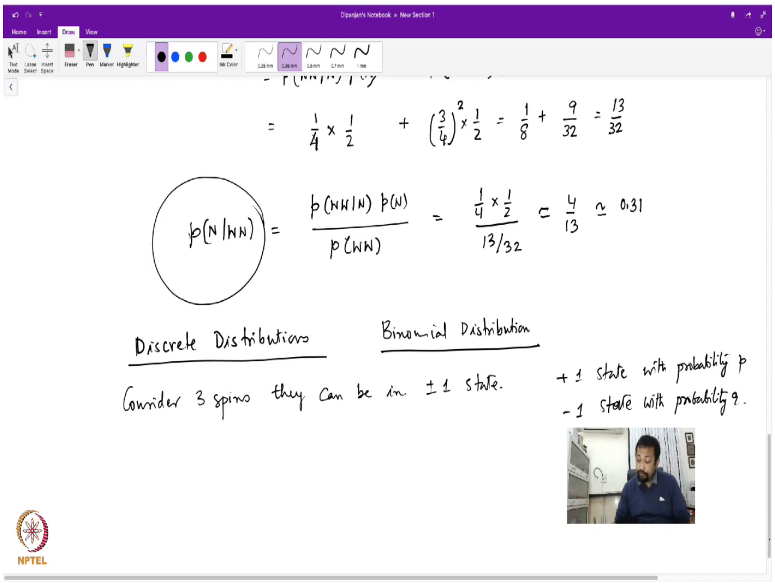
Statistical Mechanics
Prof. Dipanjan Chakraborty
Department of Physical Sciences
Indian Institute of Science Education and Research, Mohali

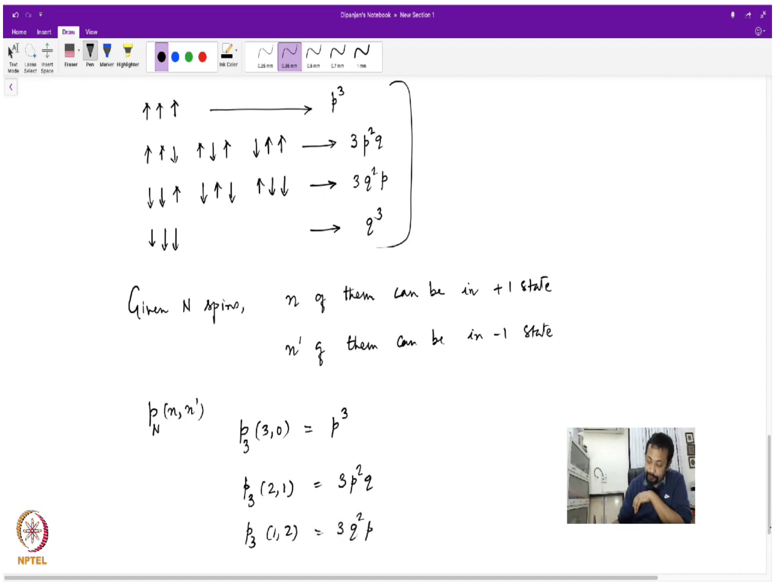
Lecture – 18
Discrete and Continuous Distributions

(Refer Slide Time: 00:15)



Now, binomial distribution is a distribution which will often encounter in physics and therefore, it is apt to start with the physical example to illustrate this. So, consider 3 spins, magnetic spins; let us say they can be in plus minus 1 state, right. So, plus 1 state with probability p and minus 1 state with probability q.

