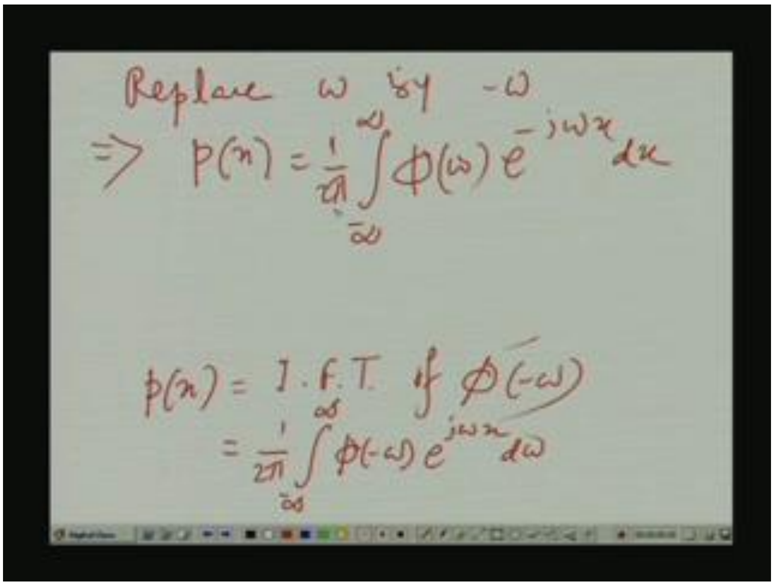
So, given the probability, given the random variable x with probability density p x, we can compute its characteristic function by taking their integral. But given a characteristic function, how to get back the corresponding probability density? So, it is… So, we can obtain that by just applying the well-known results from Fourier transform, where on one hand, you have got the direct Fourier transform, which gives you the function in the Fourier frequency domain. And from frequency domain, you can get back the time domain function by the inverse transform. So, the same thing can be applied here.

(Refer Slide Time: 15:44)



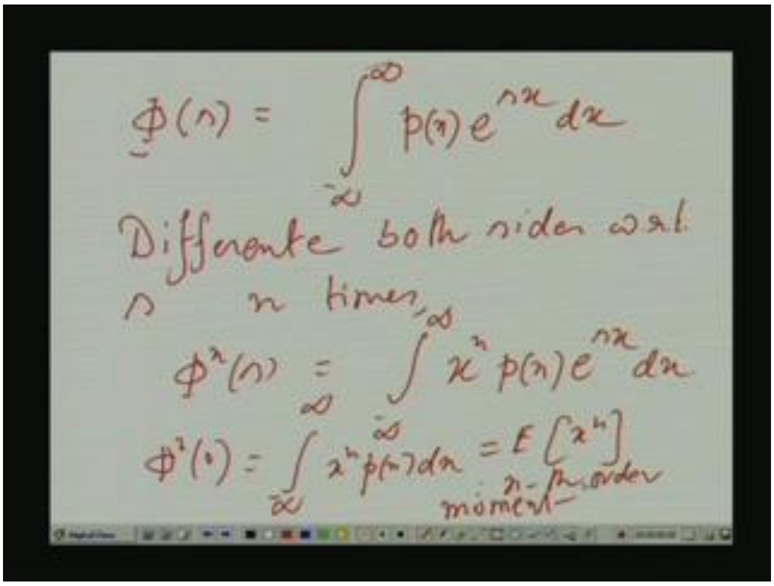
You can see one thing. I repeat we have phi omega is equal to… This was plus. So, this is nothing but Fourier transform of p x at minus omega. Or in other words, if I put a minus here, you get a minus here; so phi minus omega. Then, I put plus. If you put phi minus omega, then it becomes simply Fourier transform of p x at plus omega. So, that means, what is p x? Inverse Fourier transform of phi minus omega. So, p x is IFT – inverse Fourier transform of phi minus omega; which is nothing but 1 by 2 pi integral minus infinity to plus infinity phi minus omega e to the power j omega x d omega. Now, you replace omega by minus omega. So, d omega becomes minus d omega; a minus sign comes. Here you have got minus sign e to the power minus j omega x. This becomes phi omega; outside we have a minus, but the limits becomes plus infinity to infinity. So, if you want to again make it from minus infinity to infinity, minus will again come. So, the two minuses cancel each other. So, what you get back is the following.

(Refer Slide Time: 18:24)



That is replace… I repeat – d omega becomes minus d omega. So, minus sign would come here. But limits becomes from plus infinity to minus infinity. If you reverse it, again a minus sign comes. So, if you want to stick to minus infinity to infinity, another minus sign will come. So, they will cancel each other. So, you will get back what you had. And minus omega replaced by omega means it becomes phi omega; and this becomes… You get a minus sign here. So, this is the inversal formula. From the characteristic function, how to get back p x?

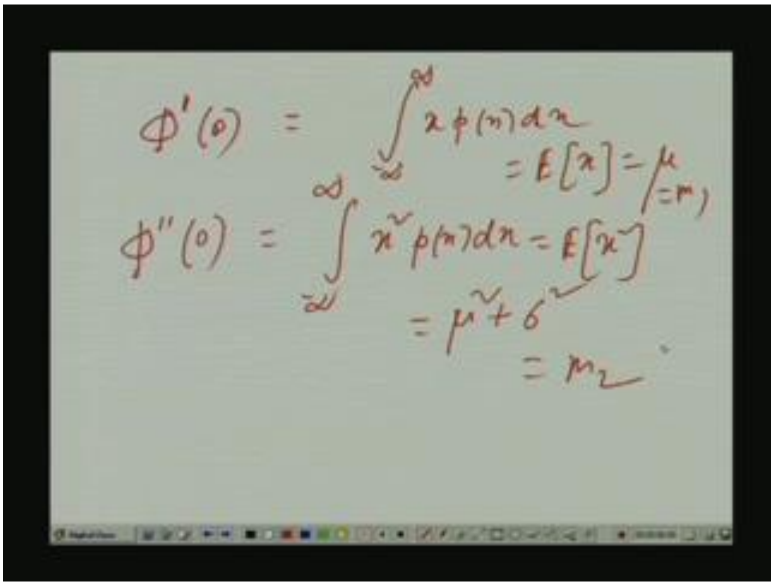
(Refer Slide Time: 19:26)



