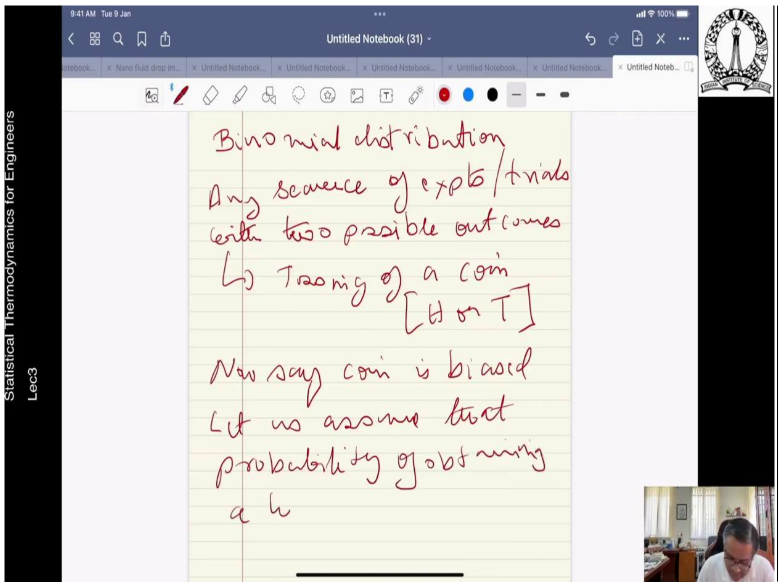
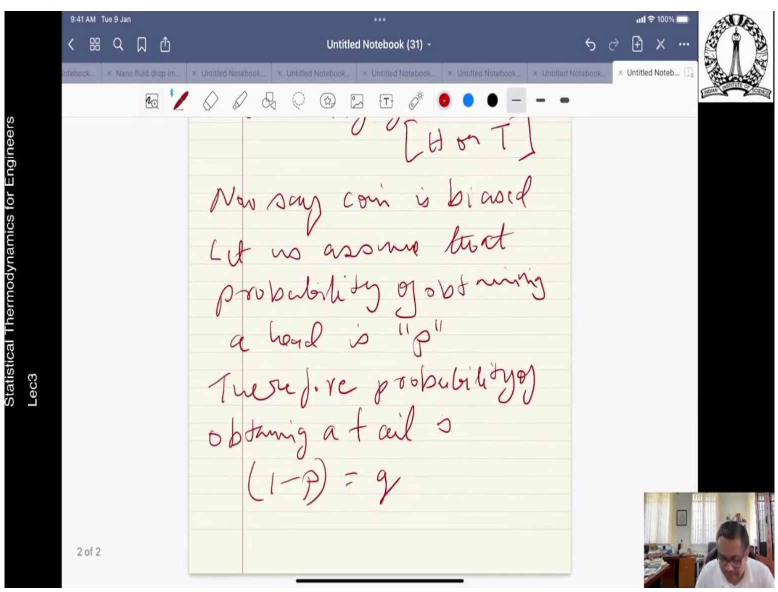
Statistical Thermodynamics for Engineers
Professor Saptarshi Basu
Indian Institute of Science, Bangalore
Lecture 03
Important Probability Distributions

(Refer Slide Time: 00:02)

