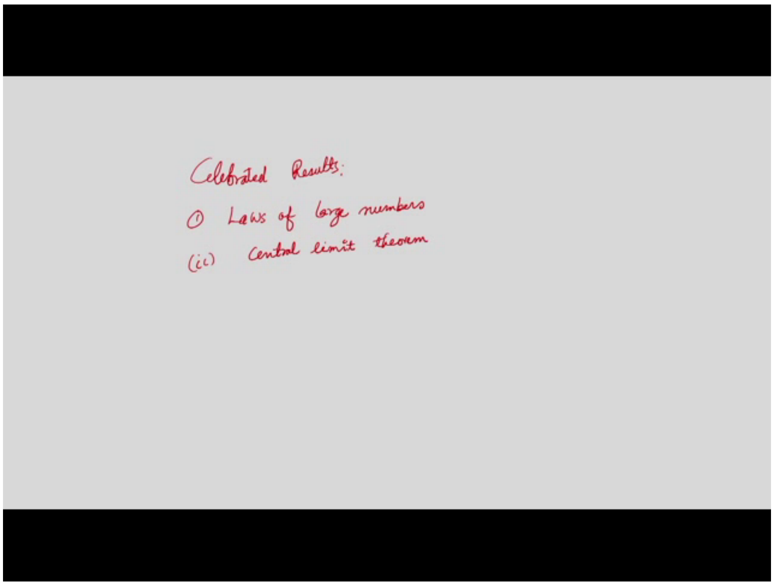
Advanced Topics in Probability and Random Processes

Prof. P. K. Bora
Department of Electrical Engineering
Indian Institute of Technology, Guwahati

Lecture - 10
Laws of Large Numbers

Laws of Large Numbers, in the last lectures we covered different modes of convergence, two of the applications of the convergence concepts are the laws of large numbers and the central limit theorem, so two important results.

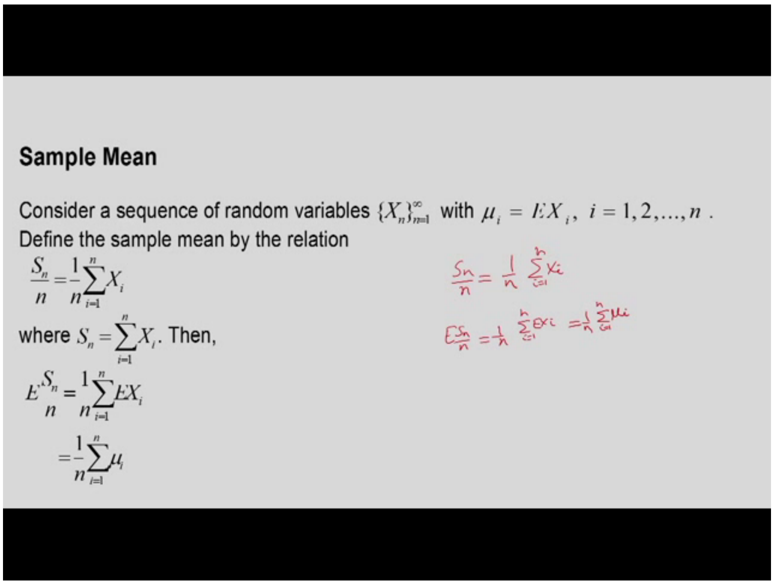
(Refer Slide Time: 00:55)



In fact, celebrated results that is number 1 is laws of large numbers, and number 2 is central limit theorem. So, laws of large numbers and central limit theorem are concerned with the behaviour of the average of large number of random numbers. Suppose we have a large number of random variables, we average them and how that average will eventually behave that answer is given by laws of large numbers and central limit theorem.

Now, these results that is theoretically very important, laws of large number and central limit theorem theoretically they are very beautiful results. And they have several areas of applications which we will be covering in this class and subsequent class. First we will discuss laws of large number.

(Refer Slide Time: 02:32)



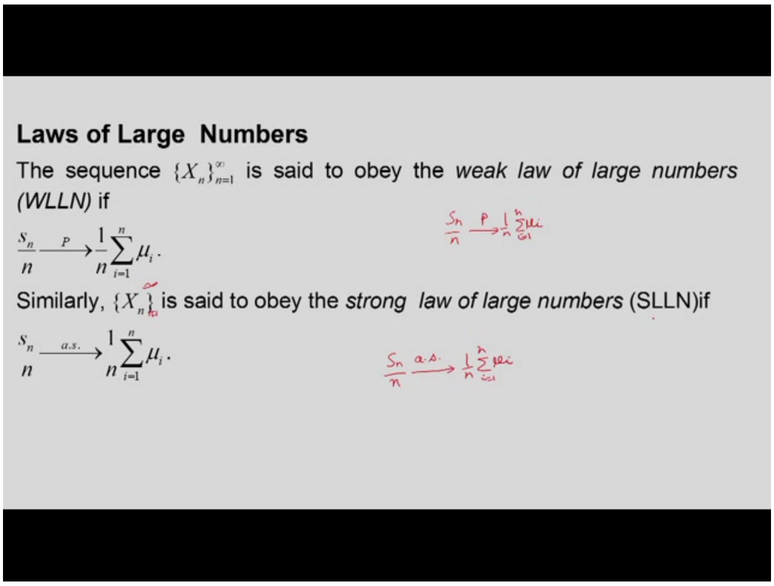
Basically this laws of large number is concerned with the behaviour of sample mean, what is sample mean? Consider a sequence of random variables and is suppose is has mean average value or expected value mu i. Then we define the partial sum S n is equal to summation X i, i going from 1 to n this is the partial sum and if I divided by n then I will get the average.

