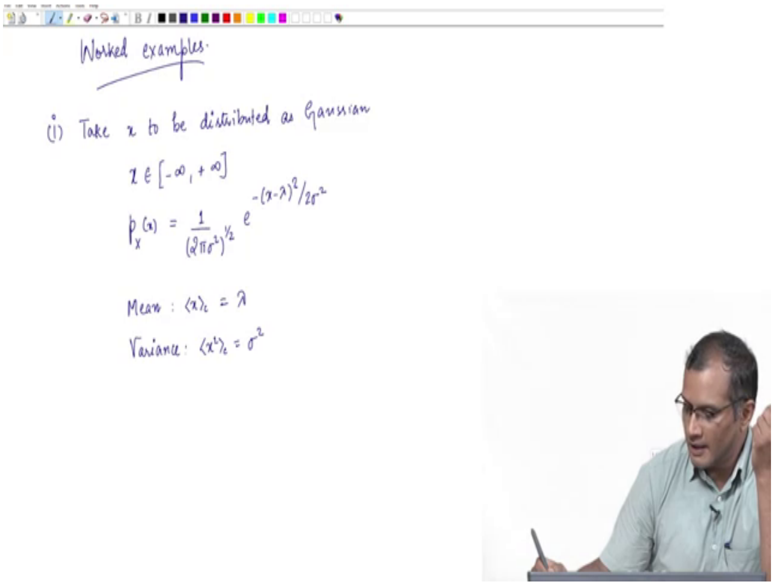
Statistical Mechanics
Prof. Ashwin Joy
Department of Physics
Indian Institute of Technology, Madras

Lecture – 12
Tutorial

So, good morning students, today we will talk about some problems and I will demonstrate how to use the concept of characteristic functions. So, basically you can take this as a tutorial on worked examples.

(Refer Slide Time: 00:34)

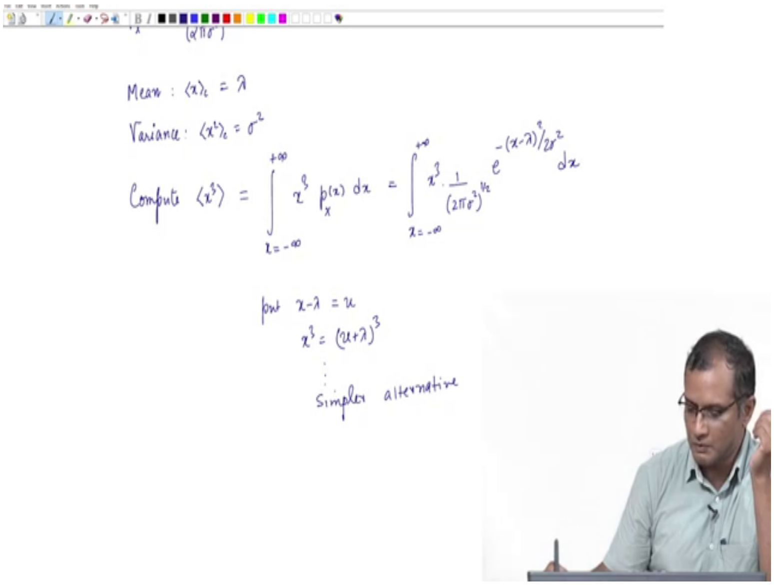


A couple of problems that come to my mind are basically applications of characteristic functions ok. So, I will take a couple of distributions that are already discussed in this course. And the fact that they are very profound and commonly found in nature makes the case for some applications and so I take one distribution to begin with which is the Gaussian distribution ok.

So, let us say take x to be distributed as a Gaussian. So, I have taken the range of x to be minus infinity to plus infinity, and the PDF is basically 1 upon 2 pi variance to the power half into e to the power minus x minus lambda square over twice variance ok. Now, you can see easily that the distribution is completely specified by the two cumulates, and this was also discussed in the class.

And I have taken the mean of this distribution to be nonzero. So, I have the mean which is basically the first cumulant as lambda and I have the variance as sigma square. So, the first two cumulants are already known to me alright.

(Refer Slide Time: 02:57)

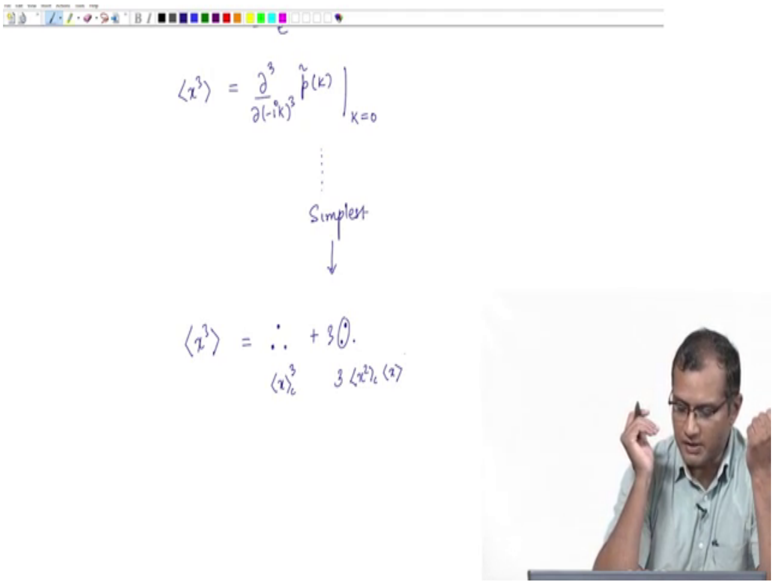


And if you are interested in computing the moments let us say your task is to compute the third momentum ok, then there are three ways of doing it at least the three easy ways of doing it, but they all take different amount of effort. So, I will list the efforts in the order of descending order of descending amount of effort.

So, the first effort that is sort of takes more time is to simply sample x cube in this distribution. So, I will take the distribution from minus infinity to plus infinity and sample x cube in it. So, this is my sampling and my distribution is p of x dx. So, basically I am forced to evaluate this integral x cube into 1 upon 2 pi sigma squared e raised to minus x minus lambda. You can solve this integral by a simple variable translation. So, if you put x minus lambda as some u, then your x cube becomes u plus lambda the whole cube. And you simply expand it in a polynomial. You will have three terms, you will have you have this is a polynomial of order 3.

