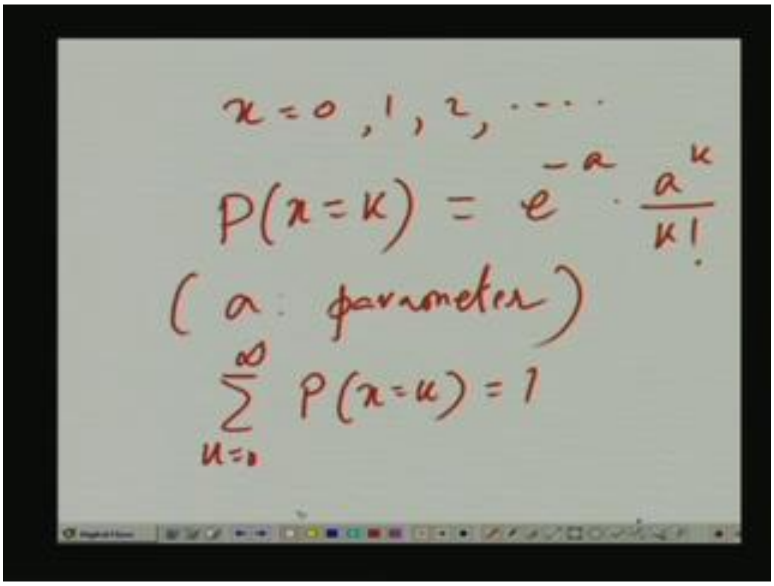
Probability and Random Variables
Prof. M. Chakraborty
Department of Electronics and Electrical Communication Engineering
Indian Institute of Technology, Kharagpur

Lecture - 10
Moments

In the last class, we gave a problem. I mean we said that x is the random variable which is poison distributed and it takes values is a discrete random variable. It takes values like, you know x equal to either 0 or 1 or 2 dot dot dot dot dot and the corresponding probabilities is given by like.

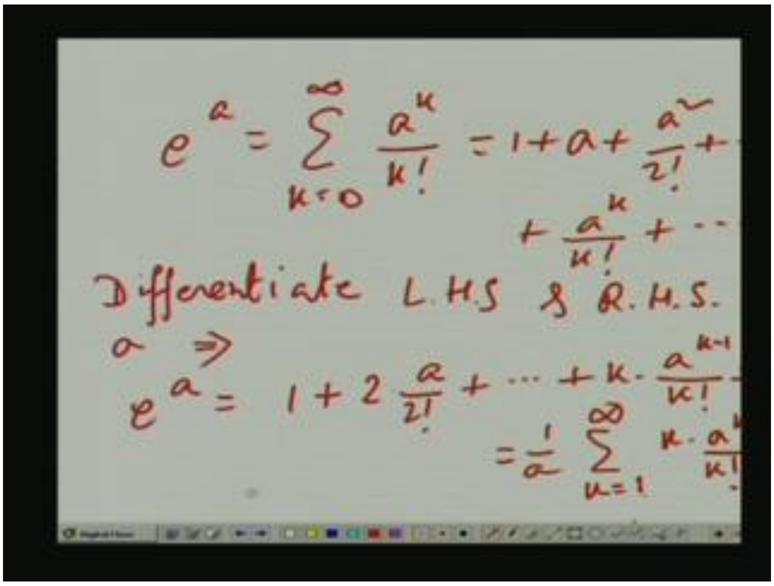
(Refer Slide Time: 01:09)



P if P equal to I mean for x equal to k the corresponding probability will be given by e to power minus a and a to the k divided by factorial k where a is a parameter. We ask you, to find out the mean and variance or rather we ask you to find out e of x expected value of x and expected value of x square. And if know expected value of x square and if you know the mean.

Obviously, you can find out the variance that was a question a is called a parameter. So, we say that x is poison distributed random variable with parameter a. can you tell us, what is the function of this factor e to the power minus a. Actually, if you just consider these parts a to the power k by factorial k this will correspond various case. If you take that values and add them this gives rise to an exponential series. In fact, we should have this thing this must be equal to 1 right, because total probability is 1. Now, because of this factor this equality will be satisfied this you can see easily.

(Refer Slide Time: 01:55)

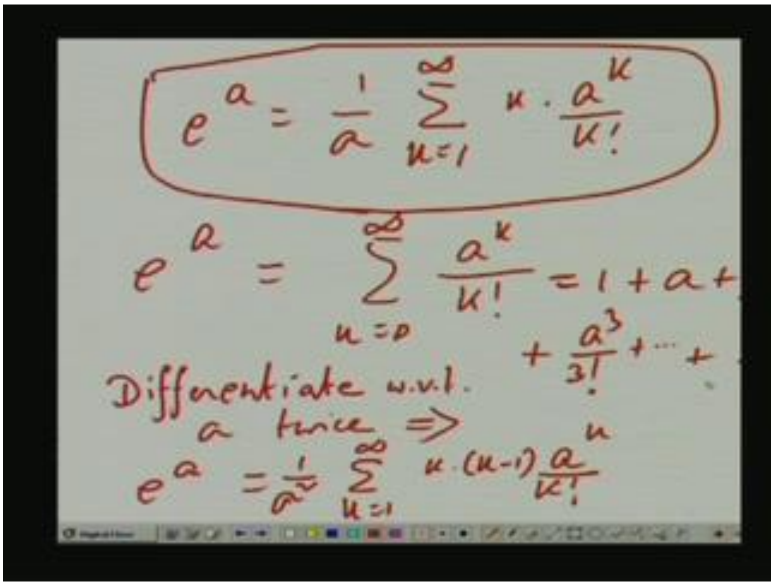


In fact, now if you put that expression take this out a equal to 0 to infinity and we have got a to the power k divided by factorial k. Now, this is nothing but e to power a. There is a exponential series e to power a when expanded gives you this. So, you e to the power minus a into e to power a is equal to 1. So, just to make this equal to 1 this factor has been added introduced.

Anyway, this is just a side issue we have to now, find out the mean of x that is e of x and e of x square to do that, we consider this exponential series. E to the power a which I have just a while we have stated is equal to a to power k divided by factorial k; k is 0 to infinity, which is equal 1 plus a a square by factorial 2 plus dot dot dot dot plus a to the power k by factorial k plus dot dot dot dot right.

