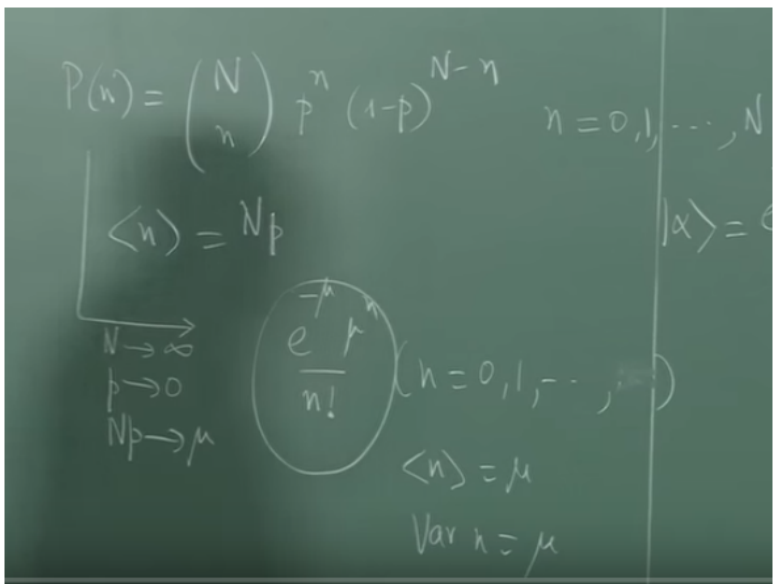
Physical Applications of Stochastic Processes
Prof. V. Balakrishnan
Department of Physics
Indian Institute of Technology-Madras

Lecture - 02
Discrete Probability Distributions (Part 2)

Now we looked at the some discrete probability distributions last time. In particular we looked at some properties of the binomial distribution and then I pointed out that the binomial distribution went over into the Poisson distribution in a certain limit, well-defined limit. We also looked at simple physical example of the binomial distribution which is density fluctuations, number density fluctuations, number fluctuations in an ideal gas, classical ideal gas okay.

(Refer Slide Time: 00:59)



If you recall the formula we had for the binomial distribution was P(n) was N n - n where n runs from 0 up to capital N. That was a binomial distribution and this is the distribution of a set of Bernoulli trials, N Bernoulli trials, capital N Bernoulli trials each with a probability little p of success and then I pointed out that if you took this limit in which by the way the average value of n was equal to N times p.

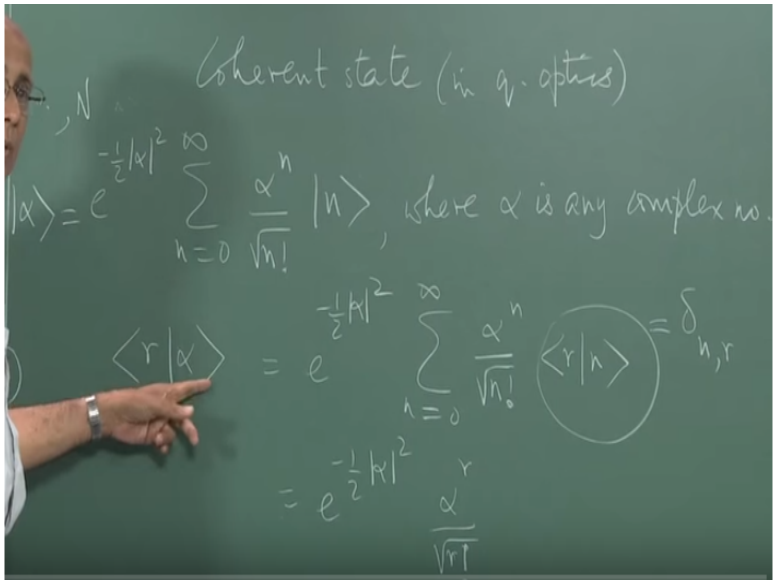
And if you took a limit in which N tends to infinity and p tends to 0 such that this number is finite at some value mu, then the probability distribution tends in this limit N tends to infinity, p tends to 0, N p tends to some value mu, some constant value mu. This tends to the Poisson distribution which is e to the minus mu, mu the n over n! and the mean value remained of course at mu and this distribution now n has the sample space 0, 1 all the way up, all the way to infinity okay alright. So we needed a simple example of this distribution and many many abound Poisson distributions abound in nature.

We will come across many Poisson distributions, Poisson processes in fact as we go along, but for the moment let me point out that if you did this elementary nuclear physics example of taking a large sample of a radioactive sample and asking what is the average number of decays that takes place, nuclei that decay in some given in a in some given interval of time, you discover it increases linearly with the time with some mean decay rate lambda.

So you discover that the number is distributed exactly as in a Poisson distribution with an average mu which is equal to lambda t where lambda is the mean rate of decays okay. So that is a classic example of a Poisson process. I also said, I would give you an example based on a number statistics exactly as in the case of the classical ideal gas and this is afforded not by black body radiation which had a geometric number distribution of photons if you recall but rather by coherent light.

So ideal single mode that is one single frequency, one given wave vector, one state of polarization if you look at ideal single laser light, it turns out that the number of photons in this radiation field is described by a Poisson distribution once again precisely a Poisson distribution.

(Refer Slide Time: 04:18)



For that let me go back a couple of steps and do a little bit of quantum mechanics to tell you where this emerges, where this comes from at least in a kind of non Rickles way and this has to do with the idea of coherent states in quantum optics. Such a state is described by a state which is expandable in a basis set which comprises photon, specific photon number states and if I call the state n a state of the radiation field in which there are n photons, exactly n photons of a given frequency and a given state of polarization, then the so called coherent state is built up of a super position of these states.

And it looks like this, it looks like a summation from n = 0 to infinity e to the minus half modulus alpha square, I will explain what these symbols mean, alpha to the power n over square root of n! where alpha is any complex number. So you give me an arbitrary complex number alpha of finite modulus.

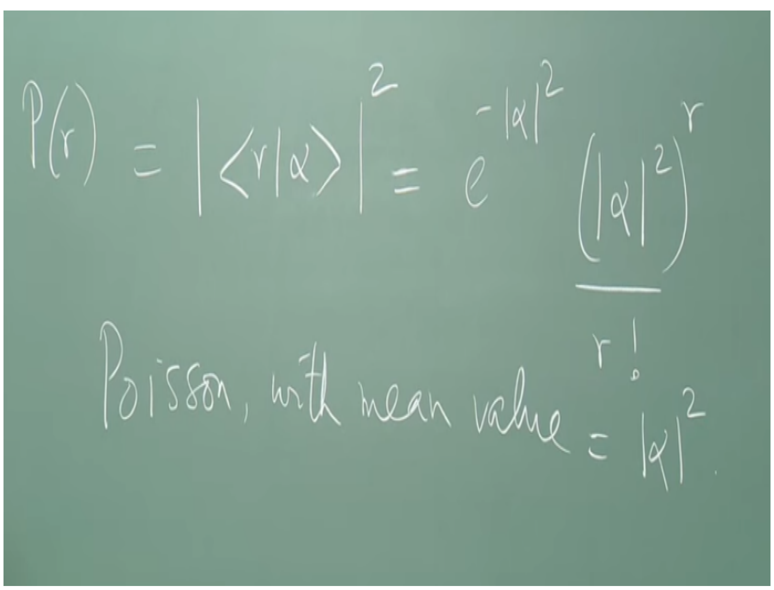
And I construct this super position of number states of specific numbers and this state here is denoted by this symbol alpha labeled by this complex number alpha as you can see when you change alpha you get a different state and so on and it is called a coherent state okay. Now there are technical reasons as to why it is called a coherent state, what is so specific, what is so special about this state and so on.

