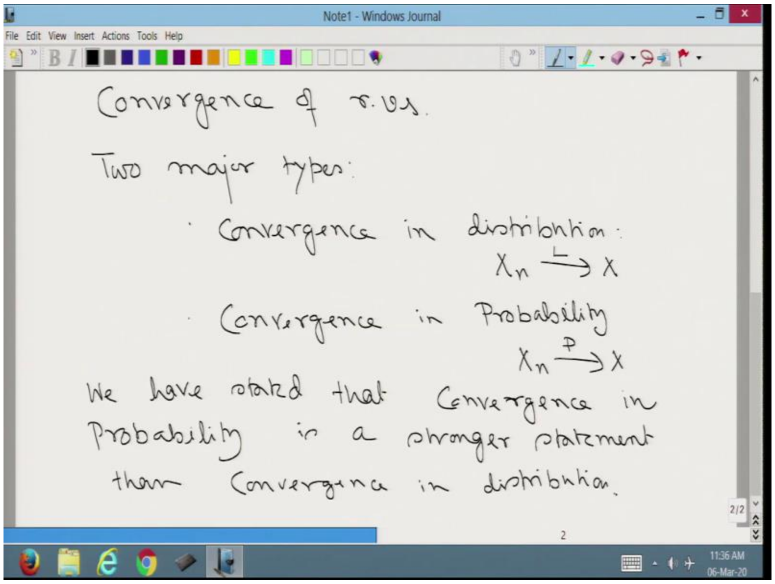
Advanced Probability Theory
Professor Niladri Chatterjee
Department of Mathematics
Indian Institute of Technology, Delhi
Lecture-26

Welcome students to MOOCs series of lectures on advanced probability theory. This is lecture number 26.

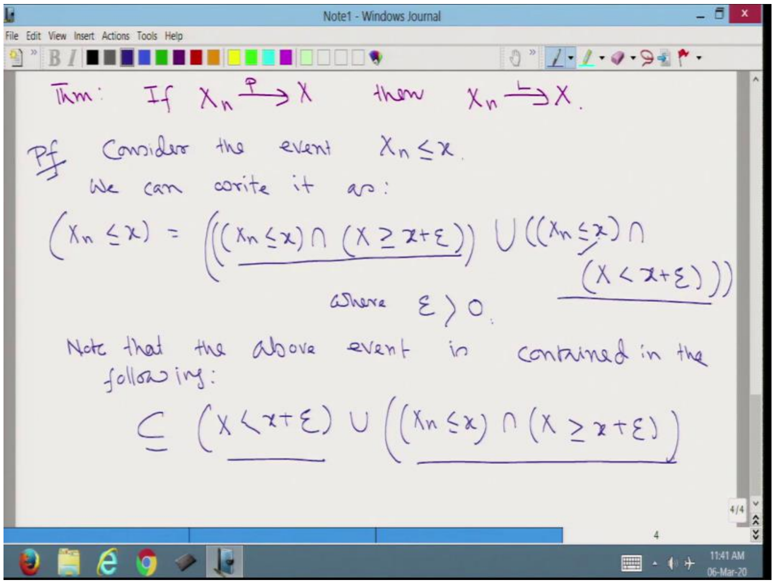
(Refer Slide Time: 00:36)



So, if you remember, we are working on convergence of random variables and we have discussed two major types convergence in distribution. When is sequence of random variables xn converges to another random variable x and also convergence in probability denote that as a sequence of random variable xn converging to another random variable x.

We have stated that convergence in probability is a stronger statement then convergence in distribution, why because there may be situation when a sequence of random variables xn converges in distribution x but not in probability. On the other hand, if xn convergence in probability to x then xn also converges in distribution to x.

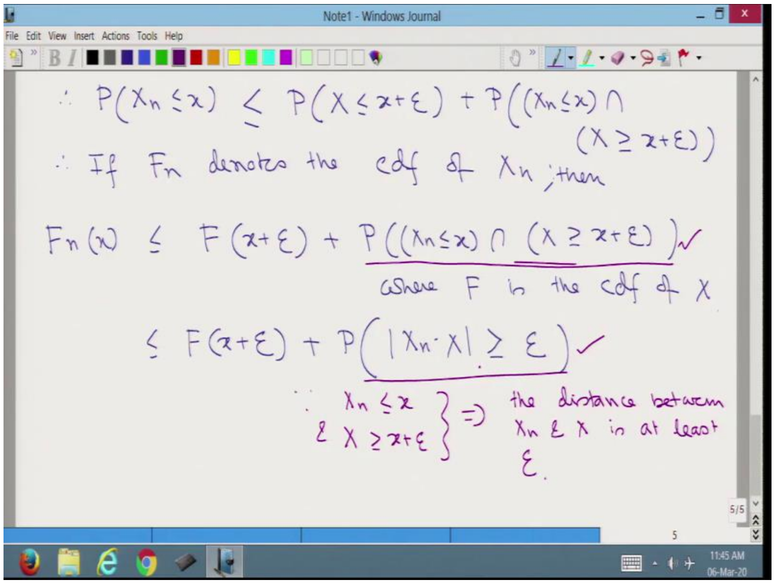
(Refer Slide Time: 03:56)



So, let us prove the result. So, theorem if xn converges in p in probability to x then xn converges in distribution to x proof. Consider the event xn less than or equal to x. Now, we can write it as xn less than equal to x is equal to union of two disjoint events xn less than equal to x and x is greater than equal to x plus epsilon union xn less than equal to x and x less than x plus epsilon where epsilon greater than zero that is it is a small quantity.

Note that the above event is contained the following that is, this is contained in x less than x plus epsilon union xn less than equal to x and x greater than equal to x plus epsilon. Why? Because I am keeping the same event here these two are same, but this is a super event of this one as we have deleted this condition.