I now differentiate these left hand side and also right side with respect to a once. And, then again 1 more time there is twice. So, first if I differentiate LHS and RHS with respect to a. This will give rise to this side as before e to power a. And this side we have got 1 plus 2 into a by factorial 2 plus dot dot dot dot plus k into a to the power k minus 1 by factorial k dot dot dot dot. Which in fact, you can write as if you take out 1 by a out, you can write it as k a to the power k divided by factorial k. This should not be this is from 1 for k equal to 1 we have got 1 1 cancelling and a to the power 1 and that a and a cancels. So, we get 1 and likewise. So, the summation is from k is equal to 1 up to infinity this. So, e to the power a is equal to this; this is 1 identity which I need I will be using later. So, as I move to the next page I rewrite this.

(Refer Slide Time: 06:20)



So, but so what we have just now seen is this e to the power a is nothing but I mean 1 can always also write like this. If we differentiate it again this is 1 result. And if I differentiate this again with respect to a left hand side I get e to the power a as before while I am not differentiating these. Let us take down that series again e to the power a that was written as a to the power k by factorial k.

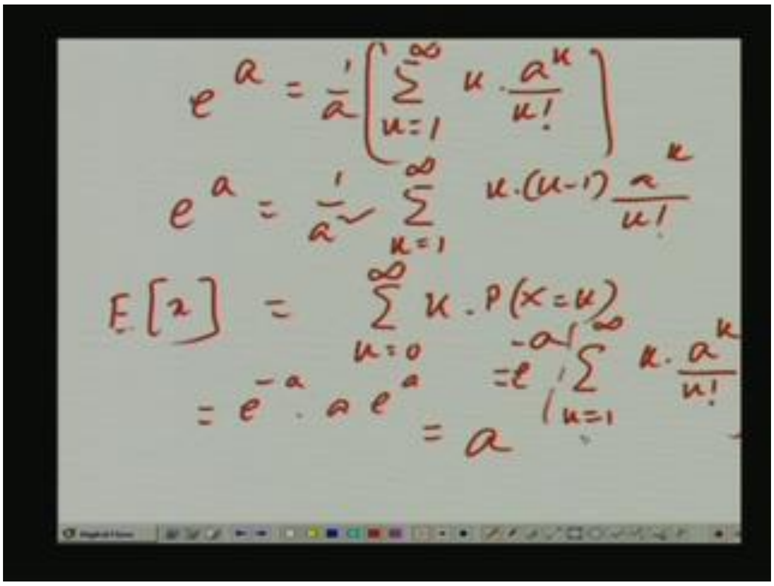
So, k equal to 1 to K equal to 0 to infinity we have differentiated it once, we got this. If we differentiate this twice what do we get. I mean you know, it is better that we expand this in expand the series term by term. This 1 plus a plus a square by factorial 2, you can have 1 more term like a to the power 3 by factorial 3 plus dot dot dot dot plus a to the power k by factorial k plus dot dot dot dot.

If we differentiate it once, you get 1 here, then 2 into a by factorial 2, 3 into a by a square by factorial 3 and likewise. If you differentiate this again out of this, you get 0 here you get 2 into 1 divided by factorial 2, here 3 into 2 into 1 into a by factorial 3 and Likewise. Here in the general, term will be k into k minus 1 a to the power k minus 2 by factorial k.

See, if I say differentiate with respect to a twice, this gives rise to again e to the power a on the side. On this side, k into k minus 1 you can put 1 by a square here. So, that you get a to the power k divided by factorial k. And k should be equal to 1 to infinity you can easily verify this put k equal to 1, obviously, this term become 0.

Because, k minus 1 is 0, and then k equal to 2 onwards. If you put k equal to 2 2 into 1 a square a square cancels 2 into 1 by factorial 2 which you get out of these. And then for k equal to 3 4 onwards you get the corresponding terms. Now, I got 2 expressions: 1 is this another is this. Let me go to the next page write them and then get you the expressions for mean and variance.

(Refer Slide Time: 10:36)

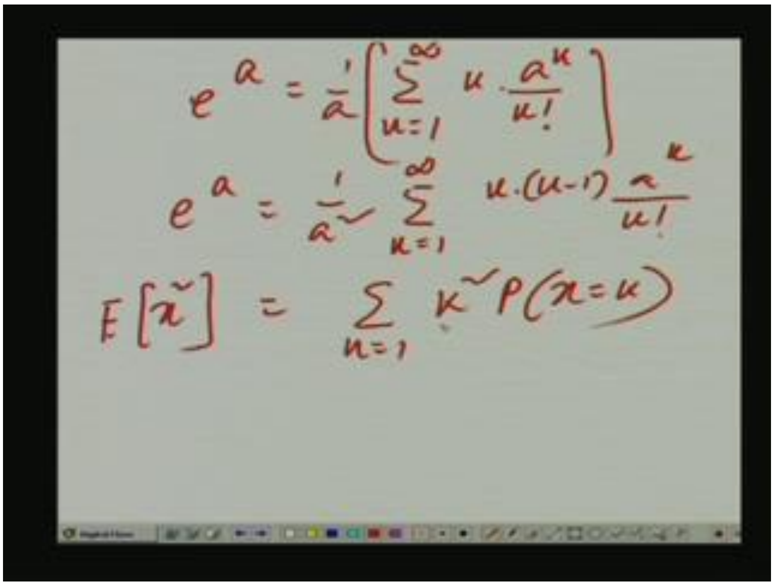


So, we have obtained these 2 expressions: 1 is e to the power a is equal to 1 by a, another is equal to 1 by a square k into k minus 1 a to the k by factorial k. Now, is now the question is what is expected value of x; x takes the value 0 1 2 3 and likewise. So, expected value will be k times p x equal to k and summation k equal to 0 to infinity. Now, you know k equal to 0 case can be ignored is, because you are multiplying by 0. I start this summation with k equal to 1. And we know the probability also.

So, this becomes equal to k into we have already seen what is the probability e to the power minus a will go out. And a to the power k by factorial k k equal to 1 to infinity. Now, what is this part from here you can see, the this this is nothing but… This part is nothing but this part which means this is equal to a times e to the power a right.