We would not get into that at the moment here because it is not, the focus is not on quantum optics, but I merely wanted to point out that in this coherent state it is not very difficult to show that you get precisely a Poisson distribution for the number because after all if you want to ask what is the probability amplitude that in this state alpha you have precisely some fixed integer number say r photos that is given by this quantity here r and alpha and I presume those of who are familiar with quantum mechanics will recognize this as the scalar product.

It is a complex number in general and it is a probability amplitude that in the state alpha you have exactly r photons okay and when you do that and you use the fact that so put that in here and this is equal to e to the minus half mod alpha square summation n = 0 to infinity, alpha to the n over root n! and then you have an r with an n. But you see these number states are all orthonormal to each other.

They form an orthogonal basis. The orthogonality of the bases implies that orthonormality of the bases implies that this quantity is a Kronecker delta. It is equal to 0 if n = r and it is equal to unity if n is 0 if n is not equal to r and it is unity if n = r, just a Kronecker delta which helps you immediately to write down what the sum is because this sum fires only when n is equal to r and therefore this becomes equal to e to the - half mod alpha square, alpha to the power r over square root of r!.

(Refer Slide Time: 08:11)



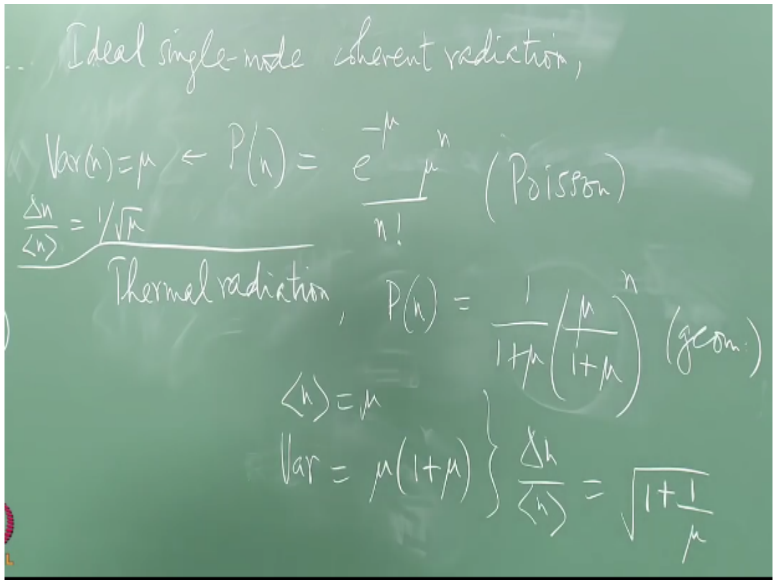
Okay and that is the probability amplitude and the rule of quantum mechanics is if you give me the probability amplitude for something to happen the actual probability is not a complex number, it is the modulus square of this complex amplitude and therefore the probability that in this state you have r photons is given by this quantity r alpha whole square and that is easy to figure out because you get another such factor here.

So it is e to the minus mod alpha whole square and then you have since you want the complex conjugate is alpha on the left and r on the right is a complex conjugate of this number here. This is a real number, that is a real number, this is a complex number. So you got to do alpha star to the power r that makes it mod alpha square to the power r/r! because there are 2 of these factors, the square root goes away.

And of course you immediately recognize that this is a Poisson distribution with mean value equal to mod alpha square right okay. So the photon number distribution in ideal single mode laser light of a given state of polarization and frequency is a Poisson distribution. Very drastically different from the geometric distribution that we had earlier because we know what the variance of a Poisson distribution is.

We know that if this is the distribution we know that the mean value n = mu and we know that the variance of n is also equal to mu. The mean is equal to the variance and that is fairly easy to derive because we know the generating function for this distribution and therefore it is trivial to compute that, show that the variance is also equal to the same, is the same as the mean. That is one of the properties, fundamental properties of a Poisson distribution. Now compare this with, so what does it mean?

(Refer Slide Time: 10:19)



If I have a number distribution that is, a photon number distribution that is Poisson, so if you have ideal single mode coherent radiation, the P(n) is equal to e to the minus some average value mu, mu to the power n over n! On the other hand for thermal radiation, the P(n) was a geometric distribution very different from a Poisson and we saw that this thing here is 1 over 1 + mu and then there was a mu over 1 + mu to the power n.

So this was Poisson, this is geometric. The variance in this case equal to mu itself, the variance in this case here. So this immediately implies the variance is mu, the standard deviation is square root of mu, the standard deviation divided by the mean is 1 over square root of mu okay and therefore delta n over n = 1 over square root of mu. That is in that case okay. On the other hand here, the mean value equal to mu and the variance not difficult to show.

This is mu times 1 plus mu. That is trivial to show from this expression here and this will immediately imply delta n over n equal to well the mu there is a square root out here and then this is going to give you a square root inside. So this is equal to square root of 1 + 1 over mu and as mu increases when there you have a large number of photons, this scatter becomes extremely small.

On the other hand this scatter is always bigger than 1, this ratio is always bigger than 1 showing that in thermal radiation you have a huge scatter about the mean whereas relative fluctuation is very small in coherent radiation, so one of the defining properties of this coherent radiation here okay. Now these results are very easily derived if you use the generating function in each case. We know the generating function for this, we know the function for this and it is trivial to do this. So I urge you to do this as an exercise.

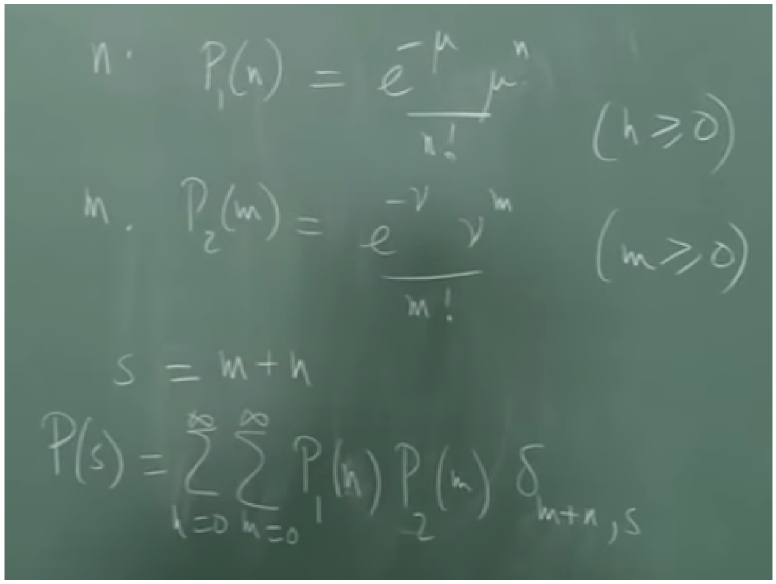
So in some sense the most coherent radiation you can have is ideal single mod laser light and the least coherent one you could have is black body radiation thermal and what is happening is that you have a distribution which goes all the way from here to here, it would be nice to have a family of probability distribution which interpolate between these 2 extremes okay.

