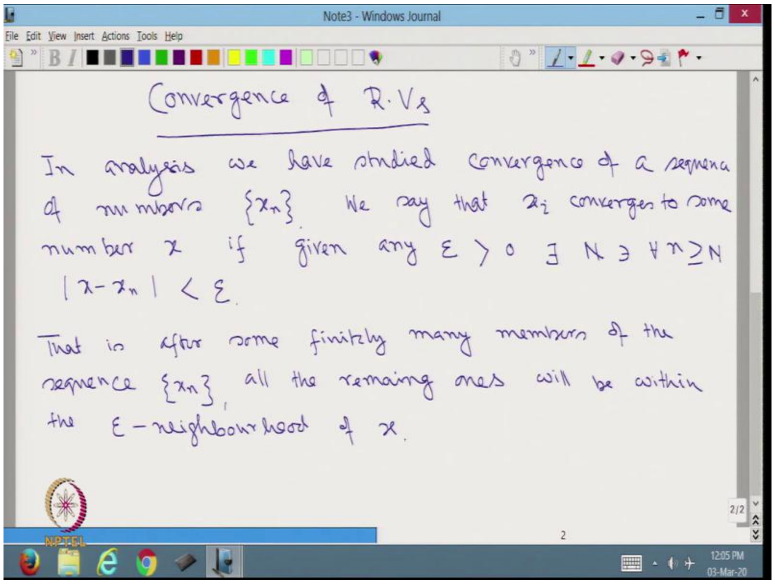
Advanced Probability Theory

Professor Niladri Chatterjee
Department of Mathematics
Indian Institute of Technology, Delhi

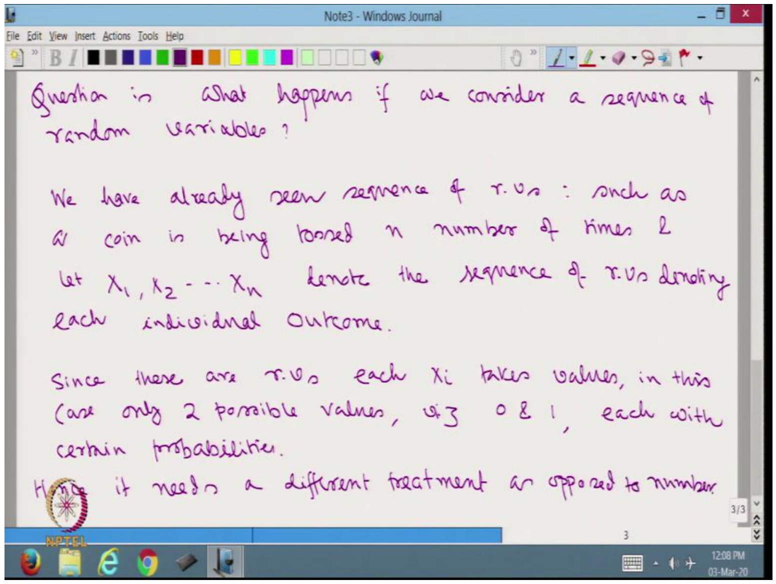
Lecture 24

(Refer Slide Time: 0:32)



Welcome students to MOOCs lecture series on Advanced Probability Theory. This is lecture number 24. As I said from this lecture, I shall start convergence of random variables. In analysis we have studied convergence of a sequence of numbers Xn. We say that Xi convergence to some number if given any epsilon greater than 0, their eXist N such that for all n greater than equal to N modulus of x minus Xn is less than epsilon. That is after some finitely many members of the sequence all the remaining ones will be within the epsilon neighborhood of x. This is what all of us know I assume.

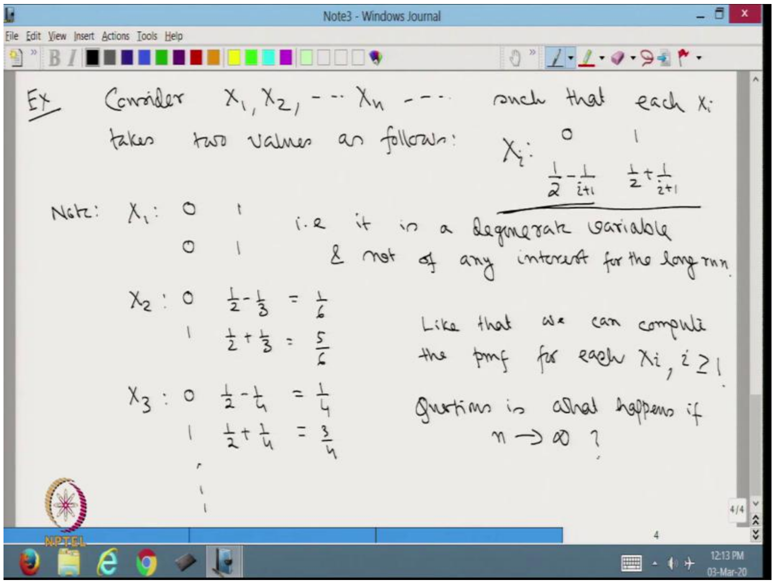
(Refer Slide Time: 3:02)



The question is what happens if we consider a sequence of random variables? We have already seen sequence of random variables such as say a coin is being tossed n number of times and let X1 X2 upto Xn denote the sequence of random variables denoting each individual outcome. Since, these are random variables each Xi takes values, in this case only two possible values, namely 0 and 1, each with certain probabilities.