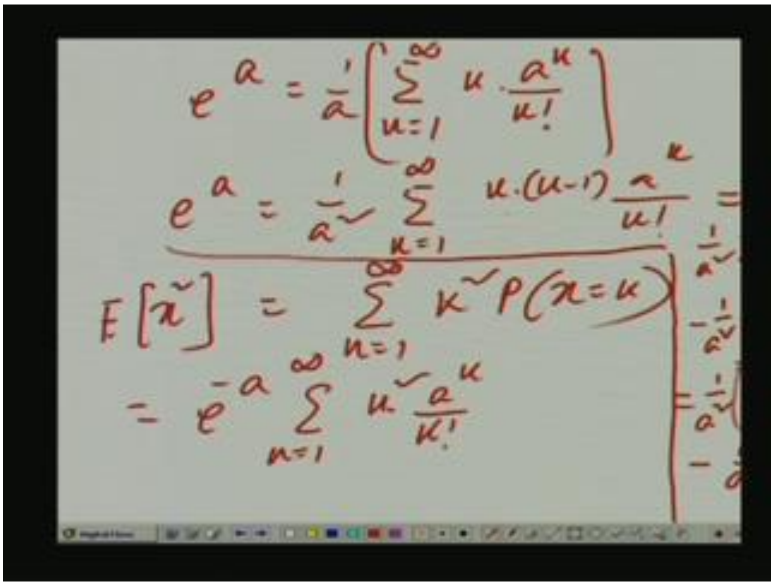
So, you have got e to the power minus a into a into e to the power a, which is equal to a. So, expected value of x here is a. And now, you go the other 1 what is the expected value of x square. So, instead of moving to the next page I erase these part here.

(Refer Slide Time: 13:11)



So, what is expected value of x square. Once again it is nothing but this should be k k square this. And k is taking value we can start with 1 directly. Because, k equal 0 means this term will 0 this finite and this is 0. So, we ignore k equal to 0 case from k equal to 1 to infinity and we know the probability.

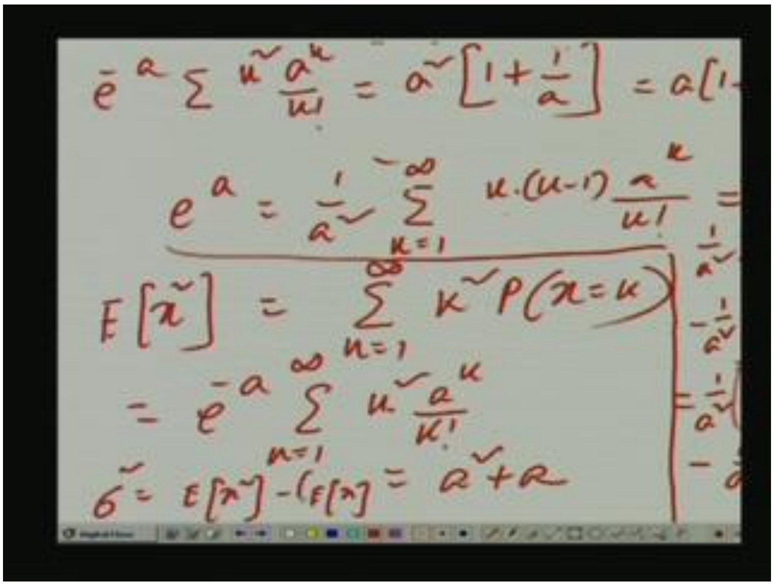
(Refer Slide Time: 14:03)



E to the power minus a comes out of the summation a equal to 1 to infinity. And we have got k square into a to the power k divided by factorial k. What is this? This is not this is coming out of this expression out of this expression this is coming. But how, you can see 1 thing if you take this expression if I write on this side from here you break it up as k square minus k.

So, 1 term you get 1 by a square k square a to the power k by factorial k. Another term you get, minus 1 by a square k a to the power k by factorial k, right. This part we have already seen what it is. It is a times e the power a. So, these becomes 1 by a square summation k square a to the power k by factorial k as before minus this part as I said just now, a to the power a. So, it becomes e to the power a divided by a. And this is equal to e the power a. So, what is this quantity; this quantity is appearing here. What is this quantity you take this to the left hand side e the power a is common. Let me write it here.

(Refer Slide Time: 16:25)



So that means, what is this quantity that is the question right. So, this is equal to. This will be equal to a square times and e to the power a is common. So, you can as well push e to the power minus a here. You take this quantity to the left hand side e to the power a is common. So, bring e to the power a on this side. On the left hand side, by I mean a square multiplying 1 plus 1 by a.

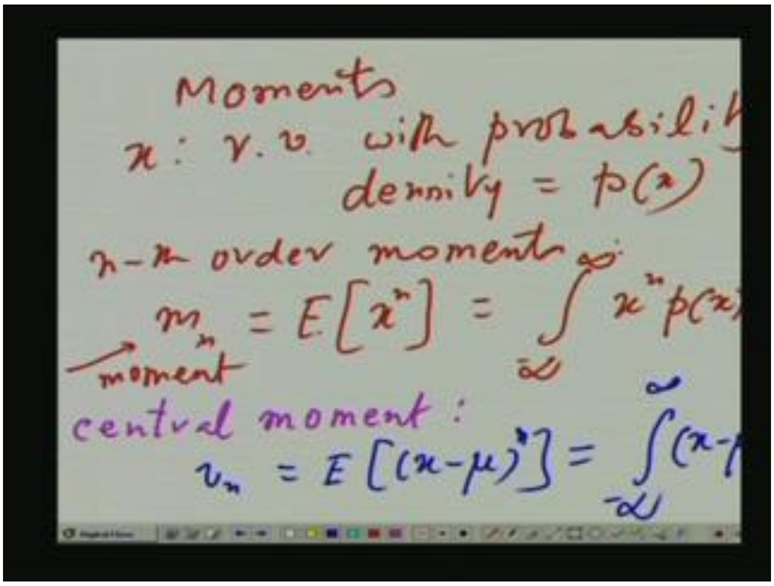
So, what is good then get then out of this let us, again revise. We had this summation that we broke it as like this 1 by a square k square, and this then minus 1 by a square and this.

And k a the power k by factorial k that was equal to e to the power a by a e to the power a into a .So, you get e to the power a by a this moves to the left hand side.

So, e to the power a is common that goes to the right hand side. So, you have got this term and this equal to this which is nothing but a into 1 plus a. So, this it this is nothing but then a square plus a. And we already, have seen what is a; a is nothing but e of x, and then what is variance; variance is e of x square minus e of x whole square which is a square.

So, variance sigma square which is equal to minus e x whole square this will be equal to a. So, for a poison distributed random variable with parameter a is very interesting that it is mean and variance both are equal to a. This is very interesting exercise. So, we now move to another very interesting topic which is called moments. Moments associated with 1 random variable x.

