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

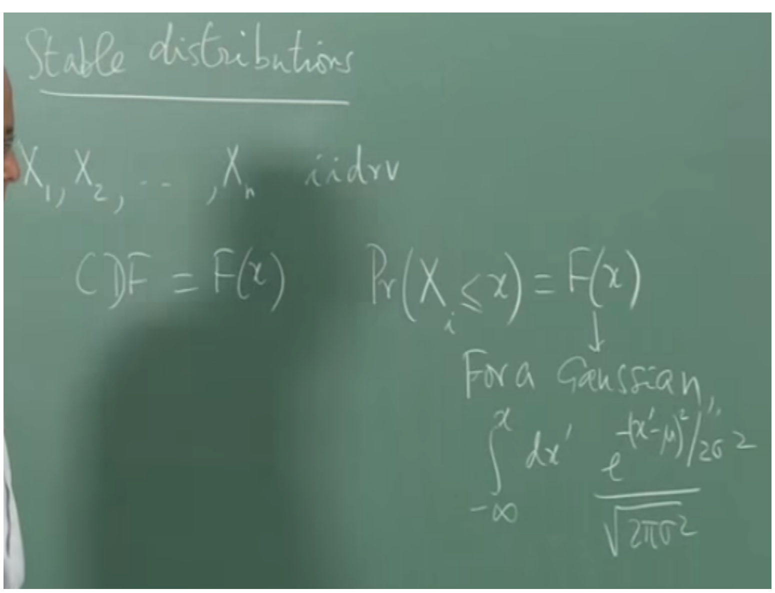
Lecture - 05
Stable distributions

Okay, we saw last time that the problem of random walks, random flights led very naturally to a Gaussian distribution for the end to end distance with displacement and this looked like it was part of a very general result namely you added up a whole lot of identically distributed random variables and you got a Gaussian in a certain limit.

This is not an accident, it is actually part of the central limit theorem which as I stated last time essentially says that if you have an identically distributed random variables then the linear combination of these random variables suitably rescaled and shifted will in the limit as n goes to infinity end up with Gaussian distribution provided each of the random variables has a finite variance.

This was the sum and substance of the central limit theorem. Now this is part of a more general class of distributions called stable distributions and I would like to talk about stable distributions to start with.

(Refer Slide Time: 01:34)



And I will try to explain at least qualitatively what this stability refers to, what exactly it implies okay. So we start by asking suppose I have a set of identically distributed independent random variables and let us call these random variables X 1, X 2 to X n and let us suppose these are iidr independent random variables and let us suppose that the cumulative distribution function of the set of each of these variables is some f of fx okay.

So the distribution function CDF equal to some F(x). What this implies is that the probability that any given random variable X i less than equal to x this thing equal to F(x) okay and then we ask the following question okay. Is there any special form or forms of F of x this distribution function such that if I add up a whole lot of these random variables iidrv's and rescale them in some suitable fashion the distribution function for the sum the resultant remains F of x, does not change at all.

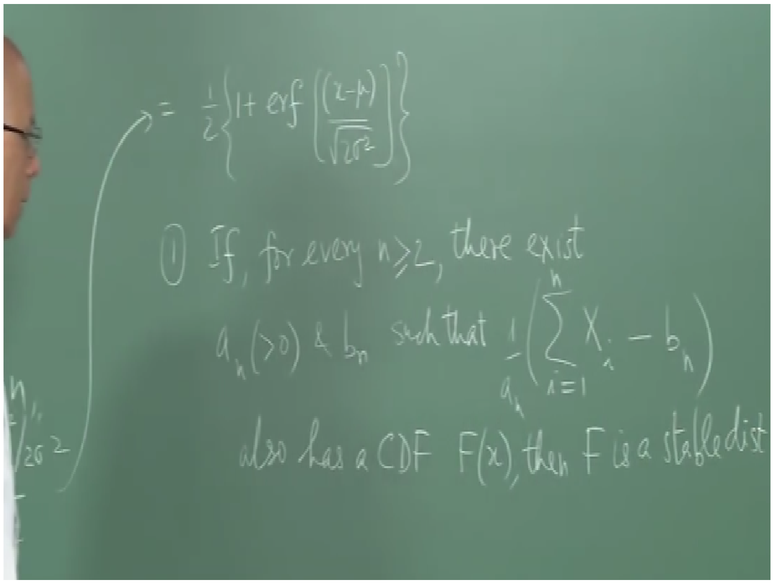
If you can do that for every n greater than equal to 2 then this F of X is said to be a stable distribution okay. So now let us formalize this definition and write it in formal terms. There are several equivalent ways of defining a stable distribution but I am going to quote a couple of them and not try to prove the equivalence of these definitions but that will become intuitively clear what we mean as we see the explicit forms possible for this F of x okay.

Just to recall to you what this Fof X is for a Gaussian distribution for instance for a Gaussian, for a Gaussian this F of x recall is integral from minus infinity up to x dx prime e to the - x - x prime - mu square over 2 sigma square over root 2 pi sigma square and that as we know is an error function this thing here. So it is minus infinity to x. I can write it as minus infinity to 0 and then 0 to X.

So this quantity will turn out to become equal to 0 to minus infinity to 0 is half the Gaussian after you shift to the origin here to x prime minus mu you set that equal to some other variable so this is a half and then there is a 1 plus an error function of we shifted the variable and therefore it is a function of x – mu divided by we scaled it with root 2 sigma square. So this is what the cumulative distribution function for a Gaussian looks like and so on.

So for each of these cases you can write down the cumulative distribution function and non-decreasing function of X and then we ask under what conditions is this F of x stable. So here is definition 1.

(Refer Slide Time: 05:23)



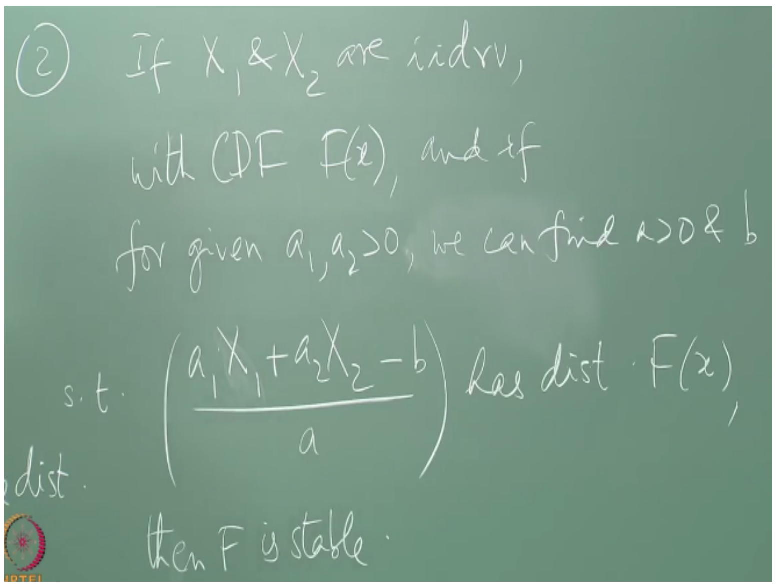
One way to do this is to say that if for every n greater than equal to 2 there exists a constant a n which is positive and b n which is just real such that this combination summation X(i), i = 1 to n that is the sum of these identically distributed random variables shifted by some amount which is an independent and then rescaled with a 1 one over a. That is a random variable 2.

If this random variable can be shown to have the same cumulative distribution function as each of the components except i then I say that f is a stable distribution okay. So that is a precise definition. You are still left with the task of finding out if this is going to work or not for a given F of x you have to find out if you can find a suitable constant a sub n and b n for each n greater than equal to 2 and if that is possible then it is a stable distribution okay.

We will see examples of we will write down all the stable distributions in some sense but we will see where this gets us. That is the first definition. It is in fact what I have said here in words. **"Professor - student conversation starts"** Ya. This is a very strong criterion because you could have n greater than equal to m where m is some finite number so. Yes. "**Professor - student conversation ends"**.

So we will talk about divisibility and so on but this is the requirement that this is should be true for every n. Should be able to do this. Then and only then is it a stable distribution. So this is a necessary and sufficient condition that this be true but is operationally not very useful as you can see although it is a formal definition not telling us how to go about finding such a stable distribution okay.