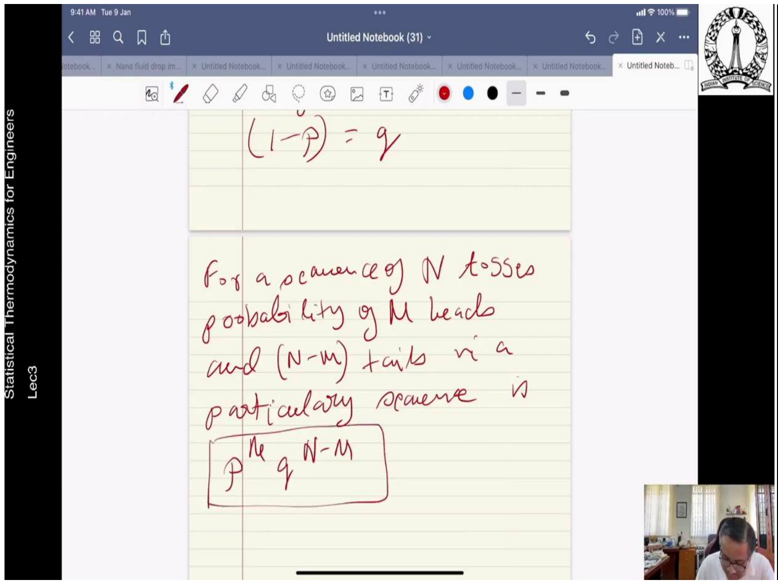


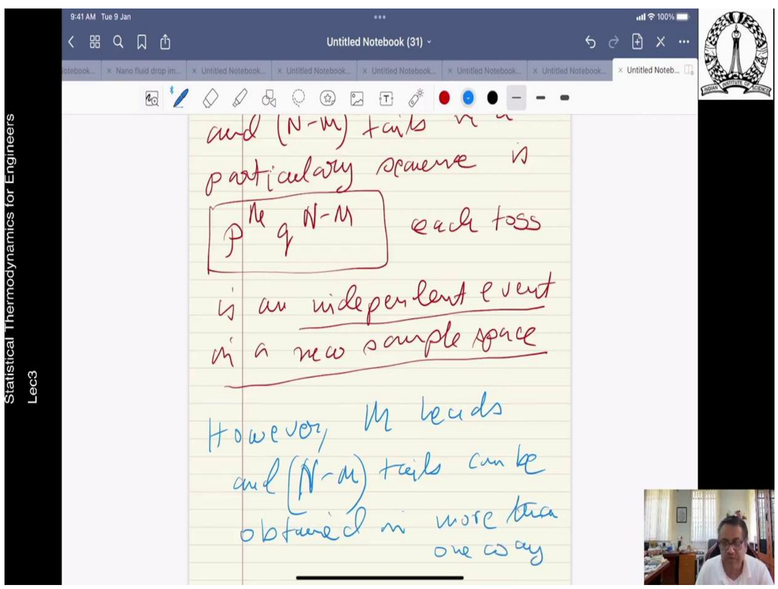


Welcome to Lecture 3 of this particular course which is statistical thermodynamics. So, in this particular course, we are going to look at this particular lecture will cover is a continuation of the previous lecture, which is basically on probability and statistics. So, now, let us take a look that we covered the different permutation combinations and stuff like that, today we are going to look at binomial distribution first. So, binomial distribution, so, it is any sequence of experiments or trials with two possible outcomes example, say for example, tossing a coin this has got two possible outcomes that outcome is head or tail, so, there is only two possible outcomes.

Now, say that the coin is biased. So, normally heads or tails should have (equal) equally likely when you toss a coin and say that the coin is biased. So, let us assume that probability of obtaining a head is say P, P is a probability of obtaining a head therefore, since there are only two possible outcomes, therefore, probability of obtaining a tail is 1 minus probability of obtaining a head and let us call it q. So, you understand because the sum total of these two probabilities should be 1. So, it could be either a head or a tail. So, this is the (())(02:33).

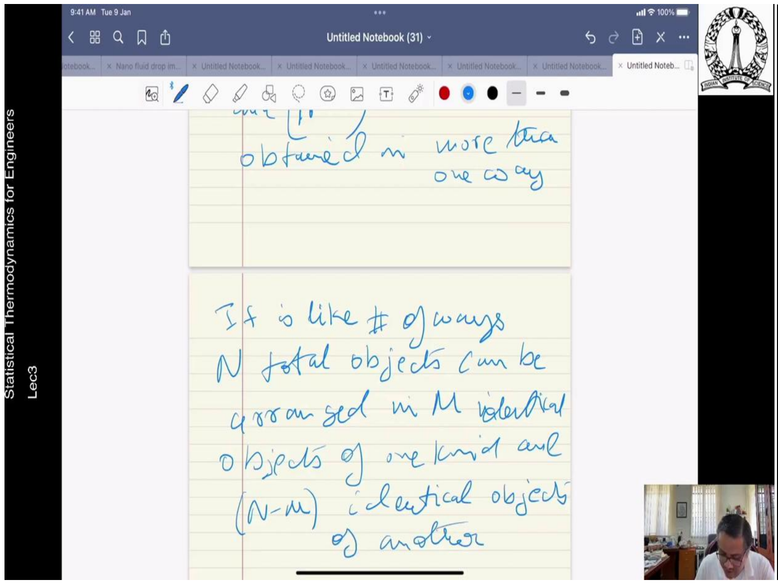
(Refer Slide Time: 02:35)