(Refer Slide Time 19:53)

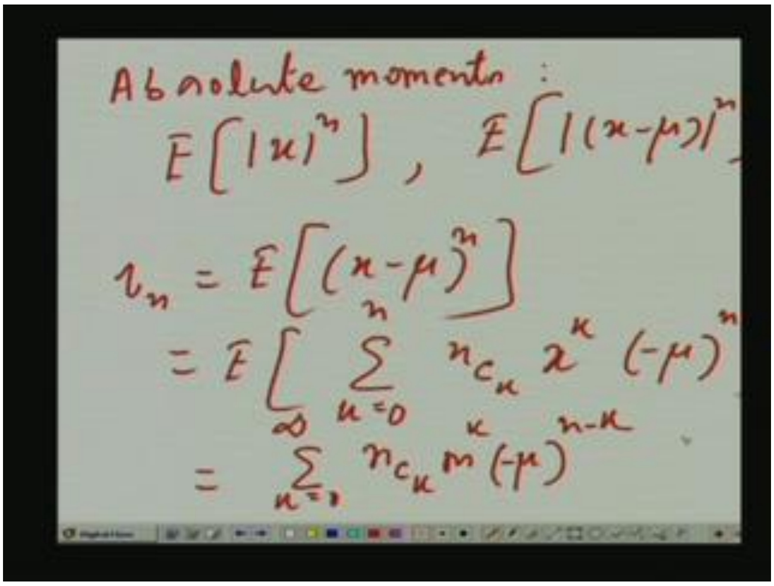


The moments x random variable with probability density p x, then, n'th order moments there are various definition of various types of moments. So, let us just define them n is an integer n could be 1, n could 0, it could be 1 2 3 or like that. So, the usual definition of moment which is actually called just moment. M n n'th order moment it nothing but expected value of x to the power n which is nothing but this.

So, you can see what is m0: m0 is just integral of px dx which is equal to 1. What is m1: m1 is x px dx integral which is mean mu and likewise. Then, sometimes some other definition also are used. Like this is moment there is something called central moment, this is moment this is called moment. Then, there is something called central moment here central moment we define as say vn.

So, which is nothing but E of x minus mu: mu is mean of x or mu is equal to m1 which is a mean actually of x. So, instead of taking x to the power n you take the deviation x minus mu raise it to the power n take its expected value that is vn. Which actually is nothing but this... Then, instead of taking x or x minus u some time we take there mod values. And those mod I mean if put a mod here with mod here mod x or mod here and those are called absolute moments.

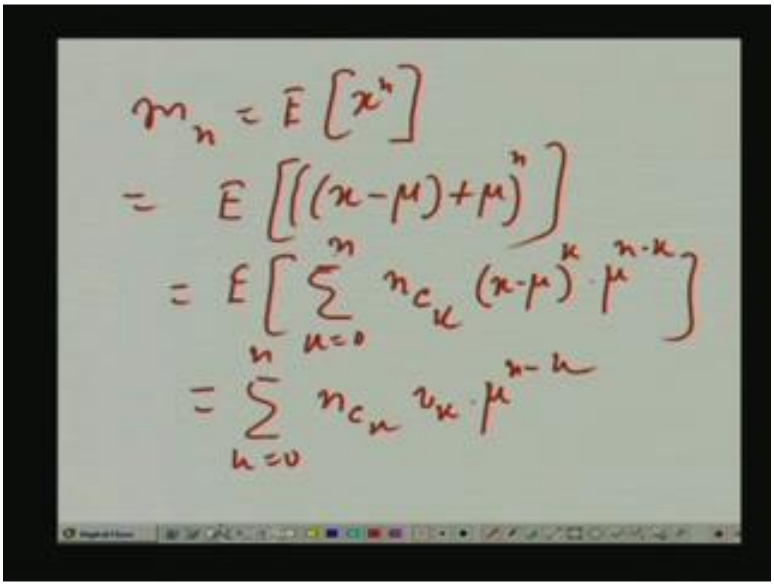
(Refer Slide Time 23:07)



So, that is absolute moments now I mean out of this usually, the moment and the other 1 that is generalize moment, not generalize moment the central moment, which actually does not take x, but whether the definition of x from it is mean mu I mean only this 2 are most commonly used. So, let us consider a relation between them.

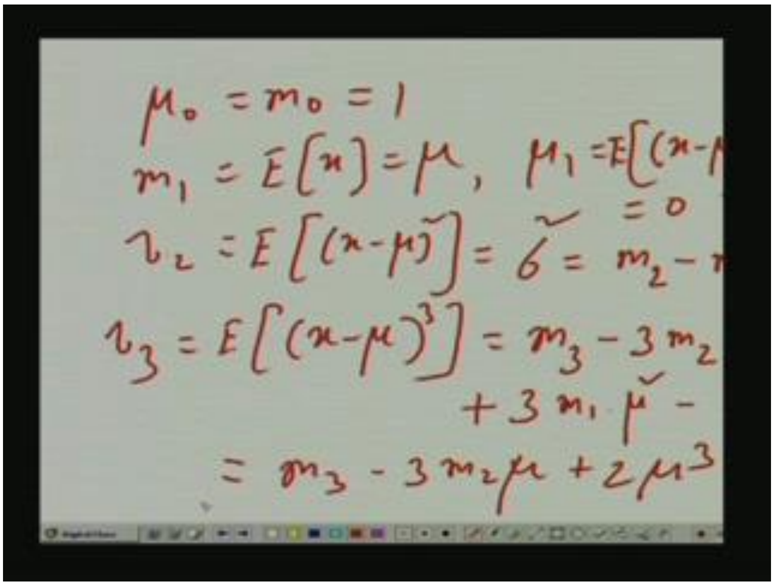
So, first you see what is vn, vn we have see the nothing but E x minus mu whole to the power n now, this is binomial series these a binomial. You can expand this a rather, you can a expand the binomial series, and then it becomes nck. Now using the linearity of expectation operator apply this E over x to the power k and E x to the power k as we know is nothing but the k’th order moment m k. So, you will get this relation this is it is these are relation between the central moment, n’th order central moment and the general the usual moments. We can have a reverse solution also which I consider next.

(Refer Slide Time: 25:59)



We have seen mn is nothing but this you can always write as plus mu whole to the power n right. And then expending binomially we have got these expression and now, I apply the expectation operator on this. This will be nothing but the central k’th order central moment and you get this term vk.