So, now this partial sum S n divided by n is the average. Now, how this average behaves that is the question. First of all let us consider the expected value of S n by n because, S n by n, how do I define? S n by n is equal to summation X i, i is equal to 1 to n divided by 1 by n. Now, this quantity is a random quantity because these values are random. So, necessarily we can find out what is the average value of this random quantity E of S n by n, so that quantity because it is a sum of random variables. So, expected value will be also sum therefore, this will be 1 by summation E of Xi, i is equal to 1 to n divided by n and as we know that E of X i is mu i. So, therefore, I can write this is equal to summation mu i, i is equal to 1 to n divided by n.

So, what we have seen that E of S n by n is equal to the average of the true means, average of the true means is of the random variable has a mean mu i therefore, we are taking the average. So, this quantity S n by n has an average value which is the average value of the true means. So, on the average that means, on the average S n by n is equal to the average value of the true means, but whether S n by n actually approach this quantity. So, for that we have to see the variance of this quantity S n by n what is the variance?

So, let us see that, so if that variance becomes smaller and smaller then we can say that S n by n a process this quantity as n tends to infinity. So, that way we get this statement of the law of large numbers; let us first give these statements.

(Refer Slide Time: 05:27)

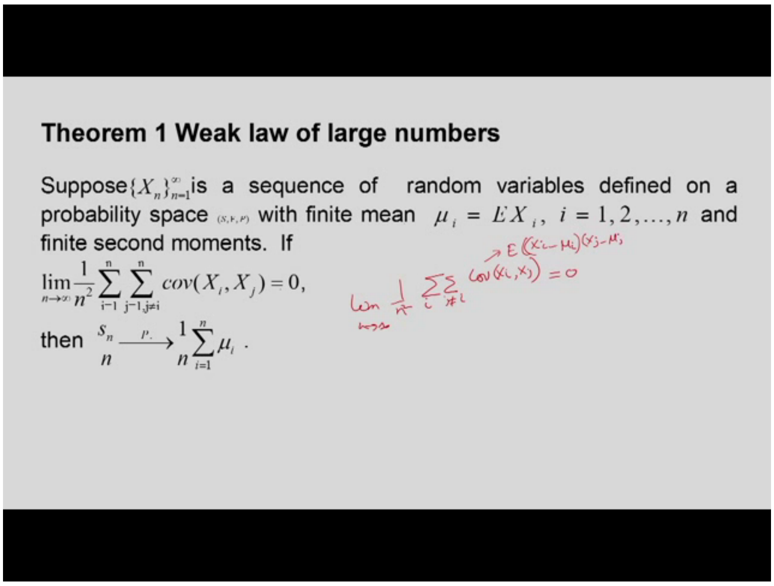


The sequence of random numbers or random variables X n is said to obey the weak law of large numbers that is we write it as WLLN, WLLN if S n by n that quantity converges in probability to the average of the true mean. So, what is the weak law of large number, that S n by n this is a quantity we are concerned with that is the sample average; if it converges in probability to the average value of true mean. Then we say that this sequence, sequence of random variables X i's they obey the weak law of large numbers. So, weak law of large numbers because here convergence is the in the weak sense that is why this is known as the weak law of large number.

Now, the sequence X i's if it obeys the weak law of large number that it means that this S n by n converges to the average of the true mean in the probability sense. So, this is the weak law of large number, similarly the sequence X n is said to obey the strong law of large, if this average sample average or sample mean converges almost 0 to the average of the true mean. So, this is the strong law of large number what does it say that S n by n that is the sample mean. Now as n tends to infinity this is convergent and it converges almost (Refer Time: 07:28) to the average of the true mean summation mu i i is equal to 1 to n divided by n average of the true mean.

Then we say that that the sequence X n obeys the strong law of large number it is abbreviated as SLLN. So, we have to see under what condition weak law of large number is true and under what condition strong law of large number is true, but both are concerned with the behaviour of this sample mean as n tends to infinity.

