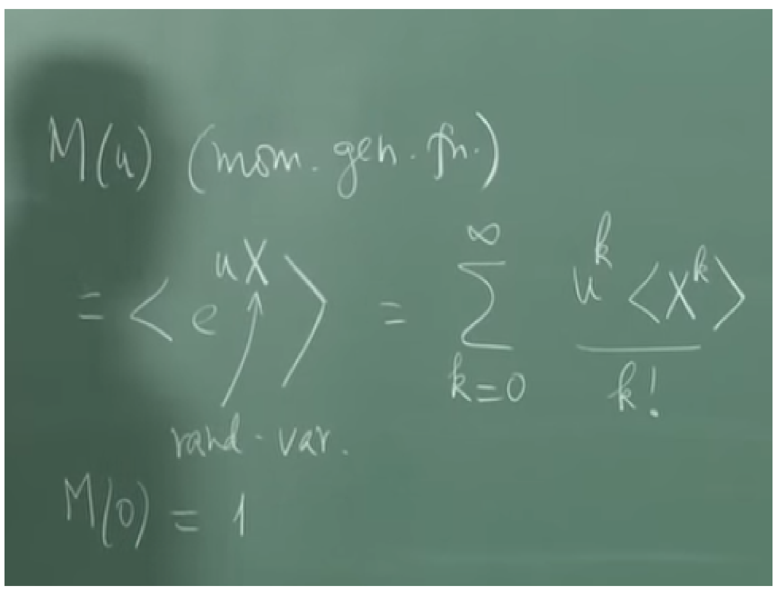
Physical Applications of Stochastic Processes
Prof. V. Balakrishnan
Department of Physics
Indian Institute of Technology-Madras

Lecture - 03
Continuous Random Variables

Alright, so we reached a stage where we were talking about the moment distributing distribution, the moment generating function for a probability distribution and then I mentioned something about the cumulant. So let us carry on with that. The idea is the following. If you have a probability distribution for some random variable and from now on let me use the symbol capital X for a random variable in general.

It could be continuous, it could be discrete. We have looked at discrete random variables, integer valued ones but we are going to extend it to continuous random variables as well.

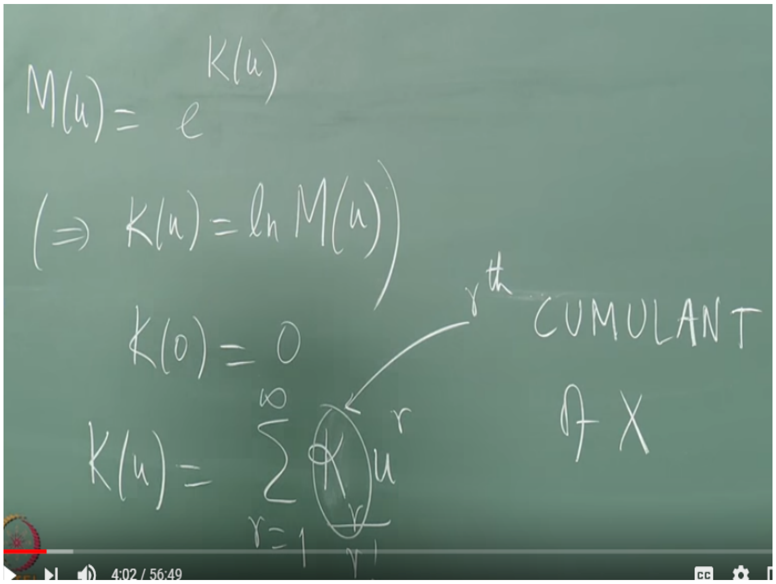
(Refer Slide Time: 00:57)



Then you define a moment generating function M of u as equal to the expectation value of e to the u X where this X is the random variable and of course that is immediately equal to a summation from k = 0 to infinity u to the k expectation X to the k/k! and these are the moments of the random variable okay. Could be discrete, could be continuous whatever with respect to, the average is taken with respect to the normalized probability distribution okay.

Now immediately it follows that M of 0 must of course be equal to 1 because we put u = 0 here, all terms vanish except the k = 0 contribution which is 1. The expectation of 1 is just 1 okay. So that is the normalization of the probability.

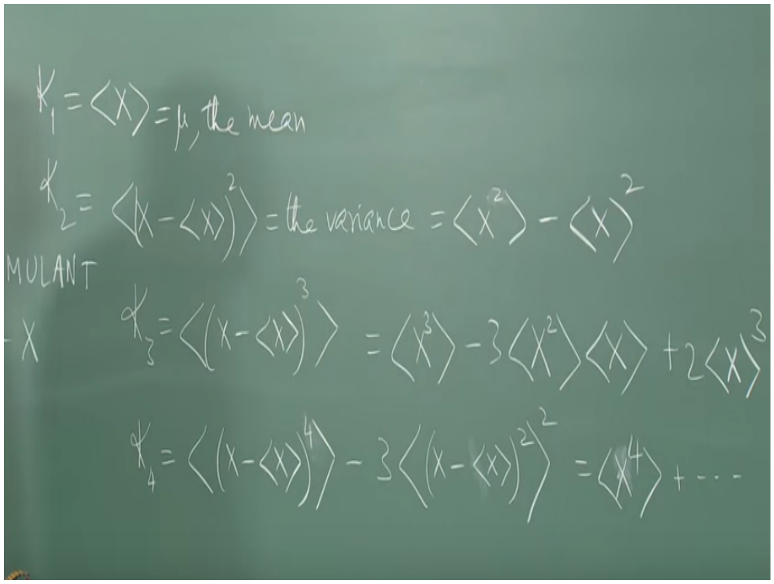
(Refer Slide Time: 02:15)



And then one can ask can this quantity, can M of u be written as e to the power K of u, some K of u okay. Can we do that? What does that mean? It says u are taking this quantity here which is in general going to be some power series in u as you can see and you are writing it as the exponential of another function out here and we will see the advantage of doing this very shortly.

Of course this immediately implies that K of u, K of u is the log of natural log of M of u here okay and K of 0 must of course be 0 so that M of 0 is unity as we have seen here. So in general summation from r = 1 to infinity some constants some kappa r mu to the power r over r! and these constants u independent quantities this quantity here is this so called rth cumulant of the random variable okay.

(Refer Slide Time: 04:06)



Now it is not hard to see that these cumulants are going to have interesting properties kappa 1 equal to it is just the mean value of X. That is trivial to do. All you have to do is to expand this quantity and pick out the coefficient of u in the power series in u and then you discover immediately it is just the first moment. That is trivially true okay. The second moment kappa 2 equal to turns out to be X minus the expectation value of X whole square equal to the variance.

It is just the variance of the random variable. As you know the second moment itself we found was a very inconvenient quantity to use. You needed to subtract the square of the mean from that and then you got a quantity which had a physical significance as a scatter of the variable about the mean and that is kappa 2, the second cumulant okay. Turns out that the third cumulant kappa 3 turns out also to be X minus X average cubed.

Namely the third central moment turns out to be identically equal to that okay. Now when you write this in a power series put that in here and write this as a product of terms and then write the expansion of each of these and collect compare with what happens here with the various moments, you discover immediately it follows from here that the highest term the leading term in the rth moment is in fact the rth in the rth cumulant is the rth moment but then things get subtracted after that.

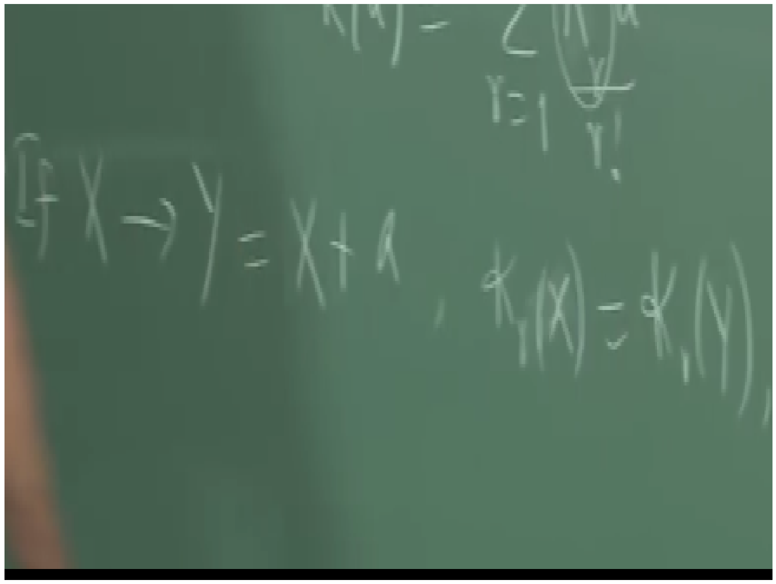
Now for instance this quantity here is just x square - x average square. This quantity here would be an x cubed minus 3 times x square x and then there would be a term which would be plus 3 times x times x average square and then a minus x average cubed but when u take the average value of that you get a cube here x average cube with a 3 and there is a - 1. So the next term must of course be twice x cubed in this fashion okay.