(Refer Slide Time: 08:02)

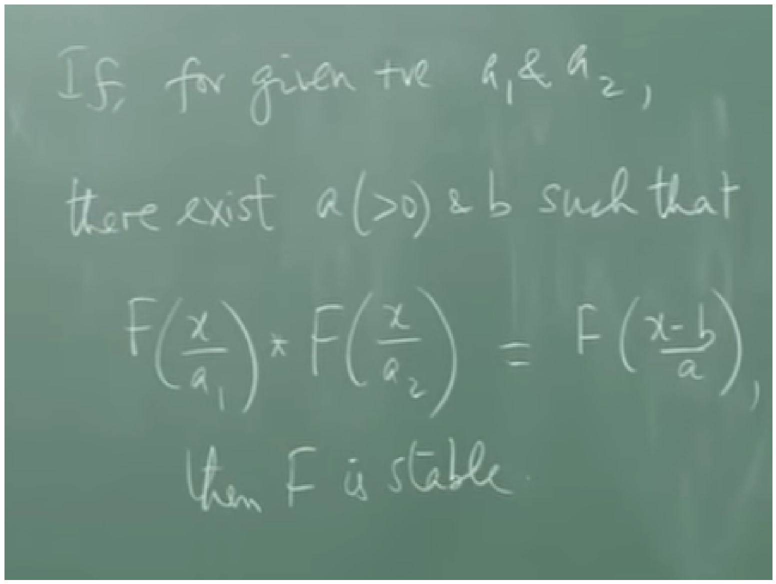


Here is the second definition which is equivalent to the first. So if X 1 and X 2 just two of them are independent identically distributed random variables with CDF F of x and if the following random variables for given any given positive a 1, a 2 greater than 0, the random variable a 1 X 1 + a 2 X 2 minus some constant b divided by the constant a greater than 0. If in fact and if for a given this thing this random we can find a greater than 0 and b such that has a distribution F then F is stable okay.

So this says okay forget about adding n of these guys trying to find out something for all n and so on just take two of them and so on and if for any given positive constants a 1 and a 2 for every set of given positive constants you can find the positive constant a and another constant b real constant such that this combination this linear combination subtracted out suitably and rescaled by a if that is got the same distribution function F of x then F of x is stable.

This two is a necessary and sufficient condition okay and with a little work one can show that these are equivalent definitions here. But you see again neither of these things is saying anything about F itself. It is saying take this random variable or that random variable and test what its distribution function is and so on. We need the condition which says something about the distribution F itself and that is the third definition and that goes as follows.

(Refer Slide Time: 10:28)



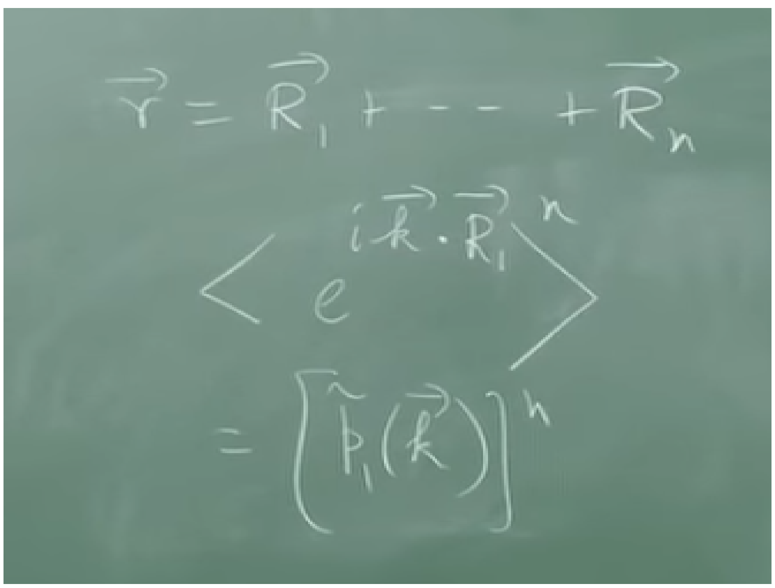
It says if for given positive a 1 and a 2 there exist a greater than 0 and b such that the convolution of F of x over a 1 and x over a 2 if the convolution of these two distribution functions is equal to F(x) - b over a then F is stable okay and now we are getting somewhere because this is now directly a condition on the distribution function itself and what does it say?

It says you scale one out, you scale the other out and then you suitably and then you have a convolution and for any given positive a 1 a 2 if you can find a subtraction constant and a rescaling constant such that this is true then F is a stable distribution okay. Now in every one of these definitions if you leave out this shifting, if this is not there if you do not need this, this constant or that constant if this is a 0 or the b is a 0 here then you say the distribution is strictly stable. Otherwise you say this is a stable distribution okay.

So strictly stable distribution is a special case of a more general definition of a stable distribution okay. Now this definition immediately suggests to us the following. It says if these two things are in convolution it means in some sense that the fourier transforms would multiply and it is immediately telling us that this fourier transform has a certain factorization property and only then would this be possible at all which sort of tells you in some sense when will it have this factorization property.

If you go back to definition 1 you need to add n of these fellows so we need a characteristic function which is the fourier transform of probability density function which should in some sense factorize which means it must be exponential in some form because what you want is the expectation value of e to the ikx. That is the characteristic function and if x is the sum of terms these exponents would multiply each other if they are independently distributed right. We saw that already working for the random walk because recall that in the random walk problem although I did not write that down explicitly this is really what it meant.

(Refer Slide Time: 13:32)

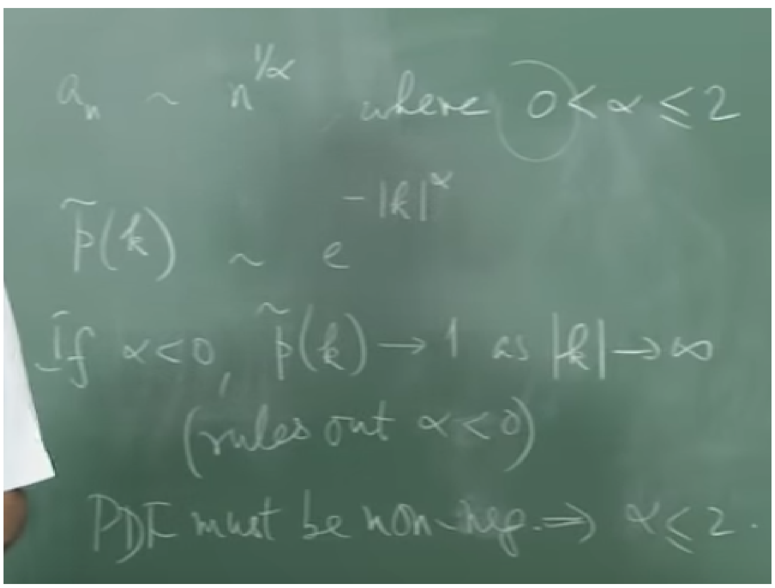


There we started with a vector r which was the sum of R 1 plus dot, dot, dot, up to R n in this fashion and then I said the characteristic function of this random variable here was e to ik dot R or whatever it is so this is the expectation value of e to the i k dot summation 1 to n R i and the next step was to write this out as an exponent here because it is a product as soon as you write it out. So this is equal to k dot R 1 e to the ik dot R 2 all the way up to e to the i k dot R n and then came the crucial observation, crucial observation that these are all independent steps and therefore the expectation value of the product is the product of expectation values.

So it immediately became e to the ik dot R 1 any one of these guys to the power n and this was the one step, this is just the fourier transform of the one step random walk which was p 1 of R p 1 tilde of k in this case and then it became raised to the power n and if you recall this was sin k lower kl to the power n and then there was all these integration variables etc. So this suggests to us that that is probably happening in general for a stable distribution and indeed it is so.

It will turn out that all the stable distributions can be classified completely and they are classified with the help of 4 parameters. The most general stable distribution is labeled by 4 parameters. I am not going to write the general form down. Text on statistics will tell you what the most general form of the distribution of the cumulative distribution function is for a stable distribution but what we need to understand is the following.

(Refer Slide Time: 15:49)



It turns out that this coefficient a sub n that we are talking about in the summation, in the first definition this coefficient a sub n in this definition here this thing here must necessarily be of the form n to power 1 over alpha where alpha is a positive constant okay where 0 less than alpha less than equal to 2 it turns out. I will explain why it is restricted by this range 0 to 2. It turns out also that all the stable distributions are unimodal distributions.

There is a single peak for every one of them and they are labeled primarily by this index alpha the exponent or index alpha okay and it is unfortunately true that you cannot write down an explicit expression in general for the probability density function for a stable distribution. But because we see that the characteristic functions must in some sense be multiplicative exponentials which get multiplied to each other it turns out that the characteristic function p tilde of k must be of the form apart from phase factors it must be of the form e to the minus some constant times k to the power alpha.

So I am merely stating these results, I am not proving this. I am merely stating these results and you can see that for a Gaussian this was e to the minus k square apart from a phase factor, we write that down explicitly. But this it went like e to the minus k square. For a Cauchy distribution it went like e to the minus mod k to the power 1 and so on. So they it looks like those guys are going to become stable distributions okay.