Next, we consider that moment function – phi underscore s. What was it? This was p x e to the power sx dx from minus infinity to infinity. If I differentiate both sides with respect to s – I repeat with respect to s; that is, s n times; what happens? If you differentiate once, x will come here; if you differentiate again, another x will come and likewise. e to the power sx – if you derived with respect to s, you get back x into e to the power sx, then x square e to the power sx; s cube e to the power sx and likewise. So, n means n-th order derivative; it is not power. If i say phi n s, it means n-th order derivative of phi s with respect to s. It is nothing but…

Then, what happens if s equal to 0? s equal to 0 means this becomes 1; you have got this quantity; which is nothing but expected value of x to the power n; which is nothing but n-th order moment. That is why this function is called the moment generating function. We can use it; that is, first you differentiate with respect to s maybe n times; then, put n equal to… put s is equal to 0. What we then get is nothing but the n-th order moment. So, once again see this formula; you differentiate the function – phi underscore s; that you put underscore; we forget to put – phi underscore s n times and then put s equal to 0. That will give you the n-th order moment.

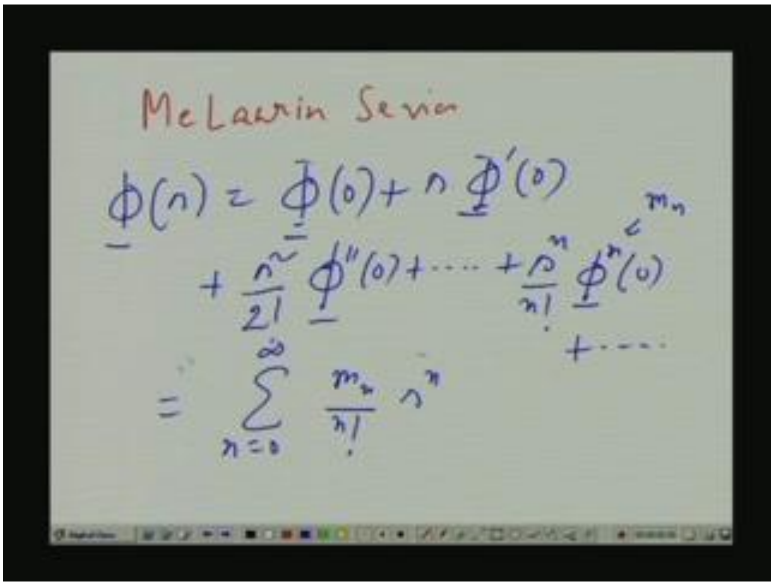
(Refer Slide Time: 22:18)



As for example, if n equal to 1, then what happens? If n equal to 1, then what happens? In the integral, you have got x to the power n; which is nothing but x at p x dx. It is further is equal to E x, that is, mu. And what is 0? That is now n equal to 2 phi – double prime means second order derivative at origin at s is equal to 0. What happens to that?

That is nothing but… As we have proved earlier, x square p x dx. And we know this is equal to mu square plus sigma square. Mu was first order moment m 1; this is second order moment m 2. So, you can see this.

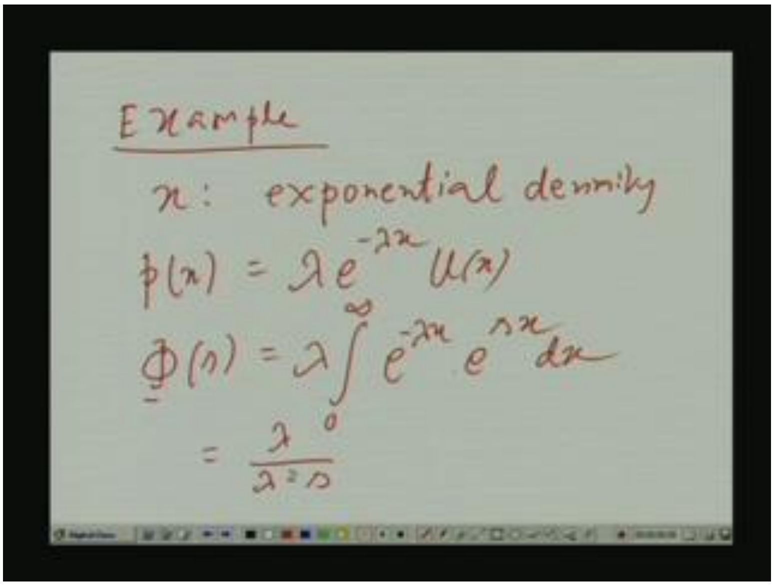
(Refer Slide Time: 23:53)



One more interesting thing we can now consider. We all know Mc Lawrin series. We can expand phi underscore s as phi underscore 0 – then s into phi prime – then s square by factorial 2 underline dot dot dot s to the power n by factorial n. This n means actually n-th order derivative plus dot dot dot dot. Of course, this Mc Lawrin series converts this near s equal to 0 if all of them – phi 0, phi prime… In fact, you can write them as… We know; we have seen earlier that, phi underscore 0 – this is equal to 1; this you have seen. Anyway, phi – n-th order derivative of this function phi underscore s at s equal to 0 – that is nothing but… This quantity is nothing m n – n-th order moment. So, this summation actually you can write as s to the power n; n from 0 to infinity. This of course converges to phi underscore s near s equal to 0 if converges absolutely if these moments are finite. If the moments are finite, the series converges absolutely near s equal to 0. But that is just a mathematical point.