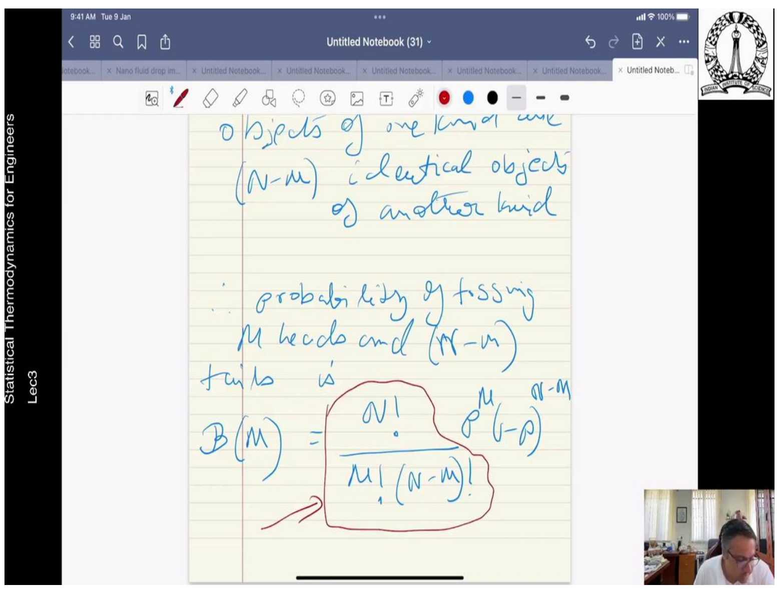


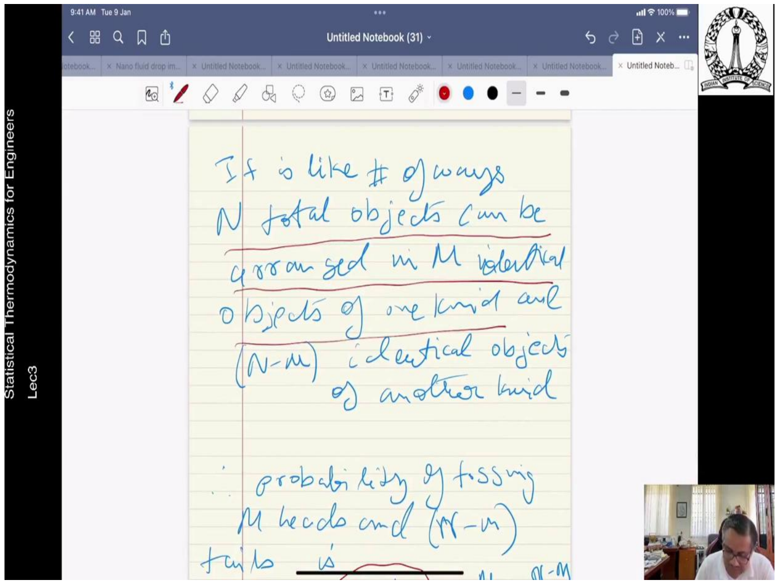


Now, let us assume now, for a sequence of N tosses, so, you toss the coin N number of times the probability of M heads and therefore, N minus M tails in a particular sequence, is P to the power M, q to the power N minus M as each toss why this is why this is possible, that is because each toss as we can see each toss is an independent event in a new sample space. So, that is the probability that for a sequence of N tosses probability of M heads and N minus M tails in a particular sequence is given by P to the power of M multiplied by q to the power of N minus M as each toss is an independent event in a new sample space. However, now, the little bit of twist however M heads and N minus M tails can be obtained in more than one way.

(Refer Slide Time: 04:54)

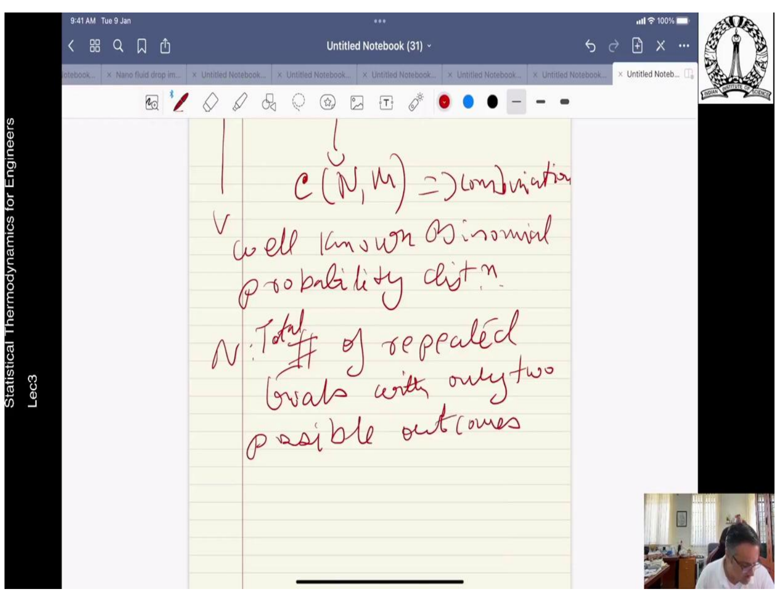


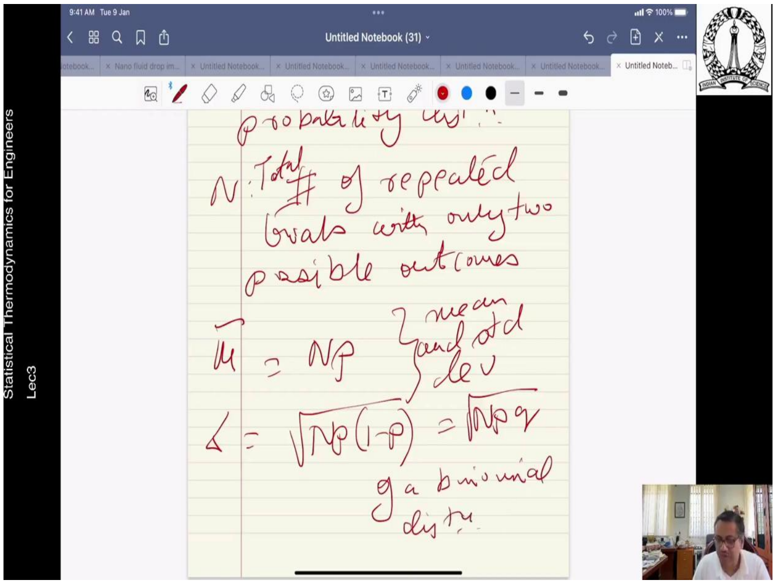




So, what is that it is like number of ways in total objects can be arranged in M identical objects one kind and N minus M identical objects of another kind? So, therefore, probability of tossing M heads and N minus M tails is given by B M given by N factorial divided by M factorial N minus M factorial and P to the power M 1 minus P to the power of N minus M. So, this should be clear this we already learned from our previous lecture. So, it is like how N total number of objects can be arranged in M identical objects of one kind and N minus M identical objects at another kind. So, therefore, the probability becomes this multiplied by that.

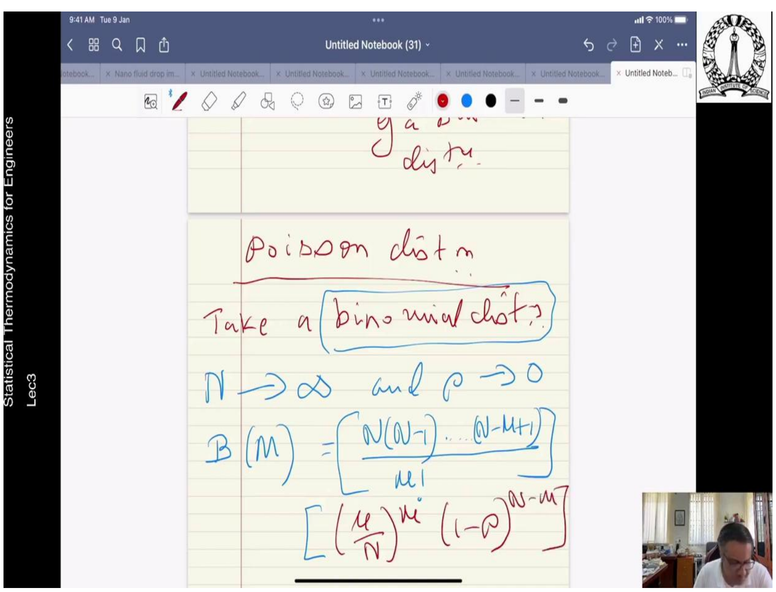
(Refer Slide Time: 07:22)





