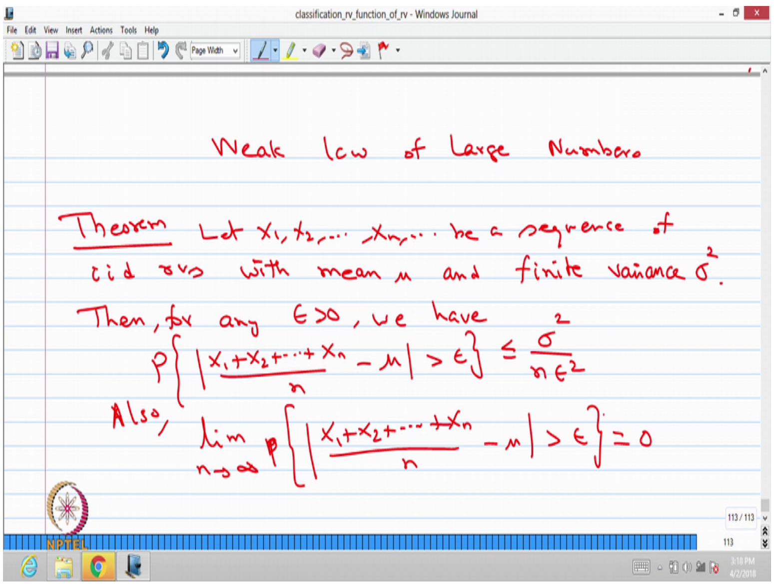
Introduction to Probability Theory and Stochastic Processes
Prof. S. Dharmaraja
Department of Mathematics
Indian Institute of Technology, Delhi

Lecture - 44

So, we are in Limiting Distributions model. In this model we have already discussed modes of convergence. In that we have discussed 4 different modes of convergence. First 1 is convergence in distribution, convergence in probability, convergence in art movement, convergence almost surely.

In this lecture, we are going to discuss law of large numbers. In that we are going to discuss two types of law of large numbers; 1 is called Weak law of large numbers.

(Refer Slide Time: 00:43)



Then later, we are going to discuss strong law of large numbers.

Let me give the definition of weak law of large numbers. As a theorem, let X 1, X 2 and so on X n and so on, be a sequence of iid random variables with mean mu and finite variance sigma square. That means, this sequence of random variable has a at least a second order movements.

Then for any epsilon greater than 0, we have probability of absolute of X 1 plus X 2 plus dot dot dot plus X n divided by n minus mu greater than epsilon. This probability is always less than or equal to sigma square divided by n epsilon square. Also, limit n tends to infinity probability of the event that is absolute of X 1 plus X 2 plus so on plus X n divided by n minus mu which is greater than epsilon. This becomes 0 as n tends to infinity.

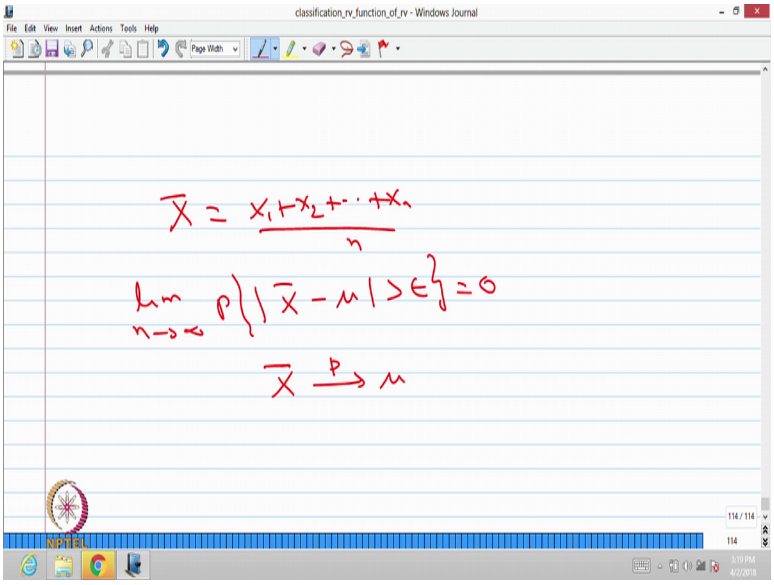
Then, we say that this sequence obeys weak law of large numbers. Here, large numbers means the sequence of random variables. When you have a many random variables and if you create sum of random variables divided by n minus the mu, in absolute sense greater than epsilon the probability of that event will tends to 0 that is what this weak law of large numbers says.