Now, see one thing – from phi underscore s, we can get back our original probability density function for the random variable x; which means any probability density function can be equivalently and exactly described by all its moments. Earlier we have seen that, given the probability density function, we can find out this moment generating function and from that all the moments. Now, this is the reverse phenomenon; given all the moments, by this series, you can find out phi bar s; and from phi bar s, you can find out phi underscore s; and phi underscore s from that, you can find out what is the original probability density function. Let us now take an example.

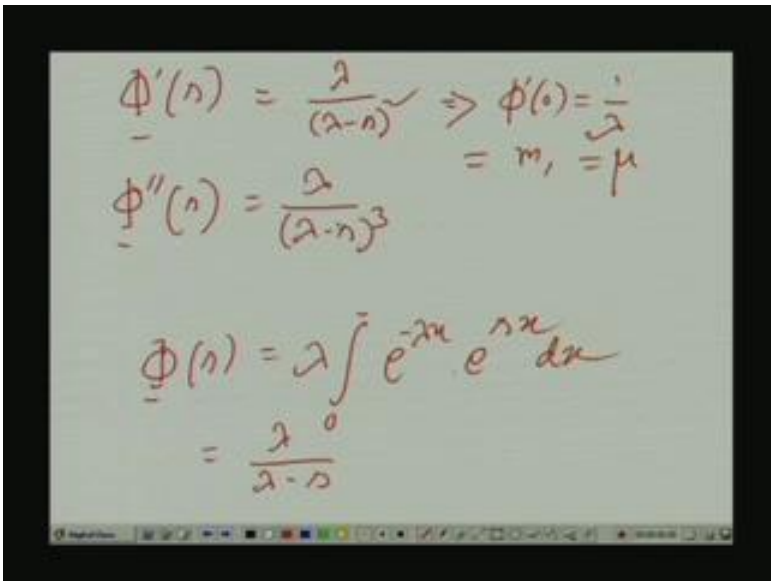
(Refer Slide Time: 27:21)



Suppose… So, this is an example. Suppose x has exponential density; p x is actually lambda; lambda is a parameter – e to the power minus lambda x u x. u x actually means – it is a unit step function; that is, for x equal x less than equal to 0 minus; for all negative x, this is 0. And for all x greater than or equal to 0, this is 1. So, this probability density has got value 0 for all negative x. And for all probability for all x greater than equal to 0, that is, for all non-negative x, its value is lambda into e to the power minus lambda x. Lambda is a parameter, which is found out by integrating this with respect to x and equating that to 1, lambda is found out.

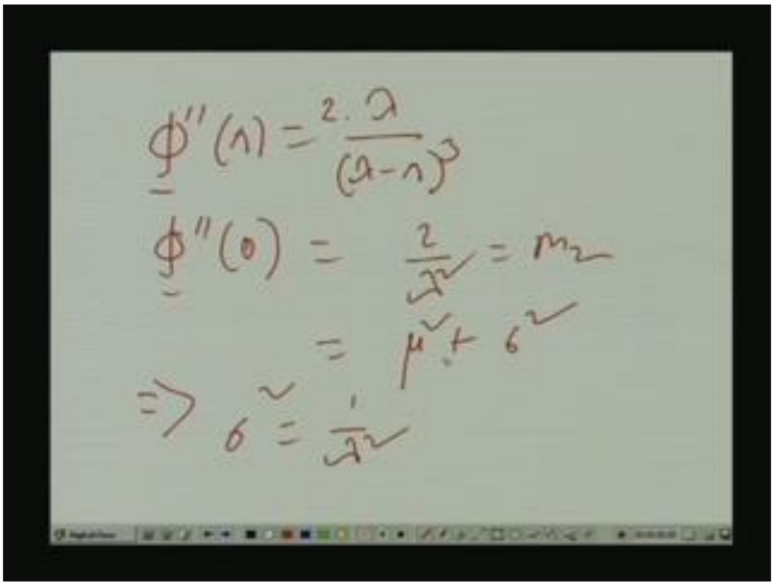
Now, for such a function, what is phi underscore s? Very simply we take the integral and take it from 0 to infinity only, because this unit step function is present. And lambda, which goes outside – e to the power minus lambda x; so p x times e to the power sx; if you integrate, this becomes… You take e to the power minus – within bracket – lambda minus s; then, x. If you integrate, then at x equal to infinity, that will gives rise to 0 at x equals to 0; we will have 1. So, there will be a factor lambda minus s below. So, we have this.