So, this is nothing but, if you look at this C N comma M which is basically the combination that we covered in the last class. So, it is a combination. So, this once again is the well known binomial probability distribution. N is the total number of repeated trials with only two possible outcomes. So, that is what it is and the mean of such a distribution which is N bar is given as N P and the standard deviation is given us N P (1 minus P) which is also written as N P q. So, this is the mean and the standard deviation mean and standard deviation of binomial distribution the distribution is the mean and standard deviation of binomial distribution.

(Refer Slide Time: 09:14)



Statistical Thermodynamics for Engineers
Lec3

Untitled Notebook (31)

$$N \to \infty \text{ and } p \to 0$$

$$B(M) = \left[\frac{N(N-1)\cdots(N-M+1)}{M!}\right]\left[\left(\frac{\mu}{N}\right)^{M}(1-p)^{N-M}\right]$$

$$\mu = \bar{M} = Np \text{ : mean}$$

$$\lim_{N\to\infty} B(M) = \frac{N^{M}}{M!}\left(\frac{\mu}{N}\right)^{M}(1-p)^{N}$$

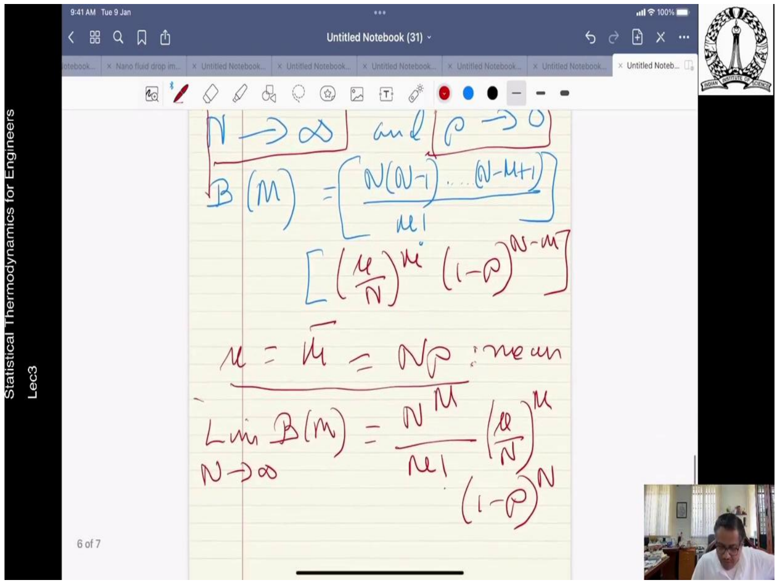
Statistical Thermodynamics for Engineers
Lec3

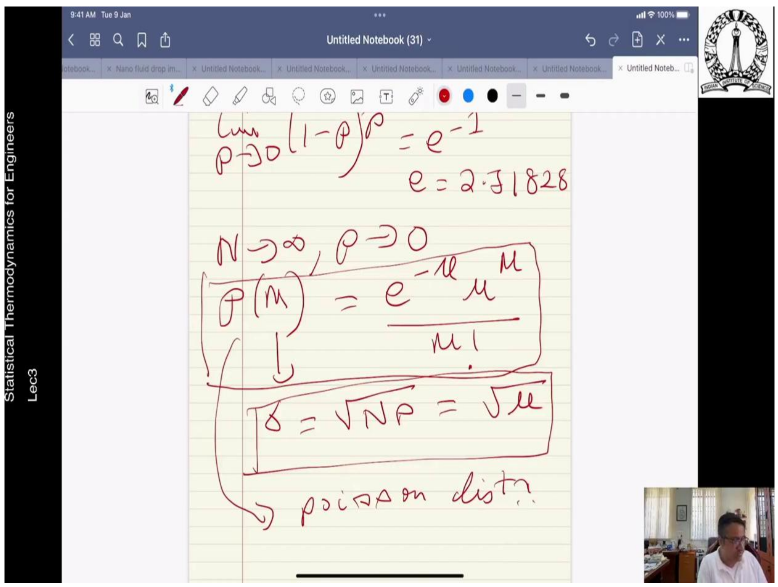
Untitled Notebook (31)

$$\lim_{N\to\infty} B(M) = \frac{\mu^{M}}{M!}(1-p)^{\mu/p}$$

$$\lim_{p\to 0}(1-p)^{1/p} = e^{-1}$$

$$e = 2.71828$$





So, this is already something that we know and we have read it earlier. So, let us look at now the next possible distribution which is called Poisson distribution. So, what is a Poisson distribution is like a binomial distribution if so, you take a binomial distribution only in this case N is very large that is it approaches infinity and very large number of trials and P goes to 0, that means the probability of M is very low with each trial.

