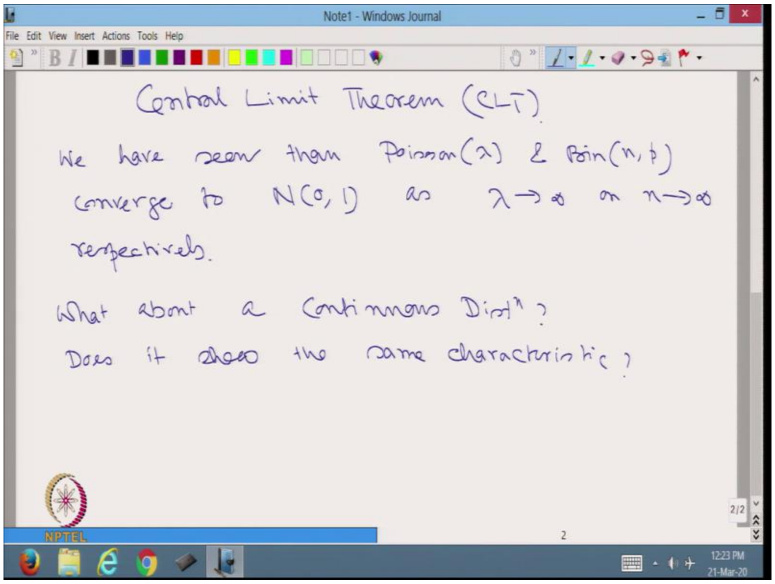
Advanced Probability Theory

Professor Niladri Chatterjee
Department of Mathematics
Indian Institute of Technology Delhi

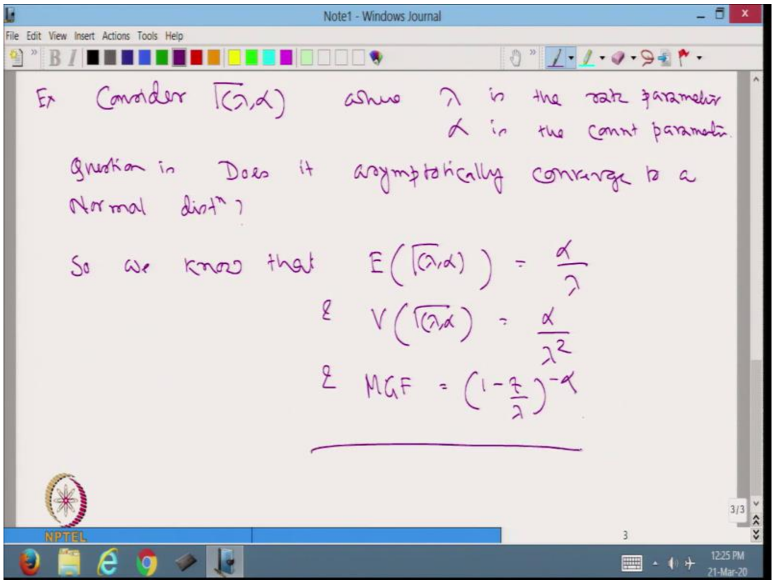
Lecture 30

(Refer Slide Time: 00:37)



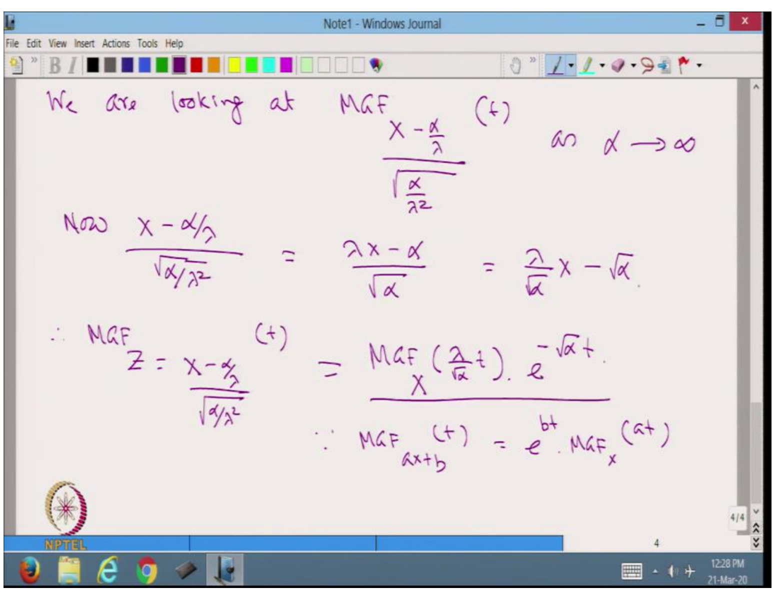
Welcome students to mock series of lectures on Advanced Probability Theory, this is lecture number 30 and also the last lecture of this series. If you remember we are working on Central Limit Theorem or CLT, we have already seen Poisson with parameter say lambda and binomial n, p converge to normal 0, 1 as lambda goes to infinity or n goes to infinity in the above cases respectively. What about a continuous distribution? Does it show the same characteristic? That is the question.