So, you will have you will have multiple terms in the expansion of x cube and so you split it up into various parts and in complete the integral and get the answer. The other way to basically and I will say a simpler method or a simpler alternative is to basically use the characteristic functions.

(Refer Slide Time: 05:18)

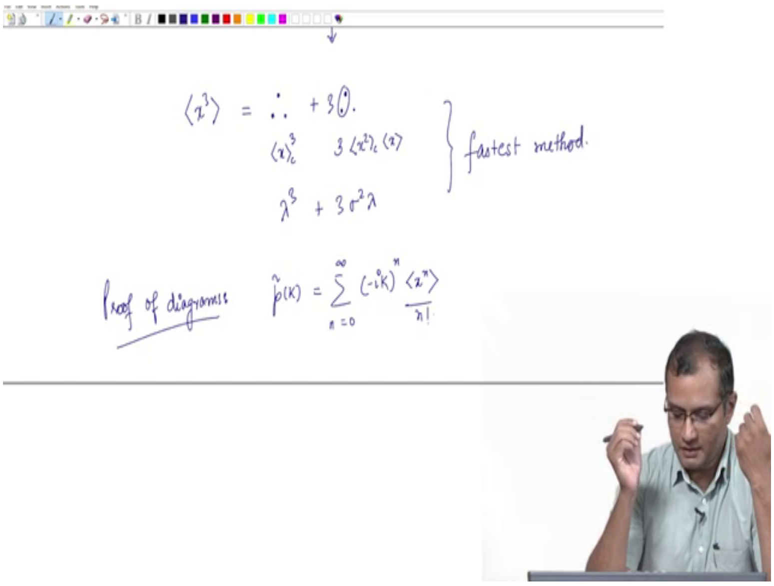


So, you can recall that the Gaussian PDF had the characteristic function given as e to the power minus i k lambda minus k square sigma square by 2 right. So, I will just rewrite it for convenience as e to the power minus i k lambda plus minus i k the whole square sigma square by 2. And, you can basically compute the third moment by invoking the appropriate derivative which in this case is the third order derivative, and what you will get here is basically the third moment. So, again this method requires some effort, but definitely lesser than the method I listed in the first you know in the previous paragraph ok.

And even simpler probably I will call the simplest method is the following. So, this was the simpler method. And I would argue that the simplest method is a combination of both characteristic functions and your diagrams. So, the diagrams that we have seen in the last couple of lectures enables us to compute it very easily in just two steps, it is that easy ok. And just from the knowledge already developed, I am going to write down the third moment in the form of our connected and disconnected clusters. So, I have third moment, so I will put three points. And by our logic of the fact that whenever you connect endpoints it is a cumulant of that order. And if you connect disconnect endpoints, then it is the moment raised to that order. So, I will with three points I can have one configuration that are all open, all three of them are disconnected.

So, this is the moment raised to the first order, moment raised to third power, then I can connect two of them and leave the third guy open. This is three times the second cumulant into the first moment and that is it. I will not connect the three dots simply because I know from prior experience that the Gaussian distribution does not have cumulants above the second order, all higher cumulants are 0. So, the third configuration of three connections will not be written here, because the third cumulant is 0. These are the only two combinations. Look how fast this is.

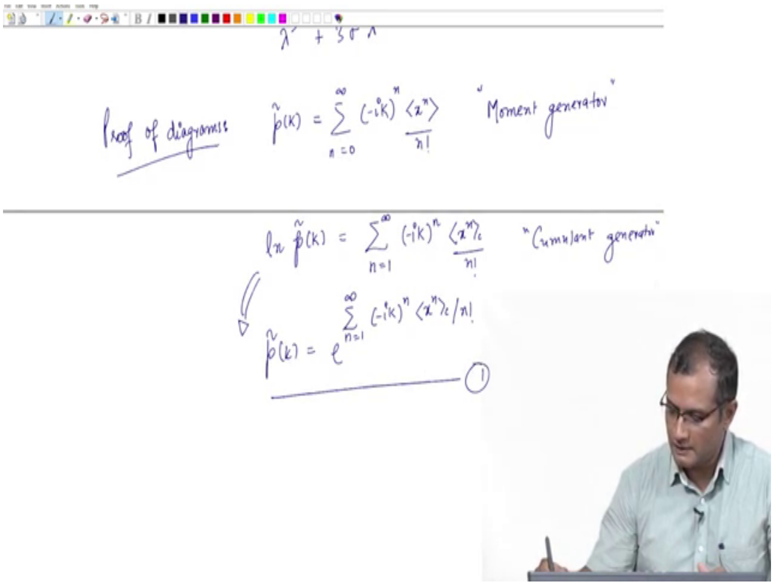
(Refer Slide Time: 08:35)



I already know the first moment is lambda. So, this is lambda cube. And this is 3 times sigma square into lambda, hardly any effort required, just one line. So, I will argue that this is the fastest method which has used the combined power of your diagrams and characteristic function fine. So, then if you want to read more about these diagrams, the only thing I would say here is that the diagrams can be easily computed by two expressions that we have already discussed. So, this is very simple, and I leave that as an exercise. Just I will give you a sort of a hint to proceed the calculation to just construct these diagrams.

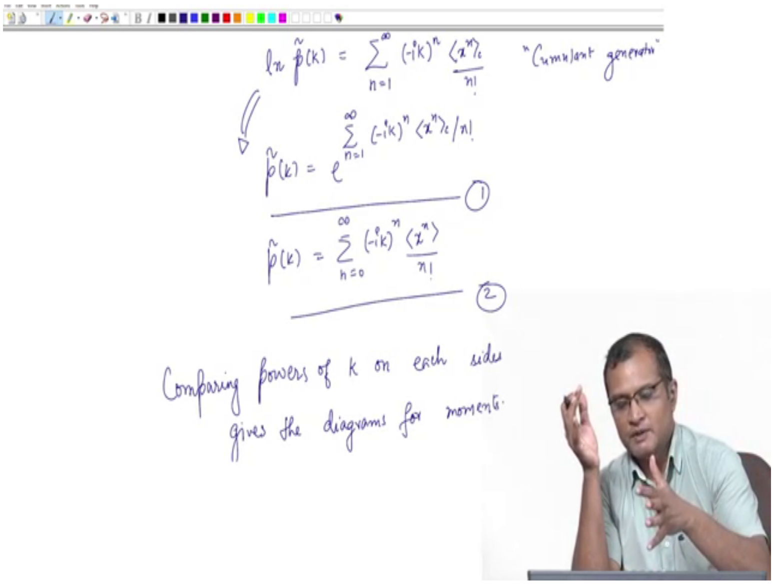
So, we know that I had discussed these diagrams without giving the proof. So, here I will suggest a very easy way of seeing how these diagrams are computed. Basically we have the form of the characteristic function which is given as summation i going from 0 to infinity minus i k to the power, so I will take summation n going from 0 to infinity just because we have a complex i already sitting inside ok. This is the cumulant generating function and or moment generating function I am sorry.