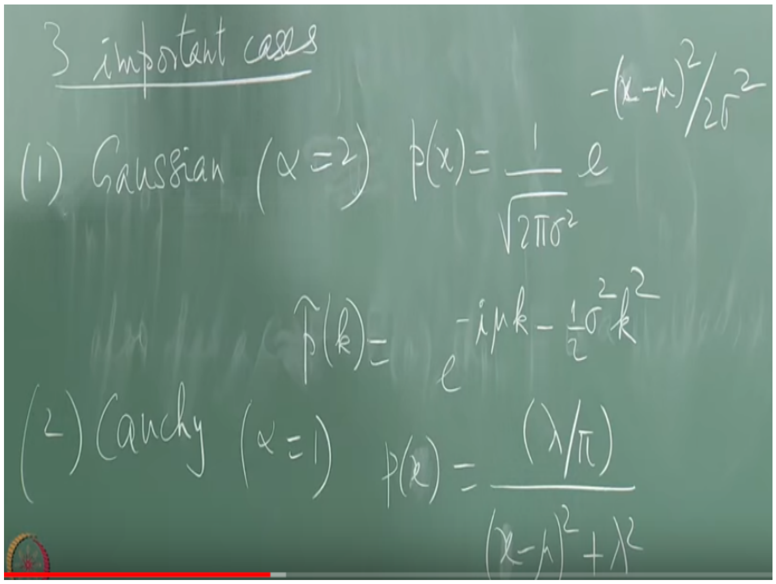
Now the restriction here is sort of understood in the following way, at least heuristically it will be the following. Suppose alpha were negative then this is e to the minus 1 over mod k to some positive power and as mod k tends to infinity that will tend to unity because it goes to e to the 0. So this means if alpha less than 0 p tilde of k will go to 1 as mod k tends to infinity plus or minus infinity. That cannot be integrated.

So you cannot find a fourier transform which will give you the probability density function, normalizable density function. So it is easy to understand why this restriction appears okay. That is immediate from this. On the other hand if alpha is greater than 2 then it is a little more subtle to show why this cannot be a characteristic function because it turns out the inverse fourier transform cannot be shown to be nonnegative.

On the other hand you know that the PDF p(x) must be nonnegative as a probability density function. So that is what puts the restriction on this side out here. So rules out, this rules out alpha less than 0. PDF must be nonnegative. Implies alpha less than equal to 2 and that is harder to prove. I have not come anywhere near proving it but this is a statement that you have to take as on faith that if alpha is greater than 2 you cannot establish the non-negativity of the fourier transform, inverse fourier transform.

So the stable distributions are characterized by this index here and the actual formal name for these stable distributions is they are actually called Levy skew alpha-stable distributions.

(Refer Slide Time: 20:02)



And for short I will just call it stable distributions. There is a little bit of confusion in terminology here because it turns out that one of the stable distributions is called the Levy distribution and it is not the general family that is being referred to here. So we will just call these stable distributions, nothing more than that. Now what are the other properties of these distributions. Well, the Gaussian is certainly a stable distribution as we will see and the most famous cases are the following 3 main cases, 3 important; they are the ones that occur in practice very often, especially the Gaussian.

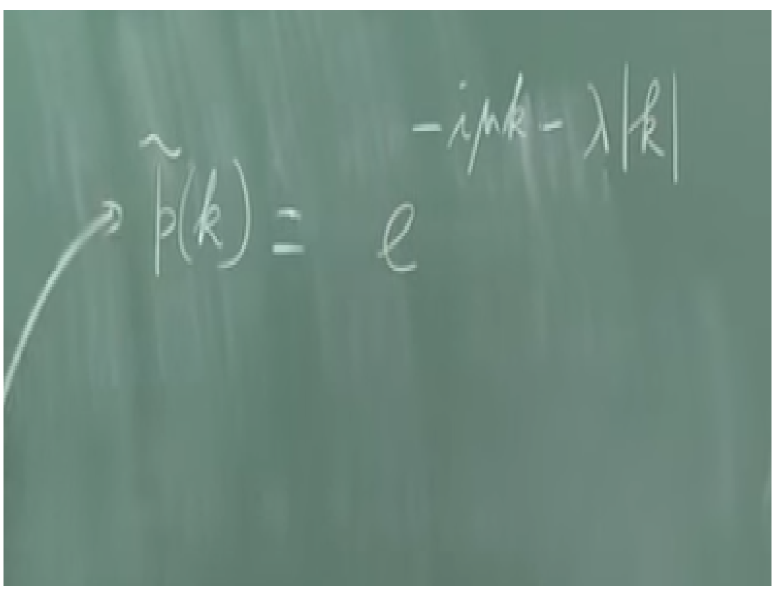
The first of these is a Gaussian. One is the Gaussian and this is alpha = 2 and we know what the density function looks like p(x). This is equal to 1 over root 2 pi sigma square e to the - x - mu square over 2 sigma square and we know the characteristic function too p tilde of k. Oh, incidentally if the distribution has a density p(x), this is the fourier transform so p tilde of 0 must be equal to 1 because that is the integral of p(x) for - infinity to infinity.

So this fellow here is e to the minus i mu k minus one half sigma square k square. Remember that the moment generating function was just a, the cumulant generating function was just a quadratic of this kind. It was mu times u plus half sigma square u square and the characteristic function is the moment generating function at the value u = - ik, so it is this; p tilde of 0 is 0 is 1 as you can see.

So it is normalized correctly and that is the Gaussian expression. What is the variance of the Gaussian, sigma square is the variance, a finite variance okay. The second case that is very important is the Cauchy distribution and in this case alpha equal to 1. Actually the general Cauchy distribution need not be symmetric between about the mean value. It is general skew but we are looking at a special case where certain other parameters other than alpha the other 3 parameters have been set equal to special values.

And the most common form of this is when p(x) equal to some lambda over x - mu whole square + lambda square is lambda over pi that is the normalization constant. This is x, curly x. Now what is p tilde of k in this case? It is got to be a proportional to ke to the power minus mod k because remember I said that the Cauchy distribution corresponds to alpha = 1.

(Refer Slide Time: 24:01)



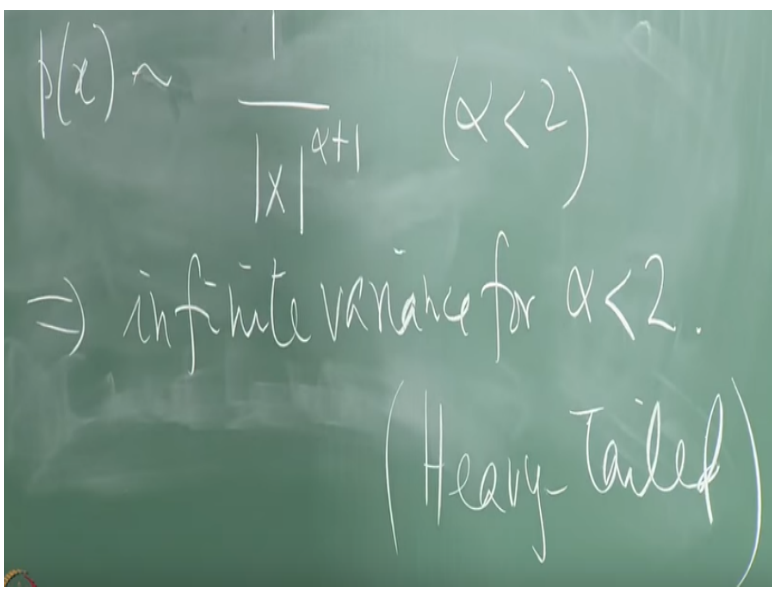
So for the Cauchy, the symmetric Cauchy distribution this guy here p(k) = e to the - I mu k - lambda mod k. Again p tilde of 0 is 1. It is normalized and it is got exponent alpha equal to 1.

The mean value is mu and that is the peak of the distribution. It is unimodal. So is this unimodal peaked about mu. What is the variance of this distribution? What do you think the variance goes like?

Well you got to multiply this by x square and integrate minus infinity to infinity and the denominator goes like x square. So it diverges. Yes, the variance is infinite. For this the variance is infinite. What is the mean value? Mu, but barely so because if you (()) (25:23) power counting you put an x here and you integrate it then the denominator goes like x square so the whole integrand goes like 1 over x which will logarithmically diverge but because it is symmetric about that midpoint if you shift to x - mu the answer turns out to be 0 the mean value. So it gives you a finite mu. But it is barely so.

The variance is certainly infinite for this distribution okay. Notice that this distribution has a tail, this guy has a tail that for large values of mod x it goes like 1 over x square. Unlike this which has an exponential e to the minus x square that goes to 0 faster than any power, any negative power of x, plus minus infinity and this is going to be a general feature. This is a general feature.

(Refer Slide Time: 26:14)

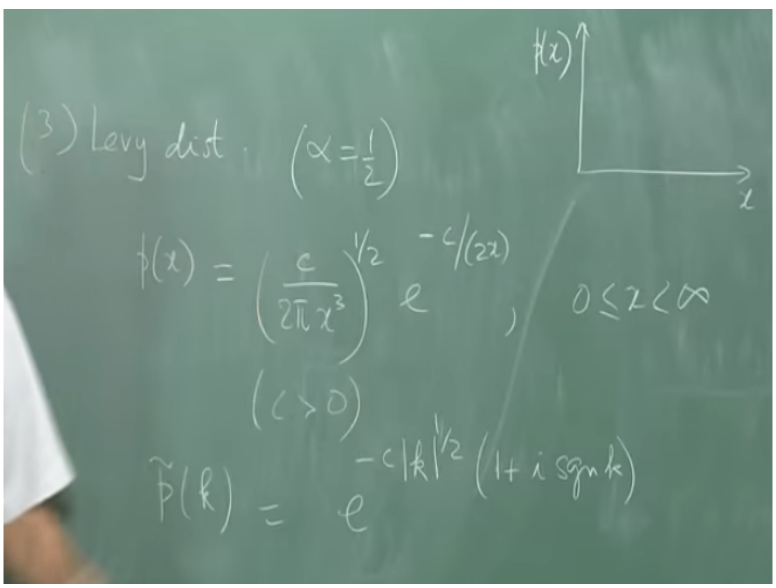


Turns out that as soon as you have this property here p(x) will turn out to go asymptotically namely as mod x tends to plus infinity. It will go asymptotically like 1 over mod x to the power alpha + 1 for alpha less than 2 and indeed when alpha is equal to 0 you see it goes like 1 over x square out there. And what does this imply? If alpha is less than 2 and the denominator goes like 1 over mod x to the power alpha + 1 it implies infinite variance.