(Refer Slide Time: 06:29)

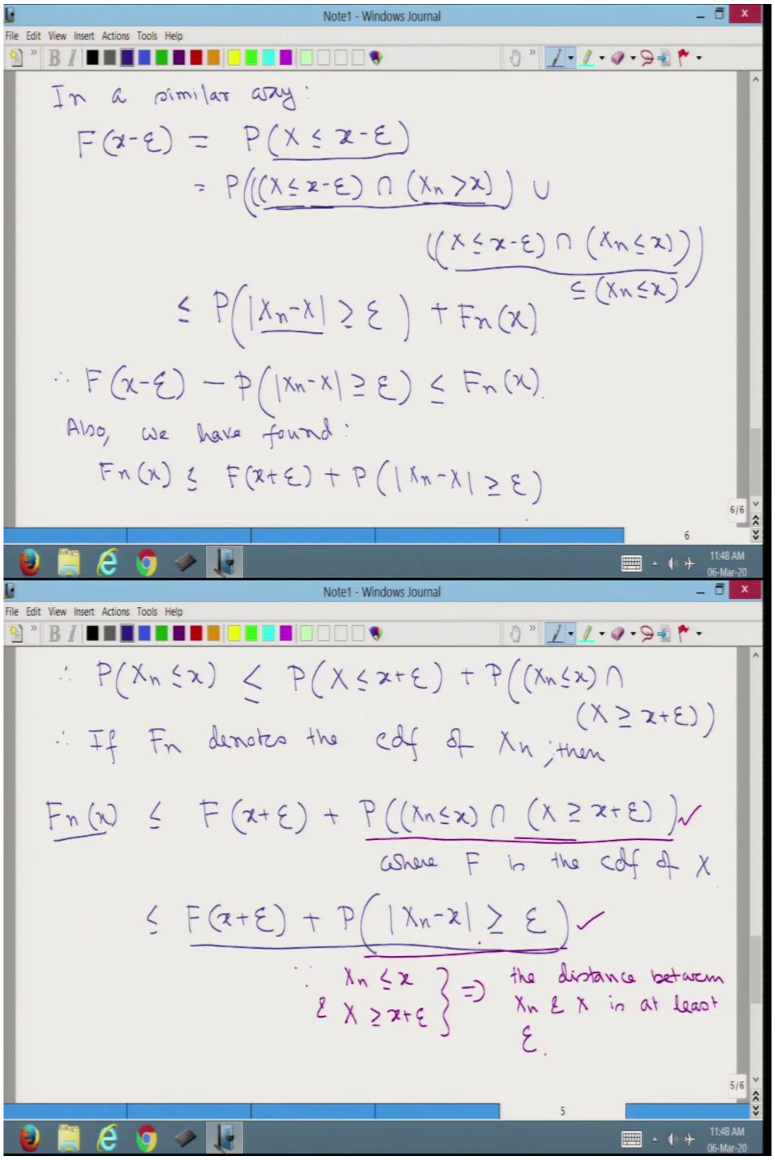


Therefore, probability xn less than equal to x is less than equal to probability x less than equal to small x plus epsilon plus probability xn less than equal to x intersected with x greater than equal to x plus epsilon. Therefore, if fn denote the cumulative distribution function of xn then we can write fnx is less than equal to f at x plus epsilon plus probability xn less than equal to x intersected with x greater than equal to x plus epsilon, where f is the cumulative distribution function of x.

Now, this is less than equal to fx plus epsilon plus probability modulus of xn minus x is greater than equal to epsilon. How, since, from this event, we can see that xn is less than equal to x and x is capital X is greater than equal to x plus epsilon they were together they imply that the distance between xn and x is at least epsilon.

And since, this can happen in some other ways also, therefore, this event is contained in this event and therefore, this probabilities are bigger than this probability.

(Refer Slide Time: 09:05)

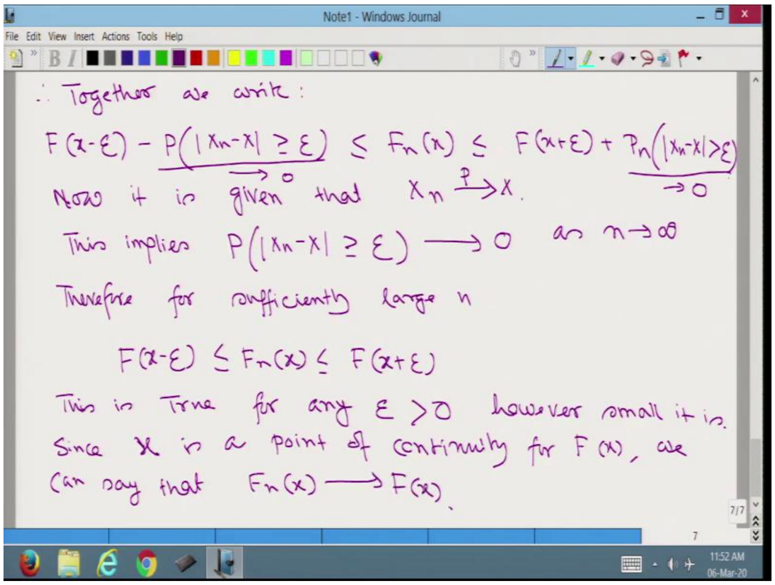


In a similar way, f at small x minus epsilon is equal to probability x less than equal to x minus epsilon is equal to probability x less than equal to small x minus epsilon. And xn is greater than x union with x less than equal to small x minus epsilon intersected with xn less than equal to x. So, again I have divided this event as a union of two different disjoint events less than equal to probability modulus of xn minus x greater than equal to epsilon.

Again by a similar logic xn is greater than x, but x is smaller than x minus epsilon. Therefore, the distance between them is greater than equal to epsilon plus this whole event is contained in xn less than equal to x. Therefore, we are writing their fnx therefore, f at x minus epsilon minus probability modulus of xn minus x greater than equal to epsilon is less than equal to a fnx also we have found if we go back.

We have found that fnx is less than equal to this entire quantity, therefore we can write it is that fnx is less than equal to f at x plus epsilon plus probability modulus of xn minus x greater than equal to epsilon greater than equal to epsilon.