The fourth cumulant turns out to be x minus x average whole square that to the power 4 but there is a correction to it which is thrice x minus x average square. This fellow is a variance, square and of course if you expand this, this is equal to x 4 and then plus etc. there is a set of corrections lower moments than the fourth are going to appear the rest of it.

So there is a systematic way by which you can write the rth cumulant in terms of the rth moment and lower moments and vice versa like this and the coefficients are standard coefficients. In a sense what is happening? You are kind of subtracting out the lower moments, an appropriate combination of lower moments when you do this and you have seen this happen in many places.

For instance when you subtract out in the quadruple moment of a charge distribution you will subtract out the lower moments in exactly the same, so that the whole thing is rotationally it has got interesting definite transformation properties. In exactly the same way, by doing the subtraction a very important property emerges among other things namely if you just shift the random variable by a constant value then the moments of course would change but the cumulants do not change.

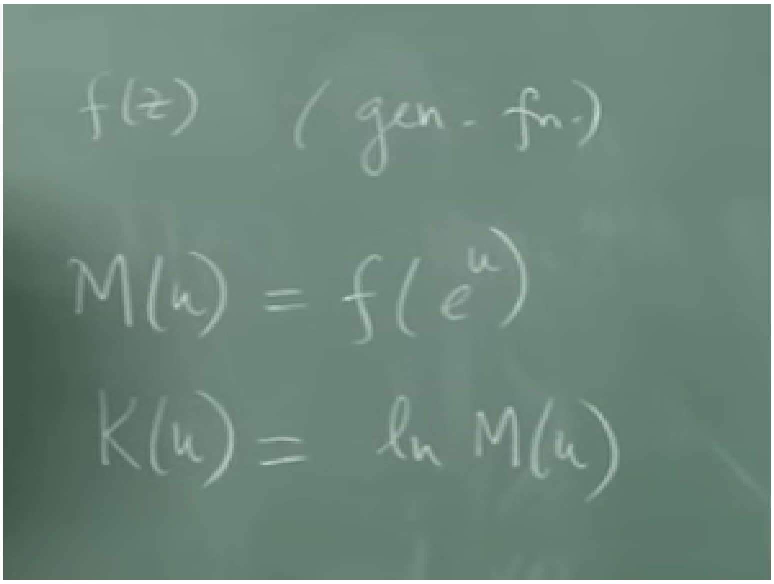
(Refer Slide Time: 08:17)



So if X is replaced by Y which is X plus a constant a, some constant value, a sure value a, then the first moment of course will change because the average of Y is now the average of X + a this constant a but all the cumulants remain exactly the same. K r for x = kappa r of Y for r greater than equal to 2. That is trivial to see. It is immediate here for example and immediately see wherever you can write it in terms of central moments this thing the mean value just cancels out the shift in the variable cancels out and then you have this invariance here.

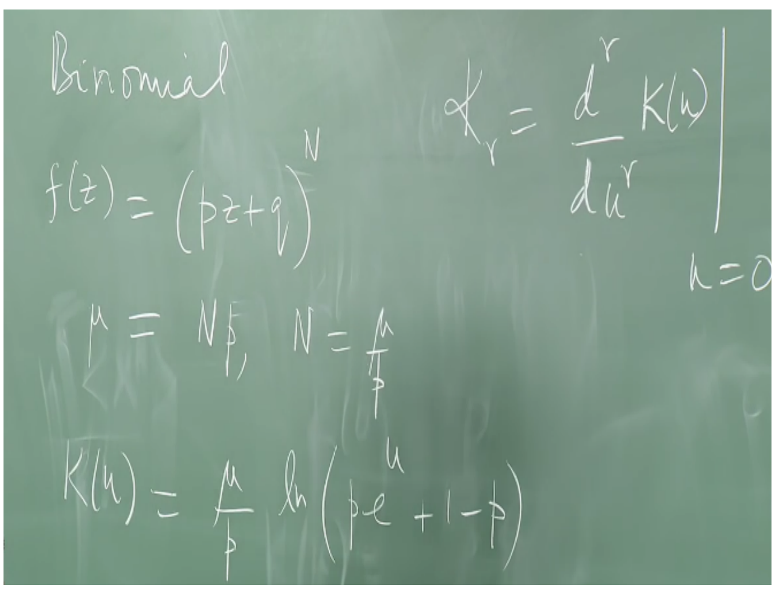
So the cumulants of a random variable are invariant under translations of this random variable. You shift it by a constant then it does not change at all. That is one very crucial property. Let us write down the cumulants for various distributions that you already know about. So we have seen a whole lot of distributions for which we know closed-form answers. So let us write it down.

(Refer Slide Time: 09:33)



If you recall, just to recall to you what happened. If you recalled I defined a generating function f of z. This was just a generating function and then I had a moment generating function M of u and we found that this was just f of e to the power u. So wherever z appears I just replace it by e to the power u and K of u the cumulant generating function is just the log of M of u. Now let us see what this turns out to be for various distributions so that you could write the cumulants down.

(Refer Slide Time: 10:13)

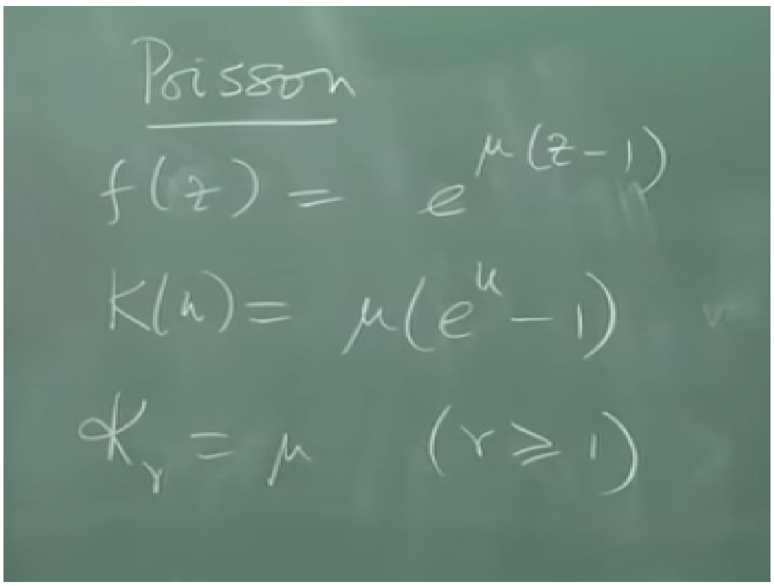


The binomial distribution the probability distribution itself we wrote down it is just the binomial. The generating function for that was of the form f of z was equal to p z + q to the power N when you had N Bernoulli trials and p was the probability of success in any given trial then it says M of u is this fellow so remember that the average value mu the mean was equal to N p.

So instead of N let us write it as mu over p and if I take logs I put instead of z I put e to the power u and then I take logs I end up with K of u = mu over p + 1 - p.

So that is the cumulant generating function for and remember that the kth cumulant rth cumulant kappa of r is d r over d mu r K of mu evaluated at mu = 0. So once you have this expression it is a very simple matter to write down what the cumulant generating function is and therefore what the cumulants are completely for this binomial distribution. What happens in the case of the Poisson distribution?