So it says the entire family of stable distributions except for the Gaussian all of them have a huge amount of scatter. The variance is formally infinite and the Gaussian is the only stable distribution with a finite variance okay. In fact this also tells you that if alpha is less than 1 between 0 and 1 even the mean value is infinite. Even the first moment does not exist for those distributions but certainly the variance is finite only for a Gaussian.

This is a very crucial observation and all these fellows are called heavy tailed distributions. Essentially it says large values of this of these random variables are possible and have a probability mass which is significant unlike the Gaussian where it just gets cut off faster than any inverse power of x okay. That is a crucial observation. The third of these we will come back to this.

(Refer Slide Time: 28:28)



The third of these special cases is this. It is called the Levy distribution and it corresponds to alpha = 1/2 and it looks like this. The distribution p(x) is characterized by a constant c. So c over 2 pi x cubed to the power half e to the - c over 2 x but here 0 less than equal to x less than infinite. I shifted the, it is a semi-infinite random variable, the semi-infinite range for the random variable and I have shifted that, the beginning of that range to 0 in suitable rescaling by a translation.

So there is a constant c positive and it is not hard to check that this is normalized to unity. You could ask what is the characteristic function here turns out p tilde of k not surprisingly is e to the minus c modulus k to the power half as promised. That is what it should be and it is multiplied by a phase factor. In this case it is 1 + i times the sign of k and it is called the levy distribution. What does it look like? What does the shape of this fellow look like?

(Refer Slide Time: 30:00)

Well, for the Lorentzian and Gaussian we have seen what the shape looks like. For this thing here, here is x. Here is p(x). As x tends to infinity, positive infinity, this factor tends to unity. This goes like 1 over x to the 3 halves in the denominator. That itself tells you that the variance has got to be infinite because the denominator goes like x to the 3 half alpha + 1. So it is 1 over 3 to the 3 halves and what does it look like near the origin?

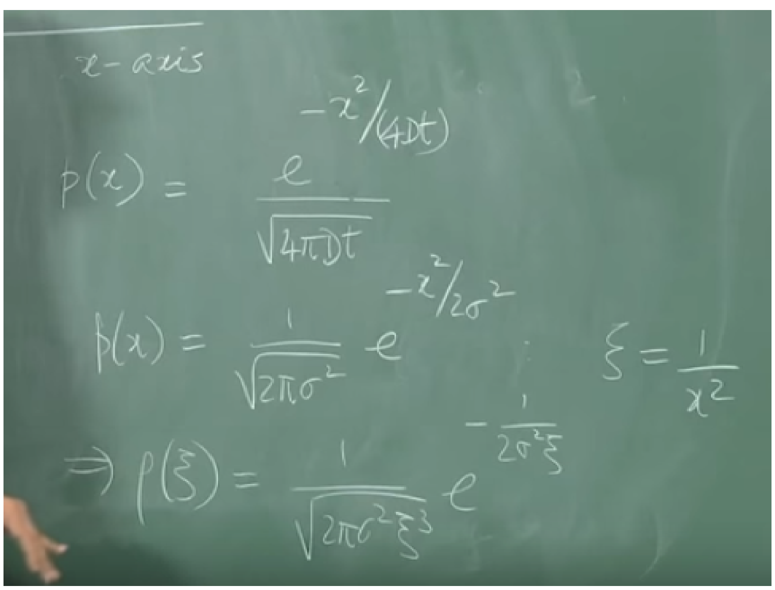
What is it going to do near x = 0? Is it 0 or infinite or finite? It is 0. It is dead 0 because this factor in the denominator is swamped by the exponential factor e to the - 1 over something which goes to 0 is going to go very rapidly to 0. So this function not only is it 0 at the origin but all its derivatives are also 0 at the origin, all its derivatives of finite order. So it looks this is the 1 over x to the 3 halves decay out here and the peak is characterized by the scale c okay.

And you could ask where do these distributions appear, where do they occur? Well, it turns out there is a very close connection between different stable distributions in a very specific sense. Oh, by the way let me before I go on mention that although I have written down explicit forms for the probability density function for these 3 special cases this is not in general possible for generic alpha between 0 and 2.

In fact turns out that you cannot write this p(x) in terms of elementary functions other than these cases, these 3 cases. You can write p(x) in terms of hypergeometric function for rational values of alpha and so on like 3 halves etc. But in general all you can do is to write down specific forms for the characteristic function for its fourier transform. But already that gives us all the information we need about these distributions.

There are examples, some physical examples of when this is going to happen; when these distributions are going to appear. For the Gaussian of course we see it appears everywhere. So let us go back to our same expression of random flights or diffusion or something like that.

(Refer Slide Time: 32:39)



When you have a particle diffusing on a line and we will do this in some detail later on. Along the x axis, if you have a particle freely diffusing in the x axis then its probability density function if it starts from the origin at equal to zero is of the form e to the minus x square over 4Dt where

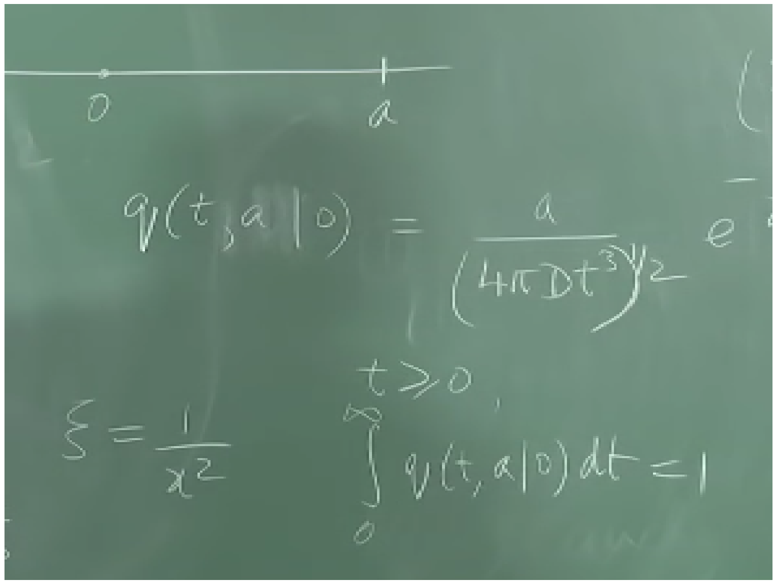
D is called the diffusion constant divided by square root 4 pi Dt. That is a Gaussian with 2 sigma square = 4Dt or sigma square the variance is 2Dt.

So it says the variance of this particle increases as time goes linearly with time. So that is a Gaussian distribution. But now if you ask what is the distribution of 1 over x square that turns out to be a Levy distribution because it is not hard to see that if you have p(x), let us write a normal Gaussian down equal to 1 over root 2 pi sigma square e to the minus x square over 2 sigma square say Gaussian is centered at the origin and ask what is the probability density function of the random variable psi which is 1 over x square that has a Levy distribution okay.

In fact the density function for rho of psi this will imply is 1 over root 2 pi sigma square psi cube e to the minus 1 over 2 sigma square psi. So the constant c is 1 over 2 sigma square, square root, 1 over 2 sigma square, c over 2 whatever it is. So this is precisely a Levy distribution with exponent half here. But that is if I took this random variable and I gave the example of the Maxwell distribution of velocities where I said the energy has a very strange distribution 1 over square root of epsilon e to the minus epsilon exponential.

That was not a Levy distribution but here we are asking for the distribution of 1 over x square and then it has this right here. Well in connection with the diffusion problem itself there is another random variable which has precisely this kind of distribution.

(Refer Slide Time: 35:09)



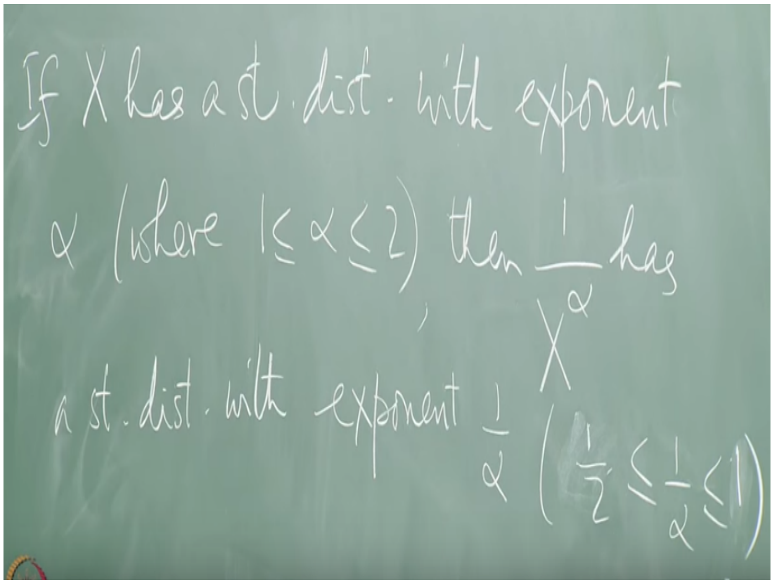
For instance if you ask alright I start with diffusing particle on the x axis. I start at 0 and I ask as it moves about what is the first what is the distribution of the time where it first hits the point x, some given point x, some given point x. Now let us call this just to be not to confuse it with that random variable x, let us call this a and ask here is this particle diffusing on the x axis starting at x = 0, at equal to 0.