For that you need a sequence of random variable should have a at least second order movement; that means, mean exist as well as the variance exist and we made the assumptions those random variables are iid random variables; that means, independent and identically distributed random variable.

Therefore, the means are going to be same and the variance is going to be same. Then for any epsilon you can have a limit n tends to be infinity probability of event in absolute sense summation divided by n that is nothing but X bar which we have denoted earlier minus mu which is greater than epsilon is equal to 0.

If this condition is satisfied, then we can conclude this sequence of random variable obeys weak law of large numbers. Why it is called a weak law of large numbers? Because if you see the different modes of convergence, you can conclude if you make a notation X bar is equal to X 1 plus X 2 and so on plus X n by n.

(Refer Slide Time: 05:18)

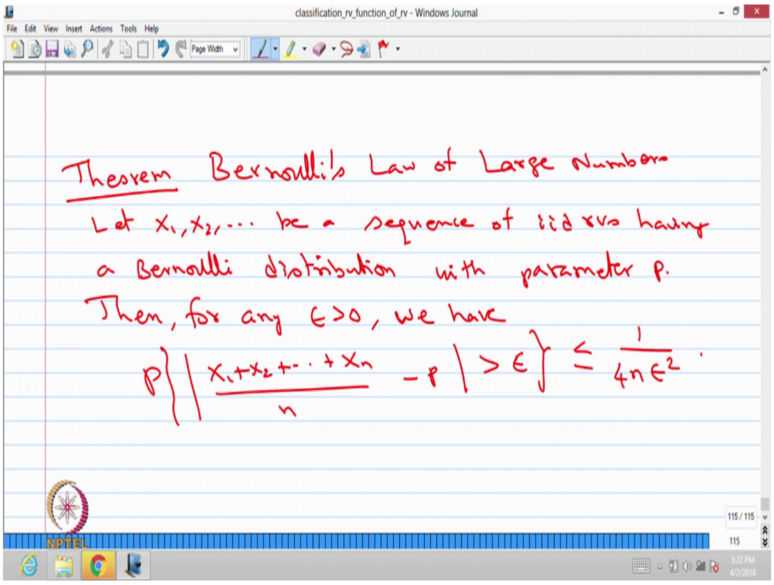


This result is nothing but limit n tends to infinity probability of in absolute sense X bar minus mu greater than epsilon that is equal to 0. If you see the definition of different modes of convergence, this is nothing but X bar converges to mu in probability.

So, since here the convergence in probability we call it as this sequence of random variable obeys a weak law of large numbers. Whereas, when we are discussing a strong law of large numbers, those sequence of random variables convergence to the some random variable and convergence in almost surely that is a strong law of large numbers; whereas, this one satisfies a convergence in probability.

Therefore, it is called the weak law of large numbers. We are not going to give the proof of this where as we are going for one Bernoulli law of large numbers that is a special case of the weak law of large numbers for that we will provide the proof.

(Refer Slide Time: 06:50)

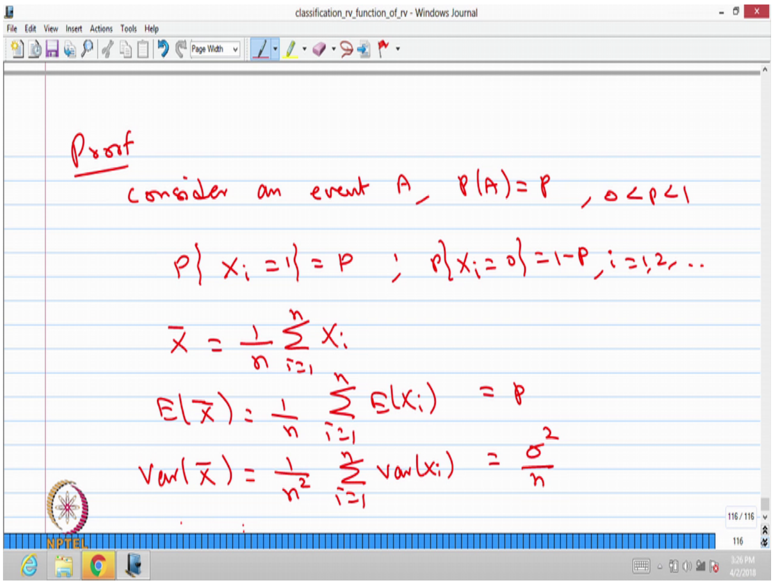


That is a next theorem that is a Bernoulli Bernoulli's Law of Large Number large numbers.