(Refer Slide Time: 12:14)

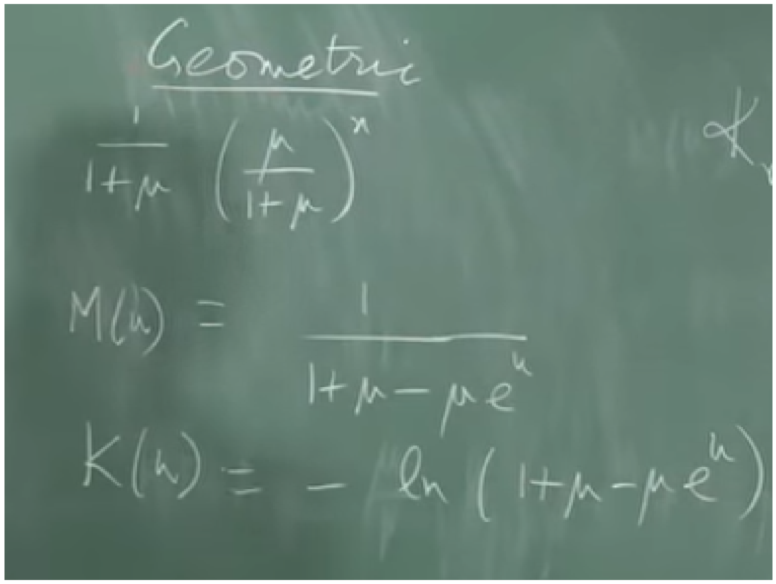


For the Poisson f of z was just e to the power mu times z - 1 okay. So I replaced z by e to the u and then take the log to get kappa of u K of u. So this immediately says K of u = mu e to the u - 1 that is it and what can we now say about all the moment all the cumulants? All we got to do is this, differentiate it z = 0. If you differentiate this fellow you can just e to the u once again you put u = 0 you get 1 right.

So it immediately says that for this Poisson K r = mu for all r greater than equal to 1. So the mean, the variance, the higher cumulants they are all the same. It is just one number. It is a very special property of the Poisson distribution not shared by others. There are other distributions which might display this property.

You can create them but the fact is that for the Poisson the variance so the statement that the variance is equal to the mean for a Poisson distribution is a special case of a more general statement that all the higher cumulants are equal to the mean value in this case okay. What other distribution did we look at? We looked at the geometric distribution right?

(Refer Slide Time: 13:50)

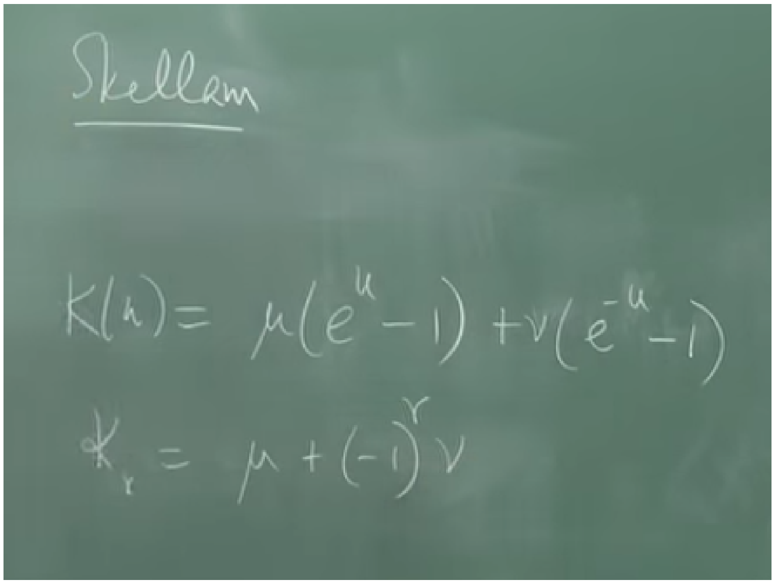


So for the geometric distribution with mean mu the distribution itself was 1 over 1 + mu, so geometric times, mu over 1 + mu to the power n. That was the probability distribution of the random variable n which took values 0, 1, 2, 3, etc. And now if I find f of z for it, it is just a summation of this guy so it is 1 over 1 + mu and then this geometric series summed from 0 to infinity which is 1 over 1 - this guy multiplied by a z.

So the 1 + mu will go away and you get 1 over 1 + mu - mu z but I got to put e to the power u here and that is M of mu. So it says K of u = - log 1 + mu - mu e to the u and that is it and now you can write down all the cumulants from this directly. A sum of Poisson random variables is again Poisson so nothing new happens.

What happens if you have a difference of the two? If you had a Skellam distribution for instance where you have the difference of 2 Poisson random variables whose means are mu and nu say. What happens then? Well the generating function was this and then in the other case for the nu because of the minus sign when we generated it you had an e to the 1 over z instead of z. So it would be just this.

(Refer Slide Time: 15:39)

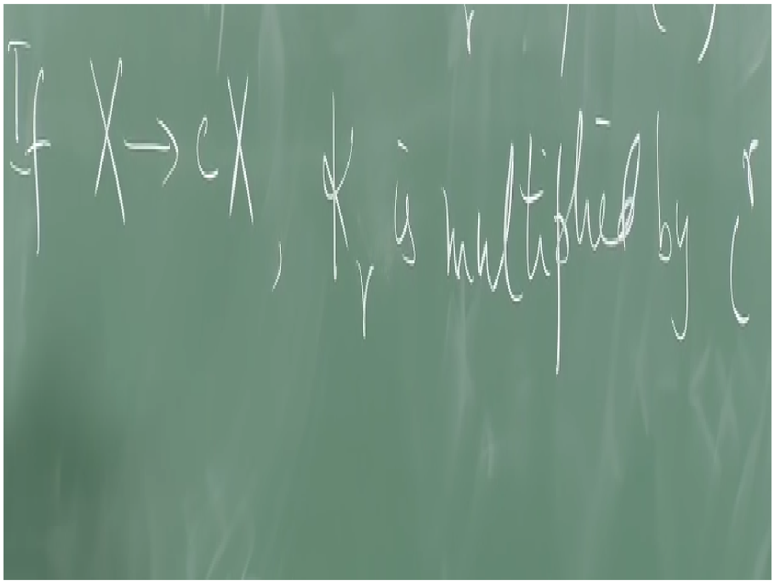


Plus mu times e to the minus nu e to the minus u minus 1. That is K of u okay and of course the first cumulant is the first derivative of this at u = 0 and that will immediately give you mu - nu which we know is the mean, the second cumulant you differentiate this twice you are going to get a plus sign again and so on. So it immediately says kappa r = mu plus minus 1 to the power r nu in this case.

So every other moment, every other cumulant is mu + nu and every other the even ones are all mu + nu and the odd ones are all mu - nu as you as you would expect in this case. We will write down the cumulants of some continuous distributions as we go. So the first important property of a cumulant is that it is translation invariant. The other property that is obvious by looking at it is that the cumulant is a homogeneous, the cumulant, the rth cumulant is a homogeneous function of this random variable in a strange way.

That is if you multiply the random variable by a constant, then the rth cumulant gets multiplied by that constant to the power r okay and that is fairly straightforward to see.

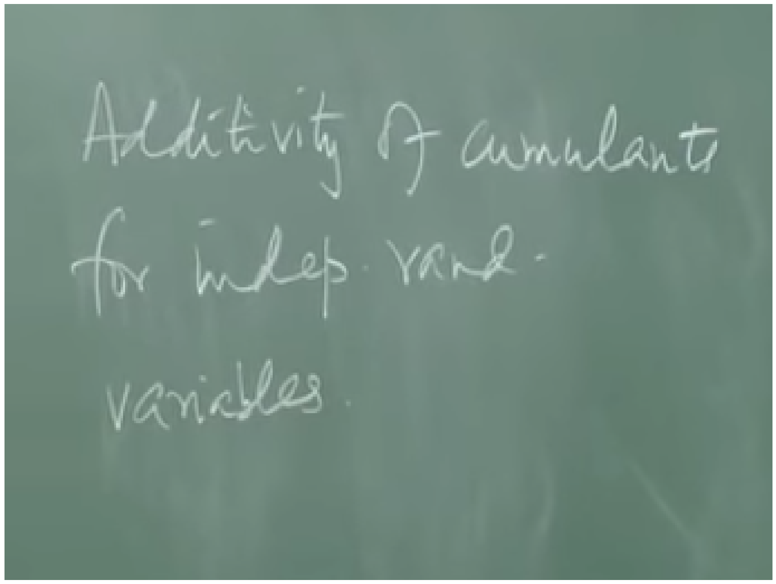
(Refer Slide Time: 17:02)



So if X goes to c times X, K r is multiplied by C to the r. So the scaling property is also very immediately obvious here. The cumulant has another very crucial property. We saw that the variance of 2 independent random variables and is simply the sum of the individual variances. This is going to happen for all the cumulants. The additivity of cumulants is a very crucial property and we can see that in many ways.