And I ask what is the probability that between time t and t plus Dt, the particle crosses this point a for the first time because it is doing a zigzag motion for the first time and that is a distribution. The random variable here is a time and if I call that q to cross the point a at time t having started at 0 and you want to I want to cross t at the point a at time t having started at the point 0, this quantity here.

This is equal to it turns out 1 over it turns out to be a over 4 pi Dt cubed to the power 3 halves e to the minus a square over 4Dt okay and that is the distribution in time so t greater than equal to 0 and integral q of t, a 0 dt, 0 to infinity = 1. That we know because this is a Levy distribution which is normalized to unity already. It is called the first passage time distribution okay and it is precisely a Levy distribution.

So that is the simplest physical example I know of where Levy distribution appears. Ya. To the power ha sorry this is the power half. I already put a t cubed in here so quite right. It is a half, yes. So it is t to the 3 halves in the denominator here okay. In general there is a connection between a random variable which has a stable distribution with index alpha where alpha is between 1 and 2 and a random variable which is a function of this original random variable has a stable distribution with index 1 over alpha.

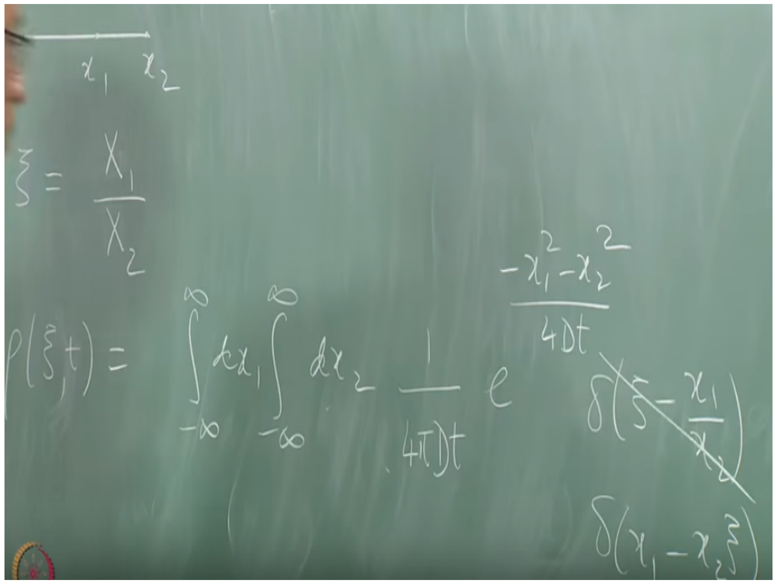
(Refer Slide Time: 38:05)



So if for instance , if X has a stable distribution with index with exponent alpha where 1 is less than equal to alpha less than equal to 2 then 1 over X to the alpha has a stable distribution and remember that this 1 over alpha therefore is half less than equal to 1 half less than equal to 1. So the new exponent is between half and that is what we used there when I said that a Gaussian which has exponent alpha a 2 = 2, 1 over the Gaussian square 1 over x to the alpha has a Levy distribution with exponent half okay.

Similarly, you could ask does the Cauchy distribution appear in a natural way in the diffusion problem. Notice that everything with alpha less than 2 has no variance. They are all heavy tailed. No variance at all. What about diffusion problem in which the Cauchy distribution appears naturally. There are lots of places where the Cauchy distribution it is called a Lorentzian in physics appears naturally but here is a very simple instance. Again let us go back to the distribution problem and look if it is a physical problem and look at a very simple function of a random variable.

(Refer Slide Time: 40:18)

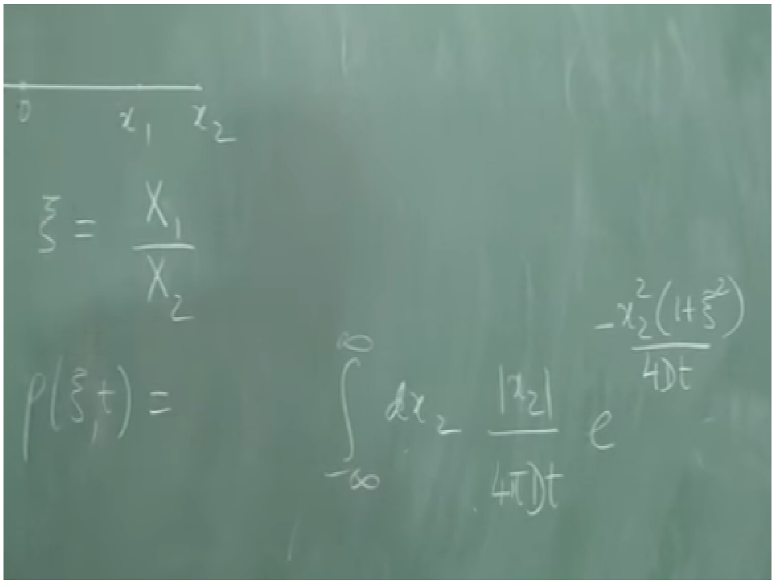


So suppose you have 2 particles, both of it start at the origin and diffuse on the x axis such that the coordinate of 1 at any instant of time is x 1 and the other one is x 2 okay and you look for the random variable psi equal to X 1 over X 2 and ask what its distribution is okay where each of these has a probability density function given by the solution of the diffusion equation right. So it says rho of psi therefore has a function of time.

This is equal to an integral minus infinity to infinity dx 1 minus infinity to infinity dx 2 and let us suppose for simplicity they have the same diffusion coefficient that need not be the case but. Then there is a 1 over 4 pi Dt and then e to the minus x 1 square minus x 2 square over 4 Dt and then a delta function of psi minus x 1 over x 2. That is the normalized density function for this whole psi. Now what is the physical range of psi?

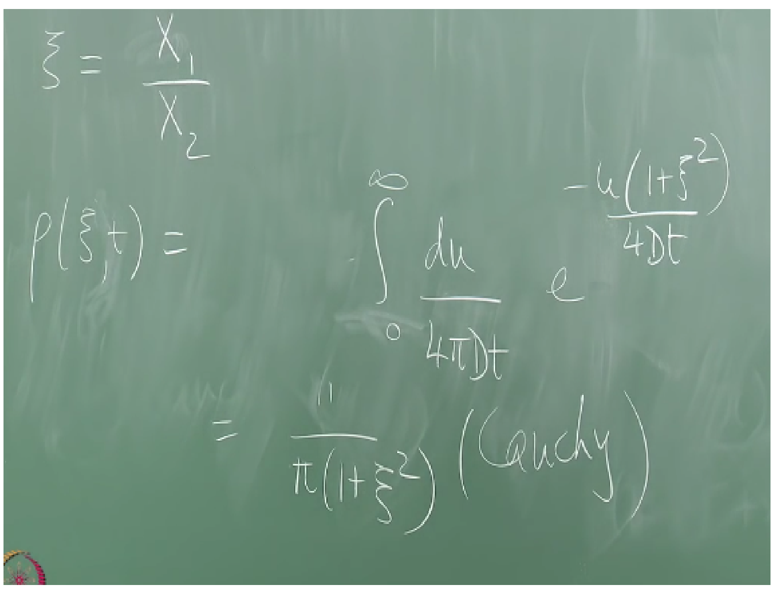
Each of x 1 and x 2 runs from 0 to minus infinity to infinity. So what is the range of psi? Again, minus infinity to infinity right. So in that sense we are spared putting extra conditions and all we have to do is to do this integral out here. Now the obvious way to do this is, is to write this is as x 2 times psi and get rid of the x 1 integral right. So let me write this as the delta function of x 1 minus x 2 psi and I have to remove this factor 1 over x 2 from there and take its modulus.