(Refer Slide Time: 02:01)



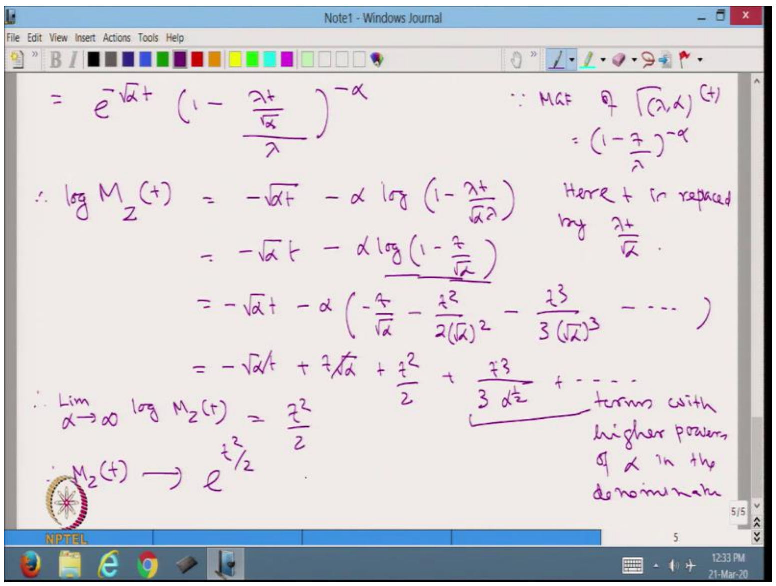
So, as an example consider gamma distribution with lambda alpha, where lambda is the rate parameter and alpha is the count parameter, question is, does it asymptotically converge to a normal distribution? So, we know that expected value of gamma lambda alpha variat is equal to alpha over lambda and variance of a gamma lambda alpha variat is equal to alpha over lambda square and moment generating function is equal to 1 minus t by lambda whole to the power minus alpha, this results we know.

(Refer Slide Time: 03:44)



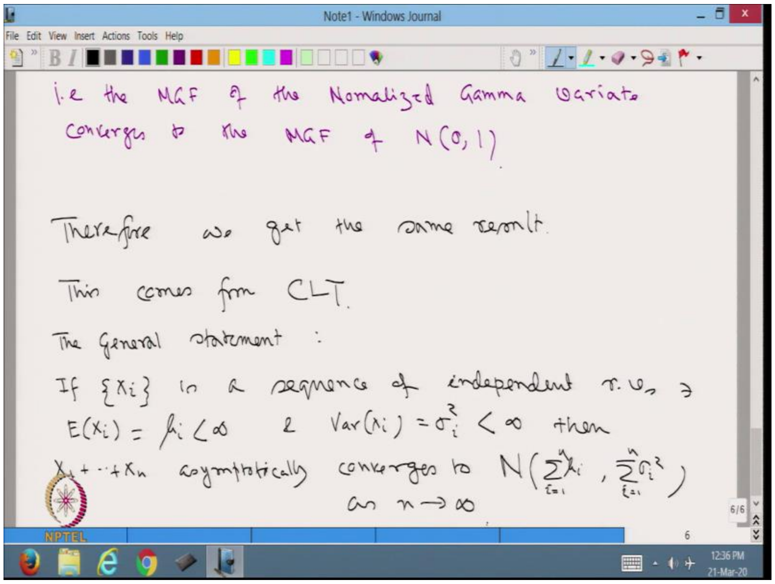
We are looking at MGF of X minus alpha over lambda root over alpha over lambda square at t as alpha goes to infinity. Now, X minus alpha over lambda divided by root over alpha by lambda square is equal to lambda X minus alpha divided by root over alpha is equal to lambda over root over alpha X minus root over alpha, therefore MGF of Z which is is equal to X minus alpha over lambda root over alpha upon lambda square at t is equal to MGF of X at lambda over root alpha t multiplied by e to the power minus root over alpha t. Since, MGF of ax plus b at t is equal to e to the power bt into MGF of X at a t. So, by applying we get this.

(Refer Slide Time: 05:40)



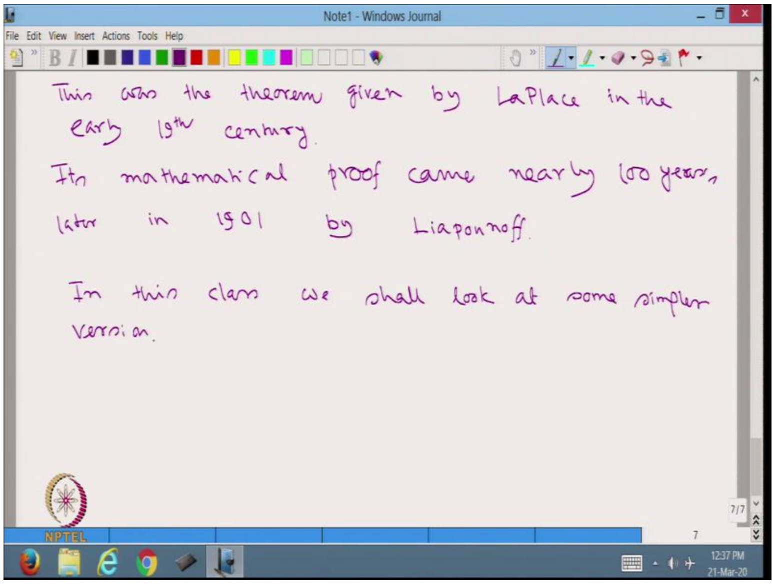
Is equal to e to the power minus root over alpha t multiplied by 1 minus lambda t upon root over alpha lambda to the power minus alpha, since MGF of gamma lambda alpha at t is equal to 1 minus t by lambda whole to the power minus alpha, here t is replaced by lambda t root over alpha, therefore log of MZ at t is equal to minus root over alpha t minus alpha log of 1 minus lambda t root over alpha lambda is equal to minus root over alpha minus alpha log of 1 minus t upon root over alpha is equal to minus root over alpha t minus alpha now we expand log of 1 minus t upon root alpha minus t upon root alpha minus t square upon 2 root alpha square minus t cube upon 3 root alpha cube etcetera is equal to minus root over alpha t minus and this minus makes it plus t upon t root alpha plus t square upon 2 plus t cube upon 3 alpha to the power half plus terms with higher powers of alpha in the denominator, therefore limit alpha going to infinity log of MZ t is equal to t square upon 2 as these 2 cancel and all this go to 0. Therefore, MZ t converges to e to the power t square by 2.