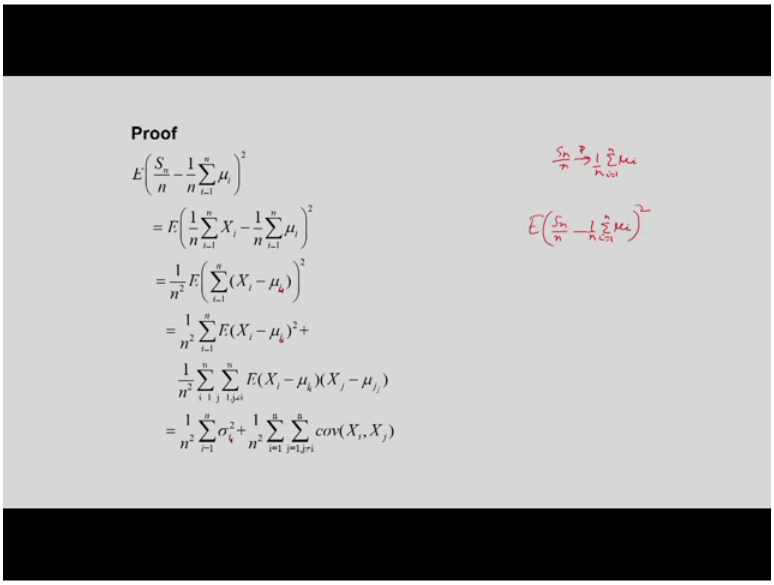
(Refer Slide Time: 08:08)



We will state one theorem, theorem one weak law of large numbers suppose X n is a sequence of random variables defined on the probability space S, F, P with finite mean mu i for each of Xi. If we impose this condition limit n tends to infinity of the double sum of covariance of X i, X j divided by n square if this limit happens to be 0, then S n by n converges in probability to the average of the true mean.

So, we impose the condition what is the condition if summation that is double summation covariance of Xi X j summation over i over j, but they are not equal to i because covariance between two distinct random variables, divided by n square if this limit n tends to infinity is equal to 0. Then we say we can say that this sequence X n of hd weak law of large numbers so that means, that sample mean converges in probability to the average of the true mean. Now this covariance of X i X j means E of X i this quantity is nothing, but E of X i minus mu i into X j minus mu j, so this quantity is the covariance.

(Refer Slide Time: 10:01)

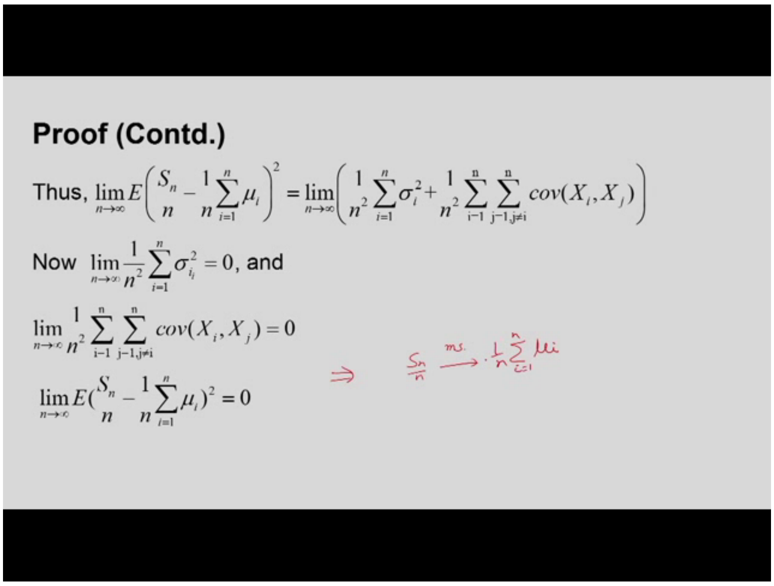


We have to prove that S n by n converges in probability to in probability to the average of the true mean, but convergence in probability directly proving is difficult. So, we will first see if it is convergent in mean square sense that means, we will see this quantity E of S n by n minus summation mu i, i going from 1 to n divided by n this is the quantity which left hand side should converge. So, we consider the convergence in the mean square sense. So, that way we consider the mean square value here and this if I write the values of S n as the summation of X i. Then I can write this as because 1 by n is here, if I since it is square value is there this 1 by n square can come out into E of now sum of the deviation. So, we are considering X i minus mu i mu i, so this sum going from i is equal to 1 to n, so this division sum square so we want to find out this quantity ok.

Now, this we see that this is a square of this sum, so we can break it in terms of the individual square and then the cross terms. So, individual square terms are there cross terms are there and then we can take the expected value. So, that way we will have this quantity first E of Xi minus mu i mu i whole square this sum plus the covariance sum and its sum is scale by 1 by n square and we notice that E of X i minus mu i whole square that is the variance sigma i square. So, I can write this as variance sigma i square summation i going from 1 to n, similarly this term is covariance of X i covariance of X i covariance of X i minus mu i, so this quantity the covariance of X i X j.

So, that way we have this quantity this mean square value is equal to summation of variance divided by n square n double summation of the covariance's divided by n square. Now our aim is what happens as n tends to infinity, we are given that this quantity is finite sigma i square that is a finite variances. So, each of the random variable has finite variance and also we are given about this summation.