(Refer Slide Time: 42:22)



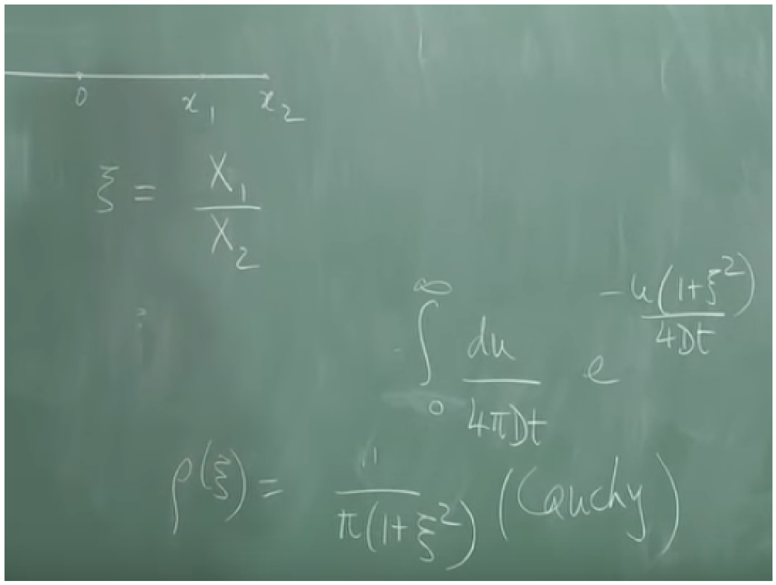
So this becomes mod x 2 times this and then all I have to do is to replace x 1 by x 2 times psi. So this becomes e to the minus x 2 square into 1 plus psi square over 4 Dt and all this goes away, the x 1 integration goes away and I have this. So this is straightforward to do. All I have to do is to write this is as twice 0 to infinity and get rid of the modulus. It is an even function now okay. But 2 x 2 d x 2 is d of x 2 square.

(Refer Slide Time: 43:17)



So I change variables to x 2 square and this goes, this goes and this becomes some du over 4 pi Dt e to the minus u times this fellow here. So that is a trivial integral and du times e to the times e to the minus au is just 1 over a and if a is positive right. So that will kill this 4 Dt and give you 1 over pi 1 plus psi square and that is a Cauchy distribution okay about mean value mu = 0 and this lambda parameter set equal to 1 in this case. What is interesting about this? I said this is distribution at time t so what really happened here to t. It disappeared. It completely disappeared. So this is true at all times.

(Refer Slide Time: 44:32)



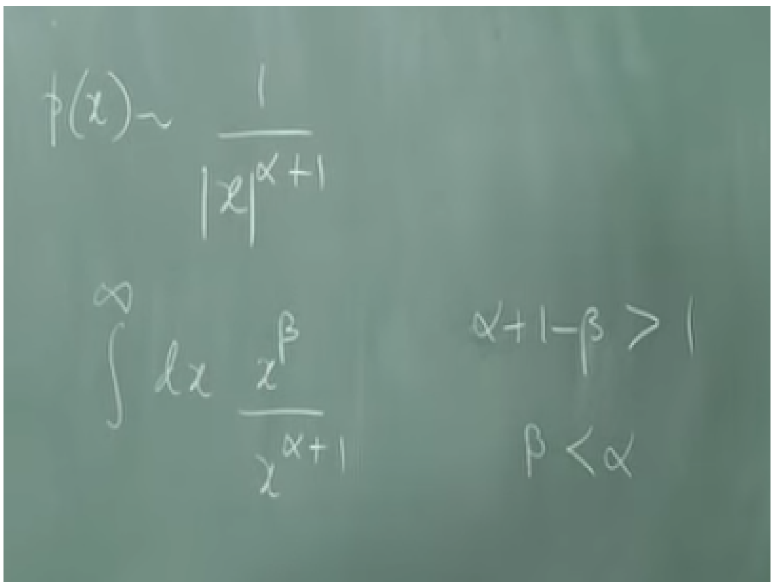
It is really true at any instant of time. So the ratio of the coordinates for a diffusion for 2 diffusing particles ratio of their coordinates is actually has a distribution independent of time and it is a Cauchy distribution okay. What would have happened if I had D 1 and a D 2? So I will leave that to you as an exercise. If the first particle has a diffusion coefficient D 1 and the second one has a D 2 then show that the result is still a Cauchy distribution except this parameter here I mean there would be a D 1 over D 2 sitting here.

There will be a lambda parameter which is equal to 1 in this special case okay. So this is one more place where the Cauchy distribution appears naturally in a lots and lots of such examples. Now we will say a little more about this Levy distribution and these long tail distributions a little later when we do anomalous diffusion when we talk about anomalous transport okay.

But the take home lesson is that you have this family of very special distributions called stable distributions and they are characterized primarily by this exponent alpha, alpha is positive, runs up to 2; 2 is the extreme case of the Gaussian, which is a very respectable distribution. It is got moments of all orders including a finite variance and all the others are heavy tailed and they do not have variances okay.

On the other hand you could ask is there a central limit theorem for them because we already said there is a central limit theorem for the Gaussian so is there a generalized central limit theorem for all the stable distributions. The answer is yes. So if you started with identically distributed variables and you said that they did not have variances.

(Refer Slide Time: 46:52)



But for instance if you if you have a p(x) going like 1 over x to the power alpha + 1 and you ask the variance does not exist because alpha is less than 2. On the other hand the sum, what is the maximum moment that exist for this distribution. So you could ask a thing like what does what kind of dx if I say x to some beta/x to the power alpha + 1 at infinity, when would this exist and I put a p(x), p(x) has a tail which goes like this.

So if I put beta = 2, I am in trouble if alpha is less than 2 but what is the maximum value of beta that you can have for which this converges. So it is clear that this denominator must go to 0, the whole thing must go to 0 faster than 1 over x. So you must have alpha plus 1 minus beta to be greater than 1 right or beta must be less than alpha.

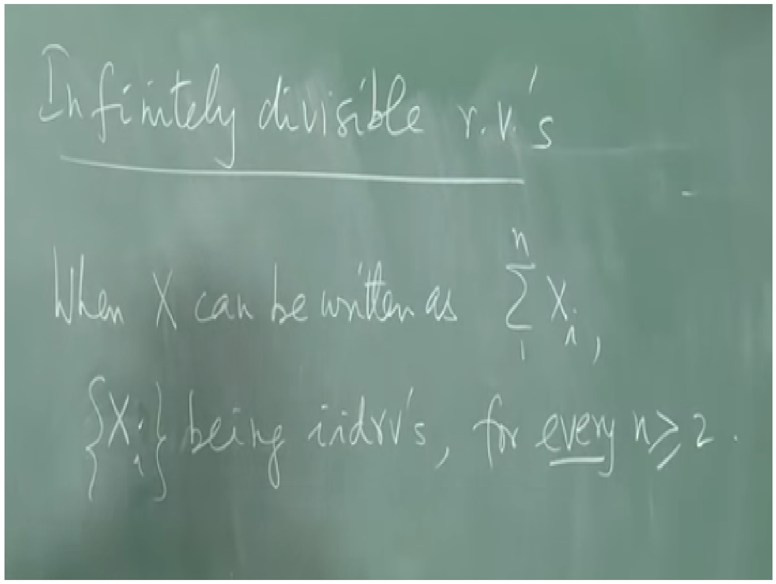
So if this alpha for example is 3 halves then although the second of this distribution does not exist the beta th moment would exist even if beta is a fraction as long as beta is less than 3 halves okay and as alpha gets closer and closer to 2 you get, the variance would be would diverge formally but beta would exist where beta gets closer and closer to 2 okay. So such moments would certainly exist.

Then this generalized central limit theorem says that if you have a whole lot of iid random variables such that the beta th moment exists where beta is just less than alpha out here then the sum of those fellows in a suitable as n tends to infinity would tend to one of these the appropriate stable distribution in this case. So this is the generalization of a central limit theorem which simply says that each of these stable distributions is the attractor for whole family of distributions all of which have moments up to a certain order.

And then the maximal one among those moments will decide the alpha value for the stable distribution to which these distributions get tend in the limit okay. So this is what the appropriate generalization is and there are further generalizations of this. I will mention this a little bit more when we do fractional Brownian motion when we talk about Brownian motion which is not the usual kind.

So, so much for stable distributions. They have a lot of other interesting properties we can discuss subsequently. But now I would like to ask a reverse question, a different kind of question. I would like to ask the following.

(Refer Slide Time: 50:04)



Given, not a set of iidrv's but given a random variable X with certain properties, specified distribution function and so and so forth. When can I write this random variable as a sum of 2 identically distributed random variables? If I can write it as a sum of 2 random variables iidrv's then I would say this variable is 2 divisible. If I can write it as a sum of 3 iidrv's I would say it is 3 divisible and so on and in general n of them n divisible.

Then I can ask other random variables for which I can write the random variable as a sum of n iidrv's for all n greater than equal to 2 no matter how large. If I can then I say this random variable is infinitely divisible. So I would like to introduce the concept of infinitely divisible random variables. So this is when X can be written. So when this can be done then I say X is an infinitely divisible random variable okay.