(Refer Slide Time: 01:12)

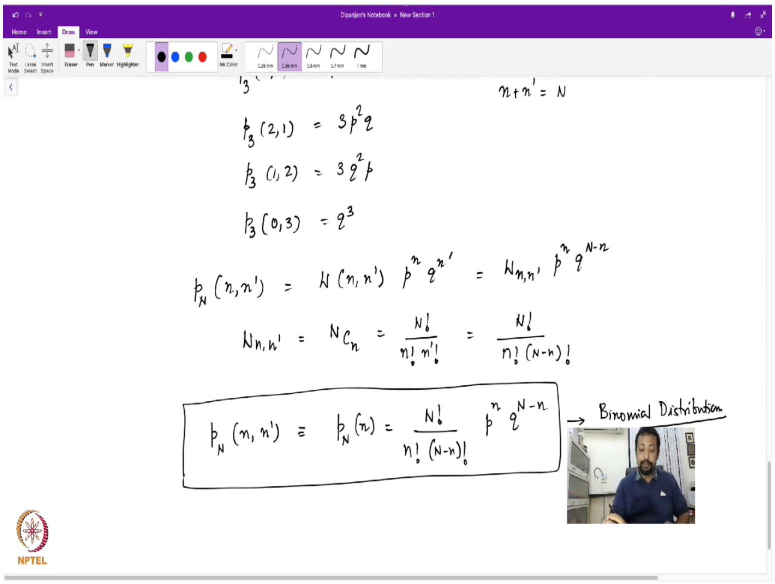


So, what are the possible configurations? Let us see all three of them can be in the upstate, two of them can be downstate. So, I think we will just let me draw it in the .So, two of them can be in the upstate, one of them can be in the down state and that can happen in three configurations as you can see, right. Now, two of them can be in the downstate.

So, which means I can have like this and I can have this possible way and then finally, three of them can be in the downstate. The probability for this to happen is p cube and the probability for this to happen is 3 p square q, for this 3 q square p, and then this q, right. So, we clearly see that there is a general structure that is energy from this.

So, we can given N spins, n of them can be in the upstate the plus 1 state, and n prime of them can be in minus 1 state, right. So, therefore, we want to ask p N of n comma n prime, correct. So, for example, in the example what we have seen in the illustrated just little while ago, I want to say p 3 3 comma 0 and that answer is p cube; p 3 2 comma 1 is 3 p square q, p 3 1 comma 2 is 3 q square p. And finally, p 3 0 comma 3 is q cube, right.

(Refer Slide Time: 03:33)

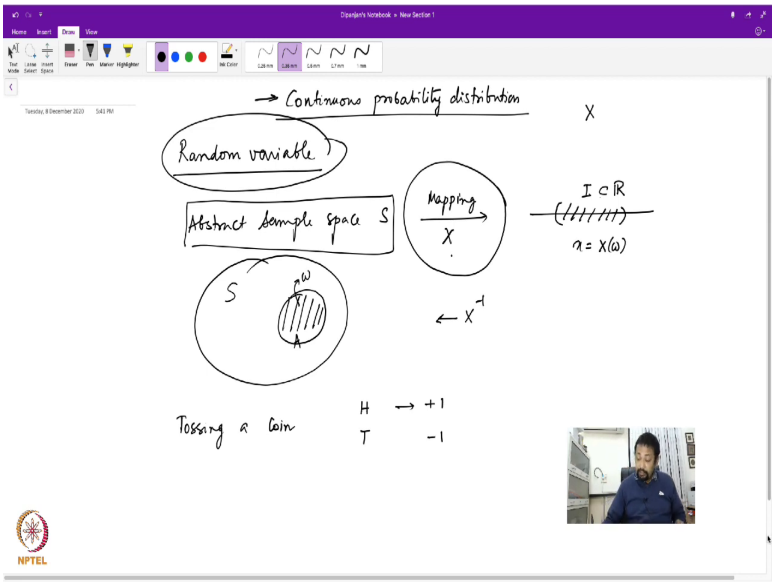


So, the general expression is p N n comma n prime is W n comma n prime p to the power n q to the power n prime; of course we also know that n plus n prime is going to be capital N. So, we can simplify this W and n prime, let us just not write it like this way; p to the power n q to the power N minus n, where W n comma n prime this you already can figure out now.

We have already got our idea is N c n, which is N factorial, n factorial and n prime small n factorial, small n prime factorial which we can simply write down as small n factorial N minus small n factorial.