Let, X 1 X 2 and so on be a sequence of iid random variables having a Bernoulli distribution with parameter P. Then for any epsilon greater than 0, we have probability of absolute of X 1 plus X 2 dot dot dot plus X n divided by n minus P which is greater than epsilon. This probability of event is always less than or equal to 1 divided by four times n epsilon square. This is a special case of the earlier theorem.

If you see the earlier theorem, the weak law of large numbers we have a sequence of iid random variables with the mean mu and the finite variance, then we concluded as n tends to infinity the X bar converges to mu and converges takes place in probability. In the Bernoulli's law of large numbers in addition to the previous theorem. We have introduced the distribution of each random variable that is Bernoulli distribution with the parameter P.

(Refer Slide Time: 09:30)



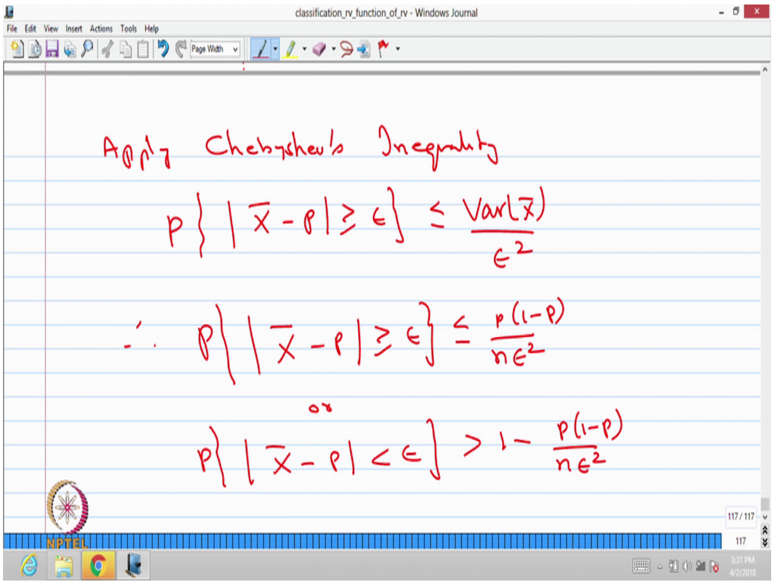
We are going to give the proof of this theorem as follows. Consider an event A whose probability is probability of A is P, where A is the success in each Bernoulli trial. A is the event in a each Bernoulli trial and the P of A, the P is nothing but probability of success in each Bernoulli trail. Since, each Xi's are Bernoulli distributed; therefore, the probability of Xi takes a value 1; that probability is that is P of A that is a event that is probability is P and the probability of Xi takes a vales 0 that is 1 minus P. This is for i is equal to 1, 2 and so on.

So, if you define a random variable X bar that is nothing but 1 divided by n summation of Xi's. We can find mu and variance of this random variable. The mean of this random variable is 1 divided by n and summation of i is equal to 1 to n mean of each random variables; each 1 is Bernoulli distributed.

Therefore, the mean is going to be P; therefore, summation of n P. Therefore, divided by n; therefore, it is going to be P. If you find out the variance of X bar that is nothing but 1 divided by n square, all are iid random variables. Therefore, it is summation i is equal to 1 to n variance of Xi's. Variance of Xi is a sigma square; therefore, it is going to be n sigma square when you make a summation. So, it is going to be sigma square by n. So, we got the mean and variance for the X bar.

Now, we may not know the distribution of X bar; whereas, we know the mean and variance of X bar. Therefore, we can apply the Chebyshev's inequality.