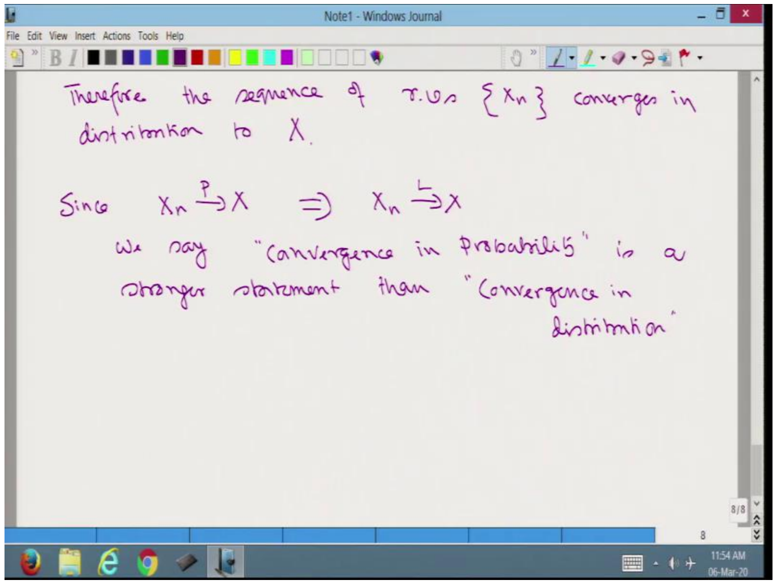
(Refer Slide Time: 11:50)



So, together we write f at x minus epsilon minus probability modulus of xn minus x greater than equal to epsilon less than equal to f and x less than equal to f at x plus epsilon plus probability modulus of xn minus x greater than epsilon. Now, it is given that Xn converges in probability to x, this implies probability modulus of xn minus x greater than equal to epsilon converges to 0 as n goes to infinity.

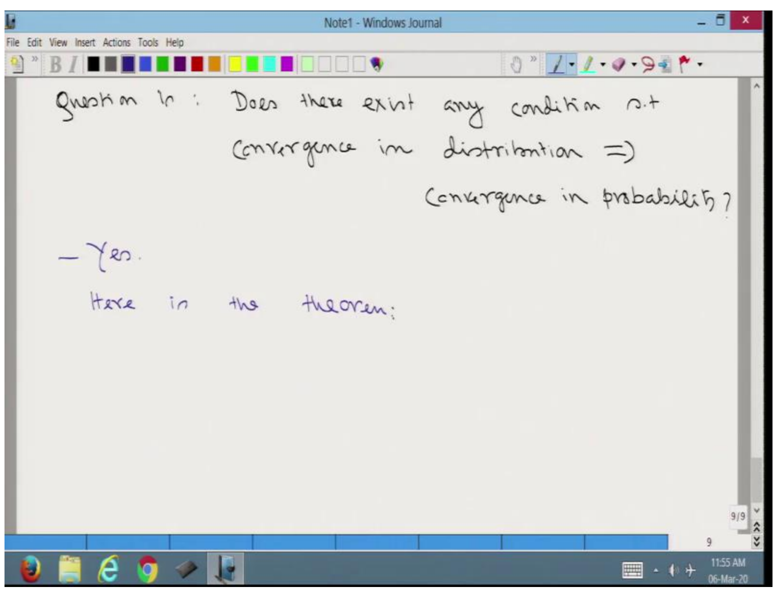
That is therefore, for sufficiently large n we can say that f at x minus epsilon less than equal to fn x less than equal to fx plus epsilon, why, because this term is going to 0. And similarly, this term is also going to 0. This is true for any epsilon greater than 0. However, small it is since x is a point of continuity for fx we can say that fnx converges to fx.

(Refer Slide Time: 14:30)



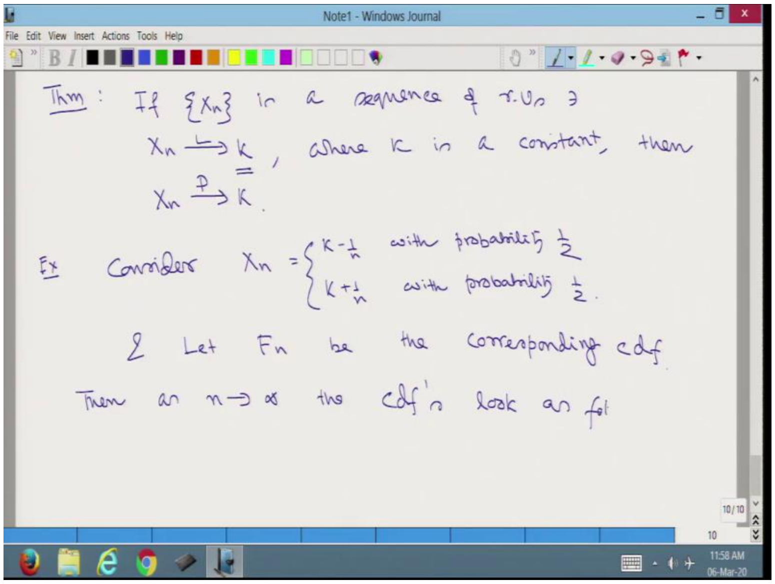
Therefore, the sequence of random variables xn converges in distribution to x. Therefore, since convergence in P in probability implies convergence in distribution. We say convergence in probability is a stronger statement then convergence in distribution.

(Refer Slide Time: 15:50)



Question is does there exist any condition such that convergence in distribution implies convergence in probability, the answer is yes and here is the theorem.