And therefore, p of N n comma n prime is going to be, well is equivalent let us write down p N comma n is N factorial small n factorial N minus small n factorial p to the power n q to the power N minus n. And this is the first example of a discrete distribution or rather this is the only example for the skill distribution that we look at and this called the binomial distribution. This is the one we are going to use in statistical mechanics and therefore, it is better that we have introduced it at this particular point.

(Refer Slide Time: 05:24)



Now, so far whatever we have been doing, we have been doing with the discrete probability distribution. So, we now want to look at continuous probability distributions, continuous probability distributions. All our discussions on probability theory has been based on discrete events, right. So, for example, if I toss a coin, I know that I am and I am bound to get either head or a tail; if I roll a die, I am going to get discrete numbers like 1, 2, 3, 4, 5, 6.

If it is an end phase state will continue all the way up to n depending on the value of n. But now my interest lies in a completely different way; I want to define a continuous probability distribution and for that, I have to define what is called a random variable, right. So, it is clear that, the name continuous itself bears certain things, that this particular random variable.

Let us say I call this variable X can take on values on the real line, right. So, the definite finding a random variable formerly is essentially done from the abstract sample space. So, we had the abstract sample space S and we define a mapping and this mapping is X, that takes it on to the real line, an interval in the real line, right. So, that X is sorry, this has to be R, such that one has to be.

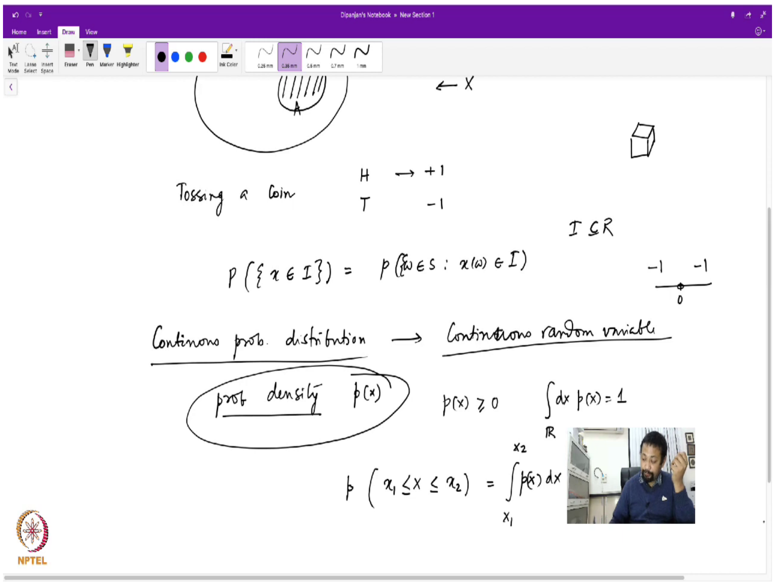
So, let us rephrase this little bit. So, the mapping is, now so far whatever probability theory we have developed or we have learned, we have discussed so far has always been concerned with discrete distributions, right. So, events are discrete, the outcomes are discrete. So, if you toss a coin, you get either a head or a tail; if you roll a die, you get a numbers discrete numbers 1, 2, 3, 4 and so on and so forth depending on the number of faces that I has.

But now my interest lies in something different; I want to discuss about a continuous probability distribution and for that, one has to talk about a random variable. Clearly the word continuous tells you that, this process, this random process of this random variable; the values of this can be will lie on the real line, it can take continuous values. So, the definition of a random variable formally happens by define a mapping, which is X from the abstract sample space S to an interval in the real life R, right.

Provided one has to be very clear about it; provided the inverse mapping of every interval will correspond to an event in the sample space. So, there is an event in the sample space and corresponding to this event, there is there are outcomes omega. So, you define a mapping, such that it takes you from the sample space to the real line; but one has to be very clear that the inverse mapping X inverse must correspond of this interval must corresponds to an event in the sample space.

A simple example is for tossing a coin, I have the sample space as head and tail; but now I can assign numbers to it, which means that I can say that this is plus 1 and this is minus 1 and this is remarkably similar to our spin system that we discussed a little while ago, right. So, having defined this mapping, one is left out by defining the probability.