(Refer Slide Time: 12:07)

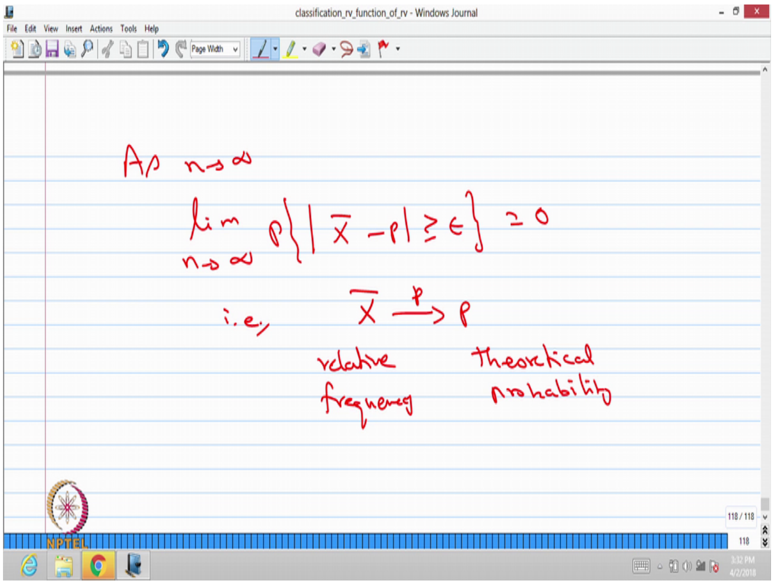


Apply Chebyshev's Inequality for the random variable X bar. So, what the inequality says? The probability of in absolute sense X bar minus their mean which is an greater than or equal to epsilon that is less than or equal to variance of a X bar divided by epsilon square.

So, just now we got variance of a X bar is sigma square by n; therefore, the probability of absolute of X bar minus P which is greater than or equal to epsilon that is going to be less than or equal to sigma square by n epsilon square. That means, the probability of X bar minus P that probability has the upper bound sigma square by n epsilon square or we can write the probability of the X bar minus P which is less than epsilon; that has the lower bound 1 minus sigma square by n epsilon square. Since, P is sigma square is P into 1 minus P; therefore, it is P into 1 minus P divided by n.

Therefore, here also I can do the simplification, where variance of that is P into 1 minus P by applying the Chebyshev's Inequality probability of the event in absolute X bar minus P greater than or equal to epsilon is less than or equal to variance of X bar divided by epsilon square. We know that each Xi's are Bernoulli distributed and variance of X bar is going to be P into 1 minus P divided by n times epsilon square or the probability of absolute of X bar minus P less than epsilon is going to be have a greater than 1 minus P into 1 minus P n epsilon square.

(Refer Slide Time: 15:00)

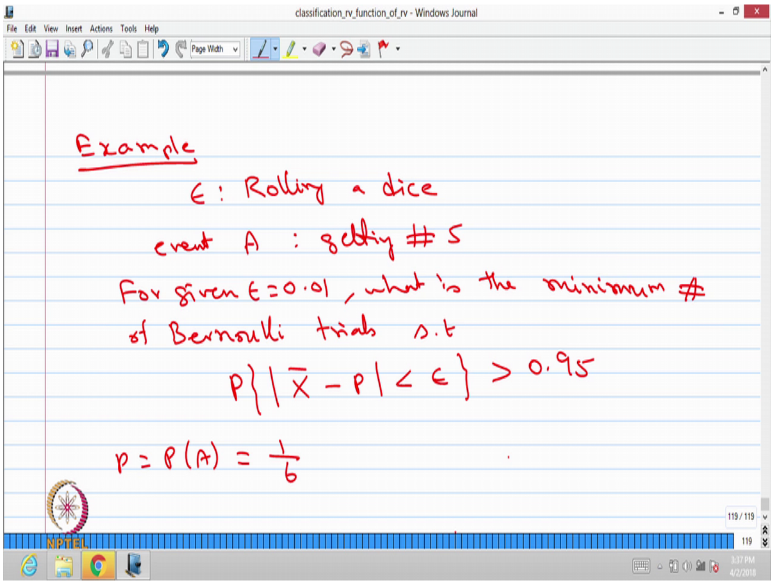


Here also we can go for as a n tends to infinity the limit n tends to infinity probability of absolute of X bar minus P which is greater than or equal to epsilon that is going to be 0; that means, the X bar tends to P convergence in probability. That means, for a larger n for a Bernoulli distributed random variable X bar 1 divided by n summation of Xi is nothing but the relative frequency.