(Refer Slide Time: 09:15)



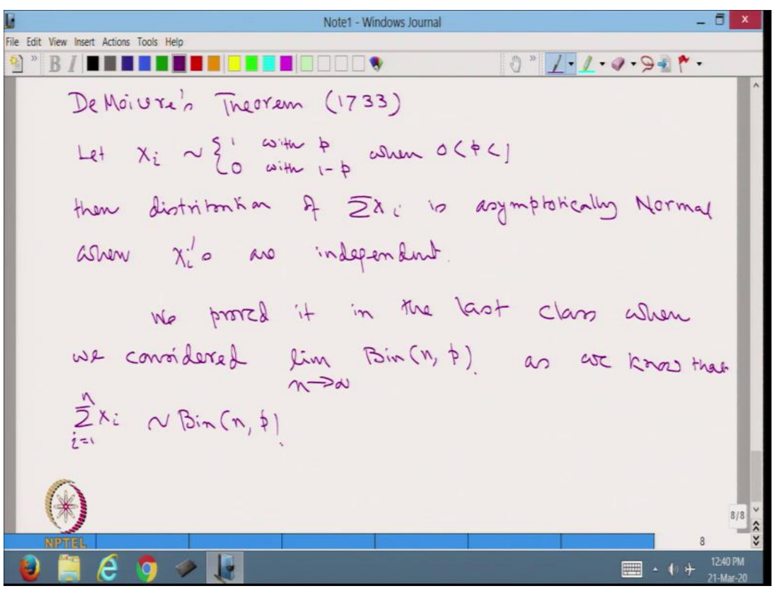
That is moment generating function of the normalized gamma variat converges to the MGF of normal 0, 1. Therefore we get the same result and therefore as I said this comes from central limit theorem as I said in the last class the general statement is, if Xi is a sequence of independent random variables such that expected value of Xi is equal to mu i which is finite and variance of Xi is equal to sigma square i which is also finite then X1 plus X2 plus Xn asymptotically converges to normal distribution with mean sigma mu i i is equal to 1 to n and variance sigma i square i is equal to 1 to n as n goes to infinity.

(Refer Slide Time: 11:39)



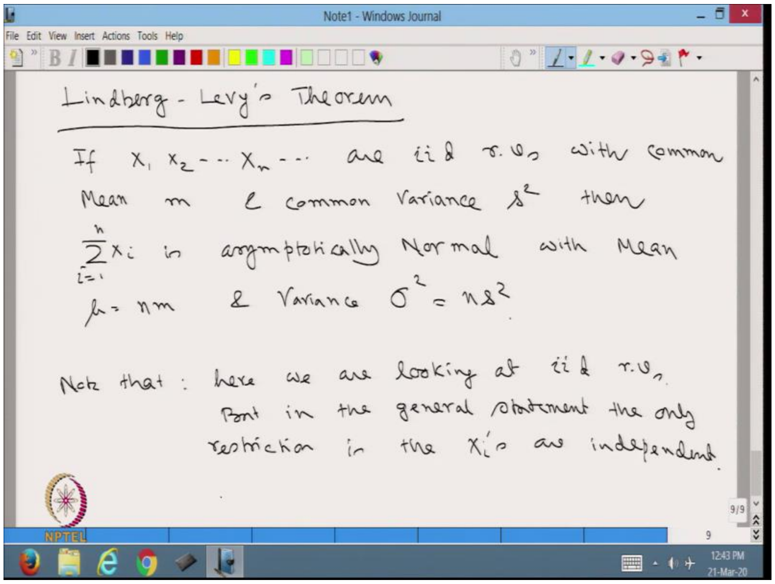
This was the theorem given by Laplace in the early 19th century, its mathematical proof came nearly 100 years later in 1901 by Liapounoff, but that is a rigorous proof in this class we shall look at some simpler version.

(Refer Slide Time: 13:06)



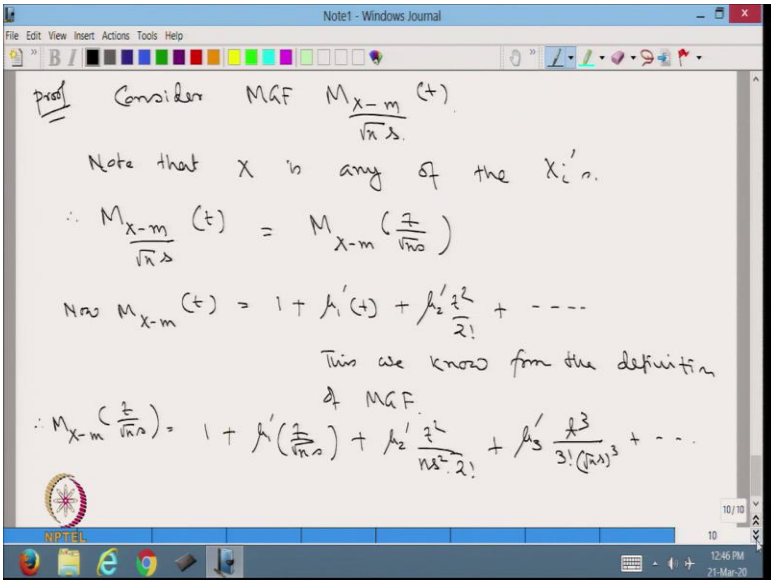
One such version is De Moivre's theorem which came in the year 1733, which says that let Xi be distributed as 1 with probability p and 0 with probability 1 minus p when 0 less than p less than 1, then distribution of sigma Xi is asymptotically normal when Xi's are independent. We have already proved it in the last class when we have considered limit n going to infinity of binomial n, p, as we know that sigma Xi i is equal to 1 to n is distributed as binomial in p.