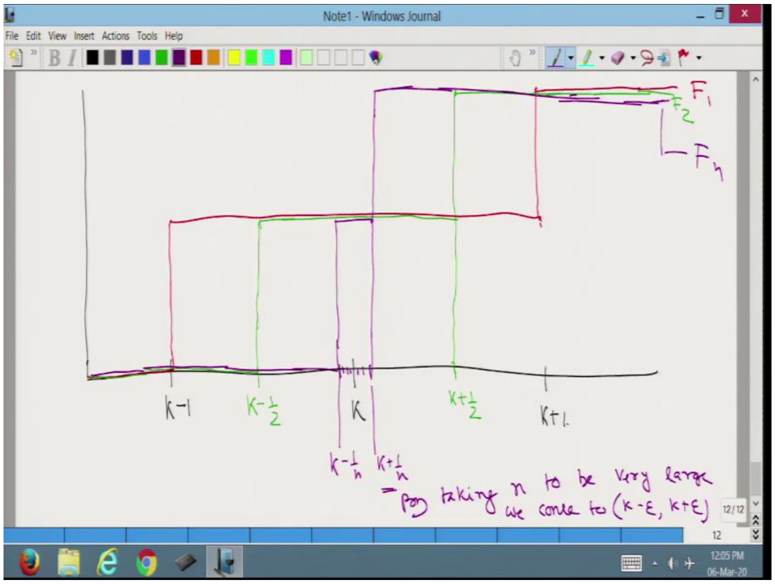
(Refer Slide Time: 16:50)



If xn is a sequence of random variables such that xn converges in distribution to a constant K then xn converges in probability to k that is, here we are looking at a degenerate case xn is a distribution, which is converging to a constant K in distribution, then xn converges in probability to K as well.

So, let me first give you an example, Suppose, consider xn to be distributed as follows, it takes the value k minus 1 by n with probability half and k plus 1 by n with probability half therefore, and let fn be the corresponding cumulative distribution function. Then as n goes to infinity, the cdf's look as follows.

(Refer Slide Time: 19:01)

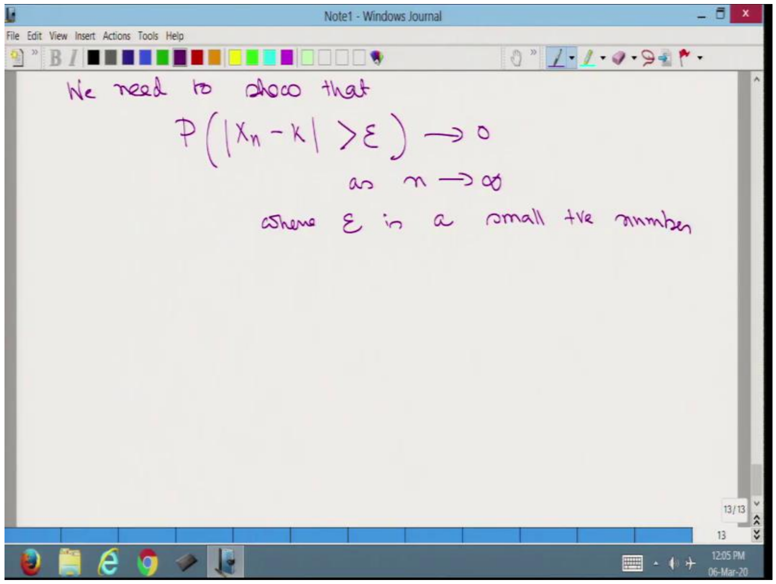


Therefore, it is cdf's look as follows. Suppose this is k, this is k minus 1 and this is k plus 1. Therefore, x1 will have a cdf which is 0 till k minus 1 at k minus 1 it jumps to half and then it continues till k plus 1 with half then it jumps to 1 at k plus half and continues like that. So, that is the cdf of x1, x2 will take the values k minus half and K plus half with probabilities half and half.

Therefore, each Cdf will be 0 till k minus half it is going to be half at k minus half and it is going to be 1 at k plus half. Therefore, this is the shape of f2 therefore what is going to be the Cdf of say k minus 1 upon n when n is large. So, let us call this value k minus 1 upon n and this value to be k plus 1 upon a then each cdf will look like this it is how 0 till k minus 1 upon in half at k plus half at k minus 1 upon n it will continue like this till the point k plus 1 upon n then it will make a jump to 1 and continue at 1 throughout.

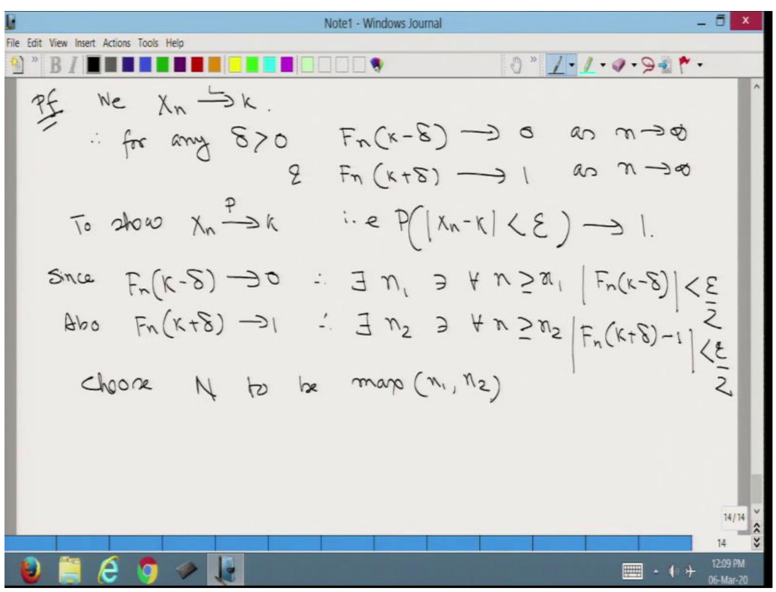
Therefore, what is happening we can understand that as n increases, we are getting smaller and smaller intervals around k where the actual jump is occurring. Therefore, this by taking n to be very large we come to the intervals k minus epsilon to k plus epsilon. If this is clear, then let us proceed as follows.