So, the relative frequency converges to the theoretical probability; that is the theoretical probability. If you have a independent Bernoulli trials for a finite n the relative frequency may deviate from the theoretical probability, but for a larger n that the relative frequency will converge to the theoretical probability and that convergence take place in probability.

Let us go for one simple problem, how to use the Bernoulli law of large numbers as a example.

(Refer Slide Time: 16:40)



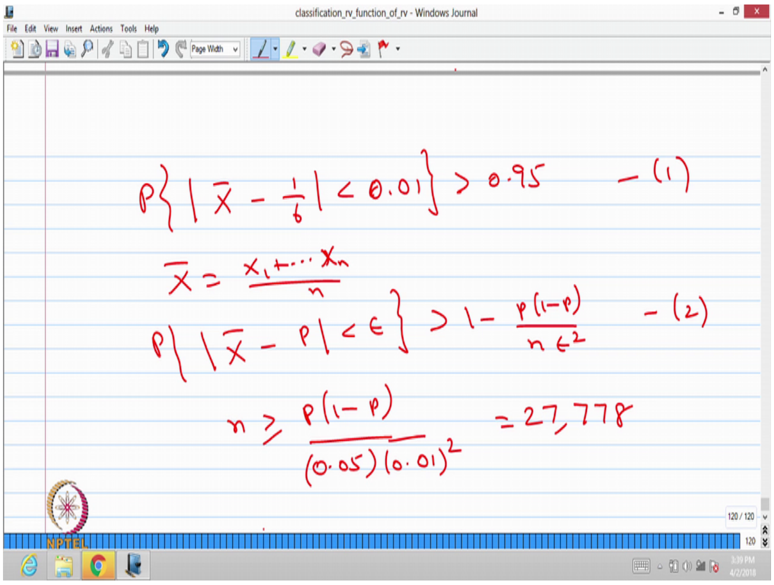
The random experiment is rolling a dice. For simplicity we assume that it is the fair dice. A event A is getting a number 5; getting number 5. Event A is nothing but the getting a number 5. We are repeatedly rolling a dice countably infinite number of times and the question is for a given for a given epsilon that is 0.01 what is the minimum number of a Bernoulli trials such that the probability of absolute of X bar minus P which is lesser than epsilon is going to be greater than 0.95.

The random experiment is rolling a dice countably infinite number of times. The event A is a getting a number 5 in each Bernoulli trial. The question is for a given epsilon, what is the minimum number of Bernoulli trials such that probability of absolute of X bar minus P which is less than epsilon is greater than 0.95? That means, a minimum how many number of times we have to roll a dice for getting minimum probability of a 0.95 within the length of epsilon which is deviated from the P.

For this problem the P is nothing but the probability of success of a event A that is getting a number 5 in each Bernoulli trial that is a 1 out of 6; I made it fair dice therefore, it is 1 by 6 that is P. So, we know P and we know the value of epsilon; therefore, you apply in the Bernoulli law of large numbers because we have a iid random variables, each are having a Bernoulli distributed with a probability of success P is 1 by 6.

So, the question is sort of reverse problem inverse problem finding the n such that this condition is satisfied.

(Refer Slide Time: 20:00)