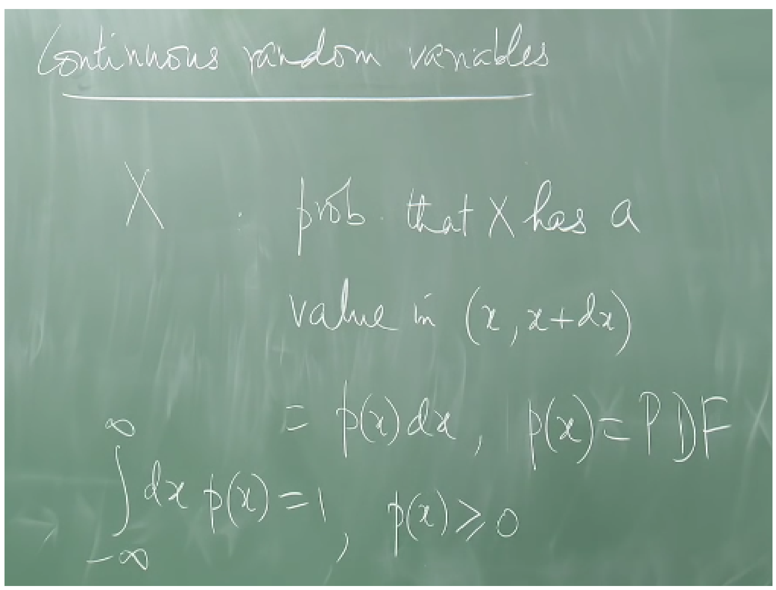
But one way of seeing it is to say that well if I take the log here, this moment generating function just multiplies for various random variables and if I take the log it simply adds up. So it is clear that if you have several random variables, independent random variables then and they are independent, statistically independent then the cumulant rth cumulant of the sum is equal to the sum of the rth cumulants.

(Refer Slide Time: 18:18)



So the crucial property is for independent this is absolutely crucial. It is a very important property. We make use of it as we when we talk about limit theorems we are going to make use of this, this part of this property okay. Now I have mentioned off and on continuous random variables, just to say a few words about it and then.

(Refer Slide Time: 19:02)



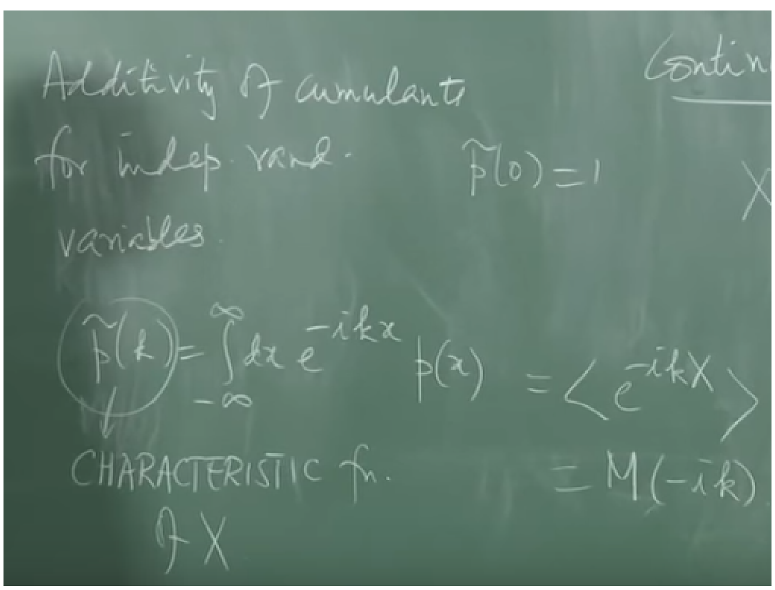
So you have a continuous let us call X the random variable, takes values in some continuous interval of the real axis reference or takes values over all real values of over all real numbers. Then instead of talking about a probability of this X for any particular value which must be defined with infinite precision now is a set of measure 0 a point in a continuum. All you can talk about is the probability density function of this variable.

That this random variable has a value probability that X has a value in some x, x + dx = p(x) dx, p(x) is the probability density function. I call that, denoted by PDF okay. That is the probability density function and it cannot be negative. This number cannot be negative; could become unbounded. All you need is normalization. So all you need is integral dx p(x) = 1 over whatever is the range of this variable, in general minus infinity to infinity say.

And we also need p(x) to be greater than equal to 0. Now we are going to be rather loose in our mathematics. If for example you have a situation where there is one particular point in the continuum where there is a finite probability and this variable has a value then we will include it in here by putting a delta function at that point and put a delta function spike in the PDF with an appropriate weight factor so that it gives you the probability of taking on that particular value.

So we will be a little casual about this. I will continue to write integrals but then in here could be delta functions out here okay. Alright, now once you have a continuous random variable of this kind it is convenient to define and once you have a probability distribution function of this kind which is integrable it is convenient to define a fourier transform for this variable.

(Refer Slide Time: 21:45)



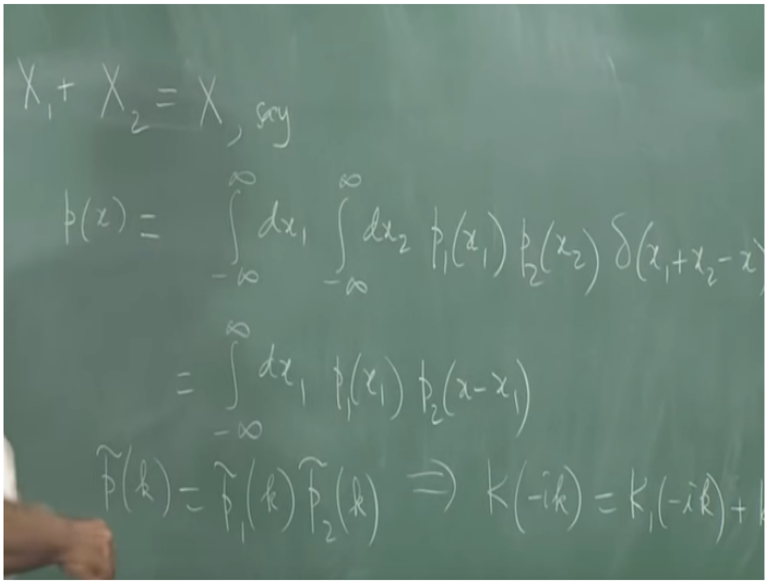
And the fourier transform of this quantity p tilde of k equal to integral minus infinity to infinity dx e to the - ikx p(x) this quantity. This is called the characteristic function of this random variable. But it is nothing new. We have already introduced this quantity in a different way. This quantity is just the expectation value of e to the - ikx with respect to this weight factor here. But we know what e to the u x is. That is M of u. So this is nothing but M of - ik.

So the characteristic function is just another way of writing or saying that you have a moment generating function okay. There are analytical properties of these variables which I am not emphasizing at the moment but you see if you give an arbitrary function p tilde of k I cannot claim immediately that it is a characteristic function of a random variable till a certain set of conditions is satisfied.

For instance if I want a normalization to be valid for this if I put k = 0 here and I want the integral to be equal to normalized quantity 1 total probability then this immediately implies that p tilde of 0 must be equal to 1 but even more strongly given a function p tilde of k it can be a characteristic function only if it is fourier transform, only if it is inverse fourier transform gives you a non-negative function p(x) okay.

So that is a very strong constraint, a real function should be real and it should be non-negative. That is a very strong condition. So all functions of k are not going to be even if they are integrable are not going to be characteristic functions okay and that is an important test alright. Now the additivity of cumulants follows trivially once I introduce the characteristic function as you can see from here because this immediately says that.

(Refer Slide Time: 24:33)



So if I have a random variable x 1 and another random variable x 2 with moment generating functions m 1, m 2 and cumulant generating functions k 1, k 2 etc. then it immediately says as you can see the probability density function of the variable, let us call this equal to x say. Then p(x) is equal to an integral from minus infinity to infinity d x 1, integral minus infinity to infinity d x 2, p 1(x 1), p 2(x 2) where x 1, x 2 are the points in the sample space of these 2 variable random variables.