And we are going to see very shortly that there is a family, it is called the negative binomial distribution which is going to interpolate between the geometric distribution on the one hand and the Poisson distribution at the other extreme here. In fact one can use that family of distributions to analyse photo photon counting statistics and see how much of thermal light is mixed with coherent radiation in an actual experiment okay.

But before we do that let me take up another question and that is the follow-up. What happens if you have a sum of 2 Poisson distributions? If you have a sum of 2 random variables each of which is Poisson distributed okay with different means in general, what would happen? Let us suppose that you have 2 species of radioactive nuclei and you are trying to count what is the total number of decays in a given time.

The rate at which one of them decays is lambda 1 and the other one is lambda 2, so each of them as I said is a Poisson distribution at any given instant of time and the question is the question you can reasonably ask is what is the actual distribution of the sum of these processes, of these 2 random variables okay.

(Refer Slide Time: 15:06)



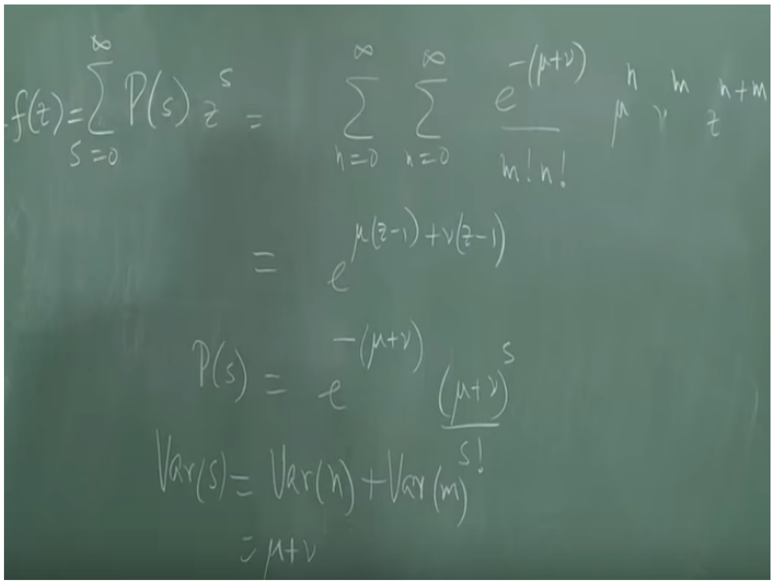
So let us suppose that you have one random variable n which is distributed, P 1 of n in a Poisson manner mu to the n over n! and you have another m which is distributed also in a Poisson distribution but with a different mu and we would like to know what is the probability distribution of the random variable s defined as m + n and I would like to know what is P(s). Now what would you say it is? How would you approach this problem?

These are independent of each other, completely independent of each other so fundamental, a basic first principle way of writing down the probability distribution of s is to say that this is equal to you try to make up s in all possible ways by summing m and n and first of all what is the sample space of s? So what is the sample space of s? Also 0 to infinity also 0 to infinity. So that way there is no change in the sample space.

So what you have to do is to say, let us suppose the first variable n has a distribution P 1 and then a P 2 of m and you sum over all possible values 0 to infinity, summation n = 0 to infinity subject to the constraint that m + n is s and that should by definition give you the probability distribution of s. Remember that m and n are dummy here, they are summed over. So the answer is a function of s and it fires only when m + n is equal to s for any given value of s.

So that is the definition and what we got to do is to put this in here and then do the summation but you cannot independently sum over m and n because there is a constraint here. So you should use that constraint, you get rid of one of the sums but it will constrain the other sum. For instance n, neither m nor n can exceed s because they are the sums of 2 non-negative integers; s is a sum of 2 non-negative integers. For any given value of s, m can at best go up to s and so can n at best go up to s. So while you can do the summation in principle the sum will get constrained here and you got to be a little careful in doing this.

(Refer Slide Time: 18:06)



On the other hand you could do the following. You could say alright let us suppose f of z, let us call the generating functions for this and this f 1 and f 2 of z and then we could do the following. We could say what is the generating function of the sum? That is equal to this by definition z to the power s, but this P(s) I put this in here and you can see that this is equal to a summation over n = 0 to infinity, summation m = 0 to infinity.

And then I plug this in, I multiply both sides by z to the power s and sum over all possible values of s okay. So there is a summation s = 0 to infinity this guy and then you have P 1 that is e to the - mu + nu and then what? So remember that yes over m! n! here and then you have e to the power you have z to the power. So mu to the power.

There is a better way of doing this and this there has got to be a much better way of doing this mu to the power n, mu to the power m, z to the power s, but subject to the constraint m + n = s. So all I have to do is to put that in here. This becomes z to the power n, nu to the power m sorry z to the power n + m in here okay and that is it okay. So what is this equal to now? I have used the constraint here. I have finished off the sum over s because I have used that constraint to replace s by n + m and what does this become mu n mu z to the power n, nu z to the power m and each of the sums can now be done completely.

There are no further constraints on it. Therefore this gives me e to the power mu times z - 1 + nu times z - 1 okay. And all we need to do is to take out the coefficient of any power that we want to find P(s). We need the coefficient of z to the power s in the power series expansion of this quantity in a power series in s okay and therefore this immediately follows that P(s) = e to the power - mu + nu, mu + nu to the power s over s! okay.

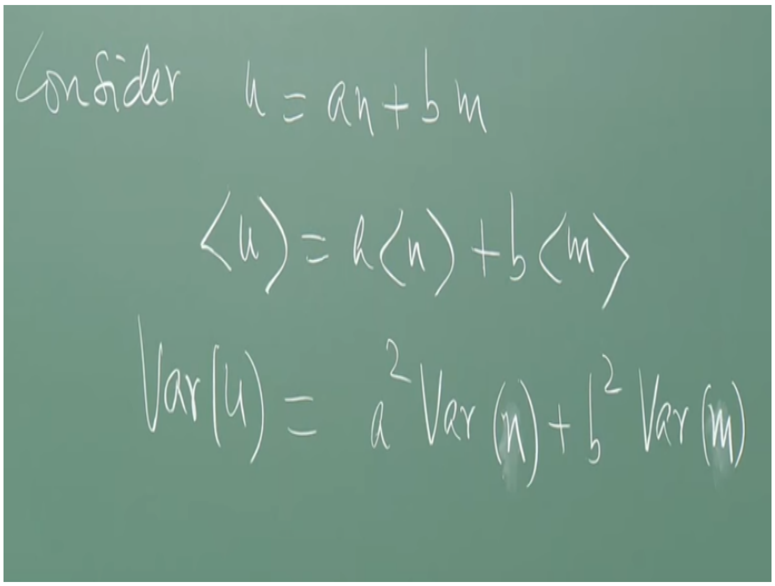
And what is that distribution? It is a Poisson distribution with the means added up okay. Therefore the variances are also added up. Variance of s is the same as the variance of this plus that. Incidentally one can show ab initio that if you have 2 random variables which are independent of each other, statistically independent of each other, their variances must add up. Their means must definitely add up, that is a trivial thing to show.