(Refer Slide Time 27:45)

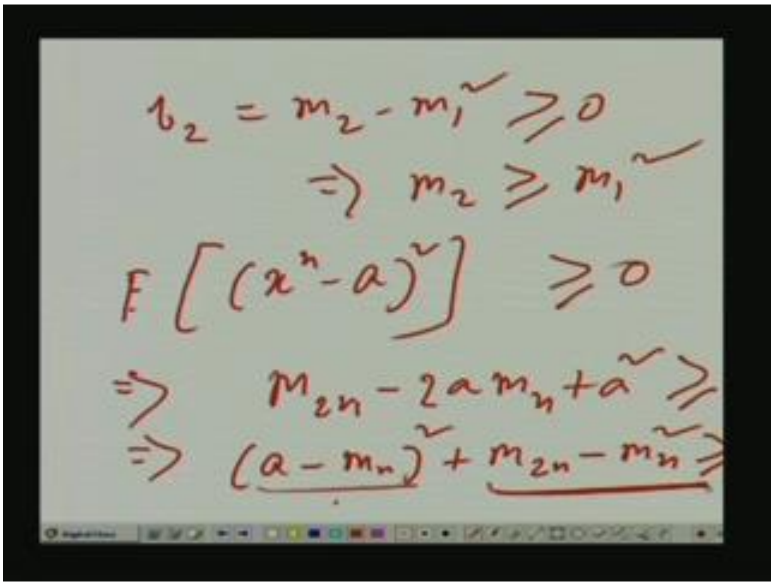


As a special case, you can see one thing that what is mu 0 or m0 both are equal to 1. That is very easy to see, I will not going doing it. Then, what is m1: m1 is which is nothing but mu, but what is mu 1: mu 1 is you can put an expectation x minus mu. And obviously, x minus mu it is average value is 0 and mu 1 is 0. What is mu 2 and this is what we denoted by I mean this is what we I mean we of term this quantity as a variance sigma square.

But we have also see that if you expand it, you get this 2 minus actually m1 square m2 minus m1 square. Further, if you take a mu 3 what is mu 3 expected value of x minus not mu this is v is just you know if you expand it binomially, x cube expected value of that will give you m3. Then minus 3 x square mu expectable value of x square is m2.

So, minus trice 3 m2 mu plus 3 m1 and mu are same. So, you get basically this term mu square minus mu cube and this is nothing but trice mu cube. Because, it is nothing m1 is equal to mu. So, 3 mu cube minus mu cube. So, which essentially leads to this thing. Similarly, you can express m3 in terms of v3 I leave it up to you this is very simple.

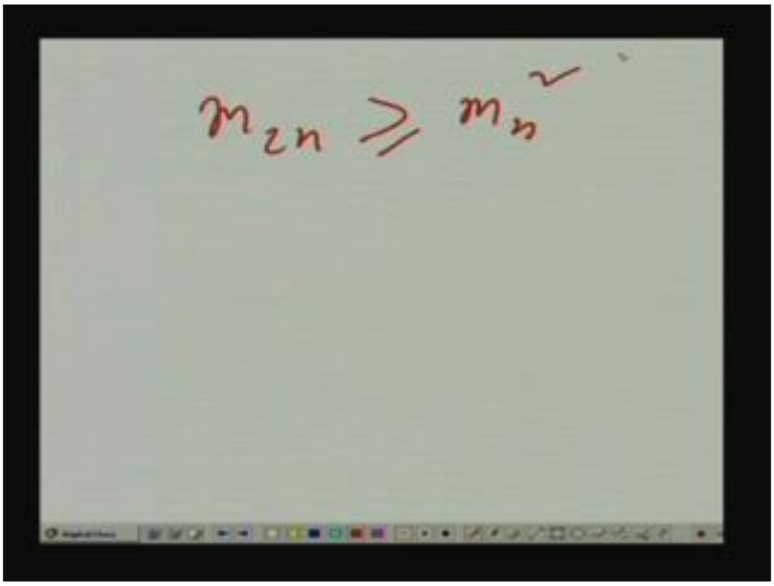
(Refer Slide Time: 30:54)



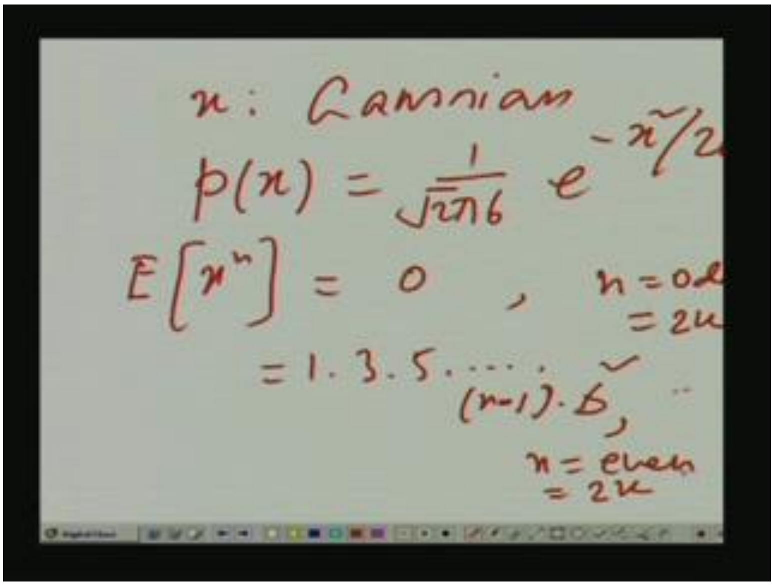
Certain things you see, you have just now seen that b2 is nothing but m2 minus m square. So, this numbers m1 m2 and all that you know there are not just some arbitrary numbers they usually, have to satisfy some inequalities. For a such we know that b2 is always greater than equal to 0.

So, m2 should be always greater than equal to m1 square. Similarly, consider this quantity as an example a any number and we take the whole square. Now, for any a I know this has to be greater than equal to 0 these cannot be negative, but what does it mean. This leads to if you expand it x to the power twice n expected value of that which gives you M 2n minus twice a x to the power n expected value of that which gives you twice a mn plus a square is greater than equal to 0 or any a that is important, which means, a minus whether it like whole square plus m 2n minus mn square greater than equal to 0. Now since, this is true this as to be true for any a this is true also for mn equal to a case the case where mn equal to a. Which means, then it is 0 for that case we have this thing that m 2n minus mn square should be greater than equal to 0 that is.