And p 1 and p 2 are the corresponding probability density functions okay multiplied by the constraint that x 1 + x 2 must be equal to x okay and where does that get us. That says this quantity is minus infinity to infinity d x 1, p 1(x 1), p 2(x) - x 1. If I use the delta function constraint then this is all it is. Now in what form is this quantity here? It is a convolution.

So it immediately follows that p tilde of k = p 1 tilde of k, p 2 tilde of k by the convolution theorem for fourier transforms. But that is the same as M of M 1of - ik and M 2 of - ik and if I take logs this immediately implies that k of u, k of u or - ik it does not matter. These are all power series so which implies that the cumulants add up because the cumulant, the rth cumulant of this quantity is the coefficient of - ik to the power r/r! or whatever right.

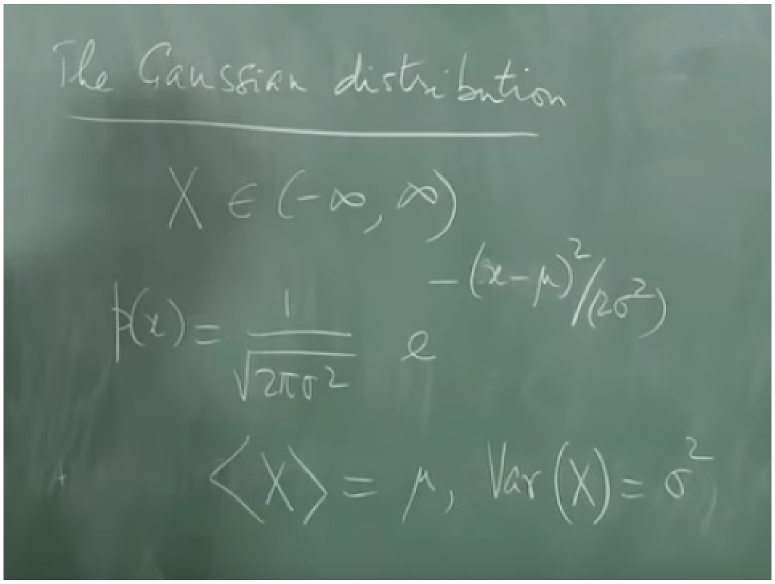
So it says immediately the additivity of cumulants is a trivial consequence of this fact here okay. If they are discrete value, random variables over some finite range or something like that you got to work a little harder to do this but it is pretty much the same here.

"Professor - student conversation starts" Ya. The x there should be both x 1 and x 2 there. I am sorry say that again. The x 1 + x 2 should be equal to x in that case when we are finding the probability density factor of x, yes, what exactly are we doing in this we are getting x 1 comma x 2 equals x on top. X 1 + x 2 is equal to x. Ha I am sorry, I am very sorry ya. Thank you ya x 1 + x2 is the sum yes. **"Professor - student conversation ends".**

Now generalization to some constant times x 1 plus some other constant times x 2 any linear transformation similar things will happen but you can check that out directly. Thank you. Now we could go back and look at various continuous distributions or even the discrete distributions and ask for sums of random variables and that is going it is going to become very important to understand how the sums of random variables behaves when you have a large number of components.

We are going to spend some time on it but let us do this in the context of a very famous example. The simplest of these example, the most ubiquitous of continuous distributions is the Gaussian distribution. So let us write down the answer for the Gaussian distribution what happens to the cumulants and a very important property will emerge.

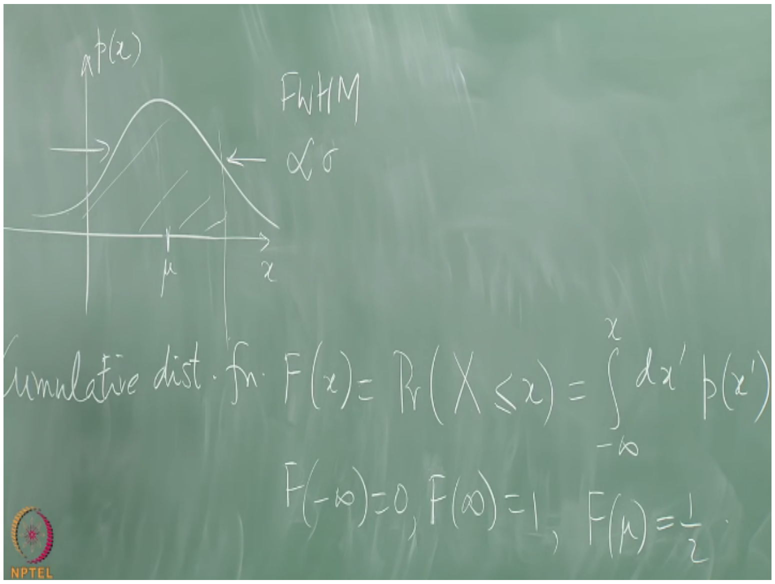
(Refer Slide Time: 28:54)



So the Gaussian, it is also called the normal distribution and it is parameterized by 2 quantities the mean and the variance. The PDF for the Gaussian, first of all you got a random variable X which is an element of minus infinity, infinity and the corresponding PDF p(x) the normalized PDF is this quantity. I should use curly x here the value. Then the mean value of x this quantity is mu and the variance of X = sigma square. So it is parameterized by the mean and the variance, the two parameter distribution okay.

Now this distribution is going to appear ubiquitously everywhere, we will see when we look at the central limit theorem how this distribution emerges in a very general context. But right now let us look at some of its properties.

(Refer Slide Time: 30:18)



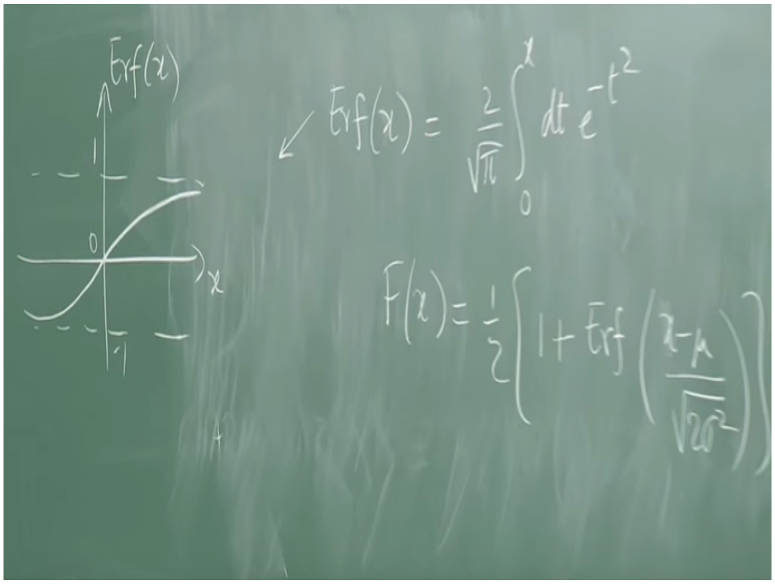
First of all the shape of this distribution, it is quite straightforward. It is something which is bell shaped here where this is the mean value. It is unimodal. The peak is at the mean. Out here is a function of little x, this is p(x) and this width here this half width is proportional to sigma. So the full width at half maximum the value at this point and this mu at x = mu is 1 over root 2 pi sigma square.

And when you go to half that value this quantity here the width is proportional to sigma and sigma is the standard deviation out here okay. You could ask what the cumulative distribution function is so the cumulative distribution function let us call it f of well there are various notations used for it. Let us call it P(x). This is equal to the probability that the random variable x is less than equal to some specified number x which is equal to the integral from minus infinity up to x of d x prime p(x) prime.

So the probability that the variable has a value less than some specified value is the area under this curve. That is the cumulative distribution function. It is clear it cannot be a decreasing function. It is got to be a non-decreasing function. In this case when you have a distribution like this as you move to the right the area keeps getting added to so it is an increasing function right and P of minus infinity is 0. P of infinity is of course 1.

Let us use another symbol for this. I do not like this P. Let us call it F out here. F of infinity is 1 and by the symmetry of this distribution it is quite clear that F of mu = 1/2. At this point you have exactly half, the area is exactly half okay to write down this F(x) in terms of some known functions.