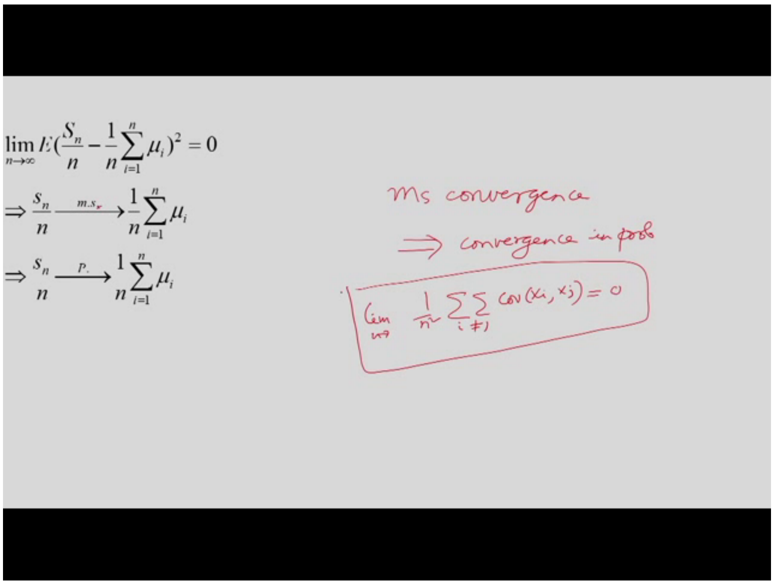
(Refer Slide Time: 13:14)



So, let us see, so if I have to take the limit of the left hand quantity this mean square value as n tends to infinity therefore, I will have the limit of this quantity and limit of this quantity. Now, limit of this quantity is equal to 0, because this sigma i squares are finite and this limit is also 0 because already given that and this is according to today condition given in the theorem this quantity is equal to 0.

Therefore, what will we will have that this quantity that means, square value as n tends to infinity limit of this mean square value is equal to 0. So, what does it imply this implies that S n by n converges in the mean square sense to summation mu i i is equal to 1 to n divided by n average value of the true means.

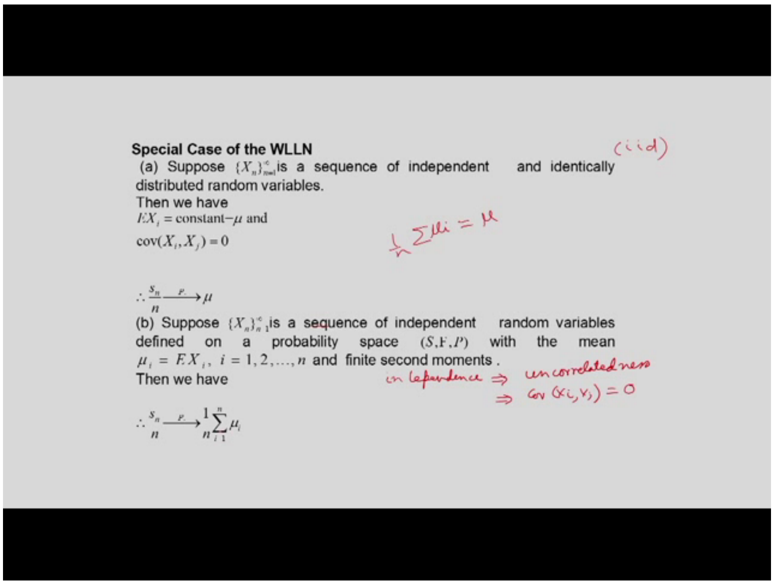
(Refer Slide Time: 14:20)



So, we get this result, so this also implies that now we know that mean square convergence ms convergence ms convergence implies convergence in probability. So, what we conclude that S n by n that is the sample mean converges in probability to the average of the true mean this is the weak law of large numbers. So, we have established one theorem according to this theorem what we require we require, suppose if we are given a sequence of random variable X i's and each has sum finite mean mu i and finite second order moment that where variance is also finite.

And if we have this condition summation of double summation of covariance of X i X j this is over i naught equal to j. So, if this double summation of covariance of X i X j scaled by 1 by n square if this is equal to 0 limit of this is equal to 0. If this is limit of this quantity is equal to 0 then S n by n converges in probability to the average of the true mean. So, this is the weak law of large numbers, so this is the condition we need to be satisfied.

(Refer Slide Time: 16:16)



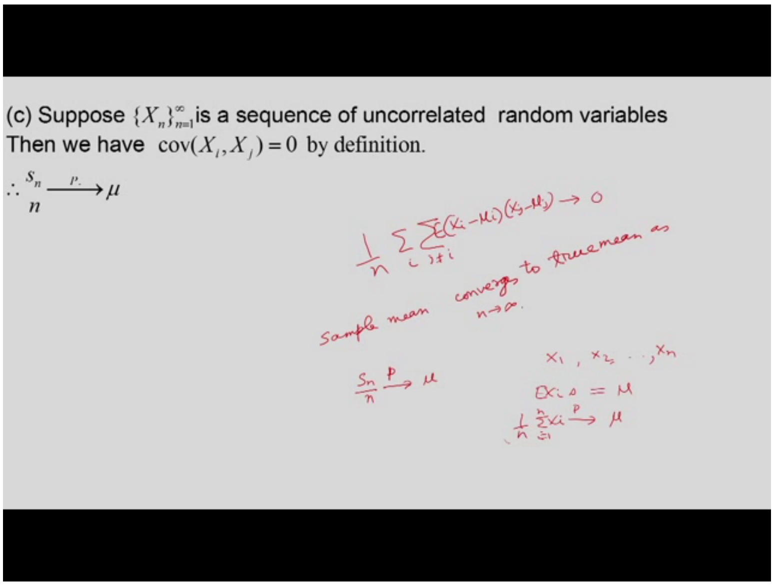
Now we what we gave me the general statement, now we can derive several simple cases suppose this special cases of weak law of large numbers are suppose a is X n is a sequence of independent and identically distributed random variables. So, if X n is known to be independent and identically distributed because suppose same random variable we are observing suppose that different instant of time. For example, if I am tossing a coin and corresponding outcome is a random variable suppose a Bernoulli random variable which takes value 1 and 0.