(Refer Slide Time: 32:49)

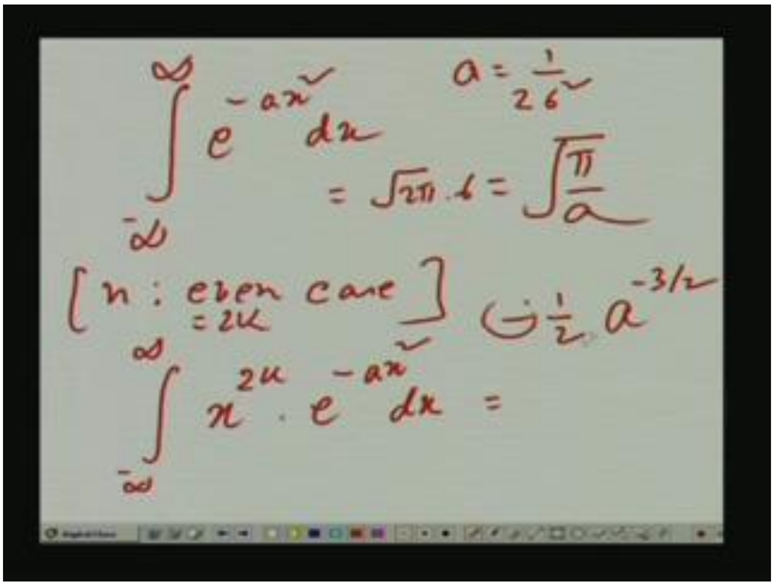


(Refer Slide Time: 32:58)



Now, as an example we consider the case where x is Gaussian to make life simple I assume that it has mean 0 and variance sigma square. So, it is given p of x is equal to 1 by root 2 pi sigma e to the power minus x square by twice sigma square. So, let us see what is I mean what is E to the power xn. I first write down the result this is equal to 0 we will show how. If n is odd that is some twice k plus 1.

(Refer Slide Time: 34:17)



Otherwise, it is equal to 1 the times 3, 5 dot dot dot dot up to n minus 1 and sigma square. When n is even equal to twice k this we will show as an example. For that we consider this integral what is the value of integral. You can always, take a 2b equal to some 1 by twice 34 32 sigma square. So, if you make that substitution if e for minus a x square by twice sigma square. And if you had another 1 by twice root 2 pi sigma here 1 by root to pi sigma, then that would have 1.

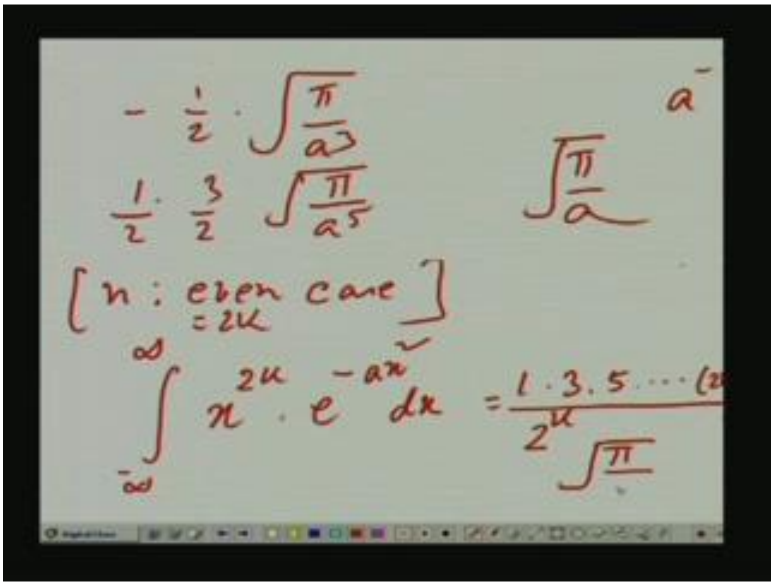
So, this is equal to root 2 pi sigma and sigma we know is 1 by square root 2a. So, this gives rise to this. This integral is essential for us before that, I mean we are taking the n equal to even case. We are taking first n equal to even I forgot to mention this n equal to even case. For this we start with this integral, my point is if I differentiate both side with respect to a say once. What I get from this side, a minus will come here also a minus will come.

If I differentiate again with respect to a another minus will come that will be the positive here also, another minus will come that will be the positive. So, either both side are minus or both side have plus. So, essentially I can retain a plus sign in both cases if you differentiate it say k times where we have seen that n equal to twice k.

If I differentiate k times, what I get x as to the power I am not bother about the sign. Because, I told you if it is minus here, it is minus here if it is plus here, it is plus here about the differentiation. So, I can always make a I mean keep a plus sign here this is x to the power twice k, because each differentiation brings 1x square. So, k times means x to the power twice k time this on this side What happens here, square root pi remains as it is. When you, first differentiate it this a to the power minus half that minus sign I am not bothered, because that will again be taken care of by the sign here. So, half then next time I mean let us, do some rough calculus an here actually.

So, for say if after differentiate once we have these thing it is a to the power minus half. So, this becomes a to the power minus 3 by 2 and half with minus sign comes this minus sign I am not bothered 1 by 2. Next time, when I differentiate may be it will better. If I erase this here and work out this part on top first.

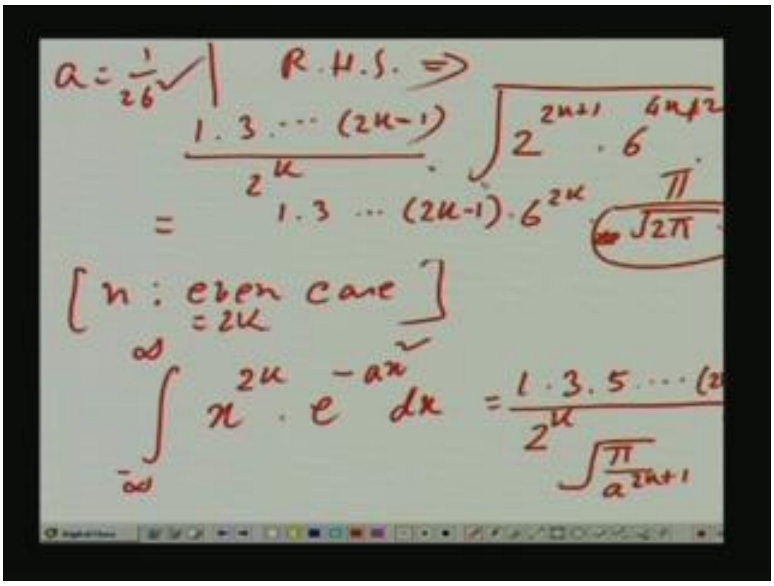
(Refer Slide Time: 38:54)



Consider this function, if I differentiate it once I get minus though the sign is not important I told you when it is minus I have got a minus from the left side also they will cancel. When I have plus sign on this plus sign here on the left hand side also I will get the plus sign. So, I am not bothered about sign both sides can be kept as plus, but still I write 1 by 2 square root pi a to the power 3 right it is a by 2.