(Refer Slide Time: 09:42)



So, starting from the probability measure in the sample space, one defines an induced probability, that is the probability that x belongs to I is equal to probability that omega belongs to S right, such that x of omega belongs to I.

So, essentially what you are saying is that the probability for the random variable x to take the values is in the interval I, which is a subset of R. So, I is a subset of R is defined as the probability of the event in the sample space, yes. Now, for a continuous probability distribution, probability distribution; the idea of absolute probability.

If it is which can take, so this variable, since it is a random variable now, which can take continuous values within the real line; the concept of the probability for random variable is very for a particular value of x is extremely difficult, right. So, for example, if you roll a die, if you roll a die which has six faces; then you can ask what is the probability that you will get a number between 1 and N, if it is an N faced die right, that is absolutely a valid question.

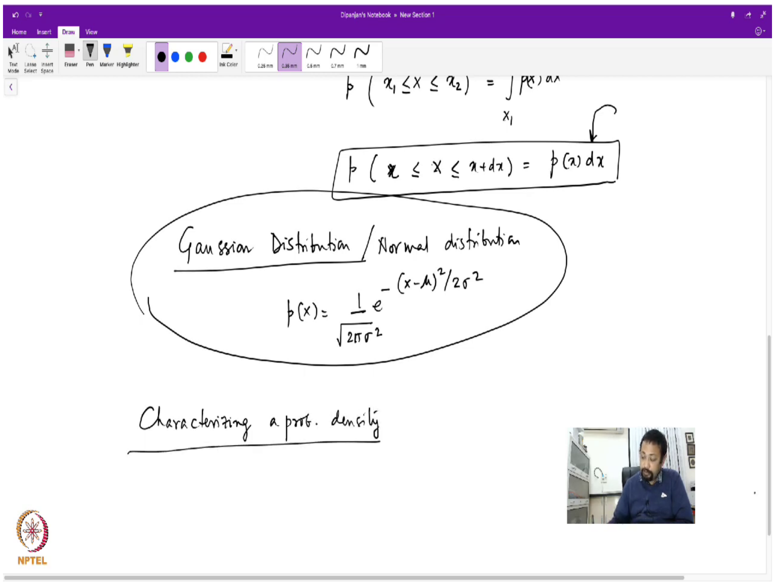
But then, if you ask a similar question for a continuous very random variable. So, we will say will. So, if you ask the same question for a continuous random variable, then that question itself does not make any sense. So, if a random variable is such that it can take values continuously between the interval minus 1 and plus 1 in the real line right; then you ask what is the probability that random variable is going to be? The probability value of x is going to be 0.

It will never be 0, it can be 0000 0.0000001, it can be plus minus 1 and 0.0001. So, if this is 0, it can be within a very epsilon neighborhood of this. So, it is strictly the answer to that question is strictly 0; rather what we do is instead of looking at defining a distribution, we define a density, a probability density p of X, such that p of X is positive integration d X p of X is 1, this is the normalization condition.

And since it is a density, density is always a quantity divided by something divided by per unit. So, if you call recall density, real density of a fluid is mass divided by per unit volume. So, here also it is a probability density; therefore it means it is probability per unit quantity.

So, since this is probability, so probability essentially of let us write down a slightly different way. We will say probability of X taking a value between x 1 and x 2 is equal to p X d X integral X 1 to X 2.

(Refer Slide Time: 13:17)