Now, this is a sequence if I go on independently tossing the coin then I will get a sequence that is a sequence of iid Bernoulli random variables. So, that way if I have an iid sequence of random variable in that case this condition will be automatically satisfied, the condition for weak law of large numbers will be satisfied because it is a iid. So, mean is in this case will be constant and because it is independent covariance of X i X j will be automatically 0 therefore, S n by n in probability it will converge to mu, because now a summation that is average of mu i is equal to nothing, but is mu i is mu, so n mu divided by n that is equal to mu.

So therefore, S n by n in this case converges in probability to mu, so this is one special case. Similarly now we can drop the condition for identically distributed only what we require is suppose mu i is given a of X i and the sequence is independent, X n is a sequence of independent random variables. So, in that case independence implies uncorrelatedness, independent implies uncorrelatedness.

What does it tell that is uncorrelated mean uncorrelatedness means covariance of X i, suppose if X i and X j are the random variables then covariance of X ij will be equal to identically 0. So, in this case also because this will become 0 because of independence condition therefore, S n by n will converge in probability to the average of the true mean.

(Refer Slide Time: 19:30)



So, this is case two, similarly we can consider (Refer Time: 19:32) because what we require is that covariance of X i X j is equal to 0, for that this sequence need not be independent this if the sequence X n is a sequence of uncorrelated random variables. Then also we will have covariance of X i X j automatically it will be equal to 0 because of the definition of uncorrelatedness and therefore, S n by n S n by n converges in probability to mu. So, these are special cases, but generally what we require that this quantity that is this covariance sequence. So, this covariance sequence has the property that if I consider the sum and then divided by n square, then this quantity limit of this quantity is equal to 0.

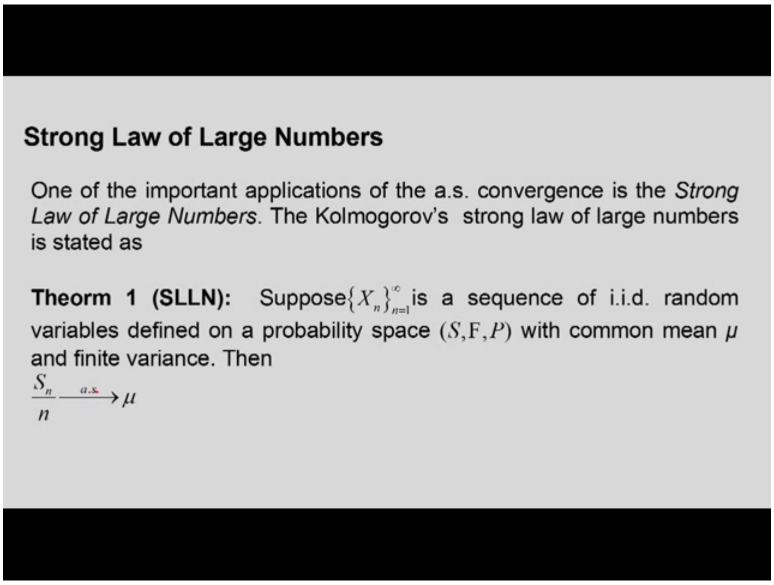
So, if we have suppose other conditions also we consider three special cases a b and c other than that anywhere where this sum becomes 0 summation that is summation X i minus mu i into X j minus mu j expected value of that that is the covariance. Summation this is a double summation over i and j not equal to i divided by n square, if somehow this sequence goes down to 0. So, ultimately depending on the property of the covariance sequence this weak law of large numbers are obeyed if this condition is satisfied then covariance satisfy this condition. And of course, other condition is if mean is of course, mean need to be finite and variance also needs to be finite in that case S n by n will converge in probability to the average of the true mean. That is the weak law of large number that is sample mean converges to true man as n tends to infinity.

So, this is a very powerful result that means, what we expect that this quantity under the condition already stated this S n by n will be as close as possible to the true mean if we are considering only iid case for example, with constant mean. Then this S n by n can be made arbitrarily close to true mean S n by n converges to mu in probabilistic sense. For example if I consider suppose measurement of a constant quantity maybe that constant quantity itself is mu by means of some noisy observations. Now these observations are suppose X 1 X 2 up to X n etcetera these are the observation noisy observation of the same constant quantity mu.