(Refer Slide Time: 33:04)



Well there is this famous integral. It is called the error function, error function of x. This is defined as an integral from 0 to x dt e to the minus t square and you want to normalize it. So 0 to infinity the answer is equal to square root of pi over 2. So 2 over root pi. Then Erf infinity is equal to 1, Erf 0 is 0 and Erf - infinity, it is an odd function so it is equal to - 1 goes from - 1 to + 1 to infinity to 1 at x = infinity.

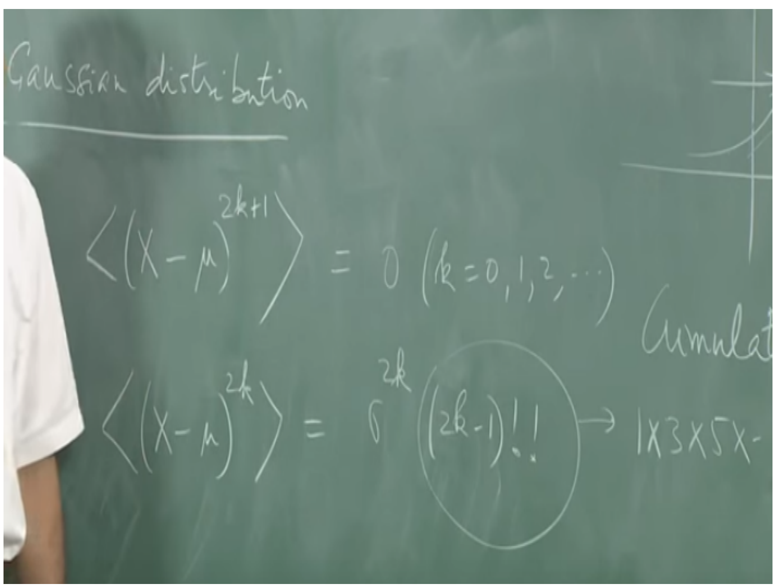
So this quantity, this f(x) it turned out to be half 1 plus the error function of x - mu because I have shifted everything to mu, the origin to mu and it should be scaled down. I used a dimensionless variable here. So what I need is in the probability density I had e to the - x - mu whole square/2 sigma square. So the length scale there was fixed by 2 sigma square right. So I got to kill that and that is all it is okay.

And this Erf x itself is a function which looks like this the function of x. This is 1, this is - 1, then 0, and the function goes like this and that is what the cumulative density distribution function of this random variable is okay. Statisticians like to use the cumulative distribution function rather than the probability distribution function itself for various reasons.

First of all when you have these atomic probabilities namely you have a given point where there is a finite measure for instance then we need to introduce things like delta functions and so on which is rather singular objects but when you integrate it out things get smooth. So people like to use this rather than using. You do not mind using step functions but delta functions are little singular. You got to define them more precisely etc. So it is convenient in many cases to do this.

Physicists generally work with densities all the time, probability density functions etc. So once we have this we could ask what is what are the various quantities associated with this Gaussian distribution, 2 parameters mu and sigma square the variance, we could ask what all its cumulants are for instance. We could ask what its moments are and what would be the moments of this distribution. We already can predict what is going to happen.

(Refer Slide Time: 36:05)



For a Gaussian we can ask what is X minus the average value, the central moment, this quantity. This is of course 0 by definition because mu is just the expectation of X but we can ask what this is for the kth moment 2k plus 1 say the odd moment. What should this be? Well the PDF is a symmetric function of X minus mu and now I am asking what is the average value of X - mu to the power an odd number in odd integer. It should be 0.

By symmetry this is 0. The integral exists for all positive K as you can see because there is a e to the - X square to take care of convergence. Definitely all the moments exist so this is identically equal to 0 etc. 0, 2 and what are the even moments like? What would this be? What is it when k is 1? It is just the variance. It is just the variance. What is it when k is 0? It is got to be 1, average value of 1.

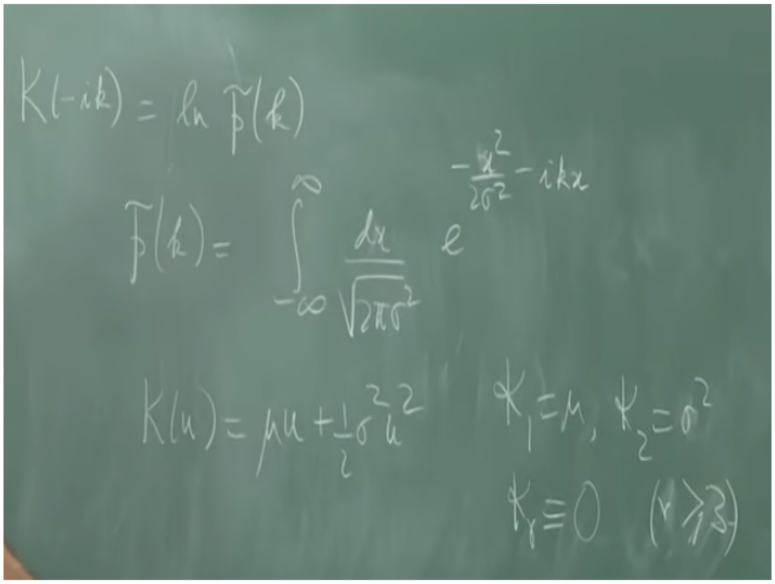
Now it is clear that the answer depends only on sigma because once you shift to mu the only parameter left is sigma and the only quantity of dimensions sigma of dimensions length is sigma in the problem. So this got to be proportional to sigma to the power 2k just on pure dimensions multiplied by some factor and that factor is not hard to find. You can write down a Gaussian integral multiplied by any even power here and turns out this thing is times (2k – 1)!!.

This stands for 1 times this fellow here stands for 1 time into 3 into 5 into 2 k - 1 okay. So this symbol double factorial is very often used for this. Or you could write it in terms of 2k!/k! times 2 to the k and so on so forth. Now there is a nice interpretation of this, combinatorial interpretation which were which is useful in places like field theory when you do what is called Wick's theorem, it is very useful.

If you took this thing here X - mu to the power 2k and wrote it out as factors, you have 2k factors each of which is X - mu then you can ask in how many distinct ways can I pair these fellows, can I write them pair-wise. How many pairs you can independent pairs can I find and the answer is precisely this. So this is really a combinatorial factor that arises from the number of ways in which you can pair 2k objects two at a time okay.

So that is what this is. Now you could ask what is so a very important property emerges that all the central moments are all dependent, just powers of this guy here, nothing more sigma square to various powers. Now you can ask what is the cumulant generating function of this distribution and a very interesting fact emerges.

(Refer Slide Time: 39:57)



K of u in this case is equal to, well first you want the moment generating function but you remember this is just a K of - ik is just log of p tilde of k and p tilde of k is the fourier transform is integral minus infinity to infinity 1 over root 2 pi sigma square e to the - x square over 2 sigma square - ikx. It is the fourier transform of a Gaussian okay and what is that going to be. It is also going to be a Gaussian.

This guy is also a Gaussian and all you can, all you need to do is to complete squares out here. Pull out a 1 over 2 sigma square complete squares here. It is also a Gaussian but what is the width of that Gaussian going to depend on? It is 1 it is sigma square itself whereas here the width was sigma square there it is 1 over sigma square. There is this very interesting property of fourier transforms that the more compact a fourier transform is the more spread out its function is, the more spread out its fourier transform is and vice versa.