Next time, when you differentiate then I have got 1 by 2 into 3 by 2 it becomes a to the power 5. Because, after all I have got a to the power minus 3 by 2 here, so minus 3 by 2 goes minus minus plus. So, 1 by 2 3 by 2 and this becomes a to the power minus 5 by 2. So, I put under square roots, so a to the power 5 and likewise. So, if you go on differentiate it k times here, I am getting 1 into 3, then 5 7 like that and here just 2 2 2 k times. And here we I have got I mean if you differentiate it once, that is k equal to 1 it is 3. So, twice k plus 1 if you differentiate it twice, it is 5 2 into 2 plus 1. So, it is always twice k plus 1. So, what I get. Then this quantity 1, 3, 5 dot dot dot dot twice k plus 1 by 2 to the power k square root pi by a to the power twice k plus 1. Then what happens 2 k plus 1 is actually n. So, you can as it put it 1, 3 dot dot 5 up to n.

(Refer Slide Time: 41:37)



Now, I replace a by a by as before 1 by 2 sigma square. So, what happens on this side 1 by 2 sigma square. So, twice sigma square whole to the power 2 k plus 1 that comes here. So, the right hand side becomes RHS that becomes this right hand side, it leads to 1, 3 dot dot dot up to 2k plus 1 divided by 2 to the power k times 2 to the power 2k plus 1 sigma 4k plus 2 into pi is square root.

So, this 2 to the power 2k I mean it is give rise to 2 to the power k after square rooting. So, that cancels a square root 2 is left, a pi is left. And what happens here, this give rise to sigma to the power twice k plus 1 this give rise to sigma to the power twice k plus 1. So, what is happening then actually I made a small mistake this should be minus.

Because, once you differentiate with respect to a once, you get 1 by 2 that is for k equal to, so 2 into 1 2 minus 1 that is 1 differentiate again which respect to a. So, now k equal to 2 you get 3. So, 2 into 2 minus 1 not plus, so it will minus 1 the small mistake I just noticed. So, what you get out of it 2 to the power k cancels sigma to the power twice k comes up here. And sigma to the power 2 square root gives rise to sigma square root pi and a square root 2.

So, this will gives rise to what, I mean a I can write it like here 1, 3 dot dot dot twice k minus 2 to the power k is canceled, then sigma 2 to the power twice k; twice k is equal to n and this additional term 2 pi and 1 sigma. So, if you take this root 2 pi sigma on this side there is 1 by root 2 pi sigma. And a as 1 by 2 sigma square then that becomes the expected value of x to the power 2k.

Then, which will then give be equal to 1 into 3 into dot dot dot twice k minus 1 sigma to the power twice k or n, which is the result that is the result. So, it is done I am not writing it formally. I am just saying that, this part you bring to the left hand side. So, it becomes 1 by root 2 pi sigma here which is and a as 1 by 2 sigma square.

So, basically this amounts to the expected value of x to the power 2k. And 2k is n say expected value of x to the power n. It is an n'th order moment which is nothing but then 1 into 3 into dot dot dot twice k minus 1 sigma to the power n. So, this is for the even case where n is even n is twice k. But when n is odd, then what happens that very simple actually. The n odd if n is odd say n is equal to twice k plus 1, then x to the power n may be I can make it little simpler, just for a consider this px which is also equal to p minus x this is important. Because, Gaussian probability density function is symmetric around its mean. And if you put x equal to minus x since, it is square there is no change.

So, what happens then when n is odd, we have got x to the power n px dx. So, what I am going to do now is valued not only for Gaussian case, but for any probability density which is symmetrical around it is mean for which this relation is valid. You can break this integral as 0 plus to infinity x to the power n px dx plus minus infinity to the 0 minus for this integral replace x by minus x.

So, what happens limits become on plus infinity I mean dx is minus dx. So, minus will come this will become plus infinity this will come 0. So, you reverse it. So, it becomes minus this integral from 0 to infinity, then minus and minus will again make it plus. So, it becomes plus 0 to I repeat again. If you replace x by minus x, so dx will give rise to minus dx a minus will come limits will become plus infinity to 0 if you want to make it 0 to infinity. Then, they integral becomes again negative. So, that another minus sign will come those 2 will make it plus. So, plus 0 to infinity x to the power n px dx. But when you put x equal to minus x here since, n is odd a minus sign will come.

So, effectively a minus will come here and we will get a term like this and do not bother about this 0 for the time being I remove this, I mean do not bother about this 0 minus or 0 plus. Essentially, the 0 minus means you are approaching 0 from the left hand side 0 plus means you are approaching 0 from the right hand side, but essential there is 0.

So, you have got again a 0 to infinity a sign is reverse we get this and they cancel and you get 0. So, I suggest that you now try and read this by Papoulis and consider other examples were not only Gaussian, but some other probability density have been considered. And the corresponding mean and you know corresponding moments expressions for corresponding moments have been derived.

The very next thing from here, is very important thing called characteristic function. It is basically, related to Fourier transform of probability density function is actually a characteristic function. Now, that I will not start now, but I just telling you a little bit about it. Since, Fourier transforms have been very effectively used for carrying out convolution and all that. I mean, you can use this result in the end where suppose you have got a summation of various random variables each having its own probability density.

Then you have to find out the probability density of the I mean resulting sum which is also a random variable. There this characteristic function, I mean is very useful also this characteristics function can be used very effectively to calculate moments. The moments which we have just considered, they can be calculated from these characteristics functions very easily. So, I reserve this for the next class and I stop here today. So, we are so far covered in this chapter function random variable and if y is gx and giving the probability density of x. What is the probability density of y, we took several examples. Then, we went into moments we went into another important things that is giving a function gx.