(Refer Slide Time: 30:02)



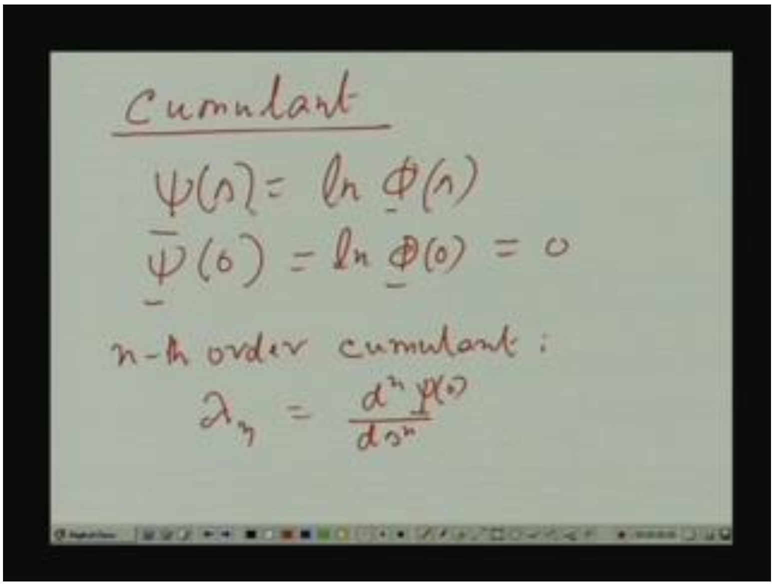
So, what is that is equal to… If you differentiate, you have got lambda minus s to the power minus 1 here. So, that becomes minus 2; minus comes; that gets cancelled, because another minus comes from here. So, meaning is 1 by lambda; which is equal to first order moment – also equal to the mean. And what is this another derivative? Again it becomes 3 – minus sign comes, but that gets cancelled with the minus sign coming from here. So, this is lambda by lambda minus s whole cube.

(Refer Slide Time: 31:20)



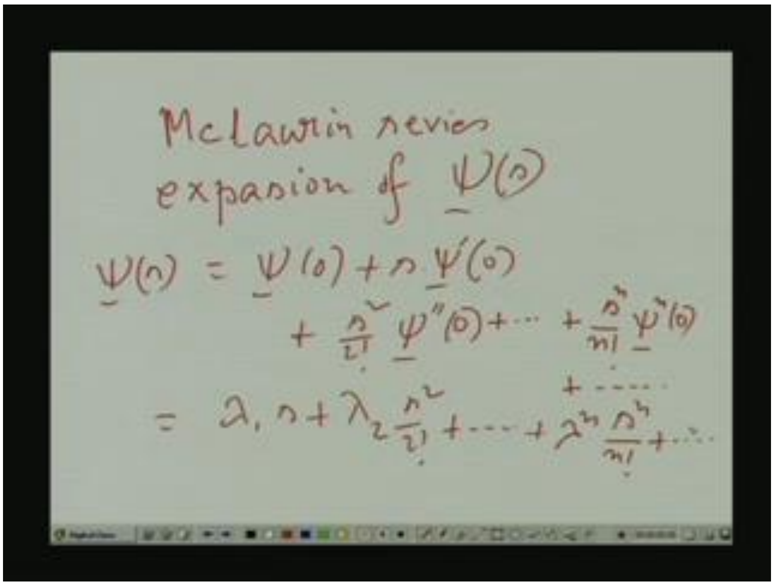
Similarly, this is equal to lambda by whole square. This should be 3 and here should be 2. So, minus signs get cancelled with each other. So, what happens to this function? This is equal to 0. This is simply 2 by lambda square. This is m 2 as we have seen; which is nothing but mu square plus sigma square – meaning mu is already… We have already seen mu is equal to 1 by lambda. So, 1 by lambda square; which means sigma square is 1 by lambda square.

(Refer Slide Time: 32:43)



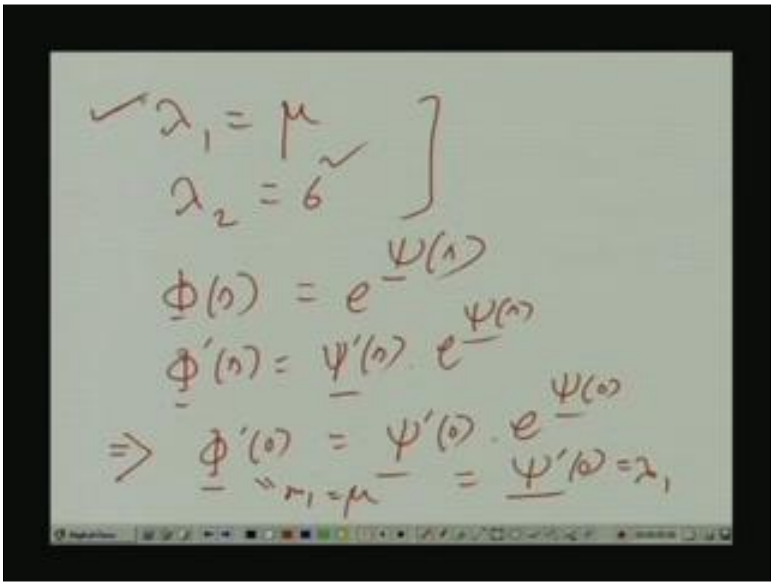
Next, we consider another very important function which is called cumulant. Now, earlier we have been… I mean so long we have been considering the characteristic function or last time we considered the moment function, that is, moment generating function. But as we have done, as we have stated in the very beginning, there is another class of function; that is called the second characteristic function, where we do not take the characteristic function as such or the moment as such, but take their logarithms. So, we will be using that. So, again we recall; we had psi bar s is nothing but ln phi bar s. And also, we recall that, psi bar 0 was ln phi bar 0. But phi bar 0 is what? That we have seen is equal 1. So, this quantity is 0. Then, we define the n-th order cumulant – lambda n. Actually, you differentiate this phi bar s with respect to s n times; then, substitute; then, put s equal to 0. The first question is what is lambda 0? Lambda 0 means no derivative; but that means we have only psi bar 0. But what is psi bar 0? That is 0. So, lambda 0 is 0. Then, we go for…

(Refer Slide Time: 35:00)