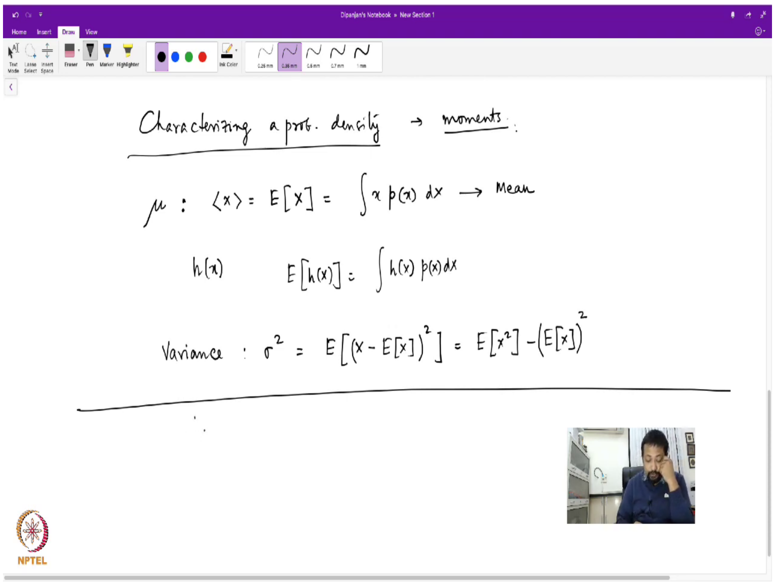


Or if you say probability of X small x capital X less than x plus d x; then this probability becomes p X d X, this is the probability. In reality when we looked at discrete distributions and discrete events, we said that the probability must be less than 1; because sum total of all probabilities for all the outcomes in these unity, but that is not true for density.

So, do not be surprised, density can have values more than 1; but this quantity should not have a value more than 1. And the most famous example of a continuous distribution is a Gaussian distribution or it is also called a normal distribution. And p of X is e to the power X minus mu whole square over twice sigma square, the normalization is 2 pi sigma square.

So, this distribution you are going to encounter plenty of times in statistical mechanics. So, characterizing a probability density and that characterization of a probability density happens by the moments.

(Refer Slide Time: 14:52)



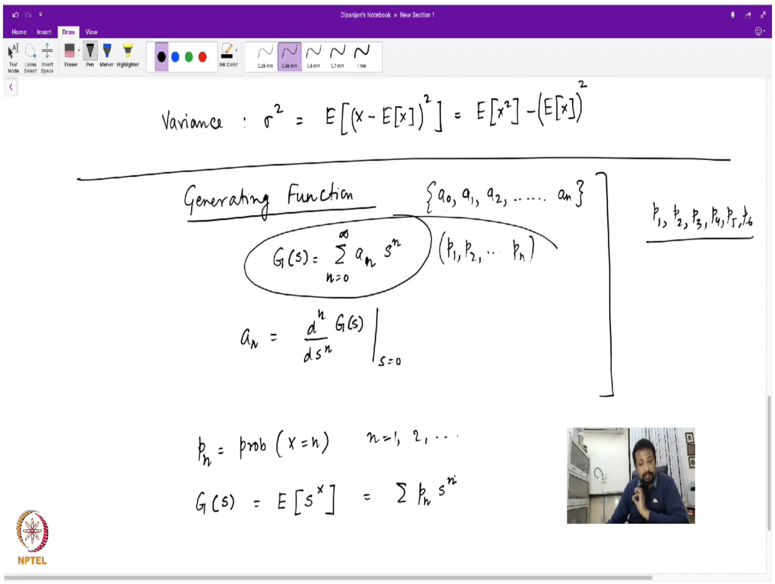
So, for example, the average of quantity random variable X; please note henceforth and all throughout our discussion when we are discussing random variable, the capital of for X denotes a random variable, whereas the small x denotes the values which this variable can take.

So, average of this quantity it is often done in this way is essentially x p x d X. If you define a function h of x of this random variable; then the average of h of X is integration h of X p X d

X, of course you have to put in the range. This one is called the mean of the probability density. The variance is average X square, which is E of.

So, this is, this is denoted by sigma square; whereas the mean is typically denoted by mu and that definition is E of X minus the average of X whole square. And if you formally expand this, you are going to get E of X square minus E of X whole square, right. One can also talk about covariances; but we will not look at them right now.

(Refer Slide Time: 16:58)



So, instead of continuing this further; I want to introduce something now, which is very interesting and it is extremely useful is what is called a generating function. And you will immediately see how useful this quantity can be. But in general, if I have a sequence a 0, a 1, a 2 all the way off up to a n; there is an alternative way.