But the variances also add up. Why is that? Why do I say that the variances add up? Because the only place they need would not add up if you try to find the mean square value of m + n for example the sum of 2 random variables, you would have the mean value of m square, the mean value of n square and then the mean value of m times n right but that factors. That factors because these are independent variables.

Since it factors, it cancels out when you subtract the mean whole square immediately. So it is trivial to see that the variances have this additive property which the mean square value of a random variable does not. So if a random variable is a sum of 2 or more random variables which are independent of each other, then the mean value is just the sum of the means of the individual components and the variance also is a sum of the variances of the individual components a property which is not shared by the mean square alone or any higher moment okay.

So this invariance is very crucial and we will see that it is actually a reflection of property of what are called cumulants which also add up and the additivity of cumulants is very very fundamental notion, we will come back to that okay. So our first lesson, if you have 2 Poisson random variables which are independent the sum is also a Poisson variable added up in this fashion.

(Refer Slide Time: 23:47)

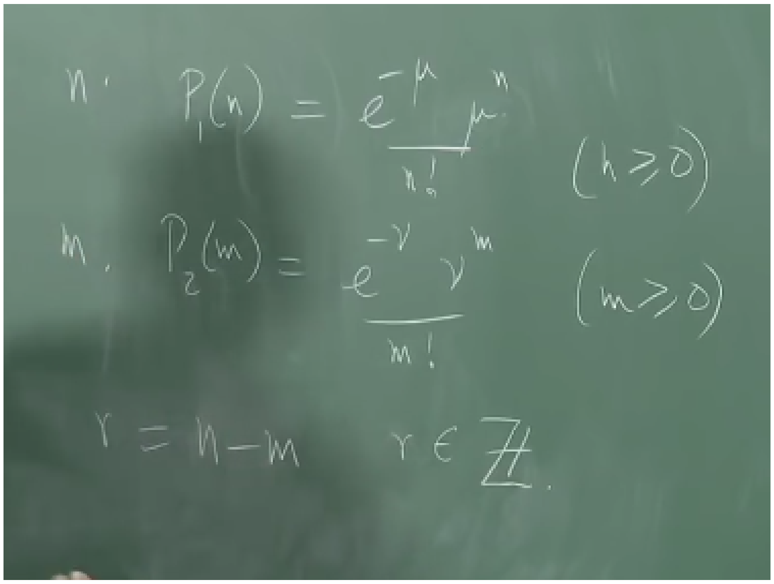


I urge you to show that in this instance for example a slightly more general property. Suppose a and b are any real constants positive say then or negative it does not matter real constants, consider the random variable, let me use a symbol for this random variable. Let us call it u. This is equal to a times n + b times m. This random variable does not have necessarily integer values in sample space because its m and n are integer valued but a and b are any constants, real constants.

Then what is the mean value of mu? What do you think is the variance of equal to, it is a square because you are now finding the variance so it is a square variance of m + b square variance of m the other way n + m in this fashion okay. Now you know this property that the sum of 2 Poisson variables is again a Poisson random variable follows very simply from the fact that the generating functions essentially multiply and the generating functions multiply and because they are exponentials because of this property here you get an additive exponent out here.

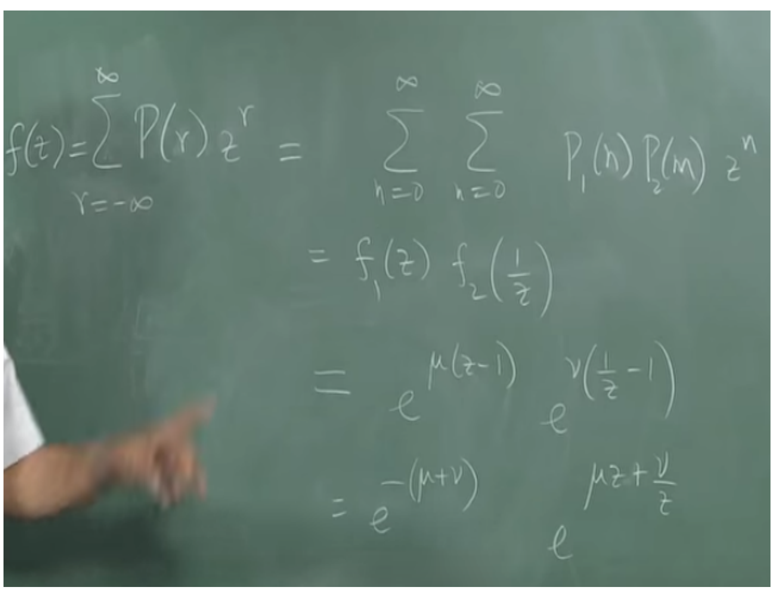
So you can see clearly that this is going to happen anyway. But now I could ask a slightly more complicated question. What about the difference of 2 Poisson random variables? What do you think is happened going to happen there? So let us go back and ask what happens if I consider not the sum but the difference.

(Refer Slide Time: 25:42)



Let us call it r = n - m. What kind of distribution would this have? What would what would this do? What is the sample space of r? Minus infinity to infinity now, not 0 to infinity, all integers are allowed right and now we try to find out what is the generating function of this thing here.

(Refer Slide Time: 26:21)



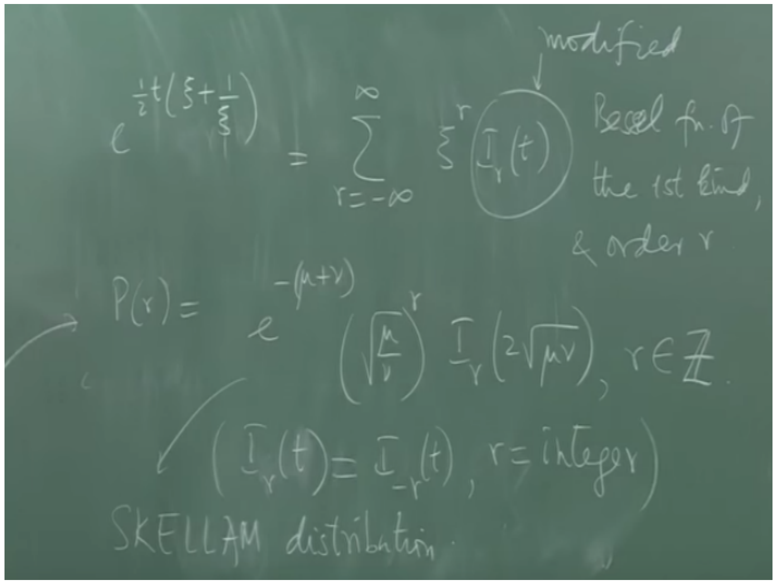
Well, we do exactly the same as we did here except we now have to write r equal to minus infinity to infinity P(r), r to the power s and it is the same as this except that you are going to have to write let us I mean put this back here. P 1 of n, P 2 of n and then we multiplied by, sorry what did I do here z to the power r, z to the power r and if I plug that in here, I get z to, there is going to be a delta function of m - n, r and I am confusing n and m all the time.