What is expected value of gx, that is nothing but gx into px dx integral we consider several examples and applications of that took us to this motion of moments not only, mean and variance to start with, but more general case of moments. And from moments we go to characteristics function that will complete this chapter. Then we will go into what is called joint statics that is not only 1 random variable, but more than 1 random variable say 2 random are present and they are related. So, we will be considering things like joint probability density and all that and do similar exercise. So, that we today we stop here in the next class as I said we will consider this characteristic function.

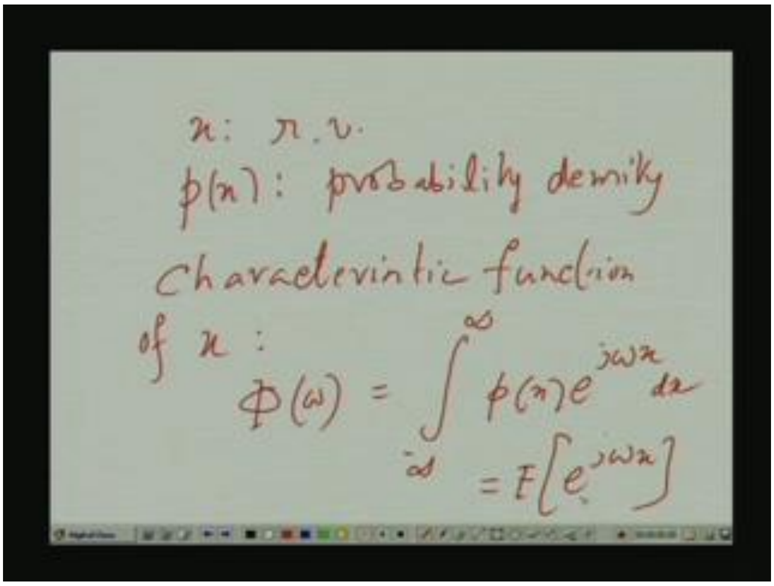
Thank you.

Prof. M.Chakraborty
Department of Electronics & 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 px, what is its characteristic function first.

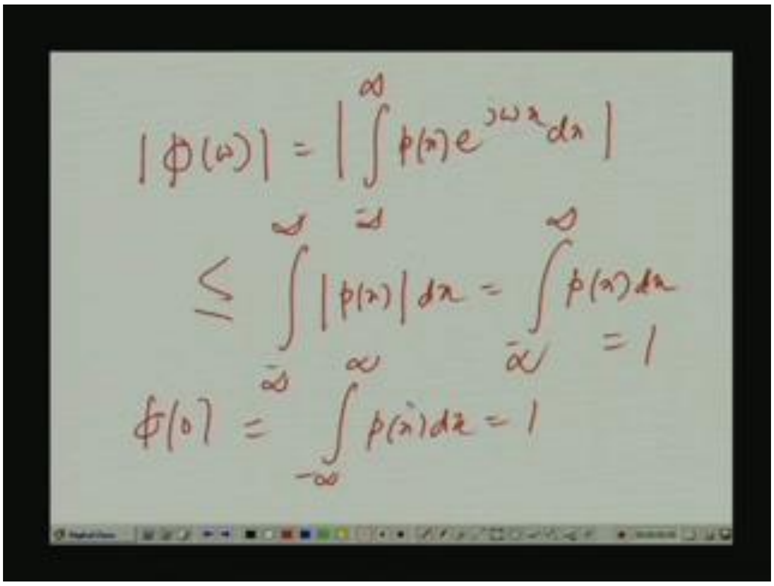
(Refer Slide Time: 54:00)



So, given x random variable px probability density, then the characteristic function of x is denoted as phi omega this integral. It is something very similar to Fourier transform only thing is that in Fourier you have a minus sign 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. So, 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 gx it is expected value is gx times px integral. So, that is what is happening here. So, characteristic function is also characteristic function of a random variable x with probability density px is actually, the expected value of the function e to the power j omega x.

(Refer Slide Time: 56:00)

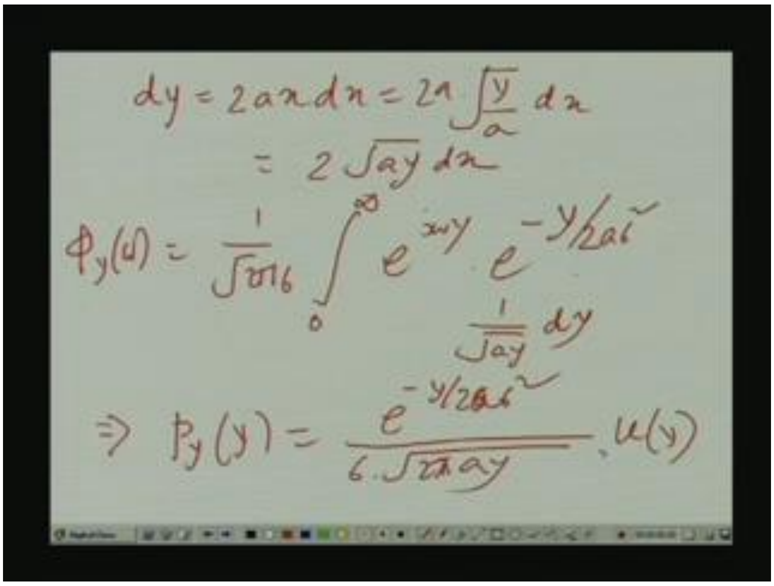


Now, as you know using the triangle inequality we all know that, but this is less than equal to you known this triangle inequality right. Sum of 2 sides is greater than equal to sum of third side, I mean third side that we generalize in the case of here it is not discrete sum. But it is an continuous 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 px and magnitude of this is 1. So, it transfers to be this and mod of px and px are same, because px is a real value real and non negative valued 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 px dx.

Because e to the power j omega x at x equal to 0 at omega equal to 0 is 1. So, you have px 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 magnitude falls after that. So, this is the property of the characteristic function.

(Refer Slide Time: 57:51)



I repeat phi omega square root 2 pi ay and of course, this is on the positive side. Obviously, you know a whether x takes negative or positive y can always take only positive values. So, for a negative y 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 next I mean new topic next time, but if there is any question.

Now, I would suggest that you all you know 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 constant I cannot carry out I can take up this worked out examples which are already there for you to read. So, that is all for today.

Thank you very much.