Hence, it needs a different treatment as opposed to numbers, I hope the concept is clear. So, we have X1 X2 Xn are sequence of random variables, each random variable can take a set of values depending upon its omega, each one with certain probability. Then how, what do you want to see is where the sequence of random variable actually converges to. So, let me give an example.

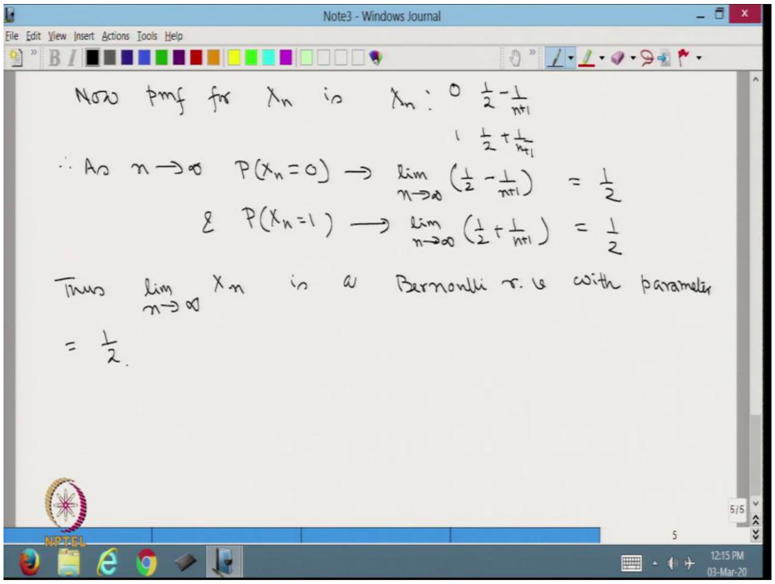
(Refer Slide Time: 6:39)



Consider X1, X2, Xn such that each Xi takes two values as follows Xi takes only 0 and 1 with probabilities half minus 1 upon i plus 1 and half plus 1 upon i plus 1. So, this is the probability mass function for Xi. Note, X1 therefore test values 0 with probability 0 and 1 with probability 1 that is, it is a degenerate variable and not of any interest for the long run.

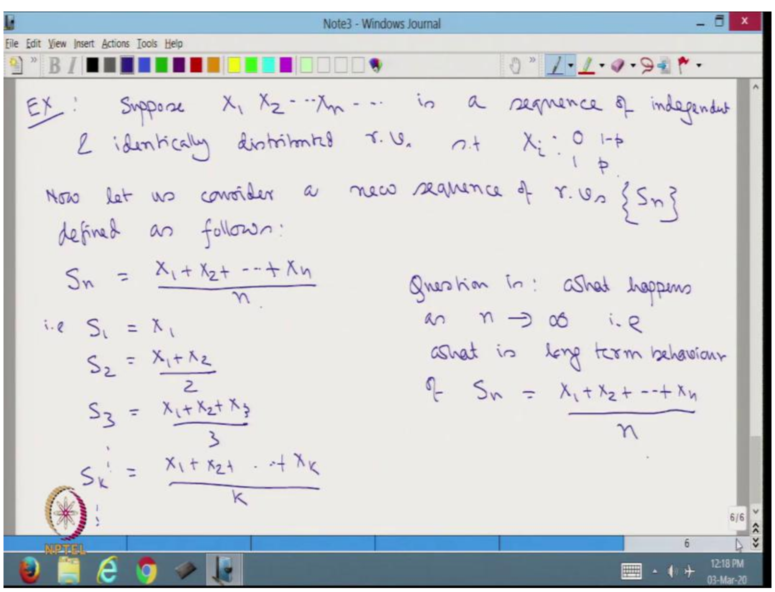
What about X2? X2 takes values 0 with half minus 1 by 3 and 1 with half plus 1 by 3. That is it is equal to 1 by 6. And this is is equal to 5 by 6. What about X3 x takes 0 with half minus 1 by 4 is equal to 1 by 4, and 1 with probability half plus 1 by 4 is equal to 3 by 4. Like that we can compute the pmf for each Xi, i greater than equal to 1. Question is what happens if n goes to infinity?

(Refer Slide Time: 9:42)



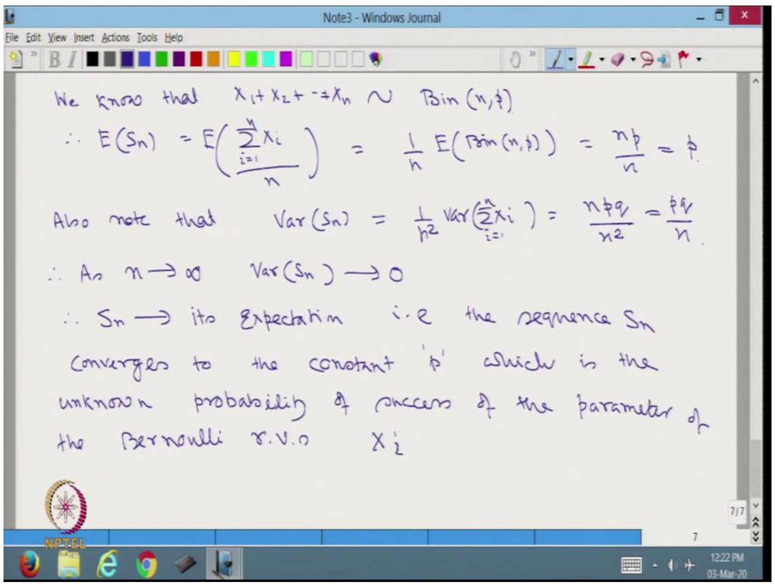
Now, pmf for Xn is, it takes 0 with probability half minus 1 upon n plus 1 and 1 with probability half plus 1 upon n plus 1. Therefore, as n goes to infinity, probability Xn is equal to 0 that goes to limit n going to infinity half of half minus 1 upon n plus 1 is equal to half and probability Xn is equal to 1 that converges to limit n going to infinity half plus 1 upon n plus 1 is equal to half. Thus, limit n going to infinity of Xn is a Bernoulli random variable with parameter is equal to half. Let us consider another example.