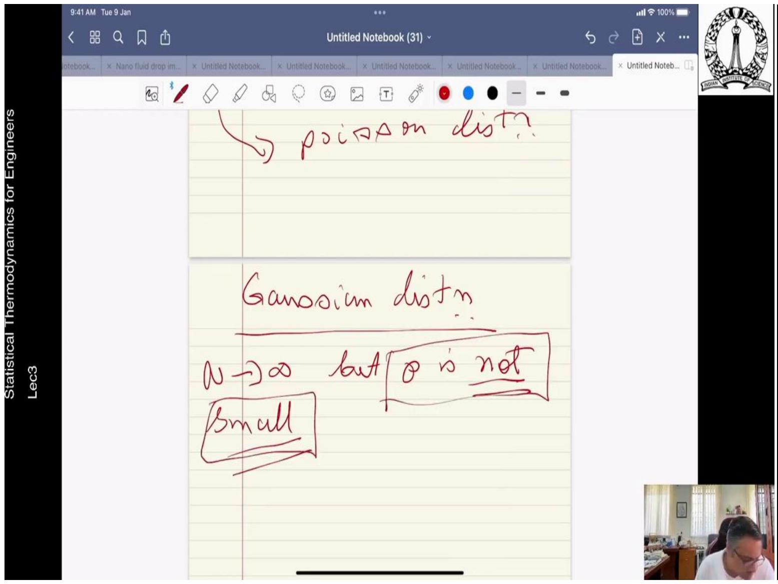
So, you can think of it as a very biased on individual that heads or tails and not half of head the probability of getting ahead is getting low. So, in this case, if you write the binomial distribution you have just open up things a little bit N minus M plus 1 divided by M factorial. So, this entire thing that multiplied by mu N to the power M 1 minus P to the power N minus M. Where mu (())(10:54) M bar is equal to NP.

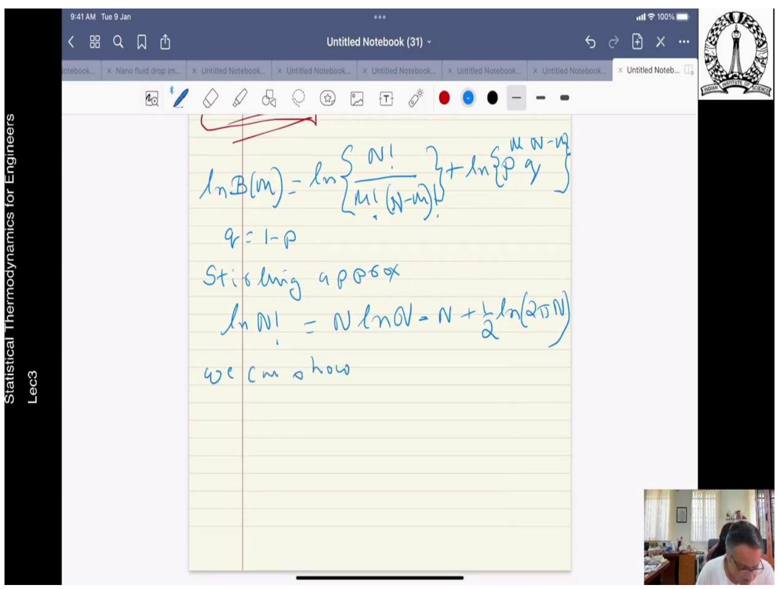
So, this is basically the mean. So, in the limit binomial distribution where N approaches infinity this becomes N raised to power M divided by N factorial mu by N raised to the power M into 1 minus P raised to the power N. So, if we revise it further N approaches infinity BM is given as mu N into N factorial 1 minus p raised to the power mu by P. In the limit P goes to 0, 1 minus P raised to the power of 1 by P is given as e to the power minus 1 is basically as you know 2.71828.

So, in the combined limit and basically N goes to infinity P goes to 0, P raised to poisson distribution is e to the power of minus mu, mu to the power M divided by M factorial. So, in the case of the poisson distribution the sigma is given as N P which is nothing but root over mu.

So, this is the poisson distribution. So, the poisson distribution if we just recap it is when N is very large P is very small. So, what we have done is that we have taken the binomial distribution applied the two limits and we ultimately get this where the standard deviation is given values. So, this is the poisson distribution**.** Which some of you may be familiar with also.

(Refer Slide Time: 13:30)





Statistical Thermodynamics for Engineers

Lec3

Untitled Notebook (31)

$$\ln B(M) = \ln\left\{\frac{N!}{M!\,(N-M)!}\right\} + \ln\left\{p^{M} q^{N-M}\right\}$$

$q = 1 - p$

Stirling approx

$$\ln N! = N\ln N - N + \frac{1}{2}\ln(2\pi N)$$

we can show

$$\ln B(M) = -\frac{1}{2}\ln\left\{\frac{2\pi M}{N}(N-M)\right\} \quad (1)$$

$$+ \ln\left[\frac{Np}{M}\right]^{M} \quad (2)$$

$$+ \ln\left[\frac{Nq}{N-M}\right]^{N-M} \quad (3)$$

Statistical Thermodynamics for Engineers

Lec3

Untitled Notebook (31)

$$+ \ln\left[\frac{Np}{M}\right]^{M} \quad (2)$$

$$+ \ln\left[\frac{Nq}{N-M}\right]^{N-M} \quad (3)$$

Define

$$y = M - \bar{M} = M - Np$$

$$\therefore \quad \frac{M}{N} = \frac{y}{N} + p$$

and

$$\frac{N-M}{N} = q - \frac{y}{N}$$

Statistical Thermodynamics for Engineers
Lec3

and $\frac{N-M}{N} = q - \frac{y}{N}$

(1) term I

$$\lim_{N\to\infty} \frac{M}{N}(N-M) = \lim_{N\to\infty} N\left(\frac{y}{N}+p\right)\left(q-\frac{y}{N}\right)$$

$$= Npq$$

$\frac{y}{N}$ scales with the relative width of the binomial distn $\frac{1}{\sqrt{N}}$ dependence