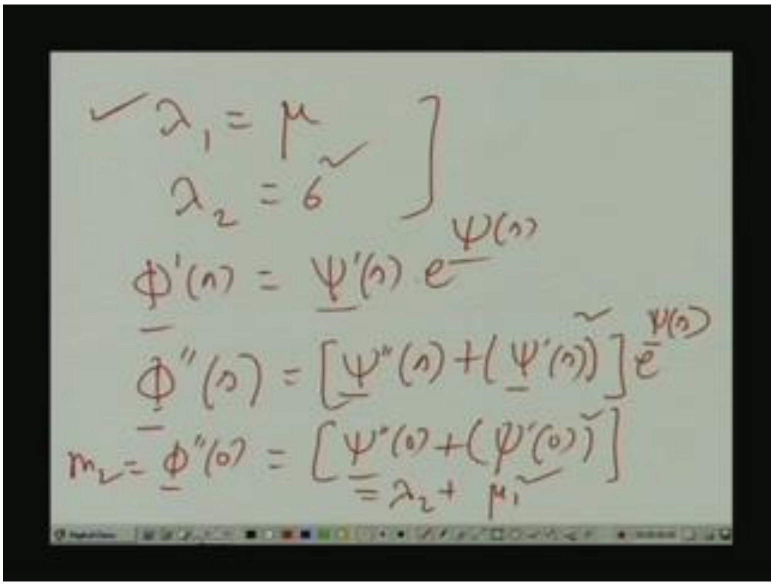
As before we go for the Mc Lawrin series expansion. This is psi bar 0 plus s times… At this quantity, psi bar 0 is 0. This quantity is psi prime psi bar 0 prime, that is, single derivative – that was lambda 1; then, we have got lambda 2 and so on and so forth. So, n-th order derivative of psi s, psi bar s. And then, at s is equal to 0, that is, lambda n. So, this you can write it as lambda 1 s plus lambda 2 s square by factorial 2 dot dot dot lambda n s to the power n by factorial n plus dot dot dot dot.

(Refer Slide Time: 37:20)



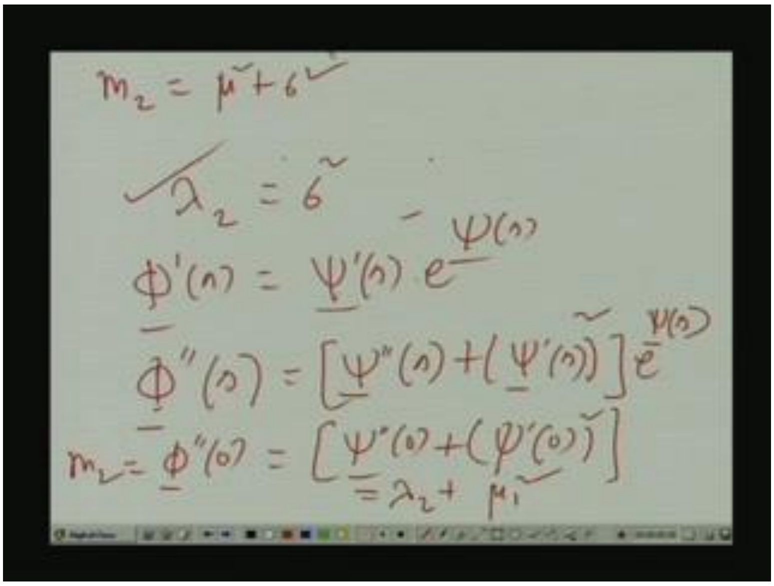
One more thing – we say that, lambda 1 will be same as mu and lambda 2 will be sigma square. How? I just give you a minute to think. I repeat lambda 1 equal to mu and lambda 2 – sigma square; just think for a minute. I do not think it is very difficult. Anyway, let me work it out. We have seen that, phi bar s – ln of that is equal to phi bar s; which means this is nothing but e to the power psi bar s. So, if you differentiate it with respect to s once, what you get? You get these things first and multiplied by the derivative of this. So, psi bar s – derivative of psi bar s multiplied by this – meaning what is this? But as you have seen, psi bar 0 is 0; e to the power 0 is 1. So, this is equal to this. But what is this quantity? This quantity is m 1; this is equal to m 1 equal to mu. And this is equal to lambda 1. So, this is proved.

(Refer Slide Time: 39:56)



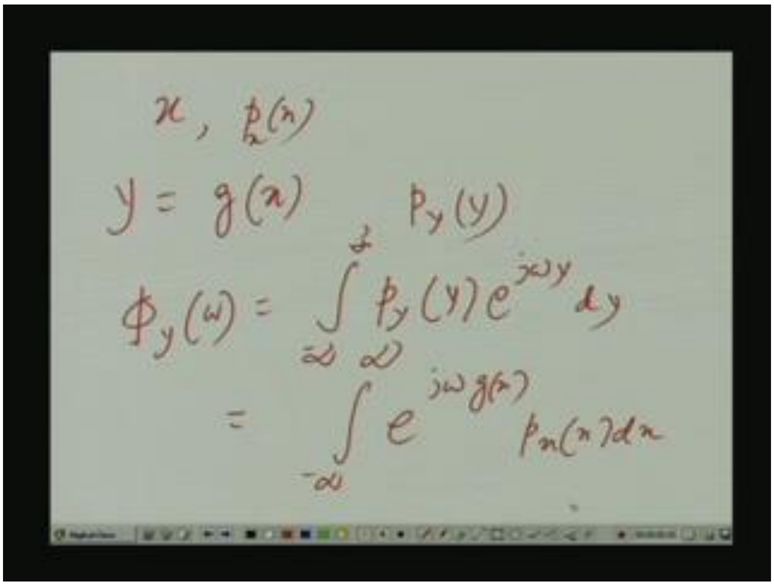
So, we come to the next thing. What happens to… We have just now seen that, phi prime bar s is this. What happens to the second order derivative? Differentiate this, hold this; and then, hold this, differentiate this. So, if you differentiate this, it become psi double prime s. This quantity goes outside. And if you hold this and differentiate these, you get back psi prime s again. So, s square; which means lambda 2; which is nothing but… This is equal to plus… This is actually m 2 whole square; and e to the power – as you have seen, e to the power psi 0 – that is equal to 1 because psi bar 0 – psi underscore 0 is 0. So, this is equal to 1. So, I do not write it if you have got this quantity; which is nothing but… This is nothing but lambda 2 and this is lambda 1 square. But lambda 1 is same as mu 1; we have seen. So, this is mu 1 square.