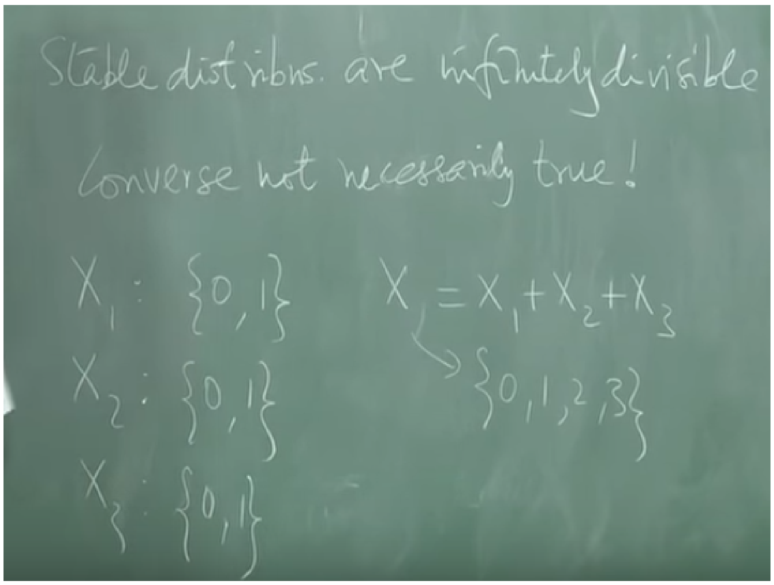
For every n I will call the X sub i's the components of X because you add them all up you get X and this is a very special property. You can see immediately it is not going to happen most of the time but when it does you have an infinitely divisible random variable. And what makes things interesting is that the distribution of every one of these X's need not be the final distribution of X itself okay, need not be so at all.

You just want them to be identically distributed random variable with a common distribution function which could be different in functional form than the distribution of the sum itself.

“Professor - student conversation starts” So also stable distributions will be definitely; right. **“Professor - student conversation ends”**. So it is clear that stable distributions are infinitely divisible immediately.

Not only that, in the case of stable distributions the distribution of each of the X i's for every n is exactly the same as the distribution of X itself and that is a very special property right.

(Refer Slide Time: 53:06)



So it is immediately clear that stable distributions are infinitely divisible. Is the converse true? No reason why that should be true at all. No reason at all. So the converse is not true. We are going to give counter examples. So converse not necessarily true. This idea of divisibility is a little subtle. You have to be a little cautious here. You may have a random variable, let me give you an example. Suppose you have a random variable which takes the value 0 or 1.

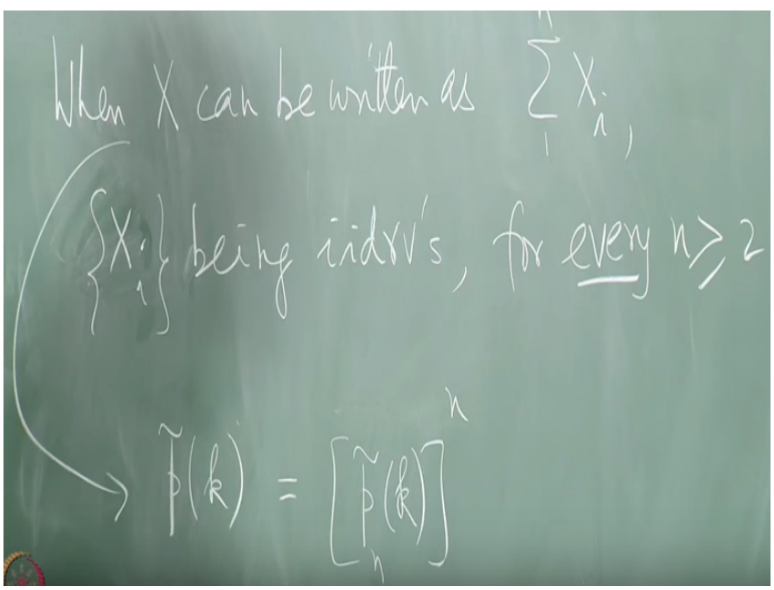
It is a Bernoulli trial let us say. So this variable can take X 1, can take values in the set 0, 1 and X 2 also takes values in the set 0, 1; X 3 0, 1 okay. Now what is the sample space of the random variable X = X 1 + X 2 + X 3, 0 to 3; 0, 1, 2, 3. So this has sample space 0, 1, 2, 3 and clearly this by inspection if you can see these appear with equal probabilities it is just heads or tails and you are asking what happens to the sum of the scores right.

Then it is clear that this is 3 divisible. This random variable has some distribution. In this case it will be a binomial distribution and it is 3 divisible in this fashion. Is it 2 divisible? Is it possible to have a random variable which takes values in the set 0, 1, 2, 3 and ask can it be written as the sum of 2 iidrv's. Is this possible? Well suppose you say 0 and 1. It is clear it would not reach 3 so that is gone. Then you say 0, 1, and 2.

Let us suppose each of the components has values 0, 1, 2 then 4 is in the sample space of the sum which is not given okay. Then you say let it suppose it is 0 and 3 halves so it reaches this. But since 0 is in the sample space 3 halves has to be in the sample space which it is not so there is no way in which you can make this 2 divisible right. So here's a random variable this fellow here which is 3 divisible but not 2 divisible.

So this divisibility is not such a trivial concept. It requires a little bit of understanding. So not everything is divisible but now we are saying something much stronger. You are saying for every n this variable is divisible, n divisible. So it puts a lot of constraints on this in the possible distributions that can have this property. And what do you think is the primary property that it has. Because this has to become iidrv's it implies that the characteristic function must be a product of characteristic function because these are iidrv's right.

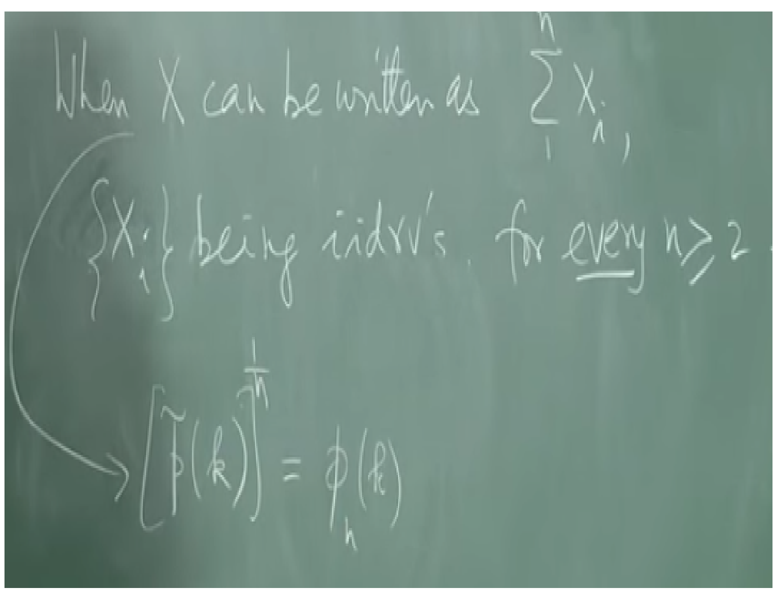
(Refer Slide Time: 56:44)



So it immediately implies that X must have a characteristic function p tilde of k which must be of the form the nth power of some other characteristic function. So this must be of the form p tilde then we put a n here to show these are different functions for different n's in general and it must be of this form okay. Only then is this variable going to be infinitely divisible, is this random variable going to be infinitely divisible okay.

So now the matter is simple. All we got to do is to look for all those characteristic functions which has this property here. So as soon as you can write this, the matter is over. Let us look at that example again. Let us look at that guy here and see what this implies for divisibility. **"Professor - student conversation starts"** Pardon me. Is the decomposition unique? So if you ya we have not answered questions like is it always going to be unique for a given n and so on and so forth. **"Professor - student conversation ends"**. No apriori reason why this should be so and so on but tell me if I take an arbitrary characteristic function p tilde of k in this fashion and ask can I not always write it as something to the power I do a 1 over n here and raise the power n. Is it not always going to be the case? Suppose that were true it would imply that this is an honest characteristic function. So it means p tilde of 0 is 1 and its inverse fourier transform is nonnegative.

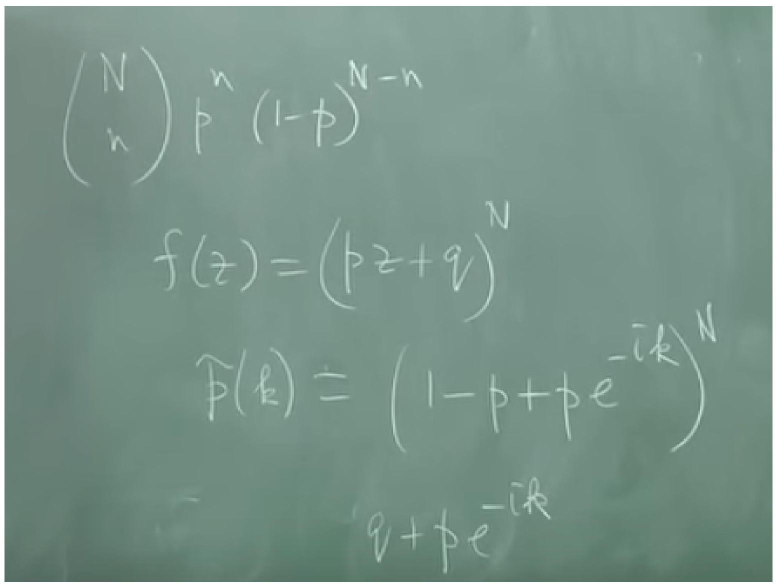
(Refer Slide Time: 58:32)



But now you are asking if I raise this to the power 1 over n and I get some function here phi n of k you are saying this too should be a characteristic function. It too must have an inverse fourier transform which is nonnegative and that is not true in general. So this means that divisibility is not a trivial concept at all. Not necessary that this is going to happen all the time. It happens only in special cases. Now let us look at the Bernoulli trials that we talked about.