(Refer Slide Time: 15:04)



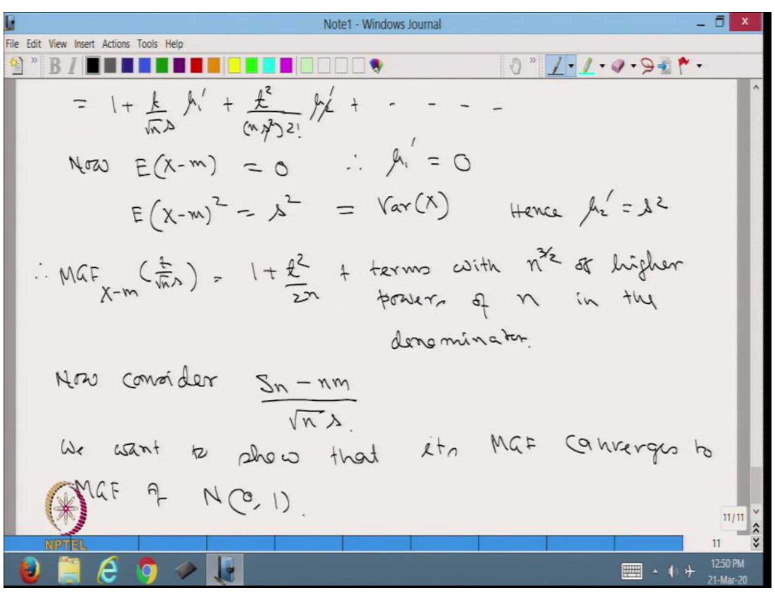
Which is called Lindbergh Levy's theorem, which states that if X1, X2, Xn are iid random variables with common mean m and common variance s square then sigma Xi i is equal to 1 to n is asymptotically normal with mean mu is equal to nm and variance sigma square is equal to ns square. So, note that, here we are looking at iid random variables but in the general statement the restriction the only restriction is the Xi's are independent. Thus we are looking at a special case of central limit theorem.

(Refer Slide Time: 17:31)



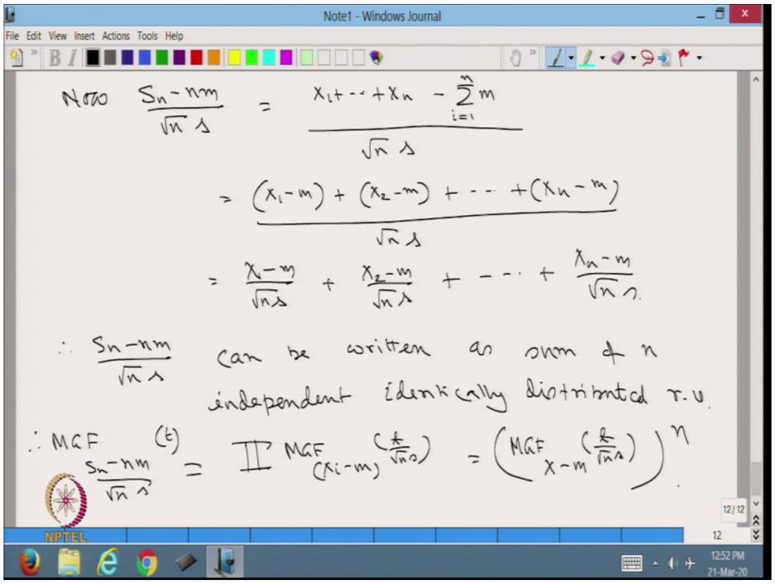
Proof, consider moment generating function of X minus m upon root over ns at t, note that X is any of these Xi's. Therefore, M X minus m upon root over ns at t is equal to moment generating function of X minus m at the point t upon root over ns. Now, M X minus m at t is equal to 1 plus mu 1 prime t plus mu 2 prime t square upon factorial 2 plus the expansion this we know from the definition of moment generating function. Therefore, M X minus m at t over root over ns is equal to 1 plus mu 1 prime t at t over root over ns plus mu 2 prime t square upon n square into 2 plus mu t prime upon t cube factorial 3 root over ns whole cube.

(Refer Slide Time: 19:56)



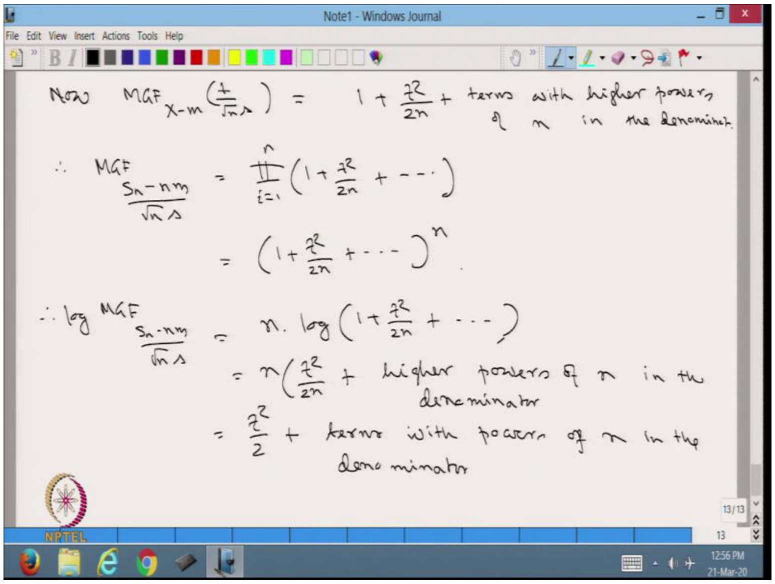
Like that is equal to 1 plus t root over ns mu 1 prime plus t square n s square factorial 2 mu 2 prime plus higher powers of t. Now, expected value of X minus m is equal to 0 therefore mu 1 prime is equal to 0. Expected value of X minus m whole square is equal to s square because it is the variance of X hence mu 2 prime is equal to s square. Therefore, MGF of X minus m at t upon root ns is equal to 1 plus t square by 2 as this square cancels with the mu 2 prime plus terms with terms with n to the power 3 by 2 or higher powers of n in the denominator. Now, consider Sn minus nm upon root over ns, we want to show that its MGF converges to MGF of normal 0, 1.