(Refer Slide Time: 11:16)



Suppose X1, X2, Xn is a sequence of independent and identically distributed random variables such that each Xi takes two values 0 with probability 1 minus p and 1 with probability p. Now, let us consider a new sequence of random variables Sn defined as follows. Sn is equal to X1 plus X2 plus Xn by n. That is, S1 is equal to X1, S2 is equal to X1 plus X2 plus 2, S3 is equal to X1 plus X2 plus X3 by 3, Sk is equal to X1 plus X2 plus Xk by k like that question is what happens as n goes to infinity, that is what is the long term behavior of Sn is equal to X1 plus X2 plus Xn upon n.

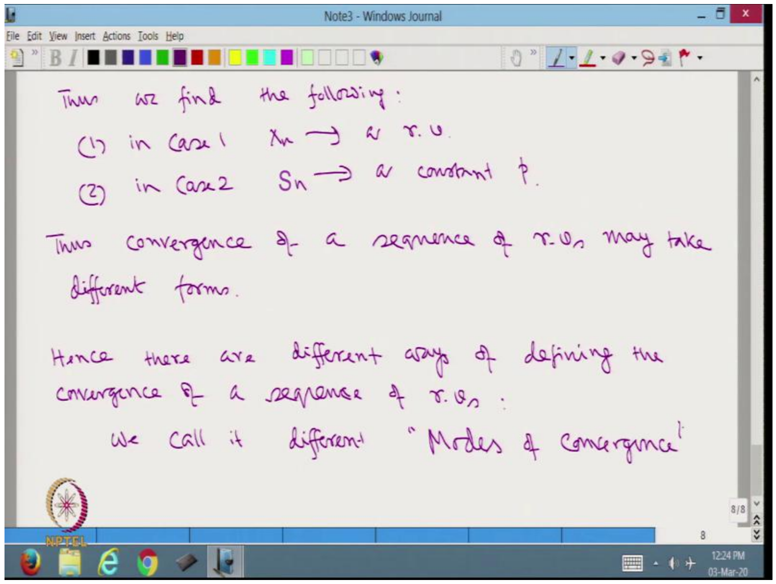
(Refer Slide Time: 14:00)



We know that X1 plus X2 plus Xn is distributed as binomial with n, p. Therefore, the expected value of Sn is equal to expected value of sigma Xi, i equal to 1 to n divided by n is equal to 1 by n expected value of binomial n, comma p is equal to np upon n is equal to p. Also note that variance of Sn is equal to 1 by n square into variance of sigma Xi, i is equal to 1 to n is equal to n p q upon n square is equal to p q by n.

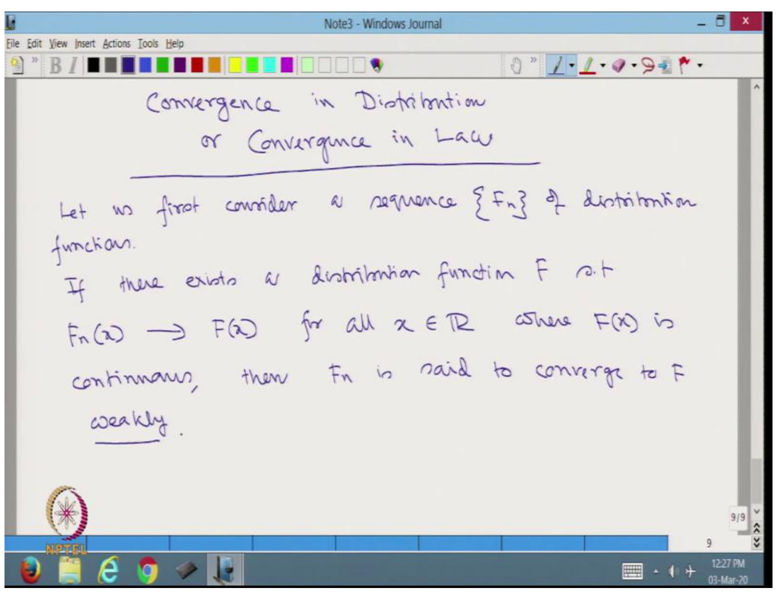
Therefore, as n goes to infinity, variance of the Sn converges to 0. Therefore, Sn converges to its expectation. That is, the sequence Sn converges to the constant ‘p’ which is the unknown probability of success of the parameter of the Bernoulli random variables Xi.