(Refer Slide Time: 22:24)



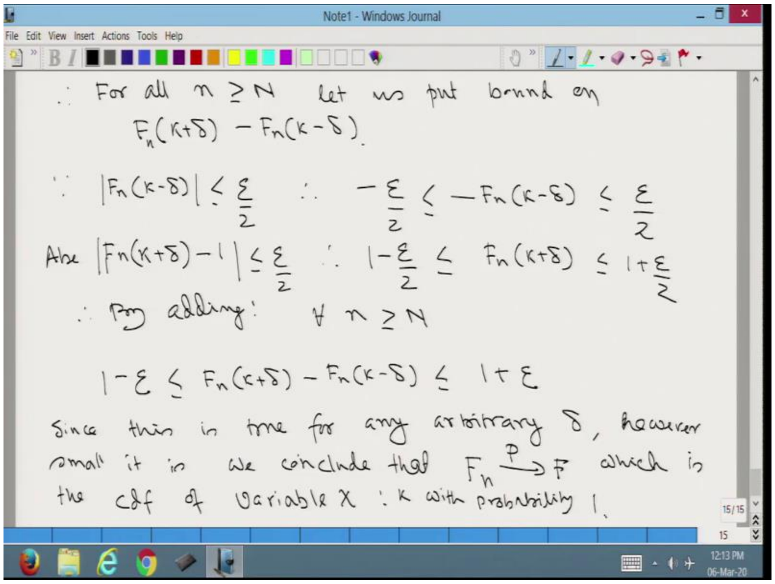
We need to show that probability modulus of xn minus k greater than epsilon is goes to zero as n goes to infinity where epsilon is a small positive number.

(Refer Slide Time: 23:01)



So, proof we have xn converges in distribution to k therefore, for any delta greater than 0 Fn k minus delta goes to 0 as n goes to infinity and Fn k plus delta goes to 1 as n goes to infinity to show Xn converges in probability to k that is probability modulus of Xn minus k less than epsilon goes to one since Fn k minus delta converges to 0, therefore, there exist n1 such that for all n greater than equal to n1 modulus of Fn k minus delta is less than epsilon by 2 also Fn k plus delta goes to 1. Therefore, there exist n2 such that for all n greater than equal to n to modulus of fn k plus delta minus 1 is less than epsilon by 2 choose n to be maximum of n1 and n2.

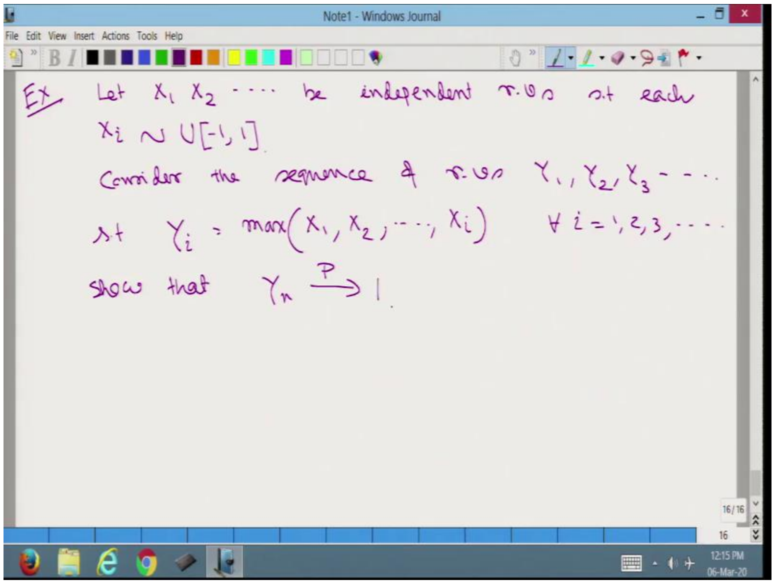
(Refer Slide Time: 25:08)



Therefore, for all n greater than equal to n, let us put bound on Fn k plus delta minus Fn k minus delta since, modulus of Fn k minus delta less than equal to epsilon by 2 therefore, minus epsilon by 2 less than equal to minus Fn k minus delta less than equal to epsilon by 2 also modulus of Fn k plus delta minus 1 less than equal to epsilon by 2.

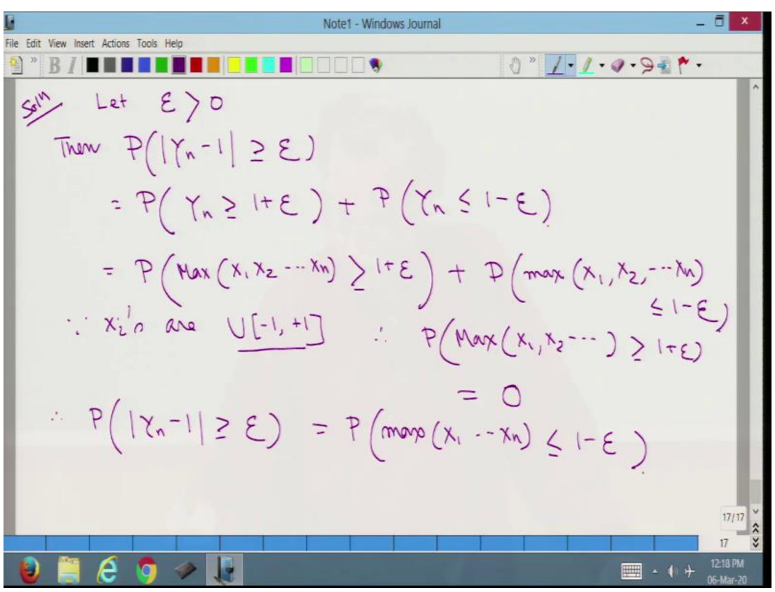
Therefore, 1 minus epsilon by 2 less than equal to Fn k plus delta less than equal to 1 plus epsilon by 2. Therefore, by adding for all n greater than equal to n 1 minus epsilon less than equal to fn k plus delta minus fn k minus delta less than equal to 1 plus epsilon. Since this is true for any arbitrary delta however small it is, we conclude that fn converges in probability to f which is the cdf of variable x defined as it takes the value k with probability 1.

(Refer Slide Time: 28:11)



So, that proves that result let me now give you an example of convergence in probability let x1, x2, xn be independent random variables such that each xi is distributed as uniform minus 1 comma 1. Consider the sequence of random variables Y1, Y2, Y3 like that such that Yi is equal to maximum of X1, X2, Xi for all i is equal to 1, 2, 3 up to infinity. Show that Yn converges in probability to 1.