(Refer Slide Time: 10:38)



Let us write it in short as moment generator or the characteristic function. We also have the cumulant generating function which is lawn of this. And this was written as n going from 1 to infinity minus i k to the power n, the nth cumulant over n factorial. So, let us call this as the cumulant generator. So, these diagrams that I am I have been using for a while now it can be very easily computed by simply equating. So, from this, I can say that from this expression I can say that p of k is nothing but e raised to summation n going from 1 to infinity, so I will call this as first equation ok.

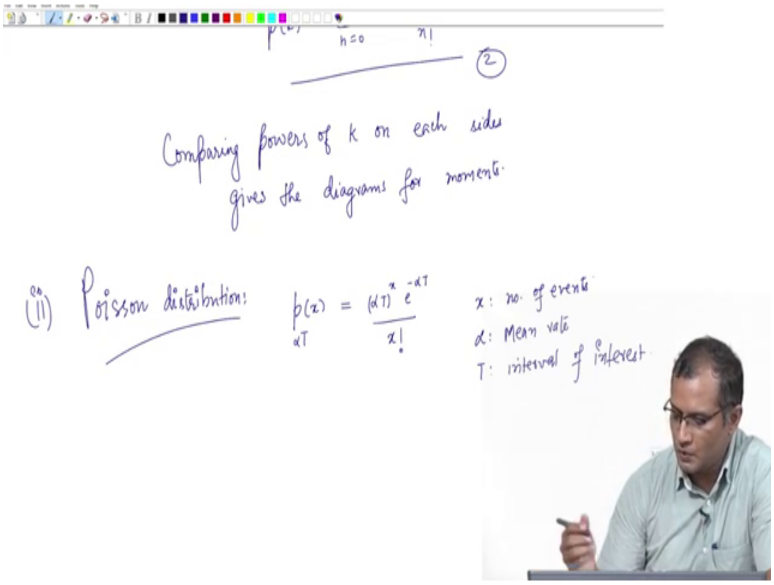
(Refer Slide Time: 12:04)



And I already have p of k which is given as n going from 0 to infinity minus i k n x to the power n over n factorial, I will call this as a second equation. So, if you compare the two equations and compare if you compare the like powers of k on each side, you will get this formula of diagrams gives the diagrams for moments ok. So, you expand e raised to x in powers of x; and on the other side you already have a series expansion from either side you pick up terms of like powers of k and in you know one side you will have coefficients in the power of coefficients in the form of cumulants from equation 1, it is e raised to some x where x is given in terms of cumulants.

And on the other side, you will have summation of terms all in the all group does moments. So, when you compare like powers of k on each side, you can get the diagrams. And this I leave it as an exercise here, you can check for yourself. At least the first two moments should be able to you should be easily able to compute. And for the higher ones, I have already given the logic of permuting combining the various number of dots. So, you can take that as a rule to generate these diagrams. This is the proof for these diagrams alright. So, I gave one example of a Gaussian distribution, and I would like to give one more example, and let us call this as example number 2 if I have already called this as first example.

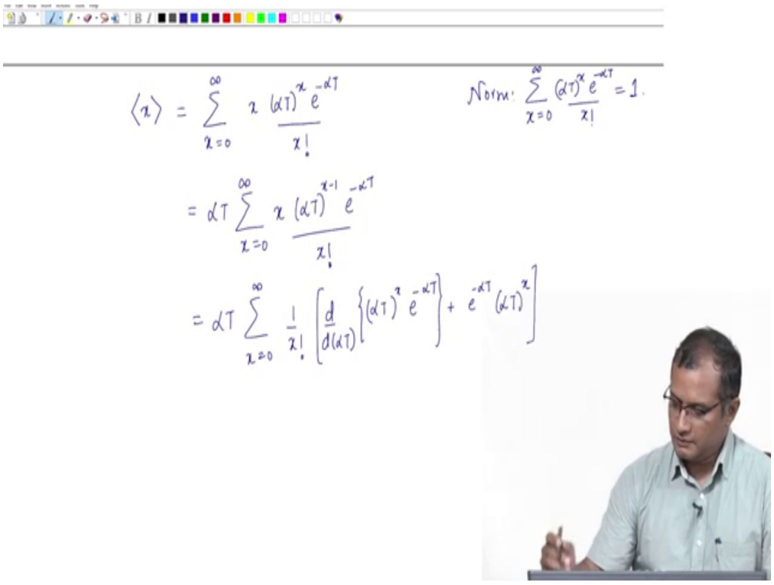
(Refer Slide Time: 14:22)



So, let me take example 2. Here I will take a Poisson distribution. Again a very important and commonly found distribution which is basically used to study processes which are completely random in time and independent. And the only thing that you know here is that the process is possessing a mean rate or a mean density. So, I have already discussed Poisson distribution in the class.

So, I am just going to write down the distribution here. But if you want let us say what is the number of what is the probability to see x events, x is a discrete number here, probability to see x events in a waiting time of t provided the process has a mean rate alpha is given as alpha t, which is to be seen as a mean number, because alpha is mean rate t is the waiting time alpha into t is mean number. Alpha t to the power of x e raised to minus alpha T over x factorial. So, x is basically number of events; alpha is the mean rate, and t is your interval of interest, interval of interest ok.