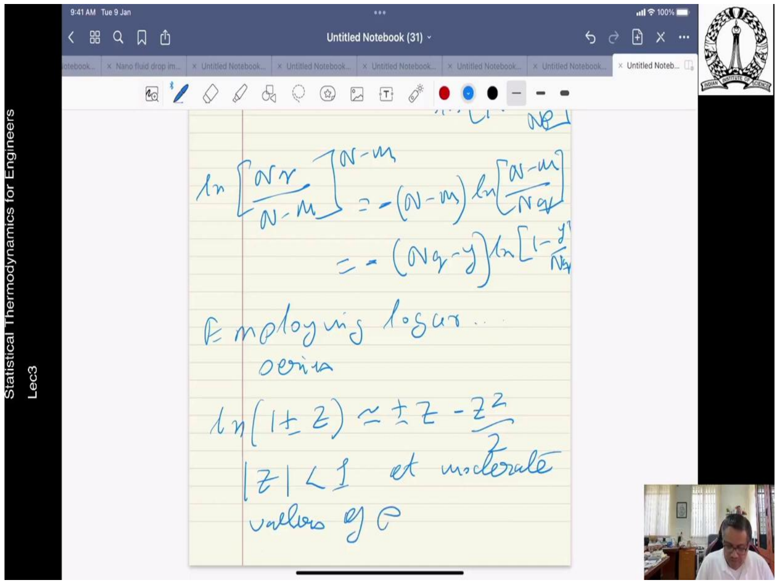
Statistical Thermodynamics for Engineers
Lec3

$\frac{y}{N}$ scales with the relative width of the binomial distn $\frac{1}{\sqrt{N}}$ dependence

For remaining two term

$$\ln\left[\frac{Np}{M}\right]^M = -M\ln\left[\frac{M}{Np}\right]$$

$$= -(y+Np)\ln\left[1+\frac{y}{Np}\right]$$



Now let us look at the next distribution which is basically called the Gaussian distribution. What is a Gaussian distribution this time your N is large but P is not, not small. So, unlike the Poisson distribution that e was also small in this case P is not small. But N is the large so what we have over here is let us take the log the (())(14:10) smaller ones (())(14:12) (log N factorial by M factorial N minus M factorial plus log P raised M q N minus M).

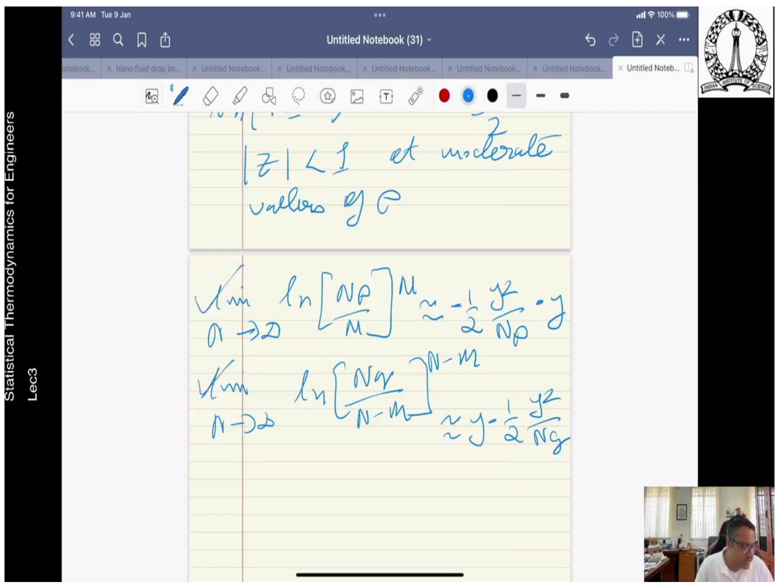
So, there is just a log q equal 1 minus P getting apply what we call a stirling approximation. Stirling approximations used that logarithm N factorial N into ln N minus N plus half ln 2 pi N. So, this is a Stirling approximation. So, we can show ln binomial distribution is given as and this is a little long half ln 2 pi M by N N minus M it is called as Term-I plus ln NP by M raised to power M it is called as 2 plus ln Nq N minus M N minus M it is called as 3. So, these are the three terms now, let us define y N minus M bar basically M minus NP. Therefore, M by N is equal to y by N plus P and N minus M by N q minus y by N. So, now, we take term by term here we define three terms remember.

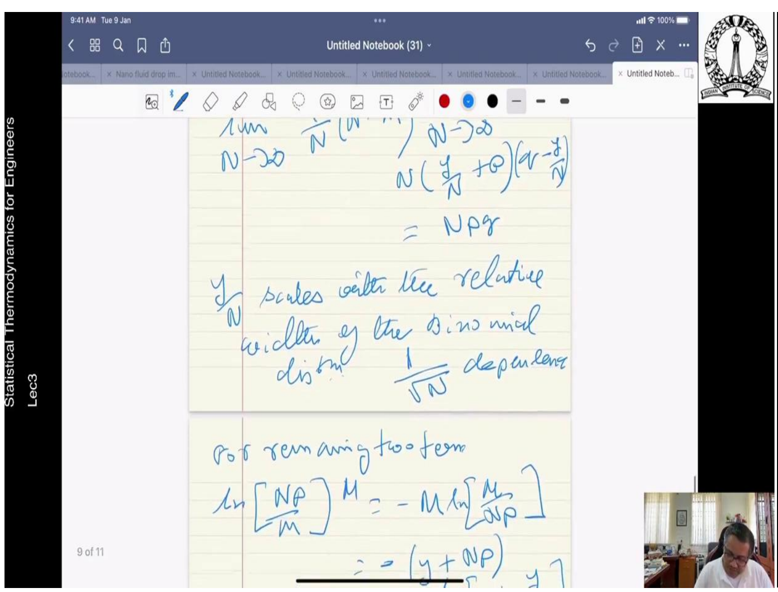
So, we take term by term let us take term 1 step term 1 term 1 in the limit as N goes to infinity mu by N N minus M is given as a limit n goes to infinity write in the next line y by N plus P into q minus y by N. This is equal to NPq. So, see if you can understand the first term (())(18:00) particular sequence this is the first one.