(Refer Slide Time: 22:23)



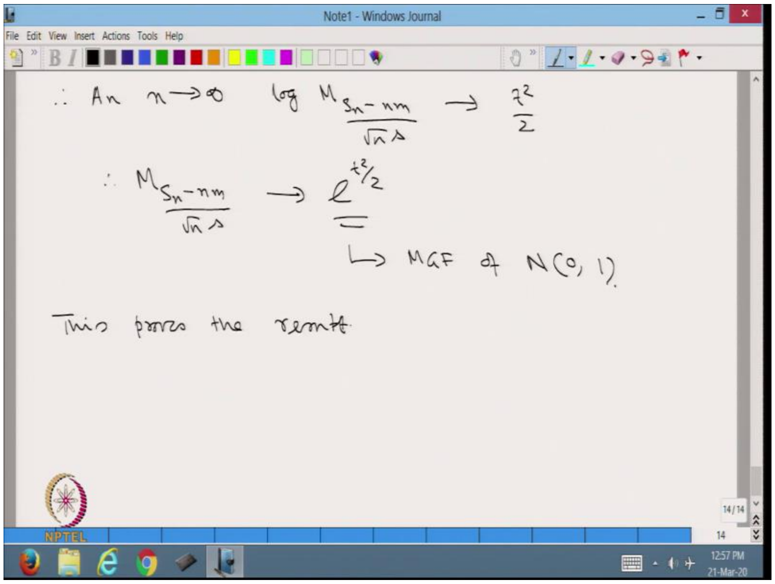
Now, Sn minus nm upon root over ns is equal to X1 plus X2 plus Xn minus sigma over m i is equal to 1 to n divided by root over ns is equal to X1 minus m plus X2 minus m plus Xn minus m upon root over ns is equal to X1 minus m upon root over ns plus X2 minus n root over ns plus Xn minus m root over ns. Therefore, Sn minus nm upon root over ns can be written as sum of n independent identically distributed random variables, therefore moment generating function of Sn minus nm upon root over ns is equal to product of MGF of Xi minus m at t upon root ns, so let me put the t here for your understanding is equal to MGF of X minus m at t over root over ns whole to the power n.

(Refer Slide Time: 24:51)



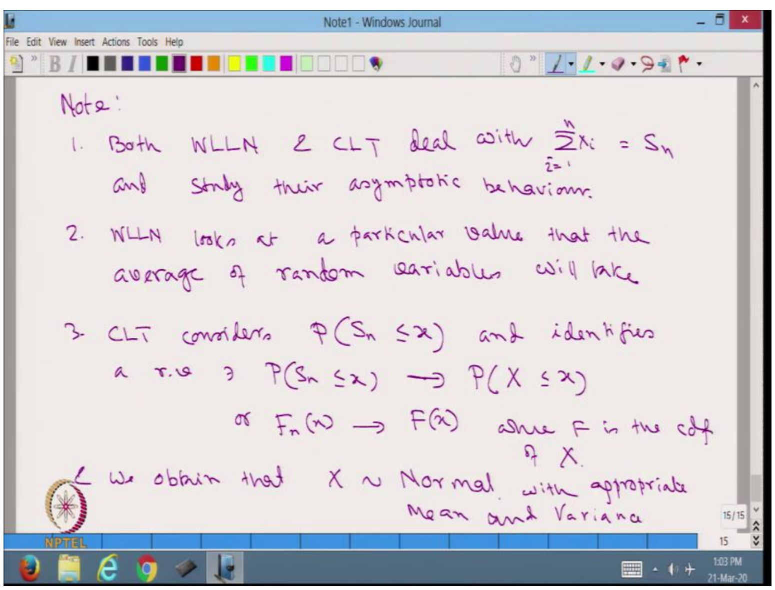
Now, MGF of X minus M at, now MGF of X minus m at t upon root over ns is equal to we have computed as 1 plus t square 2 by n plus terms with higher powers of n in the denominator. Therefore, MGF of Sn minus nm root over ns is equal to product of i is equal to 1 to n 1 plus t square upon 2n plus other terms which is is equal to 1 plus t square upon 2n plus other terms whole to the power n. Therefore, log of MGF of Sn minus nm upon root over ns is equal to n times log of 1 plus t square upon 2n plus terms with n in the denominator with higher power of n in the denominator, is equal to n into t square upon 2n plus higher powers of n in the denominator, is equal to t square upon 2 plus terms with powers of n in the denominator.

(Refer Slide Time: 27:28)



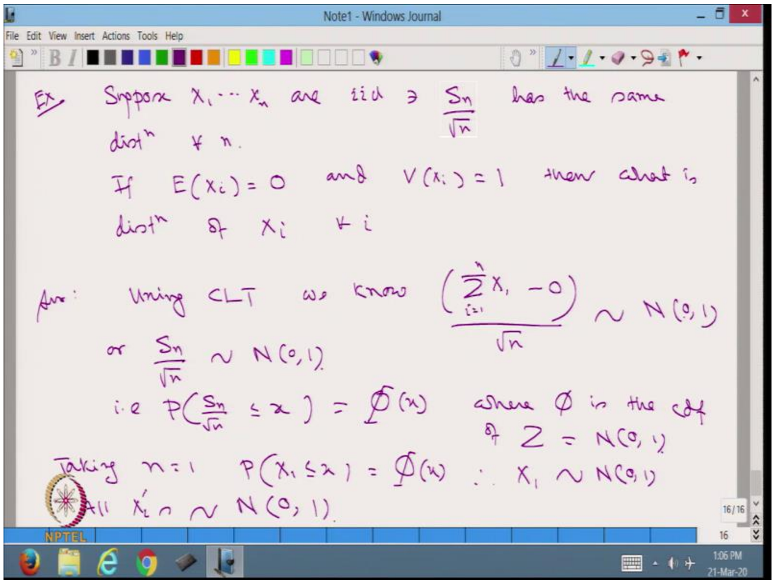
Therefore, as n goes to infinity log of moment generating function of Sn minus nm upon root over ns converges to t square upon 2, therefore moment generating function of Sn minus nm upon root over ns converges to e to the power t square by 2 that is MGF of normal 0, 1. So, this proves the result.