(Refer Slide Time: 16:34)



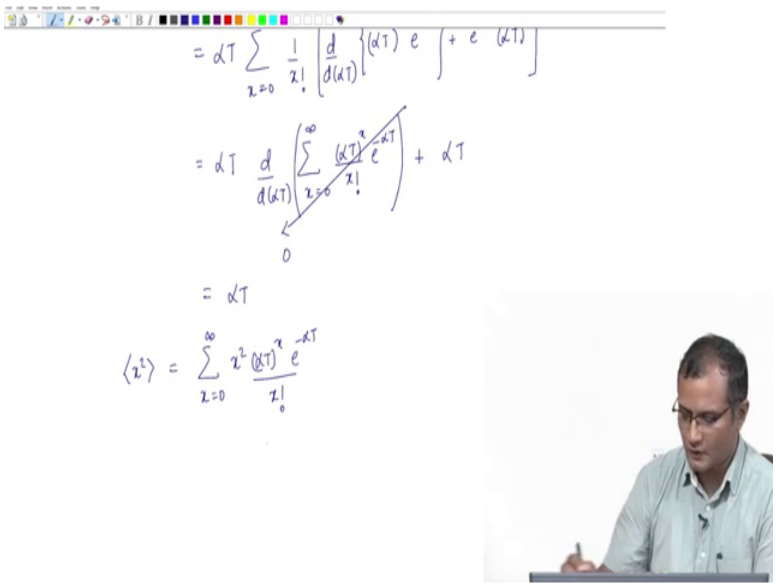
So, now if you want to know, if you want to compute averages in this distribution, so you want to know what is the average x, then you basically have to sample x in this distribution that is the brute force. There is one method like we discussed for the Gaussian distribution. So, your sample x in this distribution, now x can take values from 0 to infinity ok, and you sample x in this distribution this will give you mean x and this will require some effort. So, we can show that there is some effort not much, but there is some effort.

So, you can compute this very easily by some manipulation. So, what I will do is, I will use the fact that the power soft PDF has a is a well behaved PDF. So, so the norm of the PDF which is one is nothing but summation over all x going from 0 to infinity alpha T to the power x e raised to minus alpha T over x factorial, this is one. This is the total probability of all inclusive should be 1, you have definitely some events ok. So, I can write my first moment as simply just rewrite it as I will write an alpha d outside, because this is a constant alpha is a constant t is a constant does not depend on x.

So, I multiply it by alpha and look what this does to accommodate the multiplication by alpha I have just taken the power here as x minus 1 e raised to minus alpha T over x factorial ok. Now, this is nothing but I can write it as a derivative ok, 1 upon x factorial into derivative of and derivative with respect to alpha T of alpha T to the power x into e raised to minus alpha T. but I need to subtract something because this is a derivative of product. So, by chain rule, I will have two terms here.

So, this is basically I need to subtract or add basically e to the power minus alpha T into alpha T to the power x ok. So, if you if you like I can keep this thing inside a curly brace right. Now, you can see that the summation going from 0 to infinity 1 by x factorial into the second term inside the square brace is nothing but 1, because that is the total probability. So, I will write it as alpha T into summation.

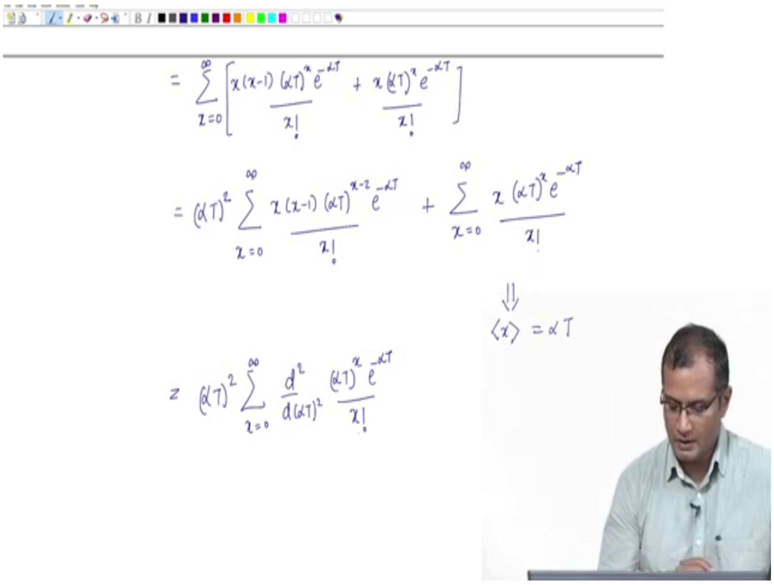
(Refer Slide Time: 20:14)



So, what I am going to do here is basically just to rewrite the things. So, let us keep the alpha T outside, and I am going to take d over d alpha T outside because alpha T and x are independent and I will say x going from 0 to infinity 1 upon x factorial into alpha T. This is going to be my first term. So, this is like I will say that this is the d over dT is acting on this plus my alpha T times 1 ok. So, then this is nothing but so this will basically go to 0 because d by dT of 1 is 0 the bracketed term is one that is the norm of PDF. So, derivative of one is 0. So, I just get an alpha T here.

So, you can see there was some effort required. To compute the second moment some more effort is required slightly more I would say substantially more very tedious calculate very I am sorry very straightforward calculation slightly tedious. So, I will say that by the same logic this is nothing but x going from 0 to infinity x squared alpha T to the power x ok.

(Refer Slide Time: 22:34)

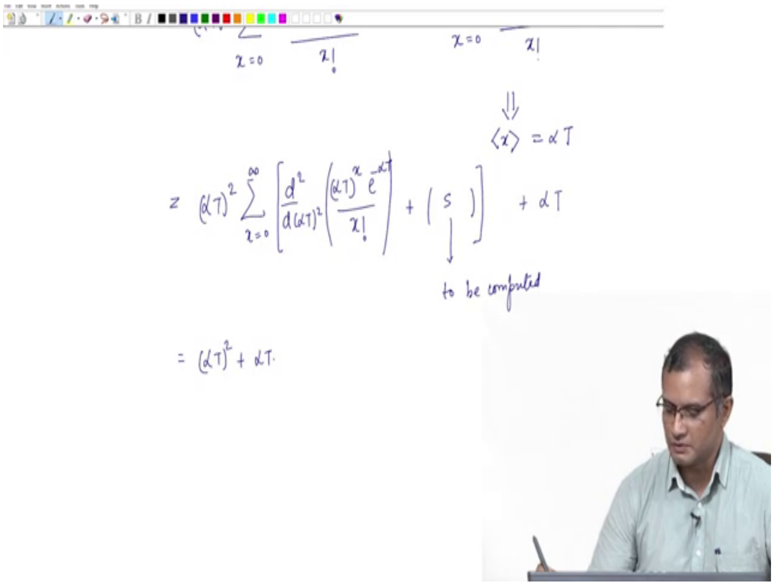


And by the same logic as before and this is purely you know a straightforward calculation; I am not going to complete this I am just going to show that this is this is time consuming ok. So, it may not be very I do not encourage you to follow this route. So, what you can do is just write down x square as x into x minus 1 into alpha T to the power x e raised to minus alpha T over x factorial. Naturally I have subtracted x times the whole distribution. So, I need to add something here. So, I am going to add x times alpha T to the power x e raised to minus alpha T over x factorial ok. And the entire thing here is under summation. I have subtracted something and I have added it back. And now to the first term I am going to do something.