(Refer Slide Time: 16:47)



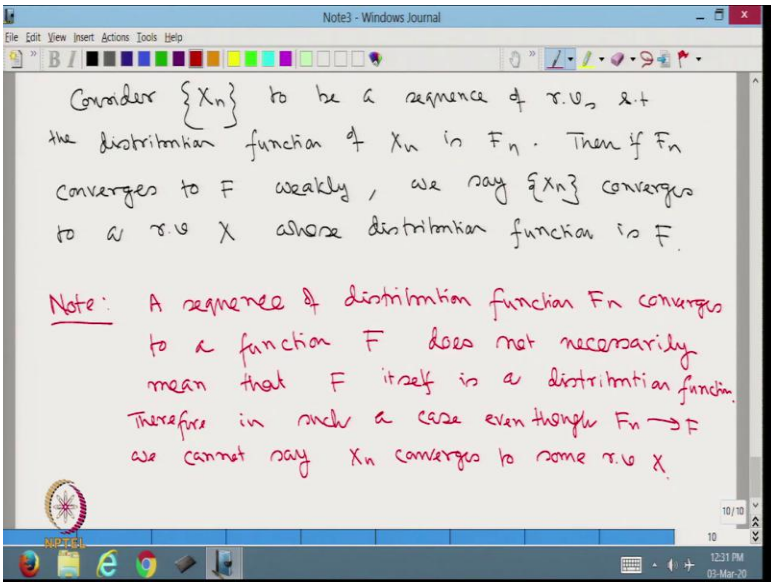
Thus we find the following in case 1, Xn converges to a random variable in case 2, Xn converges to a constant which is p. Thus, convergence of a sequence of random variables may take different forms. Hence, there are different ways of defining the convergence of a sequence of random variables or in other words, we call it different "modes of convergence". So, in this lectures, we shall study a few of the different modes of convergence for a sequence of random variables.

(Refer Slide Time: 19:07)



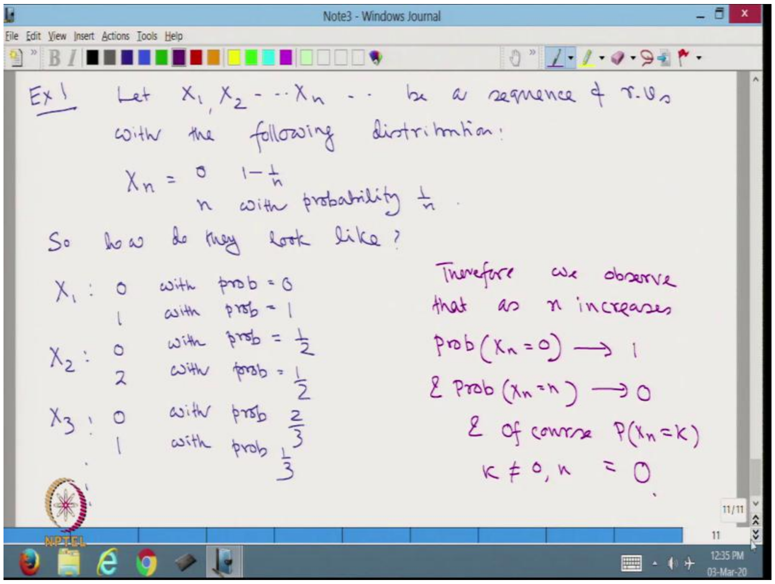
One of the weakest form of convergence is called convergence in distribution or convergence in law. To understand that let us first consider a sequence Fn of distribution functions. If there exists a distribution function F such that, Fnx converges to Fx for all x belonging to R where Fx is continuous, then Fn is said to converge to F weakly. That is it is a weak form of convergence of the distribution functions.

(Refer Slide Time: 21:17)



Now, consider Xn to be a sequence of random variables such that the distribution function of Xn is Fn. Then if Fn converges to F weakly, we say Xn converges to a random variable X, whose distribution function is F. Note, a sequence of distribution functions Fn converges to a function F does not necessarily mean that F itself is a distribution function. Therefore, in such a case even though Fn converges to F we cannot say Xn converges to some random variable X. So, this is one point that you have to understand. So, I shall give you examples so that this is going to be clear to you.

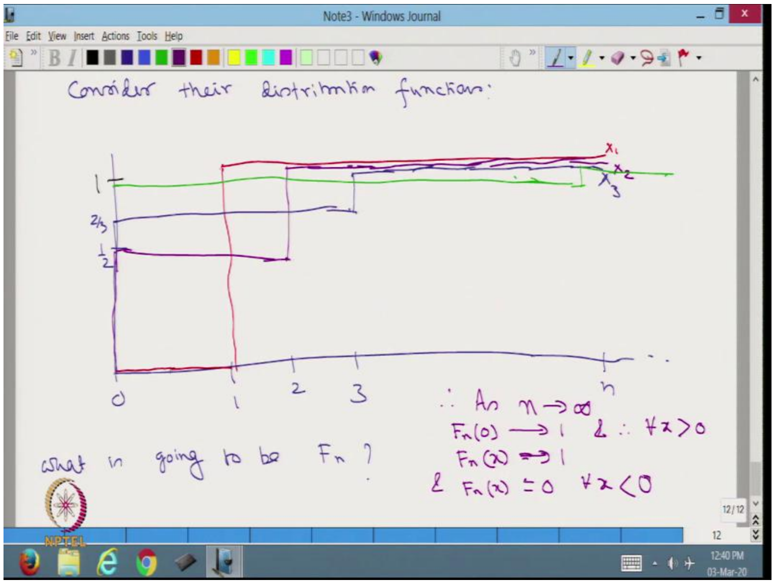
(Refer Slide Time: 24:22)