Now that means, X 1 X 2 X n etcetera they are the random variables because of the noise. Now they represent a constant quantity that is expected value of this E of X i's are E of X i's will be equal to mu is E of X i is equal to mu. Then what this now weak law of large numbers says that if I take this sample mean. So, summation X i i is equal to 1 to n divided by n this is the sample mean, what is this is the average of the noisy samples average of the noisy samples that will converge to true value to true mean.

Now, this quantity is the unknown quantity which we want to observe our which we want to measure. So, because measurements are noisy we take the sample average and this sample average as we are taking more and more number of observations then this sample average will be close to mu. So, that is our belief we generally in our any experiment we do that way if we have number of observations then we take the average. So, that that average is more close to the unknown quantity, so that is a consequence of weak law of large numbers. There are several other consequences for example, the interpretation of probability as relative frequency that can be easily explained in terms of weak law of large numbers, we will now consider the strong law of large numbers.

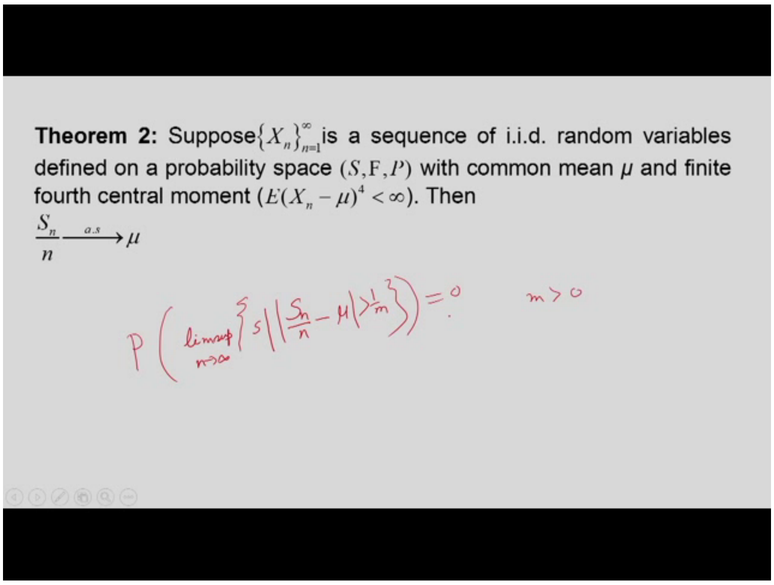
(Refer Slide Time: 25:47)



So, this is the consequence of almost sure convergence, so strong law of large number is concerned with convergence almost sure convergence with probability one. So, what this theorem says first of all when these sequence we have a sequence of random numbers. So, that sequence obeys strong law of large number if sample average converges to the average of the true mean almost sure, that is the strong law of large number if sample average converges almost sure to the average of the true mean.

So, that is the strong law of large number, but we will be considering a special case of strong law of large number that is Kolmogorov's strong law of large numbers. It says that if X n is a sequence of iid random variables here, we are concerned with only iid independent and identically distributed with common mean mu we are considering the special case common mean mu and finite variance. So, if we import this iid condition then S n by n converges almost sure to mu, so it converges almost sure to mu. So, we will try to prove is a special case of this so this is the strong law of large numbers.

(Refer Slide Time: 27:30)

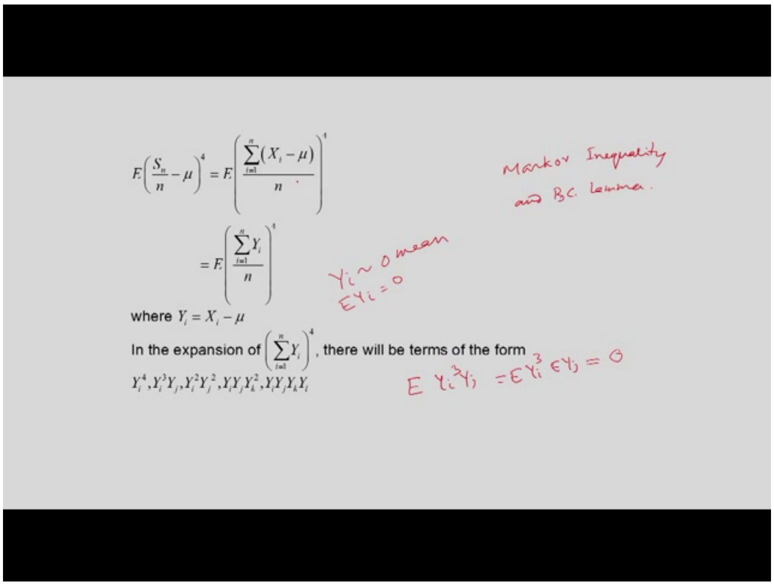


Suppose we impose this condition this sequence this is iid with constant mean mu and finite variance, but we also require that finite fourth order moments. So, this is the fourth order central moment E of X n minus mu whole to the power 4 that is less than infinity that is that is finite, we require that we impose the condition that this sequence it is a sequence of iid random variable, but it is a restricted case where the fourth order central moments are also finite.