So, what I will do is basically. So, I will just multiply alpha T the whole square, and the first term becomes summation x going from 0 to infinity x into x minus 1 into alpha T x minus 2 e raised to minus alpha T over x factorial plus summation x going from 0 to infinity x alpha T x raised to minus alpha T over x factorial right. So, summation goes in both the terms. Now, this is already known to me the second term is already known to me that is nothing but the first moment which is computed as alpha T, it is already pre-computed.

So, what I am going to do with the first term is just a small manipulation here, I am not going to complete the calculation, because this is tedious and does not prove beyond a point is basically I am going to write this as second derivative with respect to alpha T of our PDF which is ok.

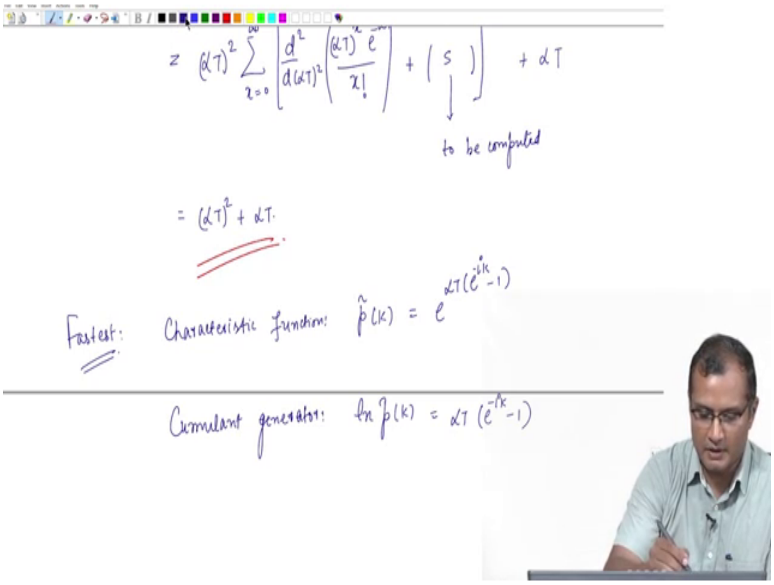
(Refer Slide Time: 25:22)



I have to add a few things here. And finally, basically I have to add an alpha T which is from the second term ok. So, you can compute what needs to be added here, let us call this as some term s ok. So, just take the second derivative which is in the first bracket here, and whatever extra term comes you simply with the minus sign put here for s. Now, that would give you the first to be computed and I leave it as an exercise.

So, there are just two terms that you may have to add here. So, that the second derivative of the Poisson PDF plus these terms gives you the first term in the previous expression. You do that you will end up with the answer which is alpha T the whole square plus alpha T. Now, this is a sort of time consuming exercise.

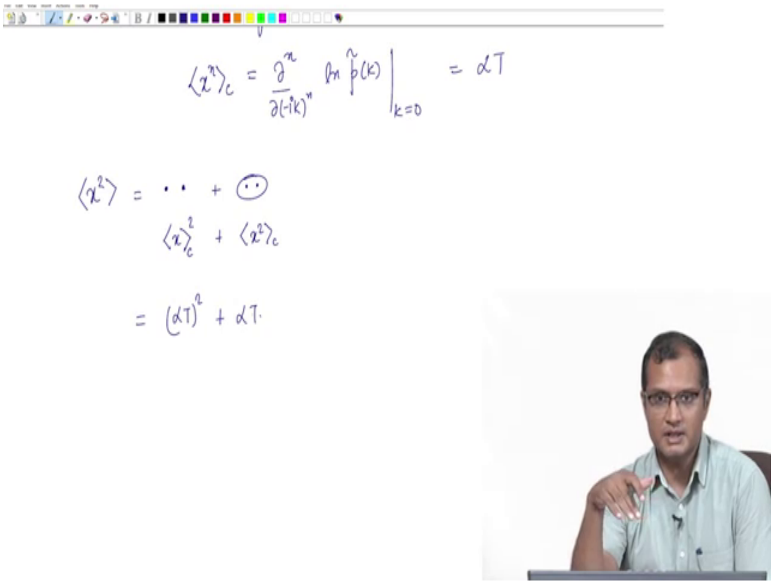
(Refer Slide Time: 26:49)



So, I do not recommend to do this way. The most you know the fastest method here would be to actually use the combination of diagrams and the characteristic function. So, the fastest method is to start with the characteristic function, which is the free transform the PDF and that is e to the power alpha T into e raised to minus i k minus 1 ok. So, you can take the cumulant generator.

In fact, if you know the cumulant generator, you can directly write it which is just alpha T into and here we see an important thing the fact that the equivalent generator has this form all cumulants of arbitrary order are nothing but just alpha T ok. So, when that happens I can compute any moment using this as a cumulantes. And all cumulants are just alpha T. So, suppose I want to have the second moment which I have just computed here this guy.