P 1 of n that is correct P 2 of m let us call n - m so that I do not confuse between 1 and 2. I call this 1 and this 2. And then I have a delta function of n - m, r and if I multiply by z to the r, I replace that by z to the power n - m so this becomes z to the power n, z to the power - m that is it. So this guy is equal to f 1 of z and f 2 of 1 over z because this power series with z to the power m is going to give me the generating function f 1 but now it is 1 over z to the power m right.

And what is that equal to? For a Poisson this becomes equal to e to the power - mu sorry mu times z - 1, e to the power nu times 1 over z - 1 which is equal to e to the - mu + nu and then e to the power mu z + s nu over z and what we seek is the coefficient of z to the power r in the expansion of this quantity in powers of z okay.

But then all possible powers, positive as well as negative powers are going to exist and it is not so easy to first you expand this and then you expand that and of course for a given power z to the power r an infinite number of terms are going to contribute because there is an infinite series and positive powers here and negative powers out here. So we need a little formula. What we need is something called a generating function for a special function.

(Refer Slide Time: 29:40)



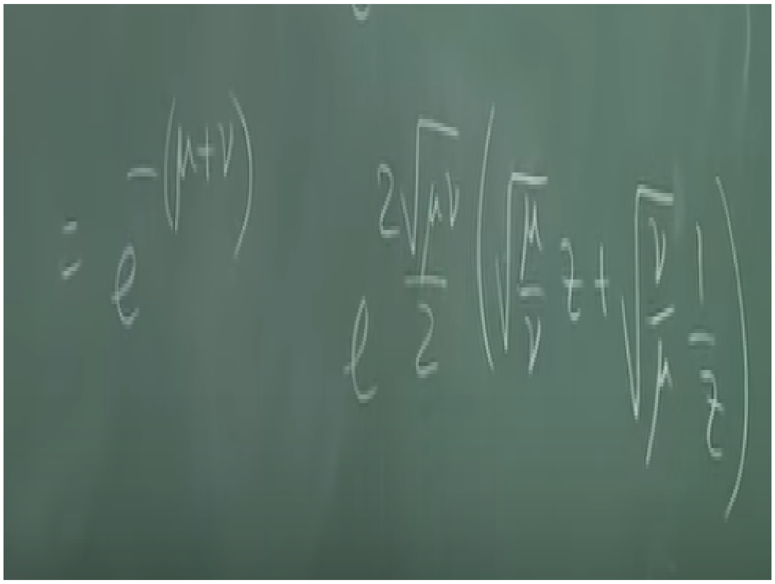
So we need an identity which is the following; e to the power one half t times some psi + 1 over psi when expanded in powers of psi positive as well as negative this guy here is a summation r equal to - infinity to infinity psi to the power r and it is multiplied by t. This here is the modified Bessel function of the first kind and order r. It has got nice interesting properties.

I will come, we will come back when we study random walk we will come back to this special function and tell you all about it but it satisfies a certain second order differential equation called the Bessel's modified, the modified Bessel equation. It is a nice function. It is what is called an entire function in the sense that it has no singularities whatsoever for any finite value of t as a complex variable. It is a nice entire function.

It can be written as a power series in t which converges absolutely for all finite values of mod t okay. Converges even faster than the exponential (()) (30:59) the power series in t, but whatever it is we can write down what the answer is based on this from this guy here because this here we can cast it in that form. What the trick is how should I cast this in that form?

I need a psi and a 1 over psi but I got a mu z and a nu over z. So what should I do? I want to make it some variable psi and its reciprocal. **"Professor - student conversation starts"** take out mu, that does not help. Root mu nu. Take out a square root of mu ya. **"Professor - student conversation ends"**.

(Refer Slide Time: 31:37)



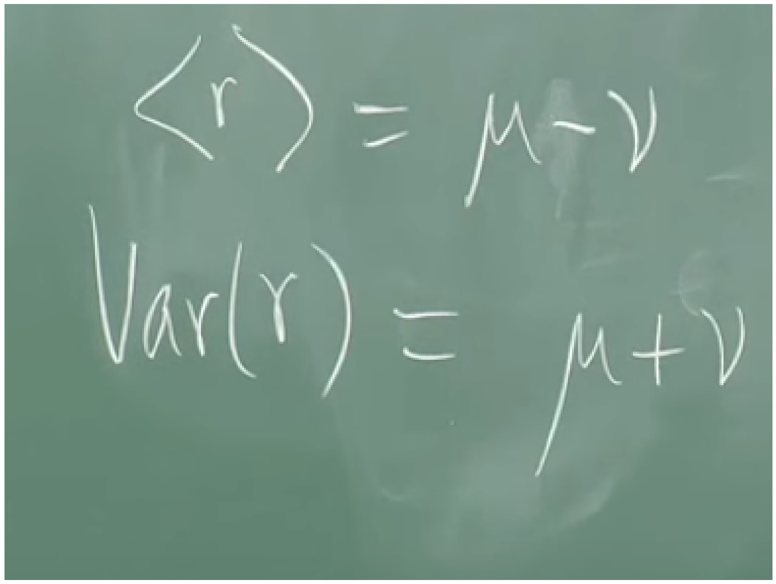
So take out a square root of mu nu and then you have square root of mu over nu z + square root of nu over mu 1 over z outside this square root, you have this, but you need a 2 out there so you put that 2 in by hand and write it in this form and that will immediately tell us that p(r) in this case is equal to P to the power mu plus nu square root of mu over nu to the power r because I want the coefficient of z to the power r in the expansion.

So that has got this guy sitting with it and then I r of twice square root of mu okay. This factor here depending on whether mu is bigger than nu or nu is bigger than mu is going to be dominated by the positive or negative r, one of the other. So it is just a power factor there. The Bessel function itself I r has an interesting property. It says I r I sub r for any positive integer r is equal to I sub minus r. It has got the symmetric property.

So the bias is entirely in this here. I r of anything of psi, the I r of t = I - r ya. So it is not at all, not at all a Poisson distribution. It is a more complicated distribution in that. By the way this guy has got a name. This thing here is called a Skellam distribution. If you like it is the generalization of the Poisson distribution to all integers to a sample space which is both positive as well as negative integers and 0.

So this is the Skellam distribution here and the difference of 2 Poisson variables is a Skellam distribution okay. Now we already know the generating function. This is the generating function for this distribution. So from here it is possible to find all the factorial moments fairly straightforwardly without any difficulty.

(Refer Slide Time: 34:28)



For example the average value of r will turn out to be mu - nu. That follows immediately if you differentiate this once with respect to z and z z = 1 you get this immediately. What about the variance? What do you think will happen to the variance? Well, remember we looked at a distribution of a n + b m and we said a and b are just real numbers. It does not matter whether they are positive or negative.

So you could set one of them equal to 1 and the other equal to - 1 and of course immediately it follows that this guy is equal to mu + nu okay. So when you have the difference of 2 Poisson variables, the mean of course is the difference but the variance will still add up okay. In fact we will see a little later that the cumulants this is called the first cumulants, this is called the second cumulant and then there are higher cumulants and so on.