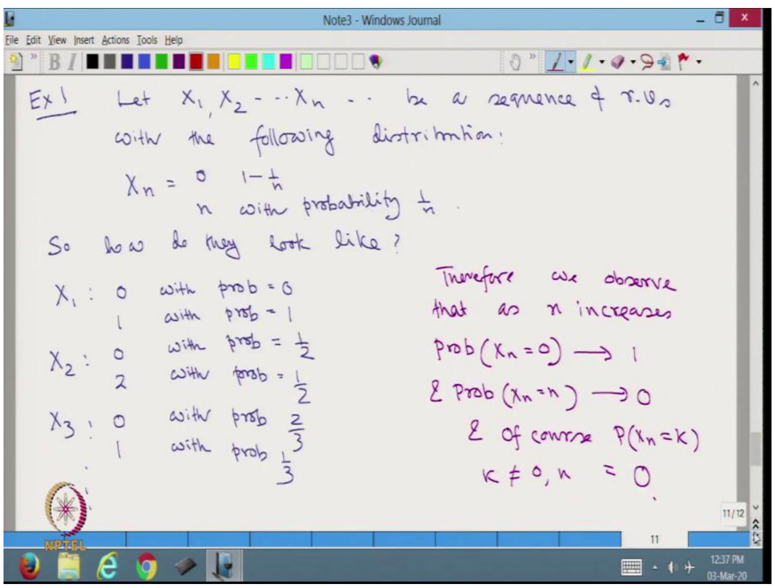


Example 1, let X1, X2, Xn be a sequence of random variables with the following distribution Xn is equal to 0 with probability 1 minus 1 by n and n with probability 1 by n. So, how do they look like? So, X1 takes the values 0 with probability 0 and 1 with probability is equal to 1, X2 takes 0 with probability is equal to half and it takes the value 2 with probability is equal to half. So, let us consider X3, which takes the value 0 with probability 2 by 3 and it takes the value 1 with probability 1 by 3.

Therefore, we observe that as n increases probability Xn is equal to 0 converges to 1 and probability Xn is equal to n converges to 0. And of course, probability Xn is equal to k, k not equal to 0 or n is equal to 0.

(Refer Slide Time: 27:50)



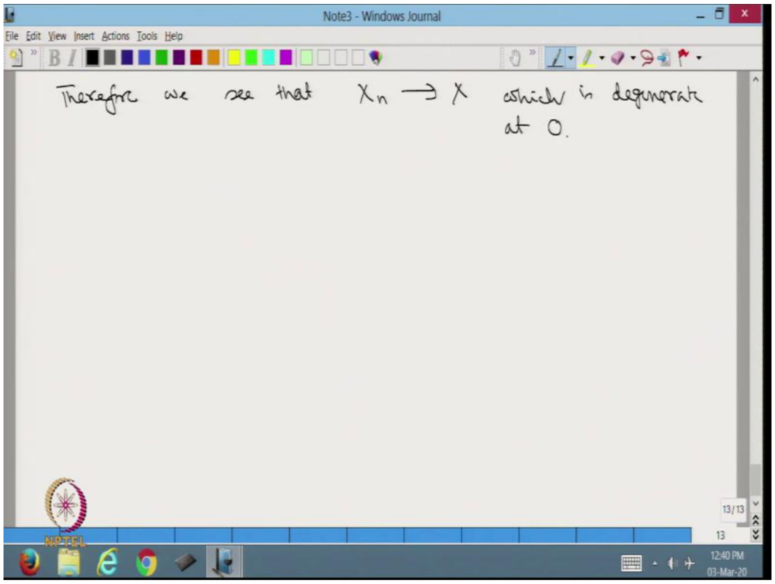


So, let us consider their distribution functions. So, let us assume 0, 1, 2, 3, n. So, what is the distribution function for X1? It takes the value 0 with probability 0 and it takes the value 1 with probability 1. So, if this is the value 1, then the distribution function of X1 is going to be something like this what about X2? X2 takes the value 0 and 2 with half and half.

So, if this is my half after 0 the cumulative distribution function goes to half and at 2 it goes to 1 and it remains 1 throughout. So, this is for X2. This is for X1. What about X3? X3 takes the value 0 with probability 2 by 3 and it takes the value 3 with 1 by 3 therefore, each distribution function will be something like this.