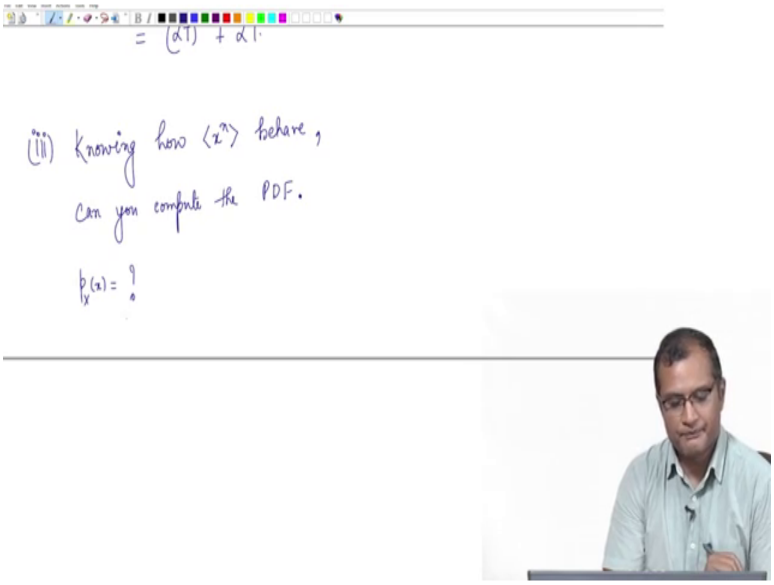
(Refer Slide Time: 28:24)



I will check it here; so, by my knowledge of the diagrams that I have just explained a little while ago. The only combination I see is basically leaving them open and connecting them. So, when you leave them open it is the moment raised to that order; and when you have connected them it is the cumulant of that order ok. Now, I have just said that all cumulants are alpha T. So, this gives me nothing but alpha T the whole square plus alpha T precisely what I had written it is just a three line problem; if you combine the powerful method of diagrams and the characteristic function.

And I encourage you to at least first try it out on problems that you are taking before you use the traditional methods. Use traditional methods only if due to some mathematical problems such as if a characteristic function depends on mod of k, then certainly you cannot take the derivative at origin, because mod k is not differentiable at origin it is continuous, but not differentiable. So, when this method does not work, restore to traditional methods; if the method works, I would highly recommend using characteristic function.

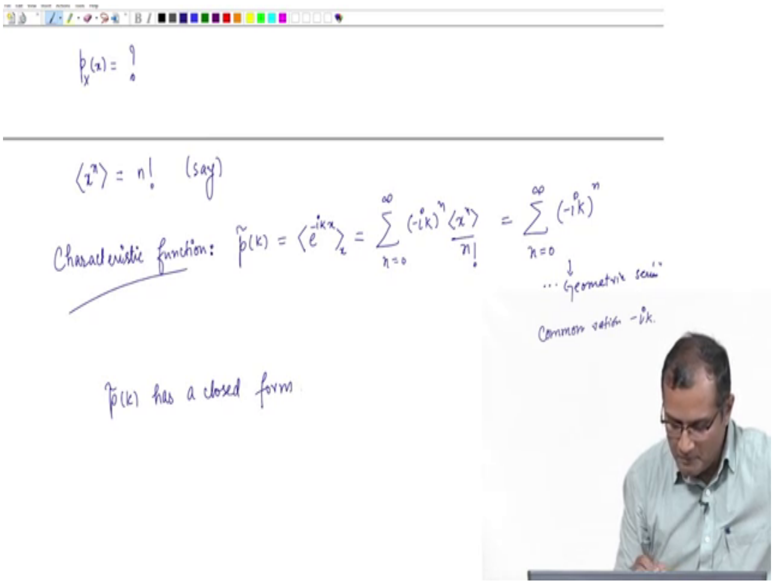
(Refer Slide Time: 29:48)



And the last problem that I would like to bring it up is basically the following you know how averages of a random variable behave. For example, knowing how averages of x behave the question is can you compute the PDF ok? So, this is like going in the reverse direction. We have always taken a PDF and computed averages. This time we are reversing the problem that I know how averages of a random variable are behaving can you give me the PDF in which the variable is distributed.

This problem has interesting applications if you know for example how you know if you do not know the form of a signal, but you know basically the averages and you know the cumulants of a signal then you can actually compute the waveform.

(Refer Slide Time: 31:41)

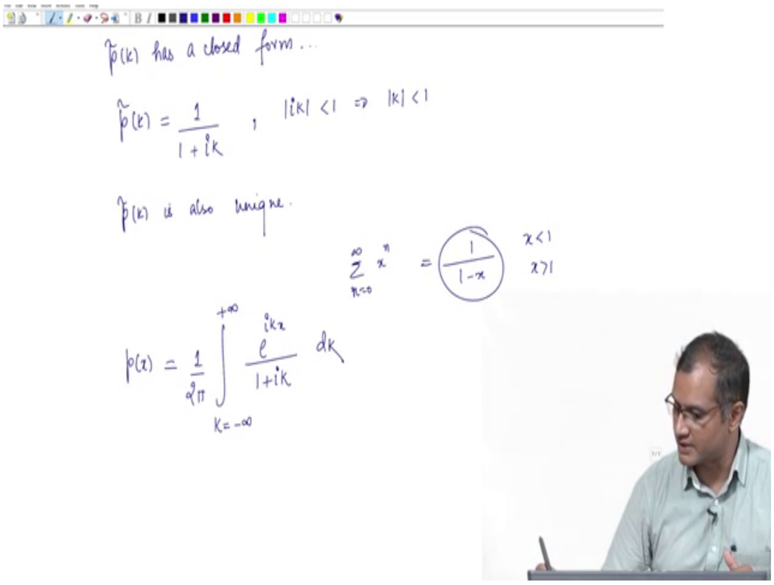


So, I would like to take one example of a random variable which is the distribution unknown. So, I will say that the unknown distribution is p of x, and I will simply take the moments that are known. And I will take a form of it. Suppose, the moments are n factorial we are taking an example I know the moments are n factor. So, what is the PDF? Now, if you have taken the first few assignments seriously, you may already be in a position to tell what is the PDF, but I am not going to disclose it here it will be revealed to you at the end of the exercise ok. So, let us change the PDF all right.

So, now I can start with the characteristic function which is a very important feature by now you should know that we absolutely you know this function has a very, very, very important place in the probability theory. And, so I am going to start with the characteristic function which I know is basically the expectation value of the complex exponential, and here the averages are angular prices are basically the averages in the x space.

So, this is nothing but summation n going from 0 to infinity minus i k to the power n nth moment over n factorial. Now, I have already given the averages. So, the nth averages are distributed as n factorial. So, this is going to be nothing but just n going from 0 to infinity minus i k to the power n. Now, you can easily see that this is a geometric series with a common ratio minus i k ok.

(Refer Slide Time: 34:02)



So, the series has a closed form. And the closed form is given as just 1 upon 1 minus the common ratio which in this case turns out to be 1 plus i k, but as we always know that the closed form for a geometric series infinite geometric series exists for mod of the common ratio less than 1 which is nothing but the statement that mod k should be less than 1.