(Refer Slide Time: 28:25)



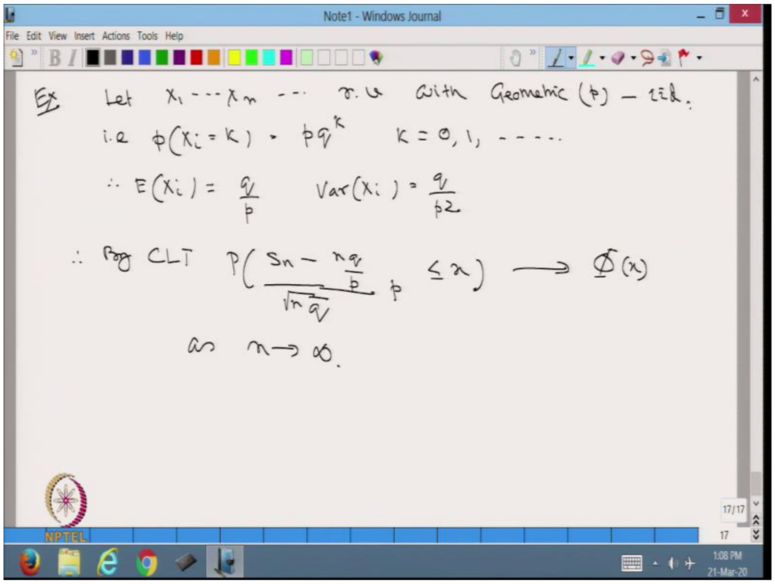
So, note that both weak law of large numbers and central limit theorem deal with sigma Xi i is equal to 1 to n is equal to Sn and study their asymptotic behaviour. Weak law of large numbers looks at a particular value that the average of the random variables will take. While central limit theorem consider probability Sn less than equal to x and identifies a random variable such that probability Sn less than equal to x converges to probability X less than equal to x or Fn x converges to Fx where F is the cdf of X and we obtain that X is actually normal with appropriate mean and variance.

(Refer Slide Time: 31:08)



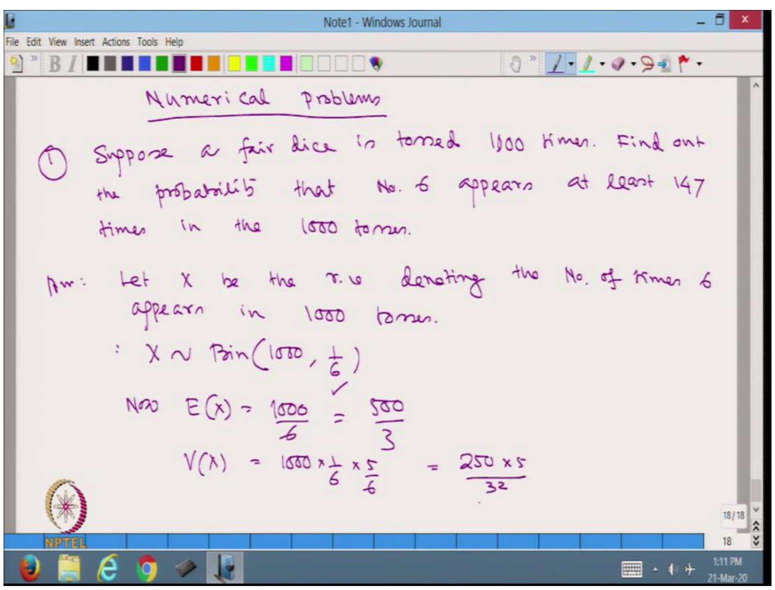
Let us solve a few problems. Suppose X1, X2, Xn are iid such that Sn upon root n has the same distribution for all n, if expectation of Xi is equal to 0 and if variance of Xi is equal to 1 then what is the distribution of Xi for all i. Answer, using central limit theorem we know sigma Xi i is equal to 1 to n minus 0 upon root over n converges to normal 0, 1. Or Sn upon root n converges to normal 0, 1, that is probability Sn upon root n less than equal to x is equal to phi x where phi is the cdf of Z which is equal to normal 0, 1. Taking n is equal to 1 probability X1 less than equal to x is equal to phi x, therefore X1 distributed as normal 0, 1, therefore all Xi's distributed as normal 0, 1.

(Refer Slide Time: 33:37)



Example, let X1, X2, Xn be random variable with geometric p, that is probability Xi is equal to k is equal to p q to the power k, k is equal to 0, 1 etcetera. Therefore, expected value of Xi is equal to q by p variance of Xi is equal to q by p square I miss that these are all iid's, therefore by central limit theorem probability Sn minus nq by p upon root over nq multiplied by p less than equal to X converges to phi x as n goes to infinity.

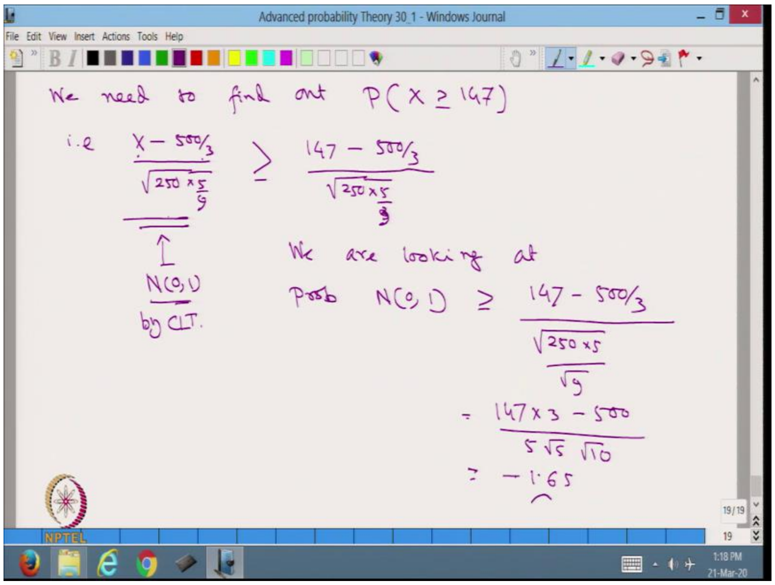
(Refer Slide Time: 35:08)



Let us now solve some numerical problems. Problem one, suppose a fair dice is tosses 1000 times find out the probability that number 6 appears at least 147 times in the 1000 tosses.