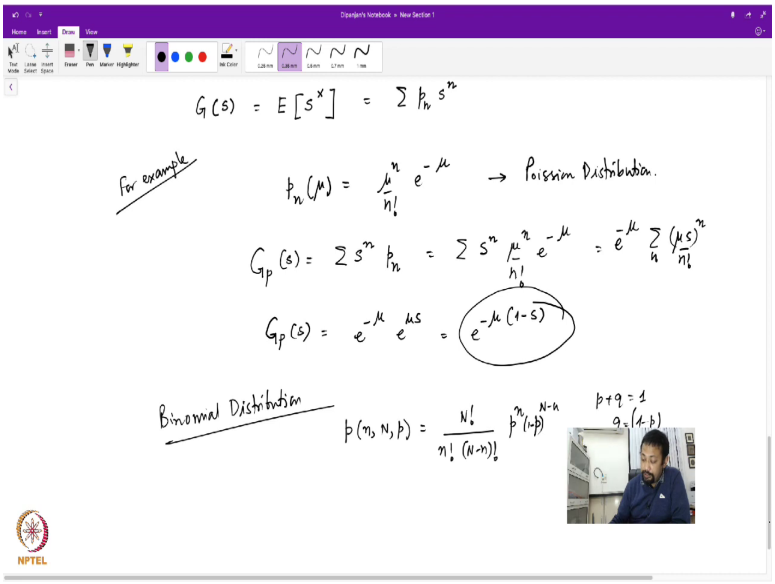
So, all of this stores, if these coefficient this sequence stores some kind of an information right; there is yet another alternative way of storing this information in the sense one defines a function G of s as sum over a n s to the power n, right.

So, let us say we will now sum over n equal to 0 to infinity, right. The sequence can be reconstructed back; how? Well, your a n's are going to be given by the nth derivative of G s with respect to s evaluated at s equal to 0. So, you can go from G s to a n very very easily.

Now, recall with this thing in hand, this sequence a 0, a 1, a 2 is very very similar to our discrete probability distributions, where we had p 1, p 2, p n right for the outcomes. For example, if you are looking at rolling a die, then you have the probability getting p 1, you have the probability of getting p 2 the number 2, p 3, p 4, p 5 and p 6 if it is a 6 face die.

If it is n face die, then of course, this sequence goes on all the way up to n, correct. So, we will define, for a discrete probability distribution we will define that, the probability p n is the probability; that your variable takes a value X equal to n, where n is 1, 2 these are discrete integers, right. Then G of s. So, I want to now use this terminology this thing that we have learned is essentially E of right, which is sum over p n s to the power n; because X can take values, integer values n, right.

(Refer Slide Time: 19:40)

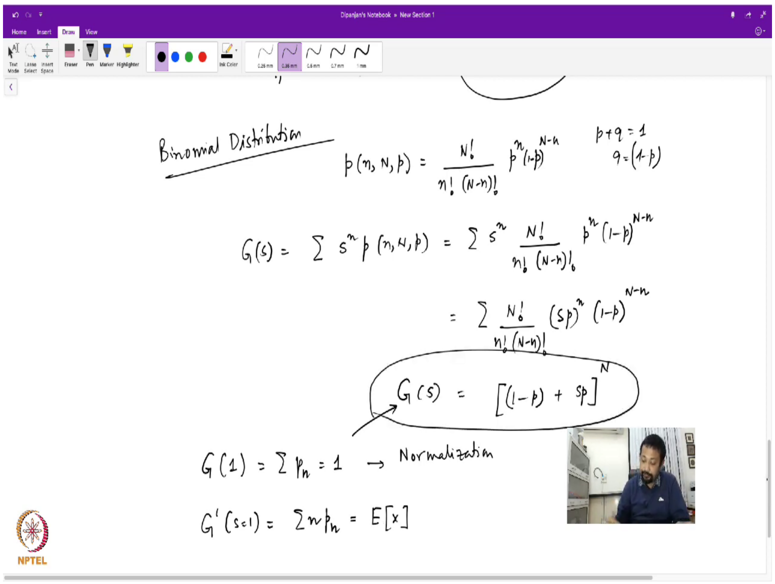


For example, let us take this as mu to the power n n factorial e to the power minus mu; this one we have not discussed, but this is what is called a Poisson distribution. And here I want to calculate the generating function. If I want to calculate the generating function, this is going to be s to the power n p n, which is sum over s to the power n mu to the power n n factorial e to the power minus mu, e to the power minus mu comes out and mu s raised to the power n n factorial, right. And therefore, this is e to the power mu.