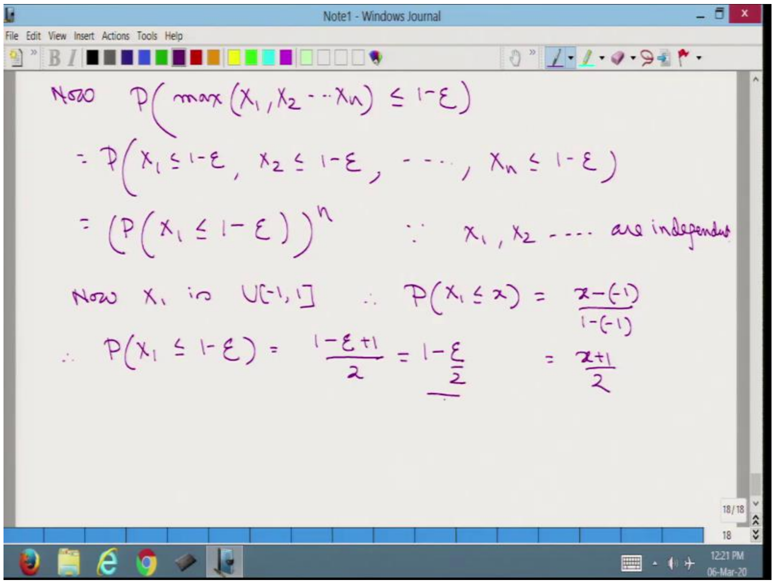
(Refer Slide Time: 29:32)



Solution let epsilon be greater than 0 then probability modules of y n minus 1 greater than equal to epsilon is equal to probability yn greater than equal to 1 plus epsilon plus probability yn less than equal to 1 minus epsilon is equal to probability maximum of X1, X2, Xn greater than equal to 1 plus epsilon plus probability maximum of X1, X2, Xn is less than equal to 1 minus epsilon.

Since xi are uniform over, minus 1 to plus 1, therefore probability maximum of x1, x2 etc greater than equal to 1 plus epsilon is equal to 0, because the upper limit of the value that the random variables can take is only 1. Therefore, probability modulus of yn minus 1 greater than equal to epsilon boils down to probability maximum of X1, X2, Xn is less than equal to 1 minus epsilon.

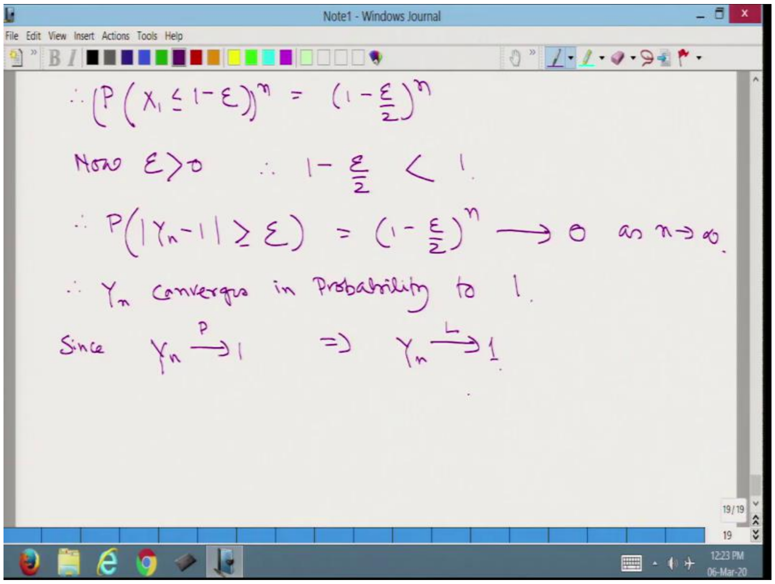
(Refer Slide Time: 31:49)



Now, probability maximum of X1, X2, Xn less than equal to 1 minus epsilon is equal to probability X1 less than equal to 1 minus epsilon X2 less than equal to 1 minus epsilon Xn less than equal to 1 minus epsilon is equal to probability X1 less than equal to 1 minus epsilon, this whole to the power n. Since, X1, X2 are independent.

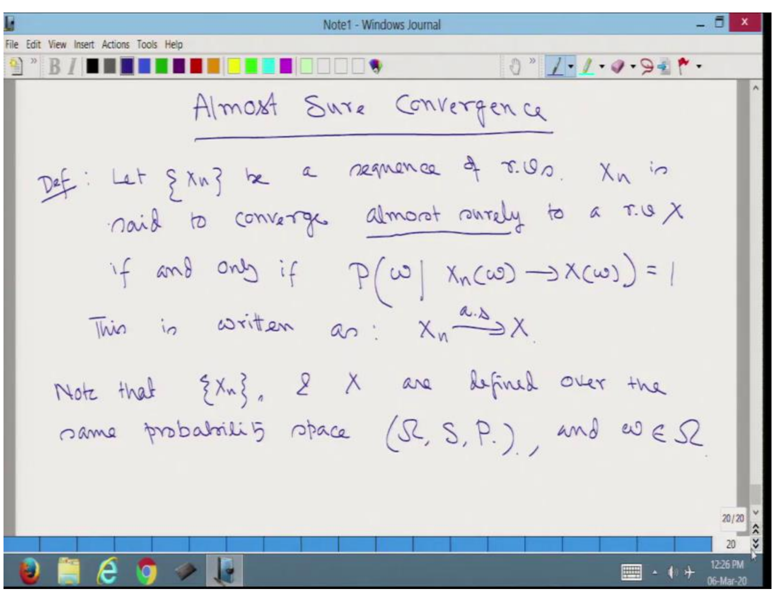
Now, x1 is uniform, in minus 1 to 1, therefore, probability x1 less than equal to x is equal to x minus 1 upon 1 minus 1 is equal to x plus 1 upon 2. Therefore, probability X1 less than equal to 1 minus epsilon is equal to 1 minus epsilon plus 1 divided by 2 is equal to 1 minus epsilon by 2.