We will see that the kth cumulant of the difference of 2 Poisson variables will be the first mean plus minus 1 to the power k times the second. So every other guy will going to be this and every odd cumulant is going to be this in this case. For the sum of course all the cumulants are going to be exactly the same. We will see that when we talk about cumulants. So this is an instance of what happens to the difference of the 2.

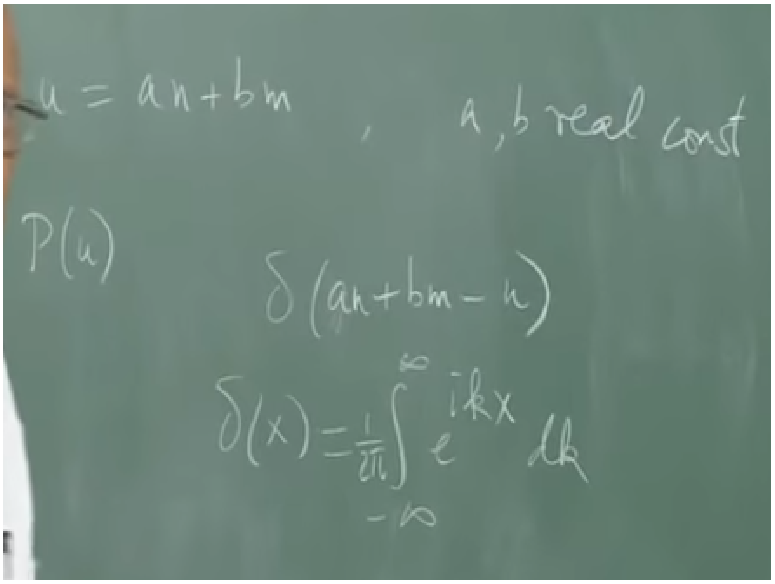
Just one final remark about a physical example once again of such a situation. We will see extensively that the solution to the simple random walk problem in continuous time is precisely a Skellam distribution. So if you did the following, if you took a coin and you had an infinite linear lattice.

You started at some origin and you tossed a coin and went to the right with some probability little p and to the left with probability little q depending on whether you got heads or tails and you did this at random instance of time with some mean rate according to a Poisson process then the probability of landing up at the point r on this lattice on the lattice point r positive or negative it does not matter is precisely this Skellam distribution okay.

So this will be the difference of 2 Poisson processes. In other words steps to the right and steps to the left would be regarded as 2 Poisson processes and then you want the difference of the 2 and we will see all the various places at which in which this random walk problem various problems map on to the solution to this random walk problem. We will see how that comes about okay.
We should probably take a little gap here and then any questions?

"Professor - student conversation starts" Sir what happens if you take a linear combination of your random variables. Will the coefficients be positive so you should take alpha times plus beta times and set a probability distribution of this variable alpha and beta being positive. Will that still follow a Poisson distribution. **"Professor - student conversation ends"**.

If alpha and beta are not integers that is not true anymore. So his question is what happens if I did that which is precisely what we have written down.
(Refer Slide Time: 38:01)



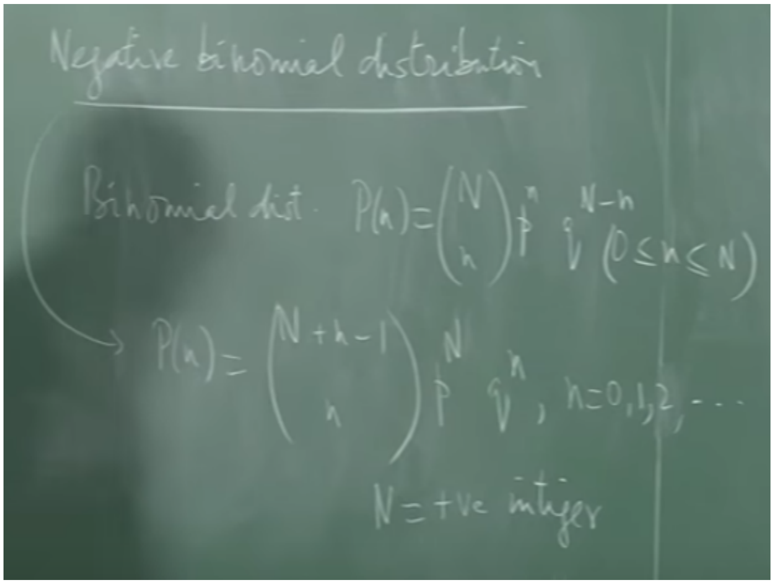
So let us call the variable u equal to a times n + b times m; a, b real constants and then ask what is the probability distribution of u itself okay. This is not Poisson. It is not Skellam or anything like that because the sample space is no longer integers right. It is any a times integer plus b times another integer. So it is again a set of points but the question is what is the probability distribution in this case.

One way to do this is to go right back to what the ab initio definition of this quantity and instead of writing a Kronecker delta you write a Dirac delta function of u - a n + b m okay and then write a representation for that Dirac delta function perhaps as an exponential which factors n and m and find the generating function of this u in terms of the generating functions of n and m, exercise for the reader to this okay.

So that is a nice interesting exercise to find out what is the probability distribution. So do the following. Write this thing in terms of so use a delta function of a n + b m - u. Put this formally inside the summation for n and m. Put this guy formally but you want to factor out. This is additive n and m. You want to factor it out and make it into some product of something or the other. When you got a sum and you want to convert it to a product what will you do?

You exponentiate, right. So we would like to find the representation for this delta function which converts it into an exponent okay and use a thing like delta of any real variable x is integral e to the i k x d k if you write a fourier representation for it. So put that in here. We are going to have one more integration to do but that is not very difficult to do and then you will end up with the generating function okay. So that is an interesting exercise to do.

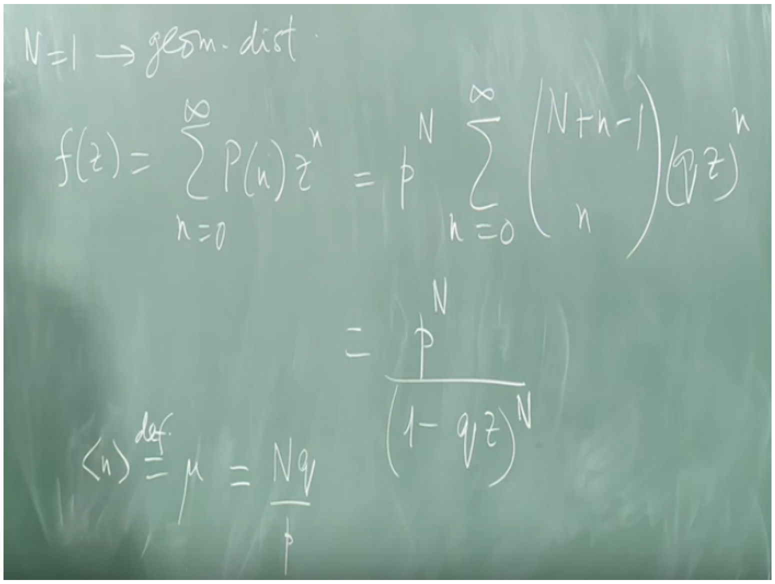
(Refer Slide Time: 40:29)



So recall the binomial distribution which was P(n), binomial distribution was P(n) = N n, p to the power n, q to the power N - n but the negative binomial distribution also has 2 parameters one of which is a positive integer capital N and the other is a little p exactly as in this case but this distribution P(n) is actually equal to N - n + 1 N plus n - 1. And then n here p to the power N and q to the power little n in this case where N equal to any positive integer.