(Refer Slide Time: 42:17)



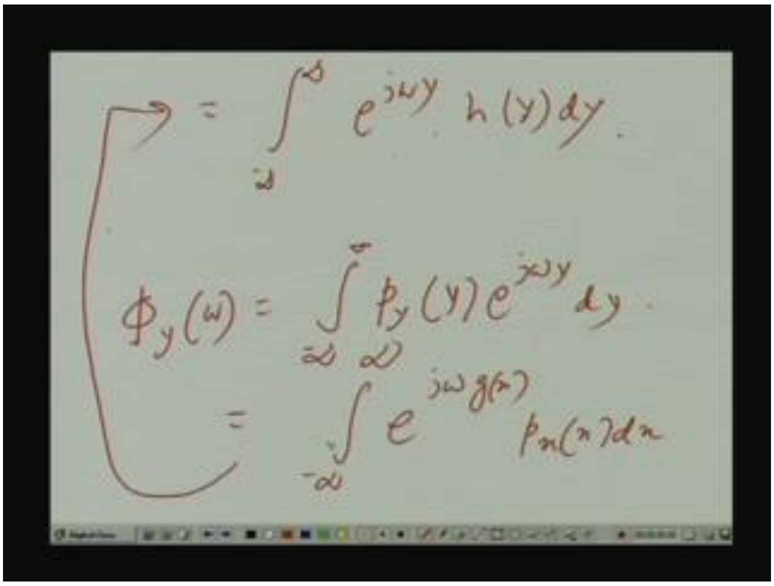
And, we also know that m 2 is nothing but mu square plus sigma square. So, if you compare, obviously, lambda 2 equal to sigma square. So, this is proved. Cumulants – I will not consider any further; but you should always… I mean remember this function, because this has got tremendous use.

(Refer Slide Time: 43:02)



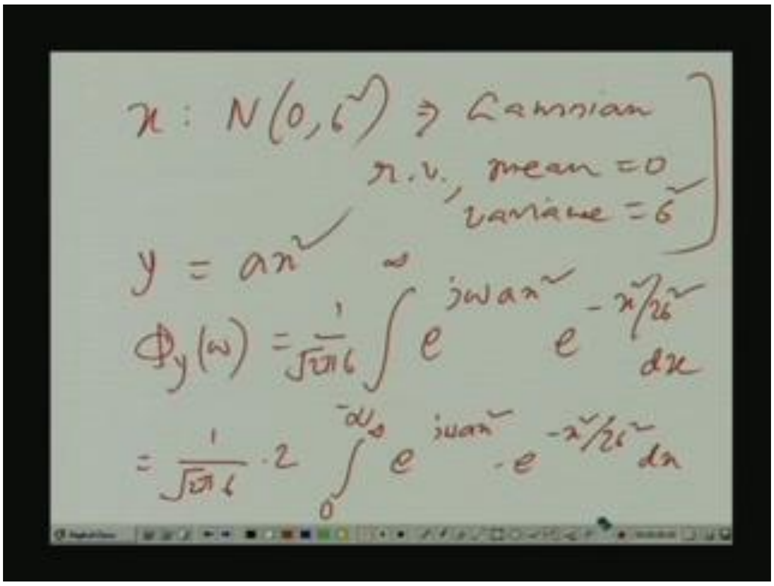
Finally, given x and p x at the random variable y, which is a function of x; we know… We have seen already that, given the probability density of x, that is, p x x; how to get the probability density of y – p y y? How to get this? But while doing this, you can also make use of the characteristic functions. What is phi y omega? But y is g x. This is nothing but expected value of e to the power j omega y. But we know y is g x. So, you can also write it this way – e to the power j omega g x p x x…

(Refer Slide Time: 44:55)



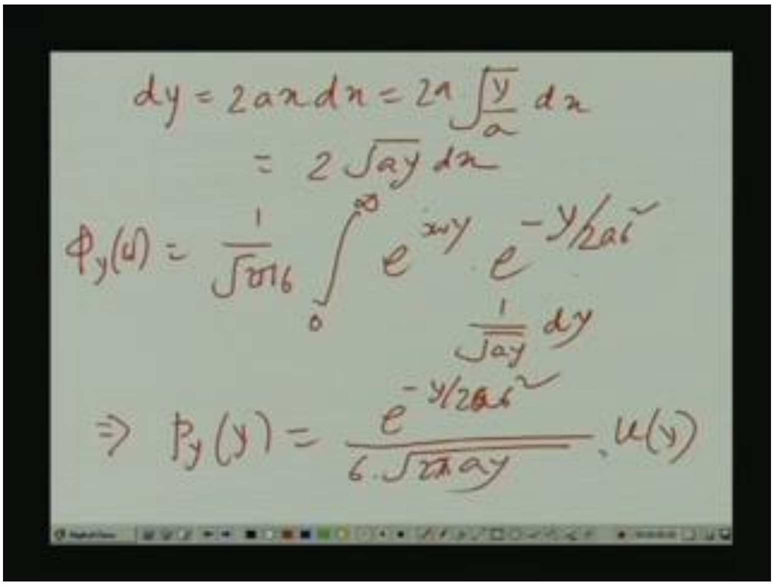
Now, suppose you start with this integral and you can write it by some manipulation; suppose you can write it in this way; you can write it like this. Suppose this integral – you can write by some manipulation like this – e to the power j omega y, which is the first quantity – e to the power j omega g x; where, j omega y are same times of h y dy. Then, just comparing this with this, you can see that, h y should be equal to p y y. So, we first write this equation; then, try to write it in this form. Whatever h y comes out, that will be equal to p y y. We just consider an example and then we call off.