(Refer Slide Time: 33:45)



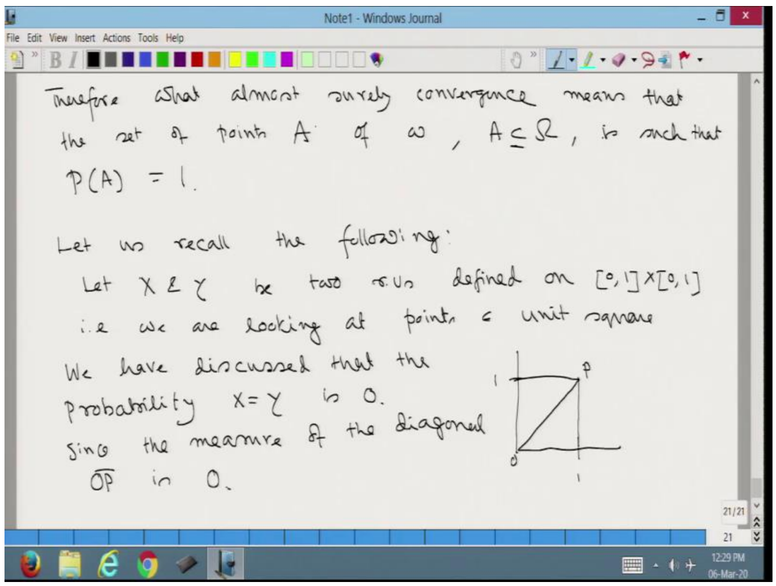
Therefore, probability X1 less than equal to 1 minus epsilon whole to the power n is equal to 1 minus epsilon by 2 whole to the power n. Now, epsilon is greater than 0, therefore 1 minus epsilon by 2 is strictly less than 1. Therefore, probability modulus of Yn minus 1 greater than equal to epsilon, which is equal to 1 minus epsilon by 2 whole to the power n converges to 0 as n goes to infinity. Therefore, Yn converges in probability to 1 and since Yn converges in probability to 1 implies that Yn converges in distribution also to the constant one.

(Refer Slide Time: 35:26)



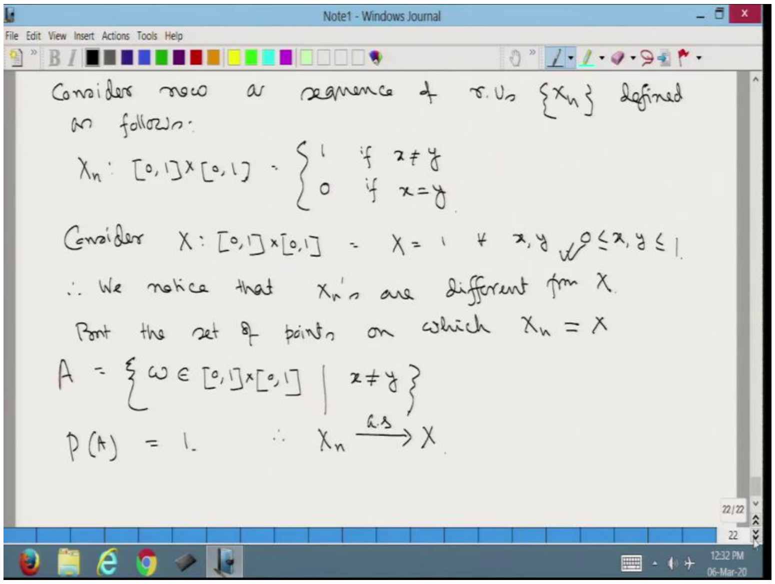
Okay friends, let me now talk about another mode of convergence, which is called almost sure convergence. Definition let xn be a sequence of random variables Xn is said to converge almost surely to a random variable x, if and only if probability of omega such that Xn omega converges to X omega is equal to 1 and this is written as Xn almost surely converging to x. Note that all Xn and X are defined over the same probability space omega is P and small omega belongs to capital omega.

(Refer Slide Time: 37:39)



Therefore, what almost surely convergence means that this set of points have A of small omega A contained in capital omega is such that probability of A is equal to 1. So, let us recall the following, let x and y be two random variables defined on 0 cross 1, 0 to 1, cross 0 to 1 that is we are looking at points belonging to unit square. We have discussed that the probability X is equal to Y is 0 because the event X is equal to Y means the points are chosen from this diagonal and since the major of the diagonal OP say is 0.

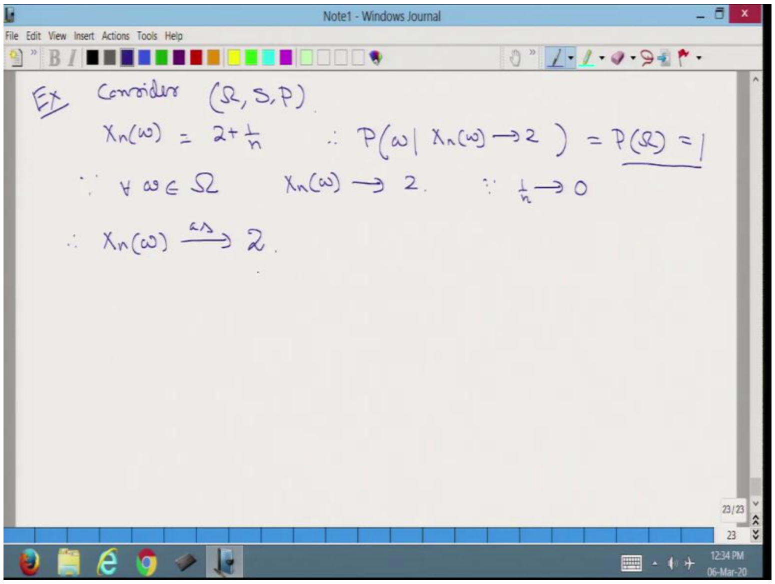
(Refer Slide Time: 40:22)



Therefore, consider now a sequence of random variables Xn defined as follows Xn is defined from 01 cross 01 such that this takes the value 1 if x is not equal to Y and 0 if X is equal to Y consider x to be defined on the same omega such that x is equal to 1 for all x comma y, such that 0 less than equal to XY less than equal to 1.