Now, as y by N scales with the relative width y the binomial distribution. We previously found to display this binomial distribution exhibits 1 over root under N dependence. So, for the remaining two terms now for the remaining two terms therefore, two terms ln NP by M raised to the power M minus M ln M by NP (())(19:17) plus y plus NP ln 1 plus y by NP.

Now, the other term which is in ln Nq divided by N minus M N minus M minus N minus M ln N minus M by Nq which is further given as minus Nq minus y ln 1 minus y by Nq. So, in for the employing logarithm series is basically tells you ln 1 plus minus z is almost equal to plus minus z minus z square by 2 for all z less than or not z less than 1 at moderate values of p.

(Refer Slide Time: 20:48)





Statistical Thermodynamics for Engineers
Lec3

9:41 AM Tue 9 Jan

Untitled Notebook (31)

$$\lim_{N \to \infty} \ln\left[\frac{Nq}{N-m}\right] \approx y - \frac{1}{2}\frac{y^2}{Nq}$$

Substituting

$$\lim_{N \to \infty} \ln B(y) = -\frac{1}{2}\ln(2\pi Npq) - \frac{y^2}{2N}\left(\frac{1}{p} + \frac{1}{q}\right)$$

$\Downarrow$

Gaussian distn

$$G(y) = \frac{1}{\sqrt{2\pi Npq}} \exp\left(-\frac{y^2}{2Npq}\right)$$

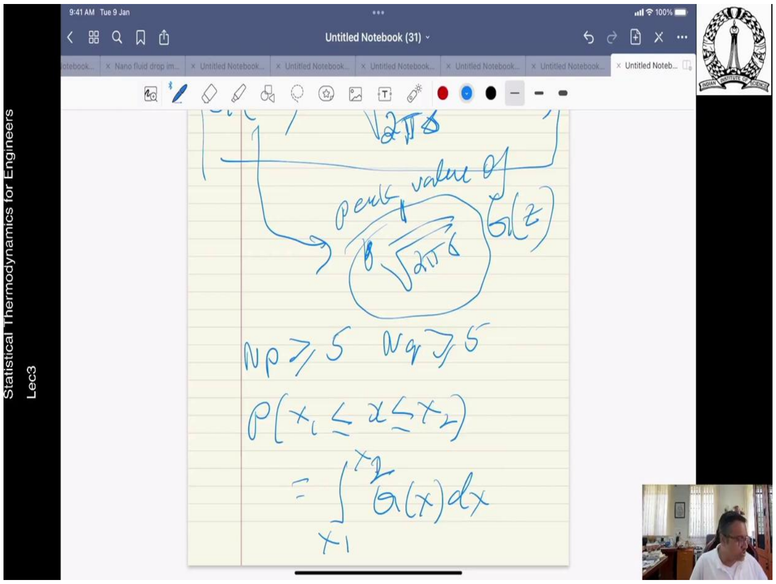
Statistical Thermodynamics for Engineers
Lec3

9:41 AM Tue 9 Jan

Untitled Notebook (31)

$$\left\{ \sigma^2 = Npq \right\}$$
$$\left\{ z = y/\sigma \right\}$$

$$G(z) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{z^2}{2}\right)$$



So, in a nutshell that translates to (())(20:52) minus 1 by 2 y square by NP minus y and in the limit ln Nq by N minus M N minus M is almost y minus half y square by Nq. So, now, if you substitute all these terms that means, this term this term so, of course, you are inputting term which is this, this, this particular term into the parent equation.

That we already know is the binomial equation in the logarithmic form we get substituting we get the limit N goes to infinity ln B y is equal to minus half ln 2 pi NPq minus y square by 2 N 1 by P plus 1 by q. So, this results in a Gaussian distribution which is given as G y which is nothing but root over 2pi NPq exponential minus y square 2NPq Gaussian distribution and of course.

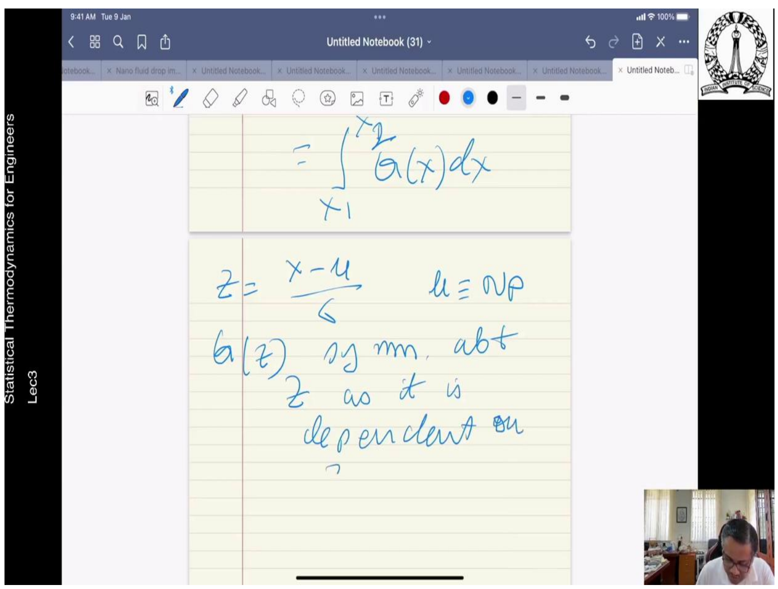
We know that sigma squared is equal NPq and define z as y over sigma we obtain you combine these two to obtain a more palatable form of Gaussian distribution which is one over 2 pi sigma exponential minus z square by 2. So, this is the Gaussian form that you may have encountered. So, for a continuous distribution that is the variable N must be replaced by its continuous analog x.