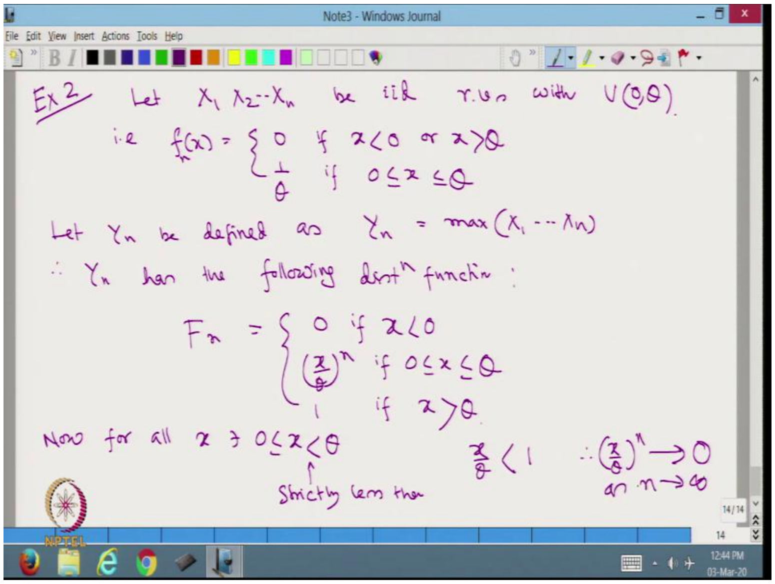
Therefore, what is going to happen? If that is the question, we can understand that it will take the value 1 minus 1 by n and go up to n and then at n it becomes 1. Therefore, as n goes to infinity if n at 0 is converging to 1 and therefore, for all x greater than 0, Fnx is going to be converging to 1 and Fnx is 0 for all x less than 0.

(Refer Slide Time: 31:40)



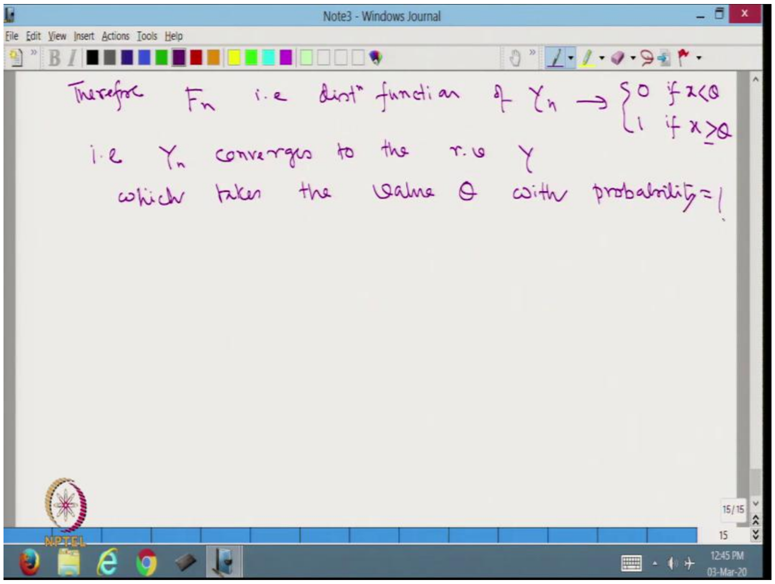
Therefore, we see that Xn converges to a variable X which is degenerate at 0.

(Refer Slide Time: 32:06)



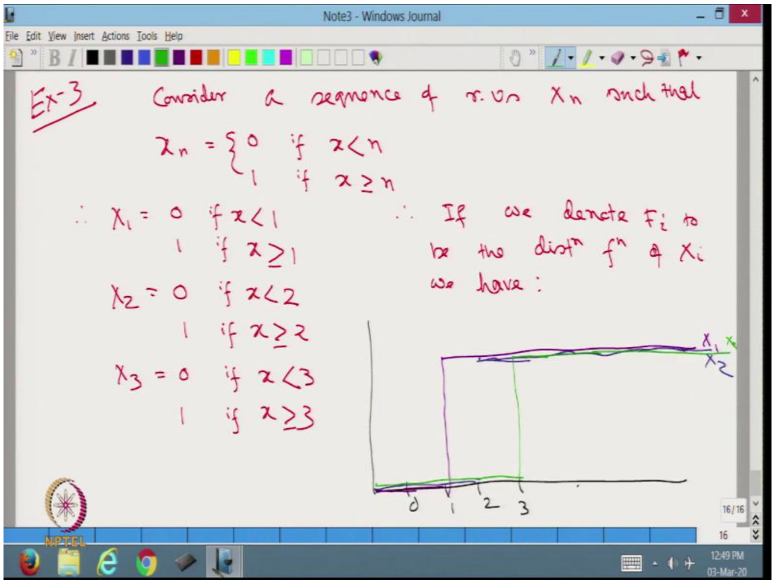
Example 2, let X1 X2 Xn be iid random variables with uniform 0, theta that is Fn of x is equal to 0 if x less than 0, or x greater than theta and is equal to 1 by theta. If 0 less than equal to x less than or equal to theta. Let Yn be defined as Yn is equal to maximum of X1 X2 Xn. Therefore, Yn has the following distribution function that is the Fn is equal to 0 if x less than 0, x by theta whole to the power n, if 0 less than equal to x less than equal to theta and 1 if x is greater than theta. Now, for all x, such that 0 less than equal to x less than theta, it is strictly less than x by theta is less than 1. Therefore, x by theta whole to the power n goes to 0 as n goes to infinity.

(Refer Slide Time: 34:42)



Therefore, Fn that is distribution function of cumulative distribution, cumulative distribution function you can say of Yn converges to 0 if x is less than theta, and 1 if x is greater than equal to theta, that means Yn converges to the random variable say Y, which takes the value theta with probability is equal to 1.