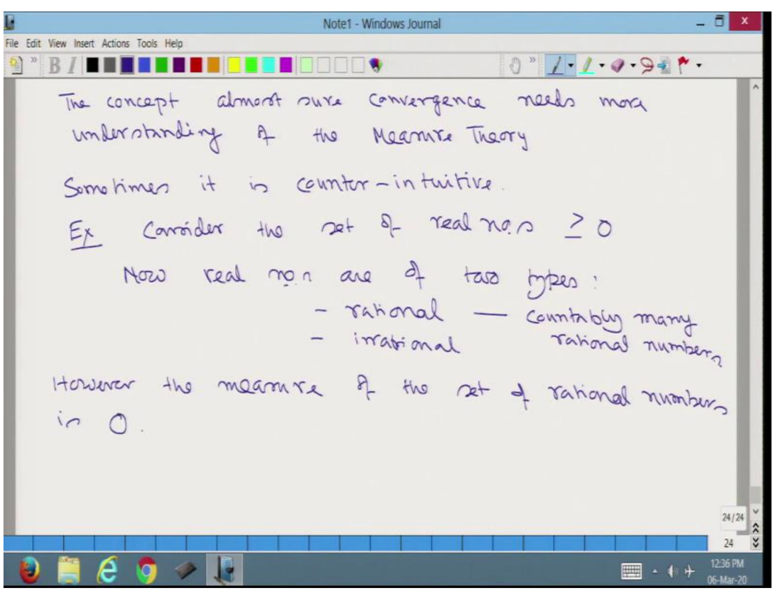
Therefore, we noticed that Xn's are different from X, but this set of points on which Xn is equal to X is equal to all this omega belonging to 01 cross 01 such that x not equal to y. So, if we call it A, then probability of A is equal to 1. Therefore, Xn converges almost surely to the random variable X, which takes value 1 throughout the entire unit square.

(Refer Slide Time: 42:54)



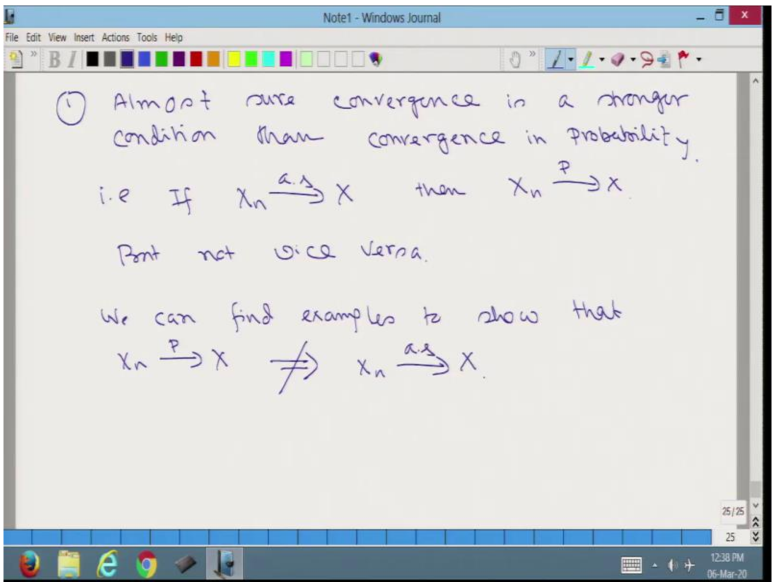
Another example consider some probability space omega SP such that Xn omega is equal to 2 plus 1 by n. Therefore, probability of omega such that Xn omega converges to 2 is equal to probability of omega is equal to one. Since, for all omega belonging to capital omega Xn omega converges to 2, since 1 by n converges to 0. Therefore, Xn omega almost surely converges to the constant 2.

(Refer Slide Time: 44:04)



Now, the concept of almost sure convergence needs more understanding of the measure theory. Sometimes it is counter intuitive say for example consider the set of real numbers say greater than equal to 0 that is all positive real numbers. Now, real numbers are of two types, rational and irrational. We all know that there are countably many rational numbers. However, the measure of the set of rational numbers is 0, although there are infinitely many rational numbers. To understand this concept, one needs to study deeper level of mathematics is I am not going much into the detail.

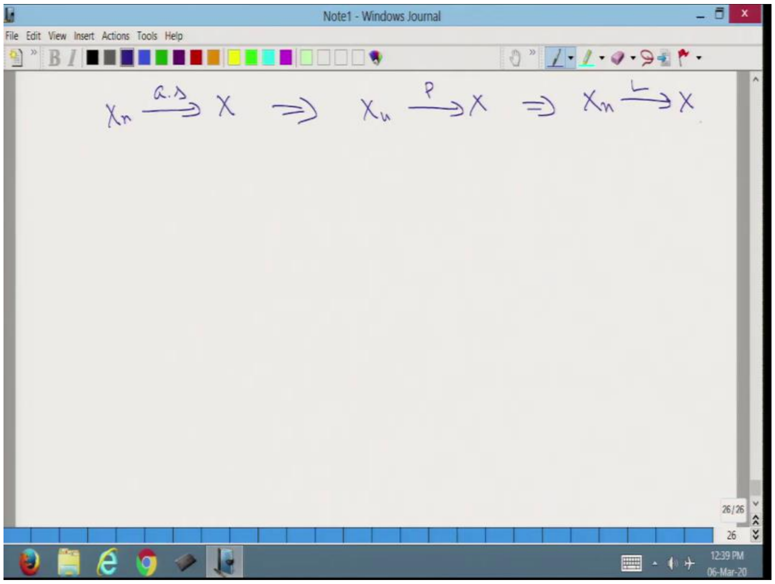
(Refer Slide Time: 46:29)



So, I conclude the talk with the following information that almost sure convergence is a stronger condition then convergence in probability That is if Xn converges almost surely to x then Xn convergence in probability to x, but not vice versa. We can find examples to show that Xn converges in probability to X does not imply Xn converges almost surely to X.

I want you to remember this fact, because we shall need almost your convergence for when we study strong law of large numbers in subsequent lectures.

(Refer Slide Time: 48:12)



Therefore, we conclude that Xn almost surely converging to x implies Xn converging in probability to x, which implies Xn converging in distribution to x. So, with that message, I stop here today from the next class. I shall start with laws of large numbers. Okay then thank you.