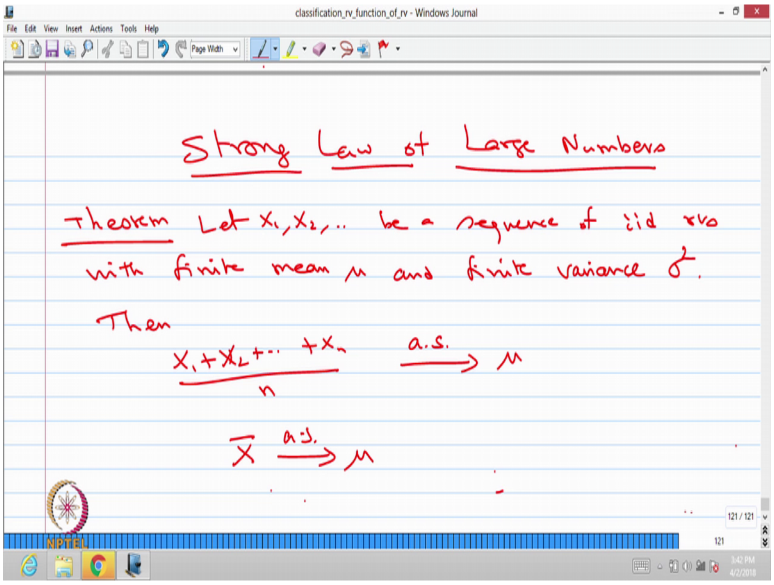


That means, probability of absolute of X bar minus 1 by 6 lesser than epsilon; epsilon is 0.01; this has to be greater than 0.95; if you simplify your X bar is here that is X 1 plus so on plus X n by n; therefore, if you compare this with the definition of probability of X bar minus P lesser than epsilon that is going to be greater than 1 minus P into 1 minus P n epsilon square.

So, now you compare equation number 1 with the 2; compare 1 and 2 we get n should be greater than or equal to P times 1 minus P divided by 0.05 into 0.01 the whole square; where, P is 1 by 6. If you simplify you will get 27778, we are finding the nearest positive integer.

That means the n has to be minimum 27778 valid. That means, if you roll a dice minimum a 27778 times we are going to attain the relative frequency deviation from 1 by 6 with the length of 0.01; probability of this event is going to be minimum point or the probability of this event is at least 0.95. So, for that the number of trials are needed is minimum 27778.

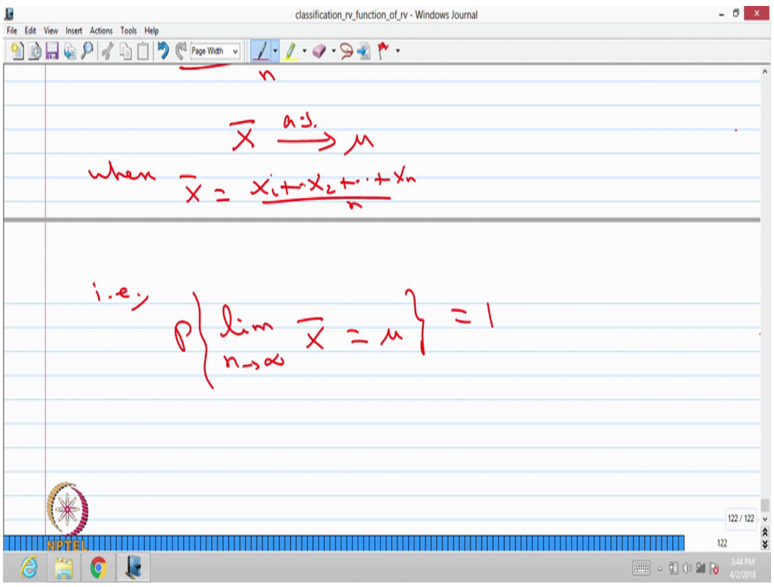
(Refer Slide Time: 22:33)



Now, we will move into the second law of large numbers that is a Strong Law of Large Numbers. Let me give the definition, the form of a theorem. Let X 1 X 2 and so on be a sequence of iid random variables with the finite mean mu and finite variance sigma square.

Then, X 1 plus X 2 and so on plus X n divided by n will converge to mu almost surely. We can define X 1 plus X 2 plus X n divided by n as the X bar. So, X bar if I define X bar as this X bar converges to mu almost surely.

(Refer Slide Time: 24:24)



Where, X bar is a X 1 plus X 2 plus X n by n; that means, if you have a sequence of random variables all are iid having a at least second order movements, then adding all those random variables divided by n that is nothing but the average of n random variables converges to the mean of these random variable that is mu and that convergence takes place in almost surely.

That means, probability of the limit n tends to infinity of X bar is equal to mu that is equal to 1. Limit n tends to infinity X bar is equal to mu that is equal to 1. That means, if you collect the possible outcomes in which X bar of w is tends to mu and if you collect those possible outcomes, whose probability put together is going to be 1.

Then we can conclude X bar converges to mu almost surely. So, without proof we are giving the strong law of large numbers. And, why the word strong law of large number is here? The convergence in almost surely that is the strongest one; whereas, the weak law of large numbers involves convergence in probability. That is a weak law of large numbers whereas, the convergence in almost surely that is going to be calling it as a strong law of large numbers.