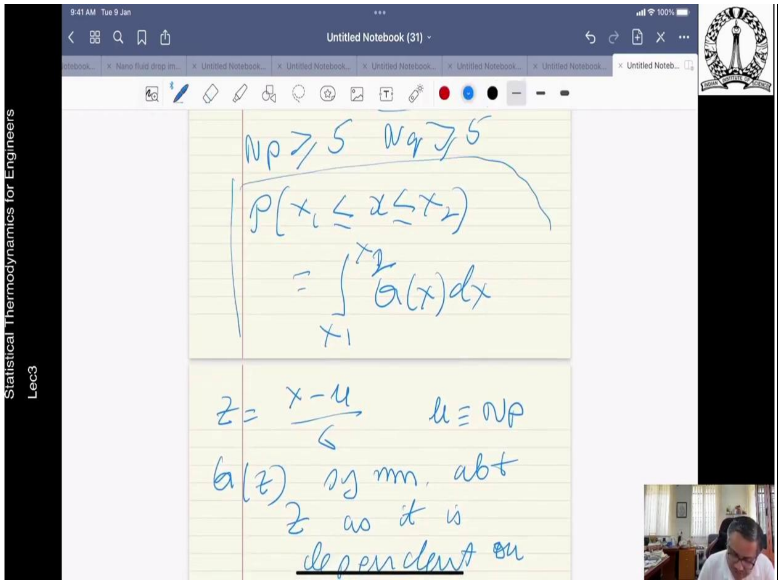
So, that the a Gaussian distribution as we know this is symmetric symmetrical about z because of its dependence on z square and unlike many of the cases for discrete binomial Poisson distributions, this Gaussian distribution indicates that the peak value of z is always at 1 over 2 pi sigma. So, this is always that is the peak value of this z.

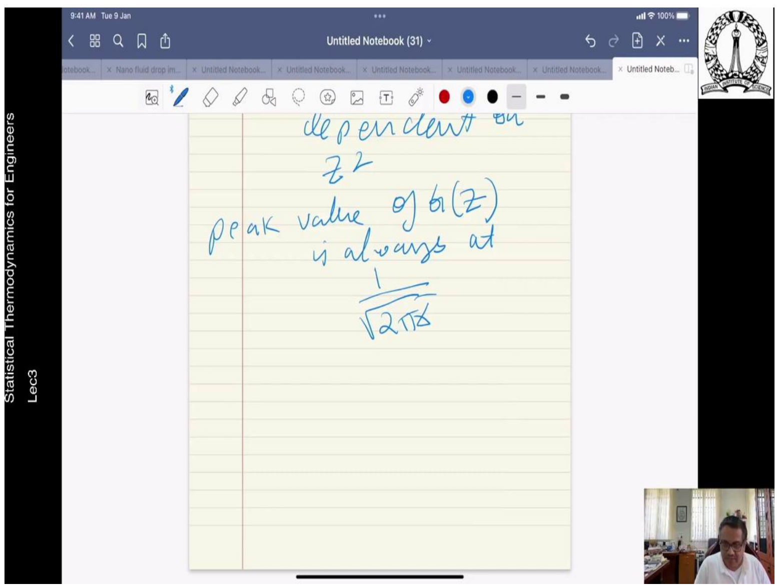
So, the Gaussian distribution is shown to be a satisfactory approximation of the binomial distribution if your NP is much is greater than 5 and Nq is also greater than 5. So, the Gaussian distribution (())(24:50) the probability for a specified range of variables therefore, is given less say if your x is between these two is essentially integrating the Gaussian distribution from over the (())(25:10).

So, that is what the Gaussian distribution is all about, and this is how we get Gaussian distribution as well. So, now that we are almost you know, so, the Gaussian distribution as you know it is very symmetric we already said that. So, we started with the binomial distribution and for the Poisson distribution, we saw that he was very small and N was large Gaussian distribution then if you go to Gaussian distribution your N is large (())(25:49) begin with.

(Refer Slide Time: 25:51)







So, few other things that for a Gaussian distribution your set is x minus mu by sigma mu is once again equal to NP. G z just write it down is symmetrical about z as it is dependent on the square value og G z as writing this explicitly is always at 1 by root over 2 pi sigma. So, let us see what else you have got.

So, this is the most convenient methods of calculating the probability this particular method when you actually have your probability follow a Gaussian distribution, which in many cases it does, it does follow the Gaussian distribution. And so, you can actually this is the most useful way of calculating a Gaussian distribution as we know Poisson distribution you may have already read in your undergraduate that is also very useful.

A few words that all of these are valid when there are only two outcomes, outcomes for a sequence of experiments, any sequence of experiments like tossing of a coin, like any other things, which has worked into outcomes, it is widely used not just in mobile economics and not in other sectors as well.

So, and also, where Gaussian distribution is applicable, it is applicable for a process which involves the photon counting, photon counting needs where the total number of photons are counted, because N is a very large number, but the possibility of observing a single photon is very low. So, that is like e approaches 0. So, that is one example of a Gaussian distribution that you may be wondering, that is what is does.

So, similarly, there would not be any other types of distributions also, but we are not going to go into the details in the next class, we are going to cover what we call the combinatorial analysis of statistical dynamics. So, that is going to be the subject of the next class which is going to where we are going to see some of the differences between distinguishable and indistinguishable objects and then go on to the purpose statistics also (())(28:40). So, thank you and this is the end of Lecture 3.