If you look at this, this is just e to the power mu s. So, your generating factor is mu 1 minus s; your generating function for the Poisson distribution is just a simple expression of e to the power mu times 1 minus s, right. For the other example that we looked at for the discrete case, which is the binomial distribution right; there we had the probability of n, N, p was N factorial small n factorial N minus n factorial p to the power n, right.

So, this has to be 1 minus p. So, here what we have done is, we have written p plus q is equal to 1; because there are only two possible states and q is equal to 1 minus p and we have replaced it over here.

(Refer Slide Time: 21:59)

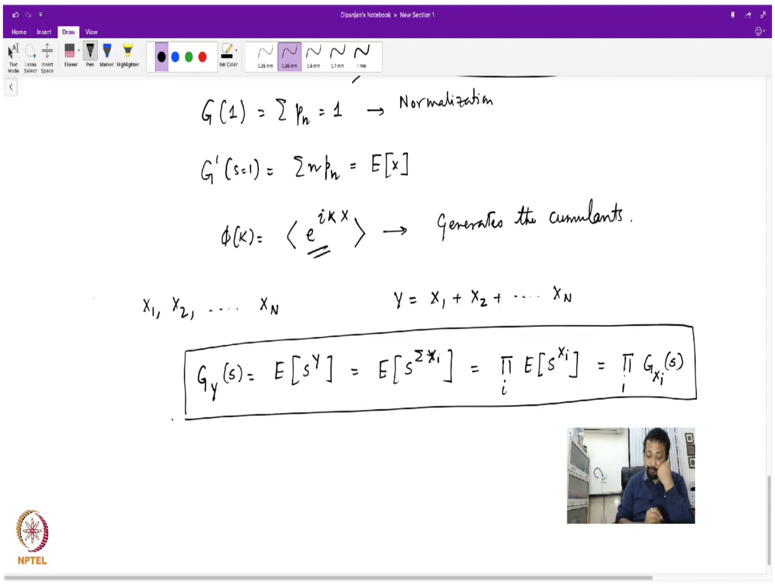


If I want to calculate the generating function for such distribution, we will be write down this as p n. So, that is going to be sum over s to the power n N factorial small n factorial N minus small n factorial p to the power n 1 minus p N minus n, correct.

If you are careful enough, so this I can simply write down as N factorial N minus n factorial N minus n factorial s p raised to the power n 1 minus p N minus n. So, this clearly is 1 minus p plus s times p raised to the power N. So, all the information that you encode is essentially contained over here. It is also worthwhile to know that for any random variable G of 1 is sum over p n, which must be equal to 1, right.

So, for example, here also if you substitute s equal to 1, you will get 1; in the formal example also if you substitute s equal to 1, you are going to get 1 and that is the normalization, right. If you look at G prime of s at s equal to 1; then you get sum of n p n right, which is nothing but the mean that you looked at. So, effectively you can get all the information, reconstruct all the information from this generating function.

(Refer Slide Time: 24:03)



For random continuous random variables, what we look at is essentially not as sum over s to the power n, p to the power n; we define what is called a characteristic function phi of k as e to the power i k X. But we shall come back to it later, we will not discuss it over here and this essentially generates the cumulants; you can easily verify this by expanding the exponential, right.

Now, suppose coming back to the generating function that we were discussing. Suppose that I have random variables X 1, X 2, X N and I define a random variable Y which is X 1 plus X 2 plus all the way up to X N right; then G of Y s is variance of s to the power Y, which is variance of s to the power X i.

And if these are independent random variables, then essentially you are going to get product of i e to the power s of X i right, which is product over i G of X i times s. So, the condition also, since we did not discuss covariance; so this is another way of checking whether your random variables are truly independent is to check whether this is satisfied, right.