So, in that case we will prove that S n by n converges in almost sure sense to mu. So, what we want to prove then we want to prove that and that is S n, suppose S n by n minus mu this deviation suppose if I consider the deviation of this is greater than m suppose we will consider the all sample points for which this happens. So, what we require s such that this deviation from mu is greater than 1 by m if we consider this event ok.

Now, we will consider the limsup of this event as n tends to infinity. So, we can define the limsup of this deviation, so we are considering all those s for which there is some deviation and here m is greater than it is a positive integer greater than 0. So, now what we require is that probability of this event should be equal to 0. So, convergence almost sure means this quantity that probability of limsup again tends to infinity of those values of s for which S n by n minus mu that deviation is greater than 1 by m. So, if I consider the infinitely often happening those deviations and probability of all those deviations, infinitely occurring deviations the probability should be equal to 0 that is the condition for almost sure convergence

(Refer Slide Time: 30:27)



So, we will prove this first at this again this also will be proving using the Markov inequality, Markov inequality and Borel Cantelli Lemma BC Lemma. So, because we are given a condition on fourth order moment, so we will be examining the fourth order moment first that is E of Sn by n minus mu to the power 4. So, this is if I simplify this then I will get an expression like this and this quantity now I will denote by Y i. So, that Y i is a 0 mean random variable, so Y i is equal to X i minus mu, so that way Y i is a 0 mean 0 mean it is 0 mean E of Y i is equal to 0.

So, but here this quantity is summation of Y i to the power 4, so it will because it is a power 4 we can expand the expand it and in the expansion we will have terms like this Y i to the power 4 then Y iq Y j then square terms then Y i Y j Y k square etcetera. And we have to take the expected value of each of the terms, but if we observe closely then we see that what will be the expected value of any term involving Y iq, because these are independent sequence we are considering iid sequence. Therefore, expected value for example, if I consider the expected value of E of Y i cube Y z, so what it will be that will be equal to E of Y iq into E of Y j.

Now, E of Y i cube because Y i is a 0 mean it is a centred around mu, so that way this quantity will become 0 this quantity will also become 0, so because of that this will become 0. Similarly, all odd moments will become wherever any odd term is there suppose this term Y i Y z. So, this is Y i there Y j is there Y k square is there because of Y i and Y j this term will also expected value will become 0. Similarly, if I consider the expected value of this will become also this will also become 0. So, with this observation now if I have to expand this expansion can be simplified I will have only this quantity.

(Refer Slide Time: 33:37)

$$\therefore E\left(\frac{\sum_{i=1}^{n} Y_i}{n}\right)^4 = \frac{1}{n^4}\left[nEY_i^4 + {}^nC_2 \times {}^4C_2\left(EY_i\right)^2\right]$$

$$= \frac{EY_i^4}{n^3} + \frac{3n(n-1)}{n^4}\left(EY_i^2\right)^2$$

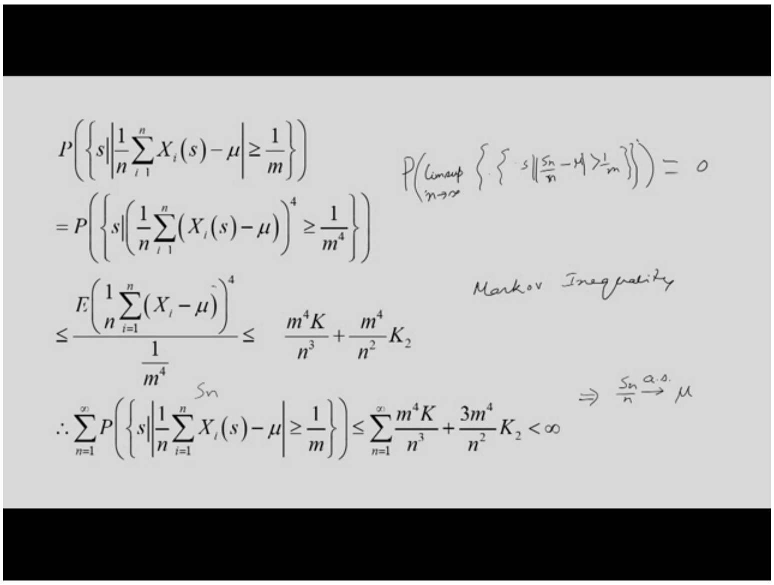
$$\leq \frac{K}{n^3} + \frac{3}{n^2}K_2$$

$$\therefore E\left(\frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)\right)^4 \leq \frac{K}{n^3} + \frac{3}{n^2}K_2$$

So, this will be 1 by n to the power 4 into n E Y i to the power 4 plus n C 2 into 4 C 2 into E of Y i square whole square. So, this we can further simplify and we can write like this we are just writing the expansion for n C 2 and 4 C 2 and therefore, we will get like this. Now, we are considering as when n is large now suppose if I consider this quantity, if I take n outside then this will be n square.

So, that way I can write this equality in terms of this inequality because this is a some constant K divided by n cube plus 1 by n square into K 2. All other terms suppose we consider this quantity n 3 etcetera all combined into this K 2, so what we get is that Y i i know that that is X i minus mu. So, that is fourth power of the average expected value of that that will be less than equal to K by n cube plus 1 by n square into K 2, so this is the result we have established.