(Refer Slide Time: 45:46)



x – this symbol means actually N for normal, which is same as Gaussian; which means Gaussian random variable, mean – 0; variance – sigma square. And given that, y is equal to ax square. Now, we know what is phi y omega. We can write it simply like e to the power j omega y, which is ax square and p x x. p x x we know – 1 by root 2 pi sigma e to the power minus x square by 2 sigma square dx. Since we have only x square here, we can take the integral from 0 to infinity and multiply it by 2. So, in that case, integral becomes 1 by root 2 pi into sigma times 2. Now, in this interval 0 to infinity, we have got of course the function y equal to ax square. But within the interval 0 to infinity, the relation between y and x is 1 to 1. For any x, you get only one y; and for any y, you get only one x.

(Refer Slide Time: 48:30)



And, there you have got dy, which is twice ax dx, which is twice a – y equal to x square; so x equal to square root of y by a dx; which means twice this. So, dx is nothing but dy by twice square root ay. And the first quantity is e to the power j omega y. We can always write it like this. There was 2, but that will cancel with this 2 – e to the power x square as you know is minus y by a; y by a; and then, you have got this function – 1 by this; which means this quantity – 1 by root 2 pi sigma e to the power minus y by twice a sigma square into 1 by square root ay. This is p y y – sigma square by sigma root 2 pi ay. And of course, this is on the positive side – u y; u y obviously; whether x takes negative or positive, y can always take only positive values. So, for any negative value of y, the corresponding probability is 0 because it just cannot take place.

So, I stop here today. And we start with the new topic next time. But if there is any question, I would suggest that you all consider all these theories and try to solve some problems on your own from Papoulis. Also, please go through the worked out examples, because due to time constraint, I cannot take up this worked out examples, which are already there for you to read. So, that is all for today.

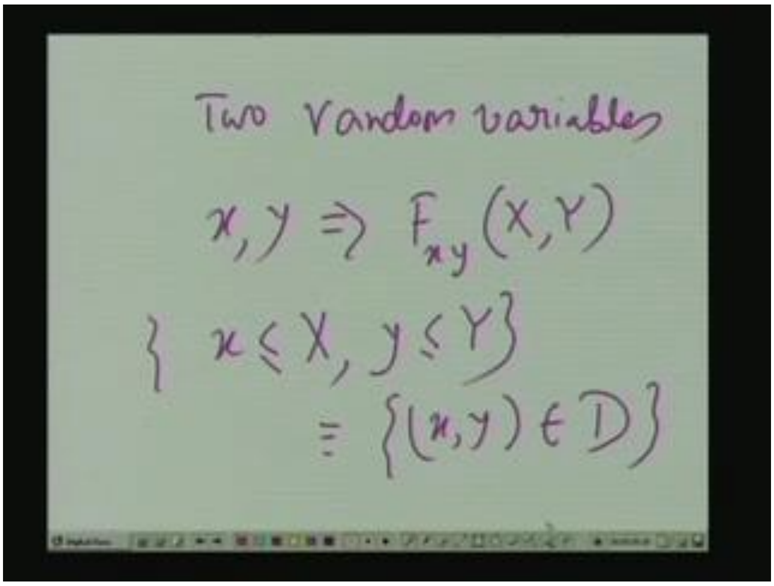
Thank you very much.

Preview of Next Lecture
Lecture – 12
Two Random Variables

So long, we have discussed just one random variable and function of one random variable. We so far did not discuss what is called joint variation. There is earlier we considered a random variable x; its probability density, probability distribution and a function of this random variable and all that. Similarly, there could be a separate random variable y; we discussed its properties and all that. But there is an important topic actually; where, x and y – they are two random variables, which are mutually related; that is, the underlying process is such that, x and y has got some kind of interconnection – I mean interdependence rather, then the question of… that individual variation, but are the joint variation and joint statistics comes. For instance, you can consider these two variables: one is say temperature of the environment; and humidity in the environment.

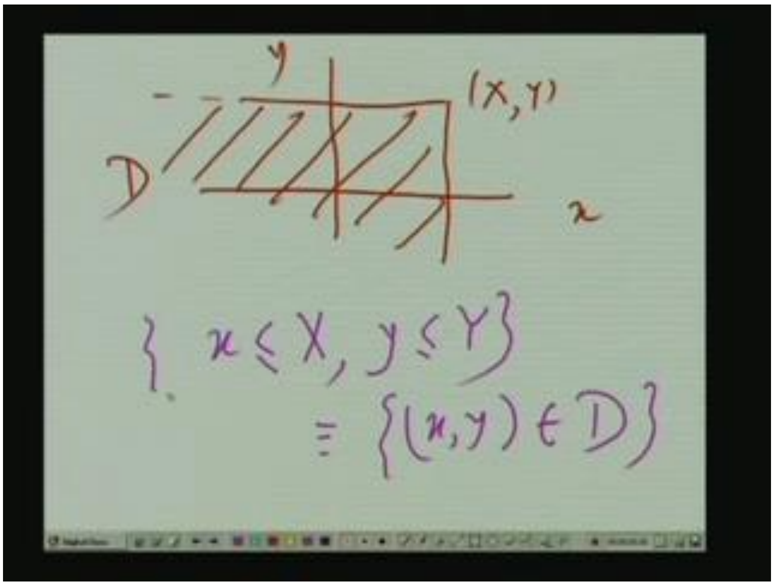
Now, temperature could be a variable t; humidity could be h. Now, you all can visualize that, the underlying physical process is such that definitely one depends on the other. Surely, if the temperature is high, there will be one type of humidity; if the temperature is less, another type of humidity and vice versa. Now, we do not go into this physical modeling because that is very difficult; but we treat both t and h as two separate random variables, which are jointly related – mutually related. Obviously, because if t is high, maybe h will be less or h will be high. Again if h is less, t could be high, less, because that is determined by the mutual interdependence. So, we define…