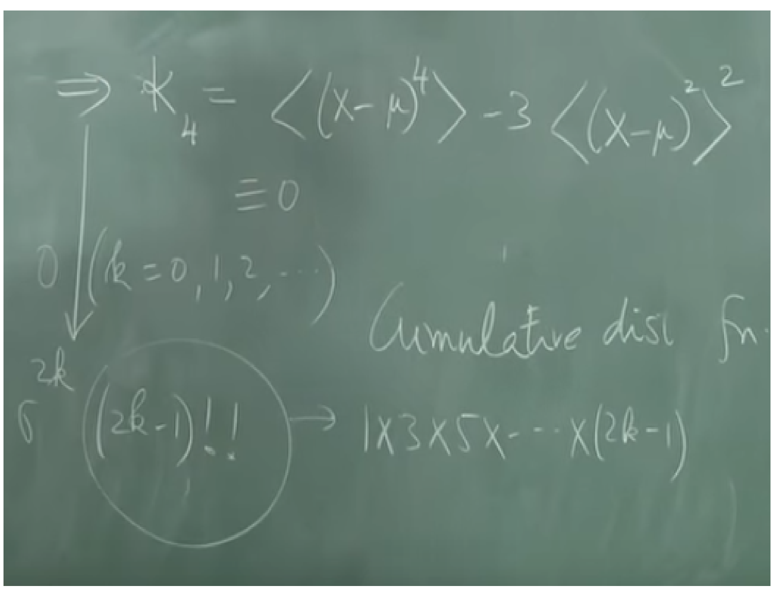
So the width here if it is sigma square the width there is 1 over sigma square, very profound implications. So this will lead to the fact that K of u = mu times u plus half sigma square mu square. K of - ik which is p tilde of k will turn out to be - mu ik minus half sigma square k square. So what does that tell us about the Gaussian? It says of course it is immediately true that kappa 1 = mu and kappa 2 = sigma square.

That is that is very clear. But what does it say about this, when it is greater than equal to 3. What does it say about it? It is 0 identically because remember it is the coefficient of u to the r over r! in a power series expansion about the origin of K of u and this is a polynomial. That is it, just the first 2 terms and everything else goes away. So an incredible property of the Gaussian is that all cumulants higher than the second one vanish identically.

There is a first moment, there is a second moment and there is a variance, mean and a variance and all the higher cumulants are identically 0 okay. So a very basic property of the Gaussian is that the higher cumulants are all identically 0 which again means that if you look at the fourth cumulant for example or ask the third cumulant that is identically 0 there is a physical meaning to these cumulants.

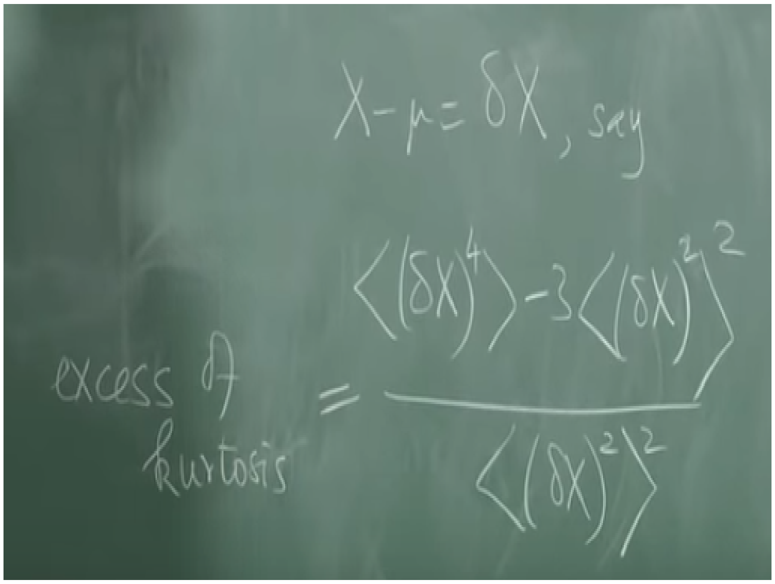
The mean is of course is going to tell you something about the average value, it is exactly the average value. The variance tells u the scatter about the mean. The third moment gives you what is called skewness. It is related to how asymmetric it is, this distribution is, and when you have a symmetric distribution like the Gaussian the fourth or any other symmetric distribution then the fourth cumulant gives you information about how much it departs from Gaussianity.

(Refer Slide Time: 43:45)



Because this relation immediately tells you, this relation here it immediately implies that for the Gaussian kappa 4 is equal to this quantity which is delta X that is the X minus average X which I call mu, mu to the power 4 - 3 times the average value of X - mu whole square square is identically 0 okay for the Gaussian.

(Refer Slide Time: 44:34)



And this quantity in general, let me call X - mu equal to the deviation from the mean. Then this quantity delta X 4 - 3 times delta X square the whole square and just to make it dimensionless we divide by this quantity delta X square square this quantity. This is called the excess of kurtosis and for a Gaussian it is identically 0. This quantity could be positive, negative or 0. If it is 0, for a Gaussian it is identically 0.

But if it is positive, what does it sort of imply? It says that this guy is dominating over that in some sense and this is the fourth moment; so it means that the higher, larger values of X about the mean are actually more significant than the smaller values. It says something about the shape of this distribution um. Similarly, if it is negative it says the large values don't dominate, the smaller values dominate right.

So in one case you got a thing which is fatter than the Gaussian, in the other case you got something that is linear than a Gaussian and these are important indicators of the deviations from Gaussiantity and the reason the deviation becomes important is because Gaussianity is what I would expect if you had as we will see a lot of random variables added up in an incoherent sort of a fashion and the limit for suitable rescaling of the sum linear combination it turns out the distribution will be Gaussian and a very -robust conditions.

So this implies that whenever you have a deviation of this kind it says something very important about the underlying physics in the problem okay. So keep that in mind that you have this is identically 0 for the Gaussian but then there are distributions for which this is not so. Now pretty much you can ask does this go on forever. After all to define the distribution completely I need information about all the moments. So is it that I need an infinite set of numbers. Only then can I reconstruct the distribution?

For instance, suppose I give you all the moments of a distribution. Can you uniquely reconstruct the probability distribution function or the density function? This important problem is a problem in the mathematical statistics. It is called the problem of moments and there are certain answers known to it under suitable conditions. It is a very important problem. We will not go into that.

But let me explain say simply say that for practical purposes very often when you actually analyze data etc. the first 4 cumulants serve to pretty much describe the random process the random variable more or less completely. So the mean gives you some crucial information about what this variable is typically likely to be if it is a simple kind of distribution. The variance gives you a scatter.

The third one tells u about skewness or asymmetry and the fourth one tells you departures from Gaussianity. So pretty much this numerical purposes this should this suffices in most cases. But of course from a theoretical point of view you need to know all the moments before you can make statements here. You could ask the following question which is an interesting one. I am not going to prove it here which is the following.

Other continuous random variables with well-defined probability density functions such that just as a Gaussian had a quadratic cumulant generating function and all the higher cumulants were identically 0 after K 3 onwards, is it possible to have a cumulant generating function which is a polynomial of some finite degree greater than 2 and everything else all the higher powers are 0.

So the distribution would have cumulants up to some n and then every other cumulant is identically 0.

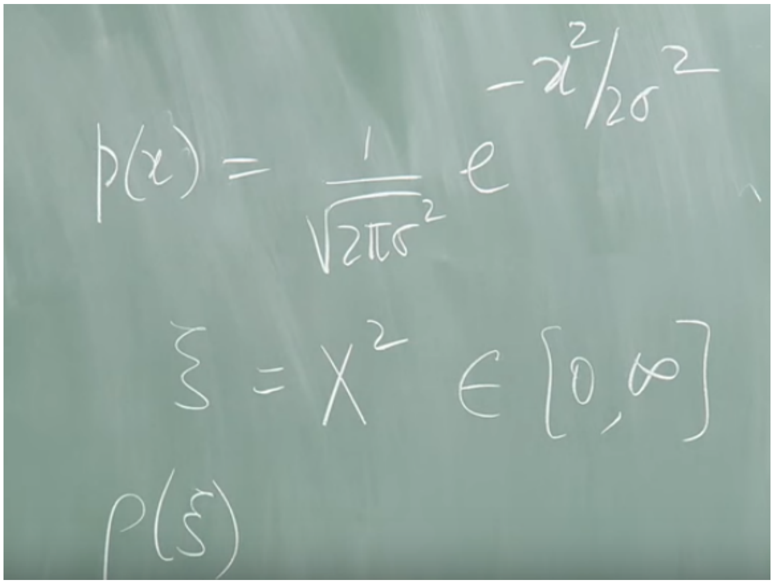
Is it possible to have such distribution? The answer is under fairly general conditions, no. Either in principle all cumulants exist barring accidents in certain cases or the Gaussian says it stops quadratic and that is it nothing more. There are other such properties which will also emerge as we will see when we talk about stable distributions.

We will see there are certain other interesting properties of this kind which will emerge that either it stops at the second order it goes on forever. They only are the 2 possibilities, we will see where this comes about okay. So the next step now is to ask I have some information about the Gaussian. Are there other such distributions, there are several but we will talk about it when we come to stable distributions.