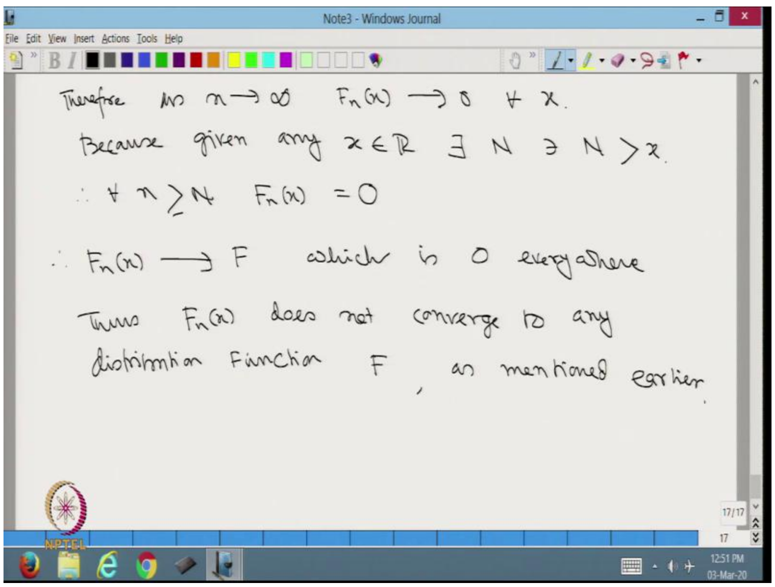
(Refer Slide Time: 35:46)



Example 3, consider, consider a sequence of random variables Xn such that Xn is equal to 0 if x is less than n and 1 if x is greater than equal to n. Therefore, X1 is equal to 0 if x less than 1, and is equal to 1, if x greater than equal to 1, therefore x 2 is equal to 0 if x less than 2 and 1 if x greater than equal to 2, X3 is equal to 0 if x less than 3 and it is is equal to 1 if x greater than or equal to 3.

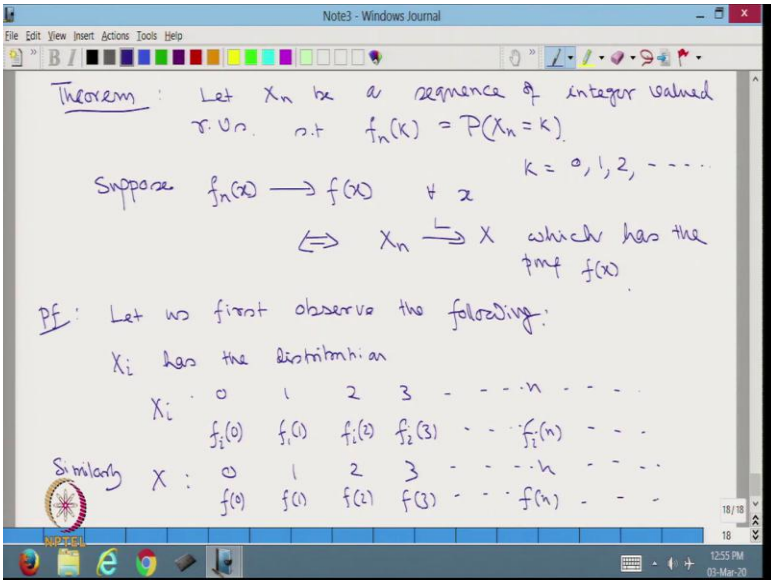
Therefore, if we denote, Fi to be the distribution function of Xi, we have the pdf of, we have the distribution function of X1 is equal to 0 if x is less than 1 and 1 if x is 1, for X2 we have it is 0 till this point and at 2 it takes the value 1 and it remains there, after that the X3, X3 will remain at 0 till 3 and that 3 it will take the value 1.

(Refer Slide Time: 38:29)



Therefore, as n goes to infinity Fn at x goes to 0 for all x, because given any x belonging to R there exist N such that N is bigger than x. Therefore, for all n greater than equal to N, Fnx is equal to 0. Therefore, Fnx converges to a function F, which is 0 everywhere. Thus, Fnx does not converge to any distribution function F as mentioned earlier.

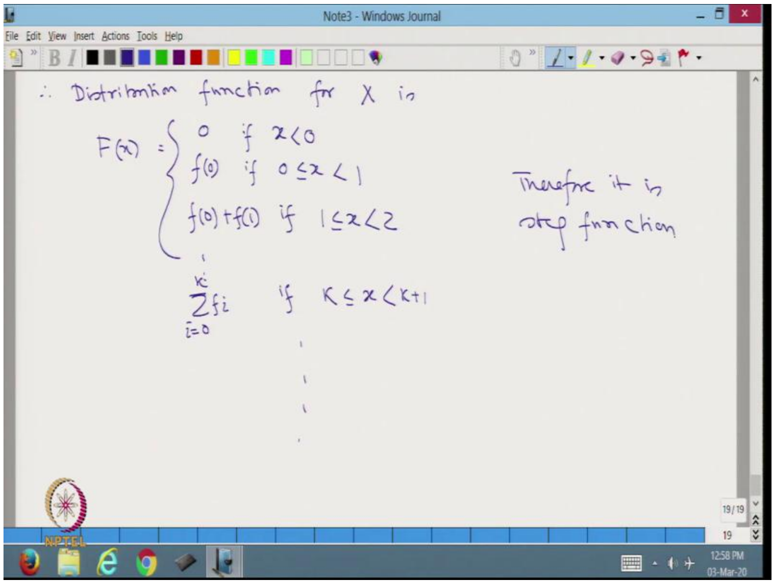
(Refer Slide Time: 40:06)



Now, let me give you a theorem, let Xn be a sequence of integer valued random variables such that Fnk is equal to probability Xn is equal to k, when k is equal to 0, 1, 2 up to infinity. Suppose Fnx converges to Fx for all x, this implies that Xn is converging by law to the random variable x, which has the pmf Fx.