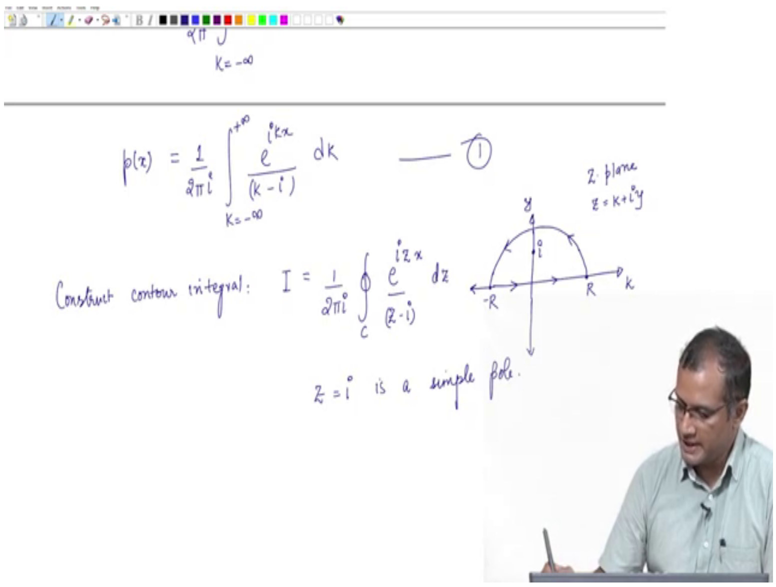
But, wait, this p of k or the characteristic function is also unique, because it is a Fourier transform which means p of k has the same form outside the circle of convergence. It is a geometric series that converges to 1 upon 1 plus i k ok; p of k applies everywhere it is like saying 1 upon 1 minus i 1 upon 1 minus x, this exists for both x less than 1 or x greater than 1. For x less than 1, it has a form which is summation n going from 0 to infinity x to the power n.

So, once you have arrived at a closed form and you know that this closed form is nothing but a Fourier transform. And by the property of Fourier transform that it is a unique function of k, it must assume the same form outside the circle of convergence ok. So, I do not care about the fact that the geometric series does not exist. Well, I do not want the geometric series specifically I want the characteristic function and that I know exists everywhere and it does a unique form.

So, we can bypass all these limitations and take 1 upon 1 plus i k is the unique form of characteristic function everywhere even outside the radius of convergence. Now, the characteristic function is available. The next stop is just to take inverse Fourier transform which is basically 1 upon 2 pi in the way I have defined my Fourier transform the inverse transform is nothing but k going from minus infinity to plus infinity my PDF, my Fourier transform which is 1 plus i k and e to the power i k x dk, this is the inverse Fourier transform of my PDF. PDF is 1 upon the inverse Fourier transform of my characteristic function which is 1 upon 1 plus i k.

So, we need to invoke some well-known tricks from the complex variables, and they are very easy tricks. I call it as I call them as absolutely essential, because they make life very easy in computing integrals. So, I can compute this integral.

(Refer Slide Time: 37:22)

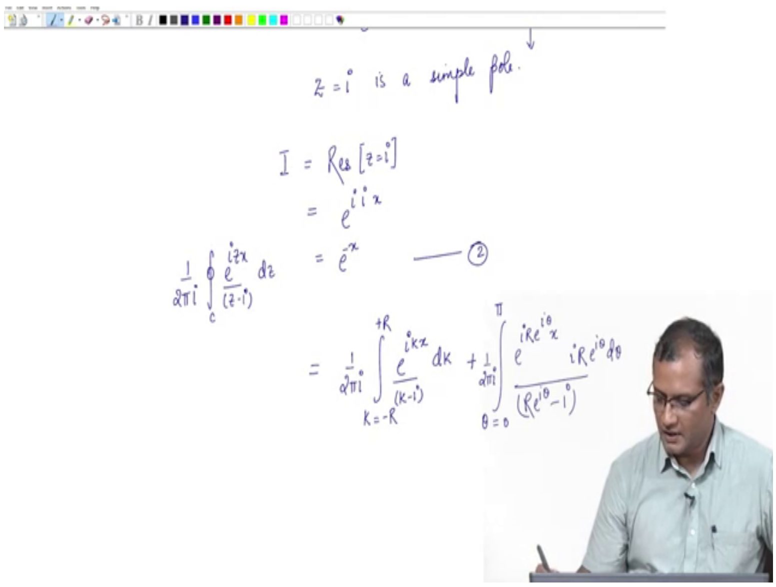


But before that I just want to recast it in slightly favorable guys. So, I will write down this as k going from minus infinity to plus infinity e to the power i k x. So, I have just pulled if you have noticed, I pulled out a i in the denominator outside. And I write it as k minus i ok. So, now, this is the integral that I am chasing ok. So, what I will do is, I will construct some new integral in the complex plane, this is a contour integral. So, construct a contour integral is different from our p of x ok. I will call it as i.

And I define this I as 1 upon 2 pi i contour integral of e to the power i z x dz over z minus i ok. Now, as per the rituals of complex variables, if you have written a contour integral, it makes absolutely no sense if you do not define the contour. So, immediately after writing down the contour, you must show your contour. So, this is my contour. I will first draw my complex plane. So, this is my complex plane, where any complex number is defined as k plus i y meaning my real axis is k axis and my imaginary axis is y axis ok.

And the function e to the power i z x upon z minus i has a pole at z equals to i ok. And I need to show a contour here. So, I am going to show a contour which is basically a semicircular contour. This is the direction of my integration as usual it is the clockwise direction where the integral has a positive value ok. These are the points big R and minus R. So, now, my contour is defined the point where the complex valued function becomes singular is also defined that is the residue. And Z equals to i here is a simple pole.