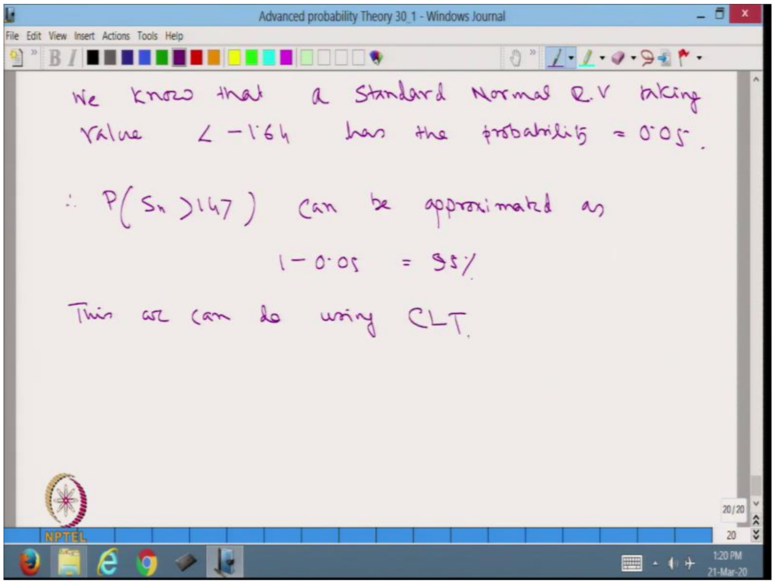
Answer, let X be the random variable denoting the number of times 6 appears in 1000 tosses, therefore X is distributed as binomial with 1000 comma 1 by 6, this is because it is fair dice, therefore each phase has the same probability 1 by 6. Now, expectation of X is equal to 1000 upon 6 is equal to 500 upon 3 and variance of X is equal to 1000 into 1 by 6 into 5 by 6 is equal to 250 into 5 upon 3 square.

(Refer Slide Time: 37:48)



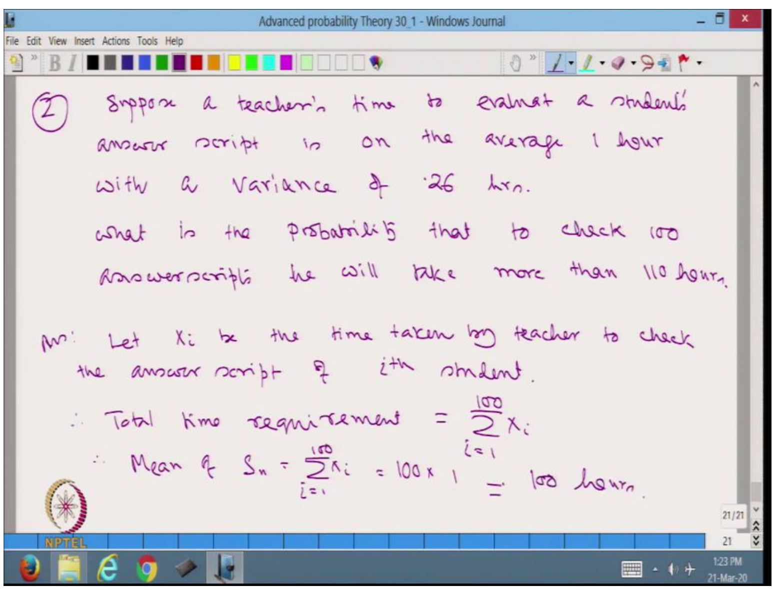
We need to find out probability X greater than equal to 147 that is X minus 500 by 3 upon root over 250 into 5 by 9 greater than equal to 147 minus 500 by 3 upon root over 250 into 5 upon 9. Now, this is standard normal because X is a binomial random variable n is pretty high is equal to 1000, therefore variable minus mean upon standard deviation converges to normal 0, 1 by CLT, therefore we are looking at probability standard normal random variable greater than equal to 147 minus 500 by 3 divided by root over 250 into 5 by square root of 9 which is equal to 147 into 3 minus 500 upon 5 root over 5 root over 10 is equal to minus 1.65

(Refer Slide Time: 39:56)



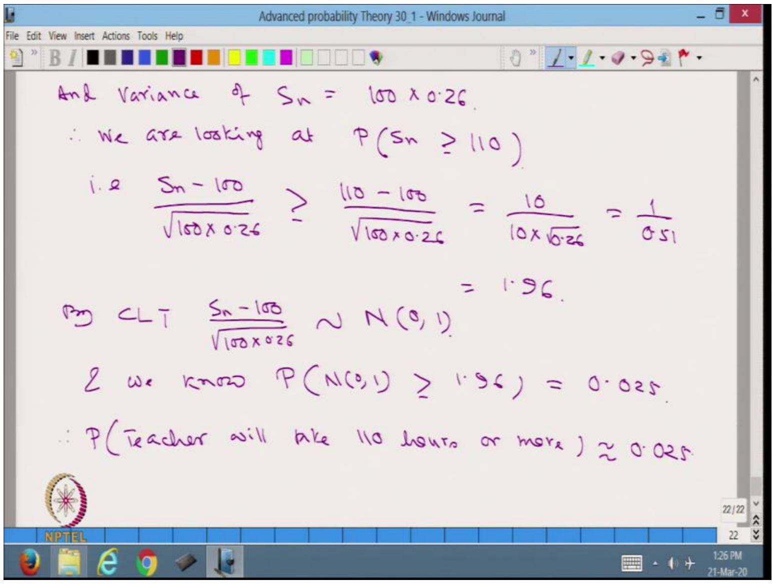
We know that a standard normal random variable taking value less than minus 1.64 as the probability equal to 0.05, therefore probability Sn greater than 147 can be approximated as 1 minus 0.05 is equal to 95 percent. This we can do using central limit theorem.