(Refer Slide Time: 35:15)



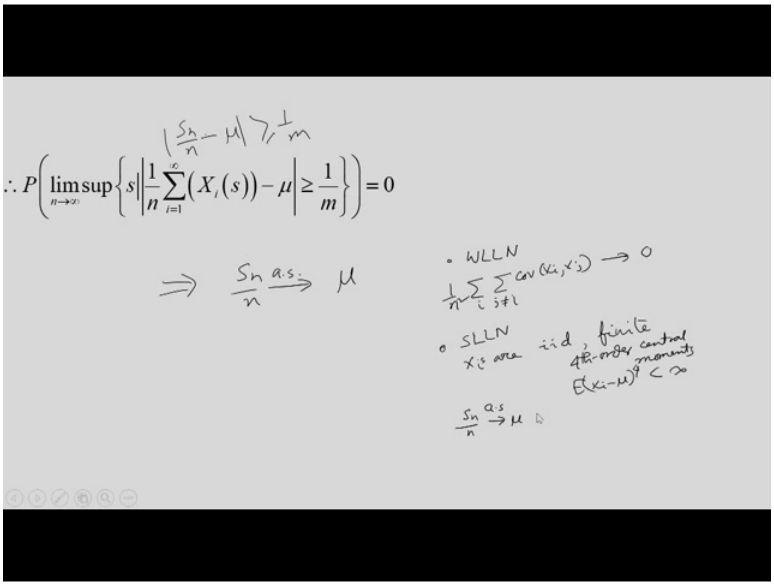
Now, let us see, so we are interested to determine the probability of the deviation. So, probability of S such that S n by S S n by n minus mu that deviation is greater than 1 by m, so this is the event we are considering the probability of the limsup of limsup as n tends to infinity of this event. So, this is a sequence of event is there because S n by n is there, so we are considering the limsup of this and probability of this limsup should be equal to 0 that we want to establish.

So, for that we consider this event, so probability of this event now we essentially we want to apply the Borel Cantelli lemma for this we want to evaluate this probability. Now, this probability is same as probability if I both side if I take the fourth power and right hand side also fourth power. So, that way this probability is same as this probability this is less than equal to now we can consider the expected value because we can apply the Markov inequality, now Markov inequality we apply Markov inequality. So, because of Markov inequality we will get that this probability is less than equal to this quantity ok.

So, now we have already established that this is less than equal to m 4 because m 4 will come because of this into K by n cube plus m 4 by n square into K square into K 2 m 4 by n square into K 2. So, now, if I have to consider the sum of this probability as n tends to 4 n is equal to 1 to infinity that is infinite sum of this probability sequence if I consider wherever there is a deviation all probability if I sum up. Now that will be less than equal to sum of this quantity and I know that 1 by n cube either convergent series 1 by n square is also a convergent series. So, this sum will be finite, so we got a condition that the sum of this deviation probabilities is less than infinity.

Now, we can apply the Borel Cantelli, Cantelli lemma and, so and prove the theorem. So, now, it is simple because this condition is satisfied, so this part is my S n by n. So, what does it say then this implies that S n by n this is a condition for almost sure convergence S n by n converges almost sure to mu. So, this is the consequence of Borel Cantelli Lemma.

(Refer Slide Time: 38:52)



So, this is this strong law of large number therefore, what we have established that probability of limsup of the event S such that S n by n that is the quantity of our interest, interest minus mu that is greater than 1 by greater than equal to 1 by m this is 1 by m.

So, limsup of that is equal to 0 this implies that S n by n converges in almost sure sense to mu. So, this is this strong law of large number we have established for a particular case where the fourth order central moments are finite for the sequence considered a fourth order central moments are finite. And in that case if X i's are iids then the sample mean converges almost sure with probability one to the true mean mu this is the strong law of large number.

So, that way we discussed about weak law of large number which requires a condition that sum of the covariance sequence divided by n square that quantity goes down to 0 as n tends to infinity. That is we considered we club large number what is the condition we establish that X i's are sequence with finite mean and variants and if in addition to that covariance of X i X j if double summation i j not equal to i. If I considered a double summation divided by n square if that goes down to 0 then X i's will already weak law of large number.

Then we consider the special cases if X i's are IID if X i's are independent if X i's are uncorrelated, then weak law of large numbers are satisfied. For strong law of large numbers we saw that we need this special condition that is independent and identically distributed sequence of random variable a sequence of independent and identically distributed random variables was considered in the case of strong law of large number SLLN strong law of large number. So, X i's are X i's are iid they are iid and we proved it that is with finite mean and variance, but we impose another condition that finite fourth central moments, moments.

What does it means that is E of X i minus mu to the power 4 that is finite in that case what we proved that this S n by n that is the sample mean will converge almost sure to the true mean. So, this is the strong law of large number, so that way we establish two of the very important result in probability that is strong law of large number and weak law of large numbers. We discussed weak law of large numbers in more details, but strong law of large numbers special case a special case of that we covered.

So, next we will cover another important result that is the central limit theorem.

Thank you.