But first we would like to do the following. I would like to take a set of ordinary, very simple distributions for random variables, add them all up and see where it goes, what the distribution of the sum looks like. In particular we will undertake a simple exercise. We are going to take a whole n random variables all uniformly distributed between 0 and 1.

So this random variable, each of the random variables takes values between 0 and 1 with a constant probability distribution function 1 and add up all these fellows and ask what is the distribution of the sum of this thing. So we will work that out explicitly here. Meanwhile, one final point. If you have functions of random variables their probability density functions can look very different from the distribution density functions for the original random variable. If you look at the Gaussian example for instance let us take a Gaussian with 0 mean.

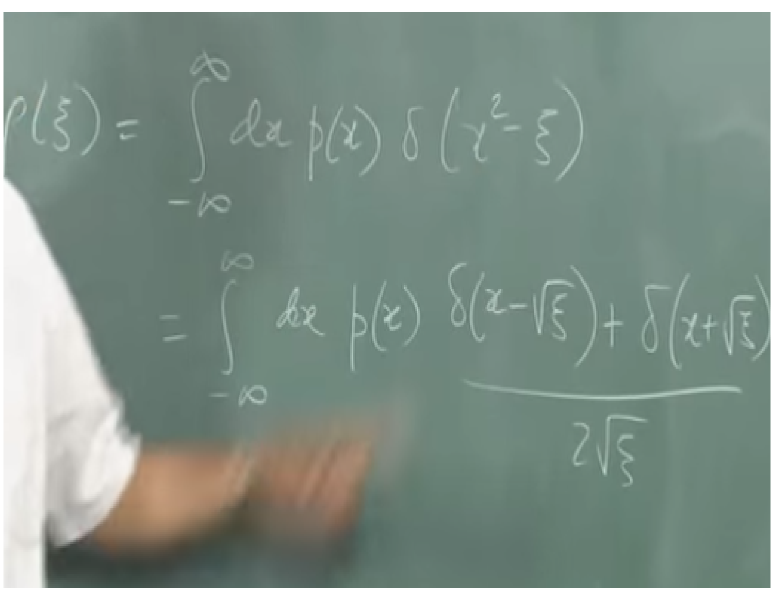
(Refer Slide Time: 50:26)



So you have a p(x). This guy is 1 over root 2 pi sigma square e to the minus x square over 2 sigma square, the Gaussian with 0 mean. For simplicity let us set the mean to be equal to 0 right and then I ask what I i the density function probability density function of a variable let us call it psi which is equal to the square of this variable X square okay and let us call it PDF rho of psi.

Now it immediately, it is obvious that this psi is an element of 0 infinity unlike the original random variable which ran minus infinity to infinity now we got 0 to infinity. What is the distribution PDF of psi going to be like? Several ways of doing this. One of them is to say alright I do it by brute force.

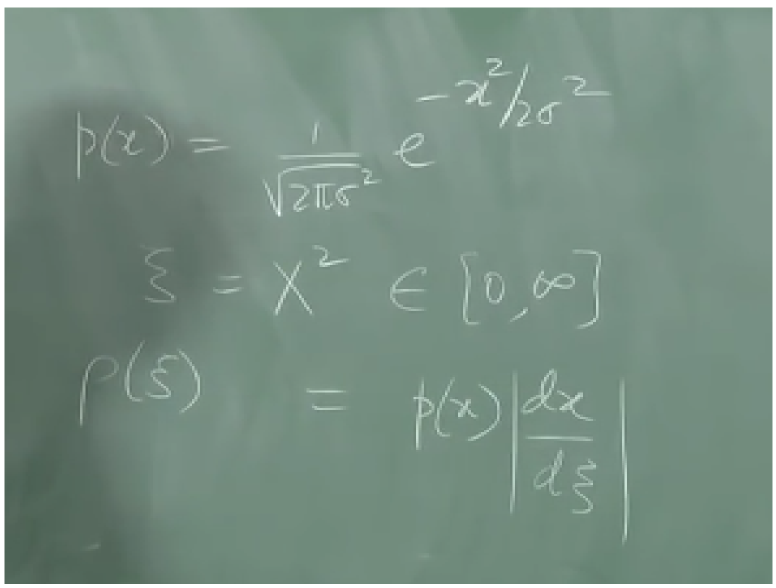
(Refer Slide Time: 51:40)



I do it by saying that rho psi must be equal to an integral minus infinity to infinity dx p of x and then a delta function which says psi - x square, x square is psi okay. But to do this integral, I got to convert this delta function over psi to 1 over x right and what's the first property of the delta function it is a symmetric function so I can write this as x square - psi in this fashion and I got the delta of x square minus constant square psi square root of psi whole square.

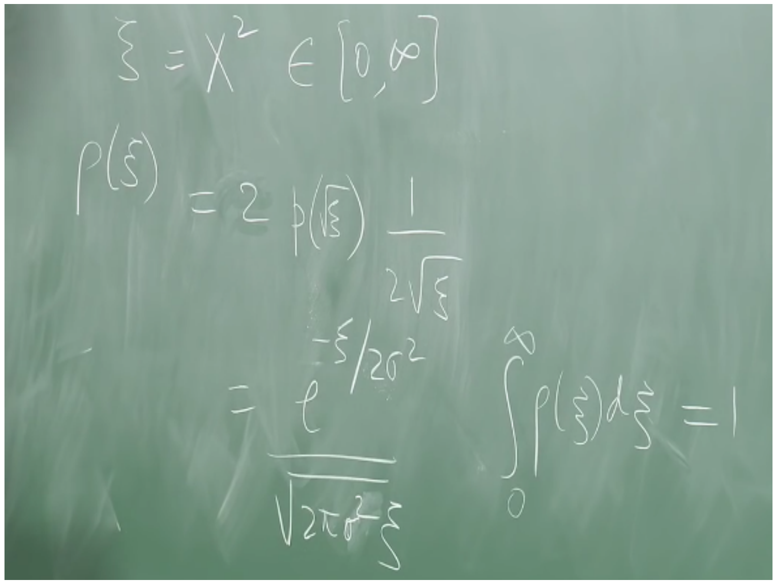
So I can write this as integral minus infinity to infinity dx p(x) delta of x - root psi + delta of x + root psi/2 root psi. That the Jacobian derivative and I do this integral. I can now do this integral because I use the delta function and plug it in. But you can also write the answer down. You see if I say that when X takes a value between x and x plus dx capital X takes a value between little x and x + dx suppose psi takes a value between psi and psi + d psi right.

(Refer Slide Time: 53:06)



Then rho of psi d psi must be equal to p of x dx. In this case we are fortunate because as X increases psi also increases. Otherwise, when you are talking about probabilities you have to make sure that they are both positive on both sides. So I can write this as dx over d psi but I must be careful to write that modulus sign okay and I must express this thing in terms of psi because that is what a function of.

(Refer Slide Time: 53:43)



So I must write this as p(x) is square root of psi um and d psi over d x is 1 over twice the square root of psi. There is no need for modulus because psi is not negative. So would that be the right answer? No because minus x contributes exactly the same amount to this right. So there is a factor 2 and that is equal to e to the - psi over 2 sigma square over root 2 pi sigma square psi. So this 2 cancels that 2 and then we get this.

Of course you got to check normalization so you got to verify that rho of psi d psi = 1 integrated from where to where, 0 to infinity. So it is an exponential but it is also got this factor sitting here. So this distribution looks rather different from what the original variable was. For instance if you wrote down in one dimensional motion if you wrote down the Maxwellian distribution of velocities, velocity component it is going to be e to the - mv square over 2kt or something like that, the Gaussian.

On the other hand if you asked what is it for the energy which is half mv square then it is going to be proportional to e to the minus the energy with some kt factor/square root of the energy in the denominator. So this factor which came from the derivative, the Jacobian sitting there too, this factor here is crucial. **"Professor - student conversation starts"** What do you call this in that example which I just talked about in the energy? The density of states. You call it the density of states **"Professor - student conversation ends"**.

It is precisely the density of states in one dimension for one dimensional motion okay. So we will see where that is when we talk a little bit about the canonical ensemble you will see that this density of states plays a crucial role and we will talk about that when I mention the characteristic function for this distribution okay. Alright, so the next exercise is to take a set of identical random variables and add them up and see if the how the Gaussian emerges magically. I will do that next time.