Now if you had n Bernoulli trials then the distribution that we got for the resultant was in fact a binomial distribution if u recall right.

(Refer Slide Time: 59:21)

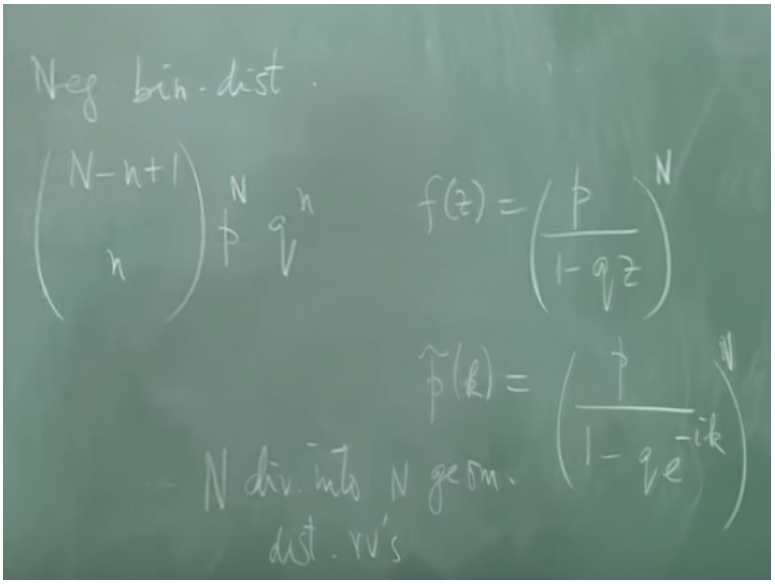


So in that case you got the binomial distribution was of the form some N n p to the power n, 1 minus p to the power N minus n in this fashion and what was the generating function for this guy? What was the f of z in this case? It was a very straightforward thing. It was just so you can remember that f of z was equal to p z plus q to the power N. That is all it was right and then the characteristic function p tilde of k in this case was pz by e to the minus ik and that was it or q plus p e to the minus ik etc. Is that n divisible?

You can see it is a product of functions, all identical functions p plus or q plus pe to the minus ik raised to the power n. So you would immediately say it is n divisible provided this fellow itself, provided q plus p e to the minus ik was the characteristic function of something or the other and it is. It is the characteristic function of a Bernoulli trial, a random variable which takes value p 1 with probability p and 0 with probability q right.

So trivially the binomial distribution with parameter capital N is N divisible into N Bernoulli trials okay, not binomial distribution binomial random variable at all but N Bernoulli trial immediately follows okay right. What about the geometric distribution? What about the negative binomial distribution? What happened in the case of the negative binomial distribution?

(Refer Slide Time: 1:01:39)

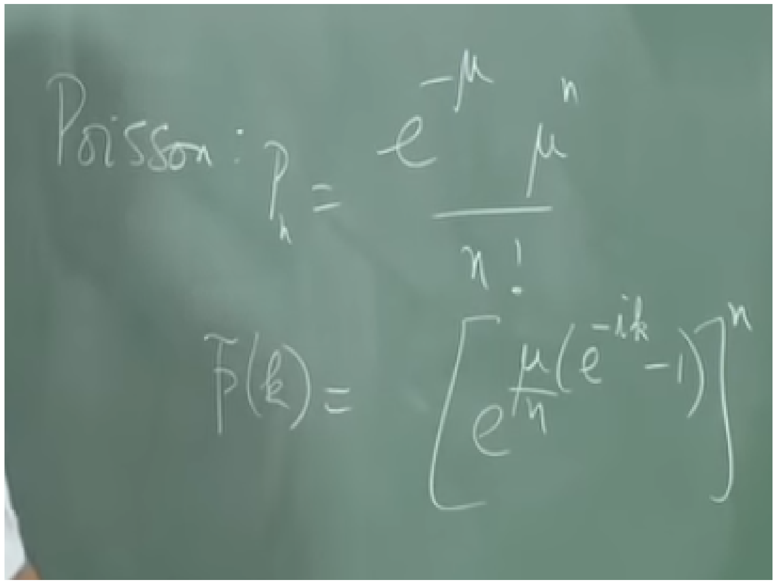


We call that the negative binomial distribution had a distribution which looked like N - n + 1 n p to the power N, q to the power little n. Little n was a random variable which took all the nonnegative integers in its sample space 0 to infinity and capital N was some given positive integer okay. This fellow here had a generating function f of z which was p divided 1 - qz to the power N. That is why it was called the negative binomial distribution okay.

Is that n divisible? It looks like the Nth power of something. Capital Nth power of something right? So if p over 1 minus qz is a characteristic function of q so in this case p tilde of k = p over 1 - q e to the - ik to the power N. So this fellow is a characteristic function or if this fellow alone p over 1 - qz is the generating function for a probability distribution then this negative binomial distribution with parameter capital N is capital N divisible into n of those distributions n of those random variables right.

Is that a generating function p over 1 - qz? Yes, it is the generating function of a geometric distribution right which had a probability density function probability distribution p times q to the power n right. So immediately this is N divisible, N divisible into N geometrically distributed random variables. What about the Poisson distribution? Is that N divisible?

(Refer Slide Time: 1:04:05)

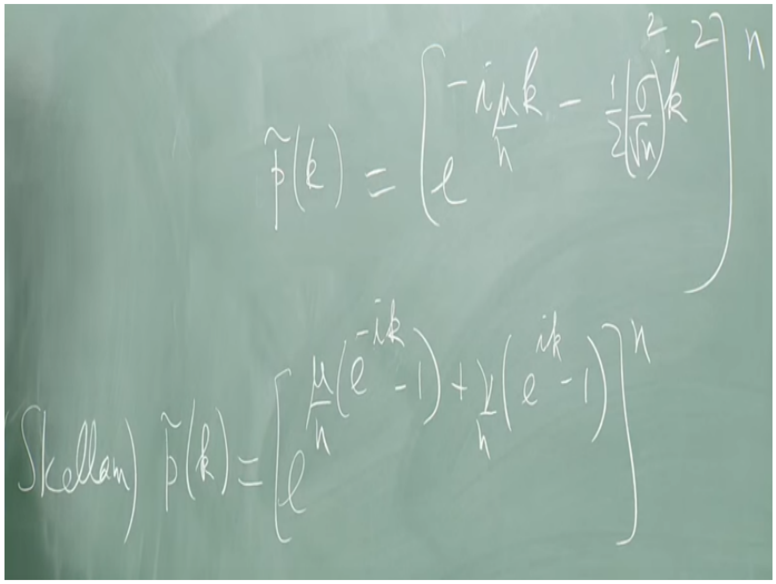


Let us write the random let us write the distribution down. For a Poisson we had e to the - mu, mu to the power n over n! was P of P n and the characteristic function p tilde of k was equal to e to the - mu e to the - e to the power mu e to the - ik - 1. It was e to the mu times z - 1 for the generating function so I put z is e to the - ik get the characteristic function okay. Can this be written as the nth power of something?

Yes, trivially so. And what sort of random variable has a characteristic function like that? A Poisson with mean value mu over n for every positive integer n right. So it is n divisible. Is it stable? It does not fall in the family of stable distributions. it is not discrete and where the sample space is discrete and so on and so forth. So it is the discrete analog of a stable distribution, but it is infinitely divisible. This guy is infinitely divisible.

It even has this property that for every little n it is n divisible into n Poisson random variables with appropriate means etc. Is the Gaussian distribution n divisible, infinitely divisible? Yes, indeed. it is a stable distribution. So it is immediately so and you see that at once.

(Refer Slide Time: 1:05:54)



Because in that case the characteristic function p tilde of k was e to the - i mu k - one half sigma square k square and you can certainly write this as mu over n and sigma over square root of n square and the whole thing raised to the power n. So of course it is. It is n divisible with mean mu over n and standard deviation sigma over root n. it is a stable distribution so it is automatically infinitely divisible as well okay. What about the Skellam distribution?

The difference of 2 Poisson random variables, is that n divisible, is it infinitely divisible? You would expect it to be so because in this case it is just the difference of 2 Poisson random variables and if you recall this had e to the mu e to the - ik - 1 + mu times e to the - u so it was e to the ik - 1. That is what the characteristic function was for the Skellam distribution okay and of course now it is very trivial matter to say this is mu over n, nu over n and I raise this to the power n.

So yes, it is also infinitely divisible okay. So the set of infinitely divisible distributions is a bigger set than the set of stable distributions but the stable distributions are a very special subset of it. In general for infinitely divisible distributions the components do not have the same distribution as the original distribution but for the stable ones they do and for the Poisson they do okay. So it becomes an interesting question to classify all such distributions.

This gets us into statistics. I am not going to go into that detail here except to show you that by these simple examples you can see the idea of the notion of divisibility and what sort of role it plays. We will try to get back to this in various other examples. So we will stop here today.