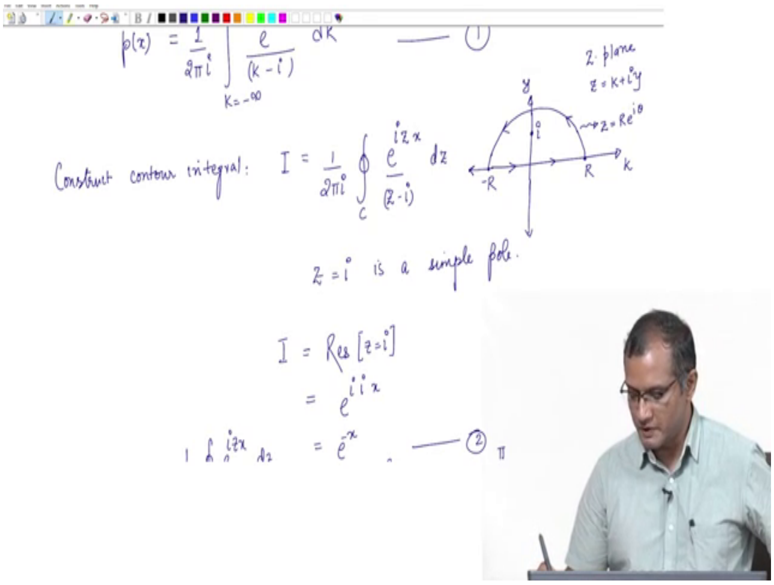
(Refer Slide Time: 40:43)



So, the integral I is then simply given as the residue at z equals to i ok, which is nothing but if I look at if I compute the residue it is nothing but e to the power i into i x which is e raised to minus x ok. So, this is my integral I. Now, I can also write down my integral I. So, this I is nothing but 1 upon 2 pi i contour integral of e raised to i into z x over z minus i dz. I can also write down the left hand side as some of the straight line segments and the circular semicircular segment on the contour.

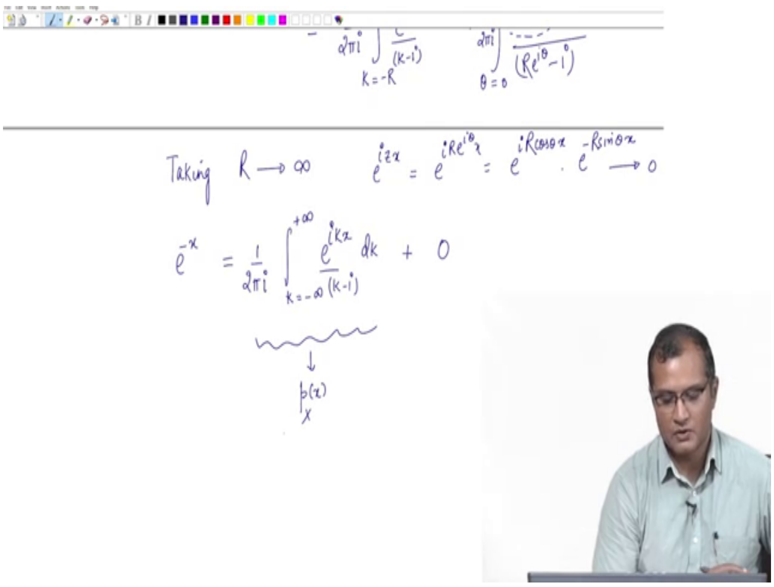
So, this is basically 1 upon 2 pi i integral k going from minus R to plus R, because on the straight line horizontal axis the complex variable is simply k. So, it becomes i k x d k over k minus i plus I need to add the semicircular part which is theta going from 0 to pi, and I have to take a 1 upon 2 pi i outside.

(Refer Slide Time: 42:34)



On the circular semicircular contour the complex variable is basically z equals to R e raised to i theta, because on the circle the magnitude does not vary, R remains constant. It is only the theta that goes between 0 and pi. So, this is nothing but e raised to i R i theta into x i ok. So, now, you can see a couple of things can be done here. So, let me just pull this slightly towards the bottom, so that we have some space for a discussion ok.

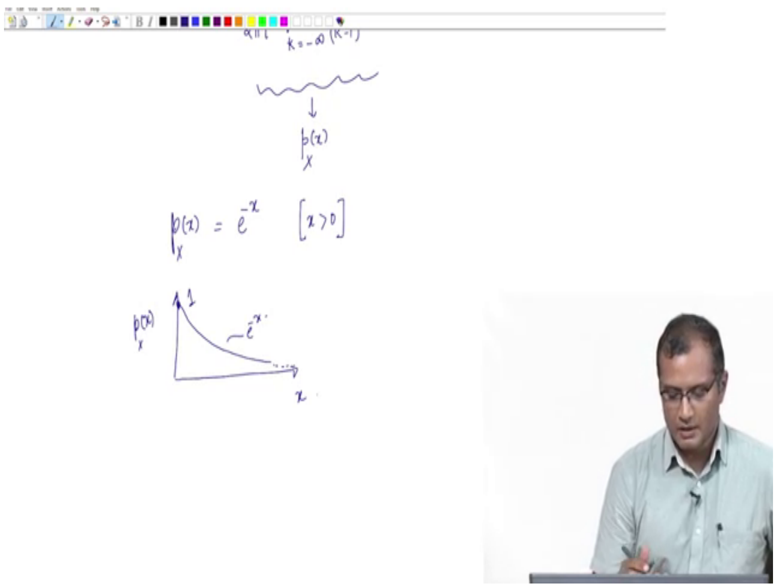
(Refer Slide Time: 44:01)



So, I can say that if I take R to infinity, two things can be seen. The first observation that I see by taking R to infinity is that my numerator e to the power i z x that remains well behaved in fact, it goes to 0 for positive values of x. So, my e to the power i z which is written here is basically i z x which is nothing but e to the power i R e raised to i theta x, this can be written as i R cos theta into x into e raised to minus R sine theta into x. Now, as R goes to infinity for positive x, this will go to 0 ok, because sine theta remains positive in the first and the second quadrant right. So, this integral is well behaved, the residue does not residue still applies here ok. The function remains analytic all the way to infinity.

So, I can write down the residue which is e raised to minus x as nothing but my first integral which is 1 upon 2 pi i integral k going from minus infinity to plus infinity I have taken R to infinity e to the power i k x dk over k minus i. The second part basically is 0 because I have just shown that the numerator goes as e raised to minus R sine theta x and that goes to 0. All other features such as R upon R goes as 1 ok. So, i can say for sure that my and this is actually my PDF p of x.

(Refer Slide Time: 46:13)



So, I have defined I am determined my PDF as e raised to minus x. And I must say here that I got a convergent to z only for x greater than 0. So, this is the PDF. You have a exponentially distributed variable, and let me sketch it for you before we end this discussion, this is a p of x. So, this is your PDF that is how you can construct PDF knowing the behavior of the averages through the method of characteristic functions.

So, we will end the class here, and proceed from the next lecture.