And the sample space here is n = 0, 1, 2 all the way up to infinity unlike the case of the binomial where 0 less than equal to n. Here it goes all the way up to infinity. It is immediately clear that if you put capital N = 1 for example this is 1 that cancels there and this symbol becomes 1 then of course it becomes p times q to the power n which was precisely the geometric distribution.

(Refer Slide Time: 42:22)



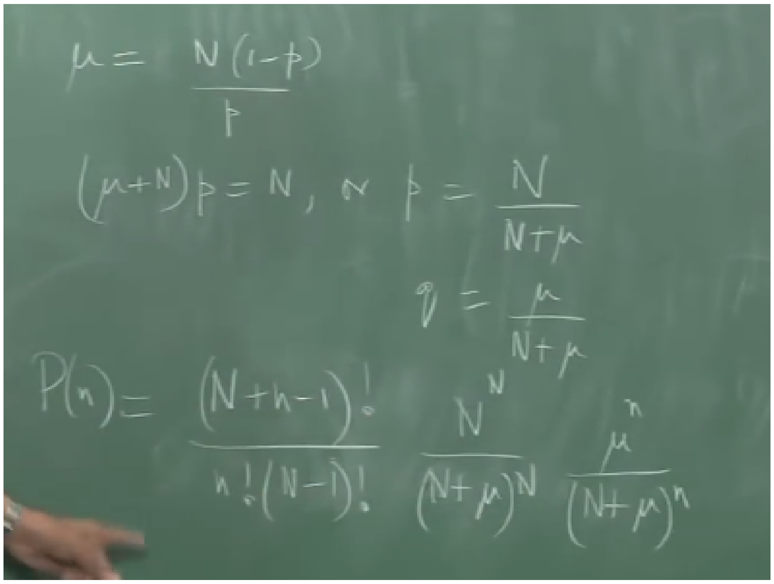
So it is clear that N = 1 corresponds to the geometric distribution. The question is what is it in general. What is it, general capital N what does it look like. Well the thing to do is again to find the generating function. So f of z is summation P(n) z to the power n and this time n runs from 0 to infinity and I plug in this in there. So this is equal to little p to the power N, N = 0 to infinity, N + n - 1, little n and then you have q z to the power n and that is an infinite series on this side.

However, it is not very hard to show that this is equal to p to the power N over 1 - q z to the power N. It is precisely the binomial expansion of 1 - q z to the power - n and that is the reason for calling it the negative binomial distribution okay. As you know when the index is not a positive integer this binomial series is an infinite series and it simplifies if you write it in terms of gamma functions and so on it simplifies to this kind here okay.

So once we have this it is easy to see what is the mean value of n equal to. If I call this mu say definition for the mean value is mu, this is equal to the derivative of this with z set equal to 1. If I differentiate it I am going to get out here a p to the N and then there is a 1 - q to the N + 1 so there is going to be a p in the denominator and an n which comes from there so it is equal to Nq by p because this is going to become N + 1 when I differentiate it.

So that is a p to the capital N + 1 and it cancels against p to the N and a q comes out on differentiation the minus sign goes away, so it is this. That is the mean value. Now of course we know already that for the geometric distribution when you put this equal to that 1 it is q over p we know that already.

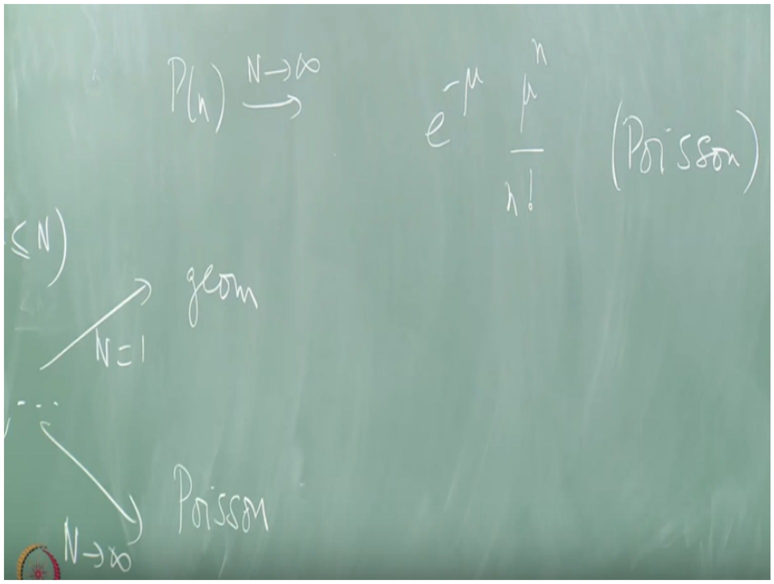
(Refer Slide Time: 44:53)



Let us rewrite this a little bit and let us write this as mu = N times 1 - p over p which means that mu + N p = N or p = N over N + mu and q of course is mu over N + mu; 1 - p is mu over N + mu. Then if I put that into the distribution into this guy P(n) = (N + n – 1)!/n! (N – 1)! that is the combinatorial part and then a P to the n that is N to the power N over N + mu to the power N; that P to the capital N q to the power little n, that is mu to the power n over N + mu to the power of n.

That is what this 2 parameter distribution looks like in terms of N and mu instead of p and q I got rid of p and n. I got rid of p and used mu instead. We have this guy here okay. Now look at what happens when capital N goes to infinity keeping mu finite at some value. What does it do?

(Refer Slide Time: 46:37)

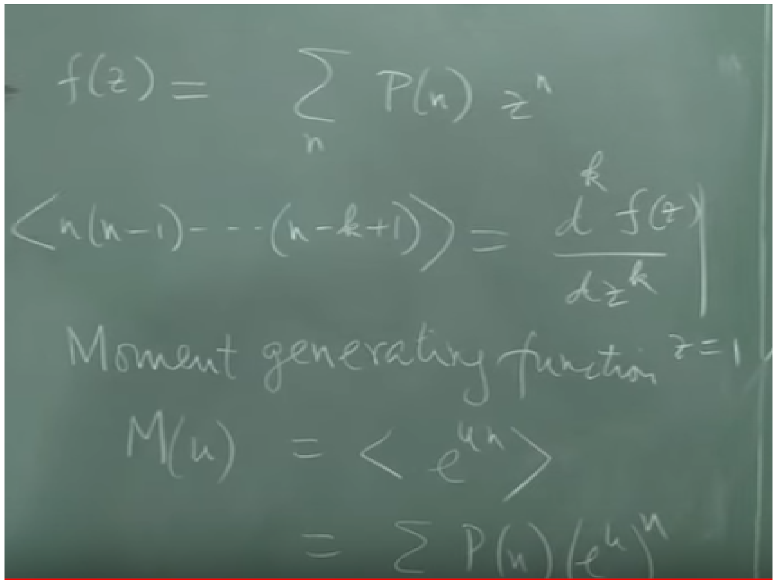


So P(n), this is going to give me an N to the power N + n to the power N + n dominant term by Stirling's approximation. This is going to give me an N to the power capital N. This guy here is going to look like 1 over 1 + mu over N to the power N if I divide through by this N to the power N it is just this fellow here and then you have a mu to the power n and you got an N to the power N, N to the power capital N and N to the power little n.