(Refer Slide Time: 41:15)



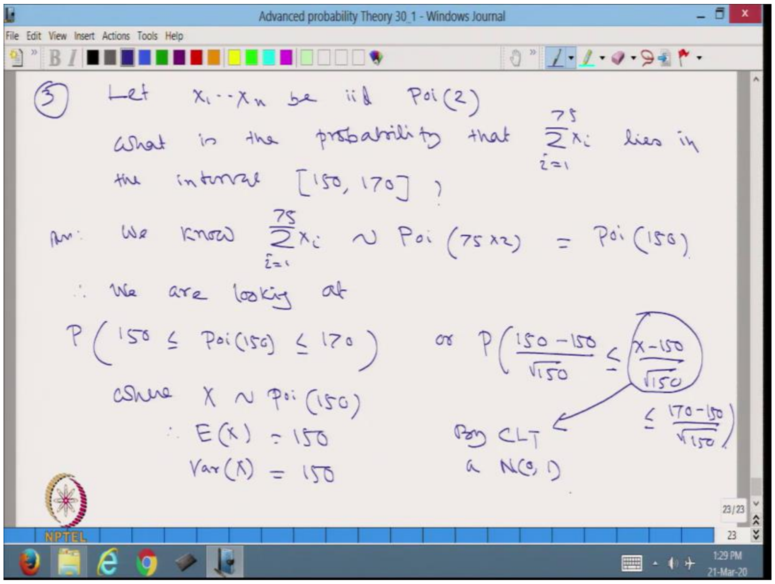
Problem number two, suppose a teacher time to evaluate a student's answer script is on the average 1 hour with a variance of 0.26 hours, what is the probability that to check 100 answer scripts he will take more than 110 hours. Answer, let Xi be the time taken by the teacher to check the answer script of ith student, therefore total time requirement is equal to sigma over Xi i is equal to 1 to 100, therefore mean of Sn is equal to sigma Xi i is equal to 1 to 100 is equal to 100 into 1 is equal to 100 hours.

(Refer Slide Time: 44:04)



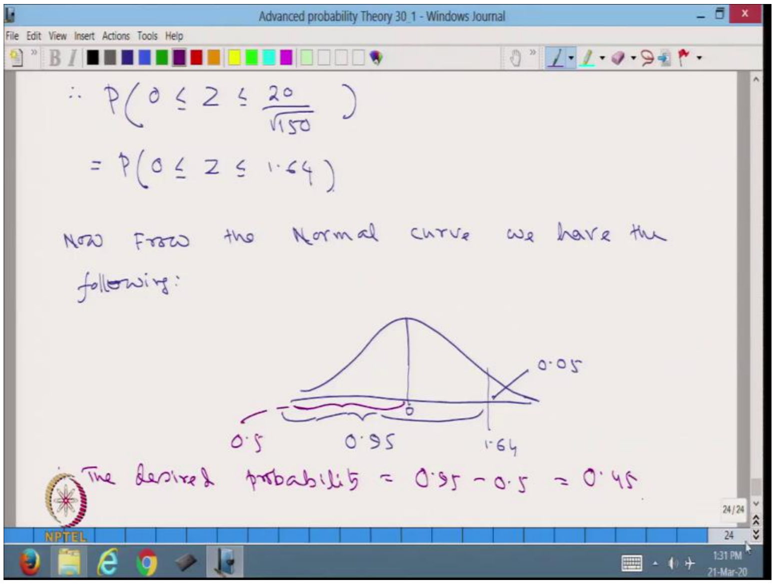
And variance of Sn is equal to 100 into 0.26, therefore we are looking at probability Sn greater than equal to 110 that is Sn minus 100 divided by root over 100 into 0.26 greater than equal to 110 minus 100 divided by root over 100 into 0.26, which is equal to 10 upon 10 into root over 0.26 is equal to 1 upon 0.51 is equal to 1.96, by CLT Sn minus 100 upon root over 100 into 0.26 is distributed as normal 0, 1. And we know probability normal 0, 1 variat is greater than equal to 1.96 is equal to 0.025, therefore probability teacher will take 110 hours or more is equal to 0.025.

(Refer Slide Time: 46:25)



So, before we conclude let us have one more example, let X1, X2, Xn be iid with Poisson with parameter 2, what is the probability that sigma Xi i is equal to 1 to 75 lies in the interval 150 to 170. Answer, we know sigma Xi i is equal to 1 to 75 is also Poisson with random variable with parameter 75 into 2 that is Poisson with 150, therefore we are looking at probability 150 less than equal to Poisson with 150 less than n equal to 170 or probability 150 minus 150 upon root over 150 less than equal to X minus 150 upon root over 150 less than equal to 170 minus 150 upon root over 150, where X follows Poisson 150 therefore expected value of X is equal to 150 and variance of X is also 150. Therefore, this by CLT a standard normal varait.

(Refer Slide Time: 49:02)



Therefore, we are looking at probability 0 less than equal to Z less than equal to 20 upon root over 150, is equal to probability 0 less than equal to Z less than equal to 1.64. Now, from the normal curve we have the following 1.64 above that this is 0.05 probability, therefore this probability is 0.95 this is 0 and this probability is 0.5, therefore the desired probability is equal to 0.95 minus 0.5 is equal to 0.45. Okay friends, I stop here, today in fact as his was the last class this is the concluding lecture of this series on Advanced Probability Theory. Although we started with from very basic definitions of probability during these 30 lecture courses, we have seen many advance topics in to the order static, theory of conversions, central theorem etcetera, I hope that you will go through the series on lectures and solve the problem and thereby will master the basics of probability, with that hope I conclude the lecture. Thank you so much.