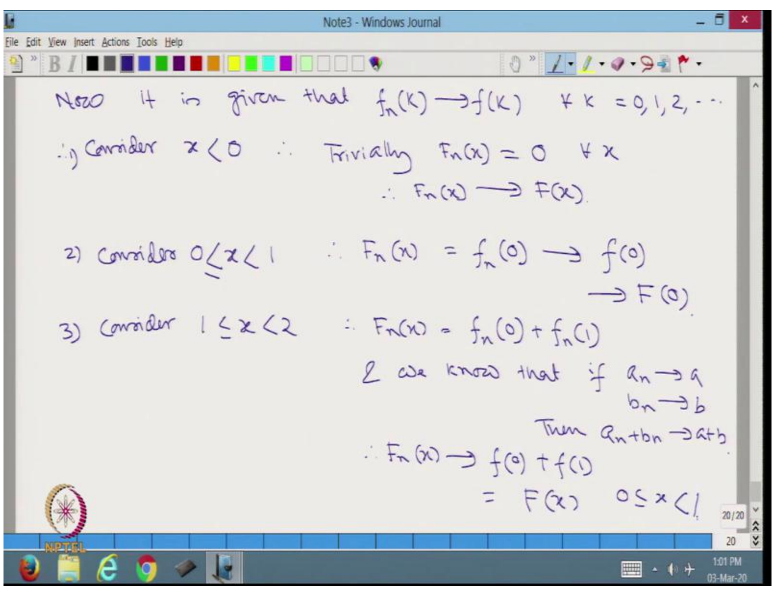
Proof, let us first observe the following, Xi has the distribution, Xi takes the value 0, 1, 2, 3, n, etc with probabilities Fi0, Fi1, Fi2, Fi3, Fin like that. Similarly, X takes the same values 0, 1, 2, 3, n with probabilities f0, f1, f2, f3, fn like that.

(Refer Slide Time: 42:43)



Therefore, distribution function for X is Fx is equal to 0 if x less than 0, it is F0 if 0 less than equal to x less than equal to 1 less than 1, F0 plus F1 if 1 less than equal to x less than 2, sigma Fi, i is equal to 0 to k, if k less than equal to x less than k plus 1 like that. Therefore, it is a step function.

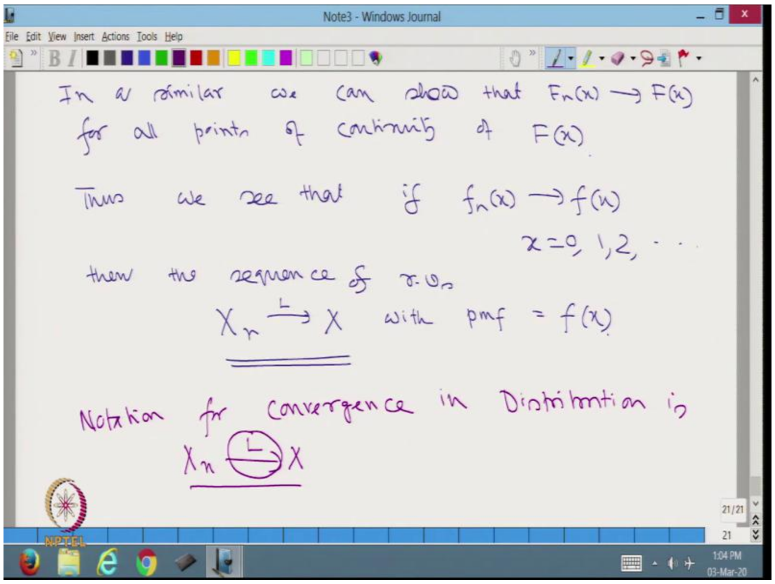
(Refer Slide Time: 43:55)



Now, it is given that Fn k converges to F k for all k is equal to 0, 1, 2, etc. Therefore, consider x less than 0. Therefore, trivially if Fn x is equal to 0 for all x, therefore, Fnx converges to Fx to consider 0 less than x less than equal to x less than 1. Therefore, Fn x is equal to Fn 0 which converges to F0, which is is equal to F at 0.

Similarly, consider 1 less than equal to x less than 2. Therefore, Fn x is equal to Fn 0 plus Fn 1 and we know that if n converges to a and bn converges to b, then an plus bn converges to a plus b. So, this we know from our high school mathematics. Therefore, Fn x converges to F0 plus F1 which is equal to F at x when 0 less than equal to x less than 1.

(Refer Slide Time: 46:27)



In a similar way we can show that Fn x converges to Fx for all points of continuity of Fx. Thus, we see that if Fn x converges to Fx, x is equal to 0, 1, 2 etc. Then the sequence of random variables Xn convergence, converges in law to the random variable X with pmf is equal to Fx. So, I want you to remember this notation.

So, notation for convergence in distribution is Xn convergence in distribution in law or Xn that is we have, we are using the (sym), we are using the symbol L to denote that it is a convergence in distribution. Now, there is a converse part of it that is if Xn converges to Xn distribution then for all X, Fn x converges to Fx, this we shall prove in the next class. Okay friend, thank you so much. Thank you.