This gave you N to the power little n plus capital N. So this leading term cancels against this and that and leaves you with just this and you cannot forget this, so that is sitting there and the limit of this of course is e to the power mu so that goes up on top and you get e to the minus mu okay which is precisely a Poisson distribution. So we have this family of negative binomial distributions which tends N = 1 geometric and N tends to infinity Poisson okay.

So that is the great advantage of this distribution and it has got a very nice very simple looking generating function we wrote it down it is just an algebraic function here okay so we can write down all its moments and so on okay. As I said one uses this in the study of photoelectron, photo photon counting statistics for mixtures of thermal light and coherent light among other things in many other places in population dynamics etc. where the negative binomial distribution is used very very frequently okay. Just one final remark and then we will take it up from here.

(Refer Slide Time: 49:12)



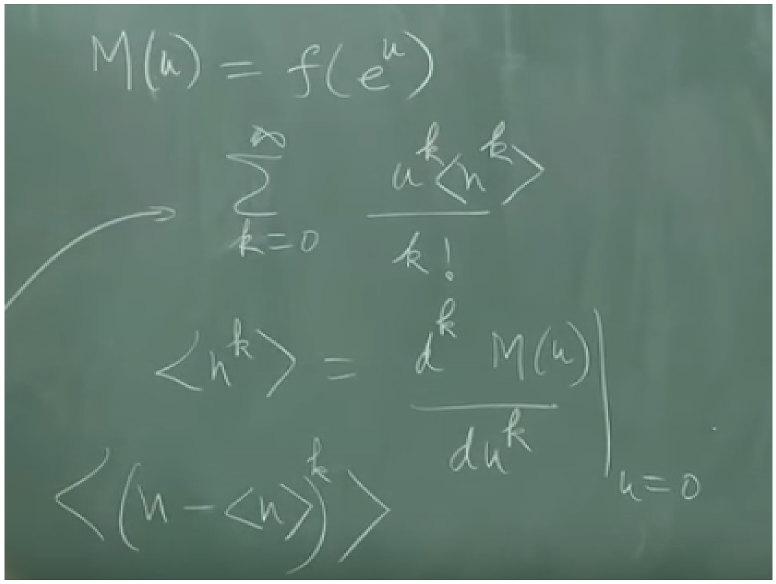
We have talked about the generating function throughout okay and while it has advantages, it also has disadvantages. So our definition of the generating function was equal was that if you had a probability distribution P(n) of a random variable n then I define the generating function as this summed over all allowed values of n. That was my definition of the generating function.

On the other hand I could also ask oh the disadvantage of this was that if I want various moments of this distribution with various powers of n here I have to go on differentiating this but if I differentiate this k times then I actually end up with plus 1. This average was d k f of z over d z to the power k at z = 1. So I actually find the factorial moments rather than the moments themselves okay.

To overcome that disadvantage it is advantageous to define what is called a moment generating function and that is defined as M of u, I will use a different variable u instead of z so as not to confuse the two. This is equal to the average value of e to the power u n okay.

But the average value of e to the power u n is summation over the allowed values of n P(n) e to the u to the power n okay; u to the power n is e to the n u.

(Refer Slide Time: 51:03)



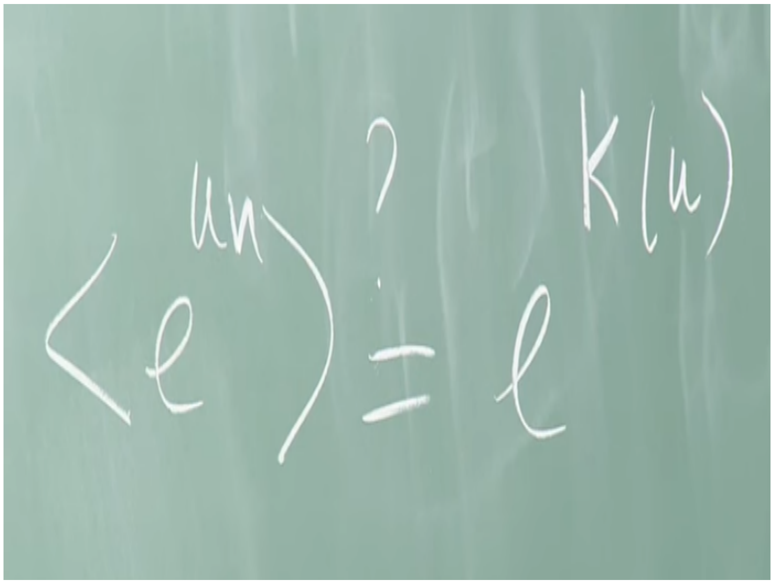
So this is all it is and therefore we have our relation which says the generating function and the moment generating function are related by M of u is f of e to the power u. So instead of z to the power n I got e to the u to the power n okay. So if I know f I know M and vice versa. The great advantage of this is that if I differentiate this after all if I take this and write this as summation n = 0 to k = 0 to infinity out here u to the power k n to the power k over k! expectation because that is a random variable n here then immediately you see that n to the power k is the derivative, the kth derivative of M of u over d u k at u = 0.

So directly by differentiation generates the moments themselves and not just the factorial moments. So that is why it is called the moment generating function and it directly gives you the various moments out here. So that is an advantage and I will stop here but we have seen we have seen that the moments themselves are not that helpful first of all. There is an average which you have got to remove first of all.

So it is not advantageous to use the moments but it is much better to find out what is the average value of n - n average to the power k and take its expectation value. These are the central moments. These are just the moments but these are the central moments. The moments about the mean value so any systematic drift or shift is got rid of. For example the second of these fellows when k is 2 is precisely the variance as you can see.

Even this is not very useful when you have many random variables because the sum of the central moments of independent random variables is not the sum of the kth moment of the sum is not the sum of the kth moments, central moments. So we need to do exactly what we did for the variance, get rid of those extra contributions and make it purely additive and that is where the cumulants comes in. This is a little bit like if you are used to statistical mechanics it is a little bit like going from the partition function to the Helmholtz free energy.

(Refer Slide Time: 53:43)



So what you have to do is to ask can I write, can I write e to the power un as equal to e to the power some function of u which is in a power series here in general okay. So this is like saying I take the partition function and then I take its you know the free energy is defined as the log of this guy this partition function, can I do that trick or not and the k of u will be called the cumulant generating function.

In fact that is the basic idea in statistical mechanics when you are writing the free the partition function you are writing a generating function actually for some distribution and when you find the free energy you are writing the corresponding cumulant generating function and that is what is additive etc. So we will talk about this next time okay.