(Refer Slide Time: 54:55)



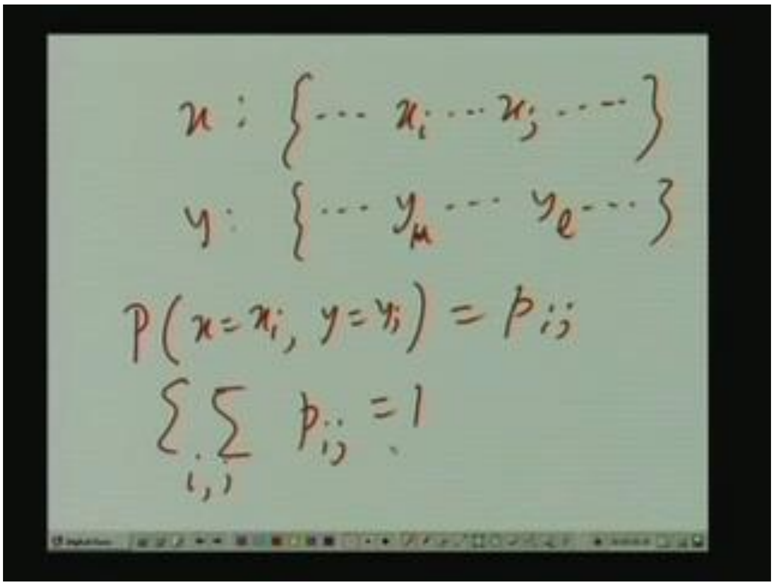
So, our topic today is… In fact, now we are considering two random variables; but in general, it will be say n number of random variables. So, we will extend this concept to more than two number of random variables. So, we now consider always a pair. Under this topic, what is a pair? x is random variable; y is a random variable. But since they are related, we consider x, y. In that case, we have to define the joint probability distribution of x and y. Now, what does it mean? Let us give a definition actually. We consider the event. This random variable x taking value less than equal to some given number capital X; y taking value less than equal to some given number capital Y; which actually means this pair is taking value for a particular region D in the xy plane. What is that region?

(Refer Slide Time: 56:28)



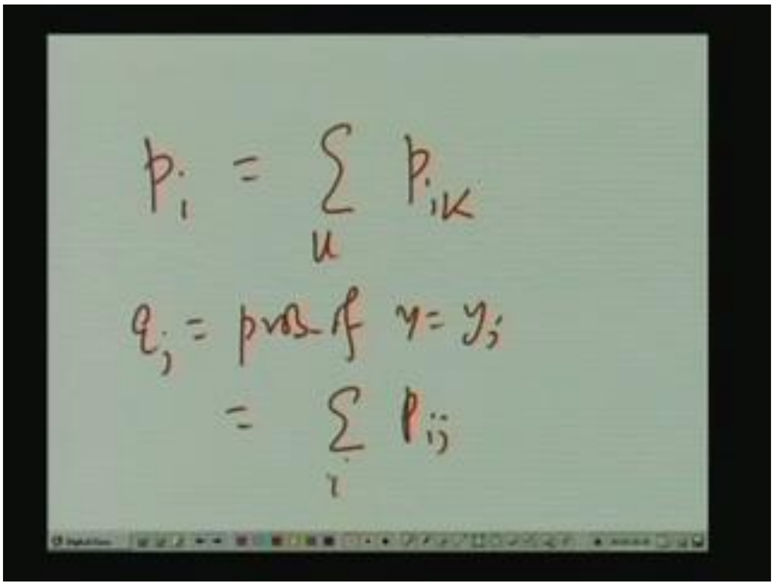
Suppose this is xy plane – x, y; and this point is the capital X, capital Y; that if you draw this line, this area is D. Obviously, within this area, both these things for discrete variables also.

(Refer Slide Time: 56:55)



Suppose x and y – they are two discrete random variables; x takes values probably set; I mean it may not be finite. So, let me not put it this way. x i, x j – like that. It could be a finite set; it could be an infinite set. Similarly, y k, y l dot dot like that. In that case, the joint probability that x takes the value x i, y takes the value y j is denoted by p i j. Obviously, i and j. So, we put double summation; that should be equal to 1.

(Refer Slide Time: 58:05)



And, p of i – that is obviously p i k and summed over all possible k's – w p i. If q j is the corresponding probability of y equal to y j, this will be again equal to… p i is the probability of x equal to x i; that is, this summation. p i is the joint probability. And here it is q j, is nothing but again p i j, but summed over i. This is pretty obvious. So, in today's class, we conclude here; we will stop here. We continue to expand on this – on this theme getting into what is called probability mass. There is a very important concept called statistical independence, then joint characteristic functions and things like that. That will be for the next time.

Thank you very much.