Probability and Stochastics for finance
Prof. Joydeep Dutta
Humanities and Social Sciences
Indian Institute of Technology, Kanpur

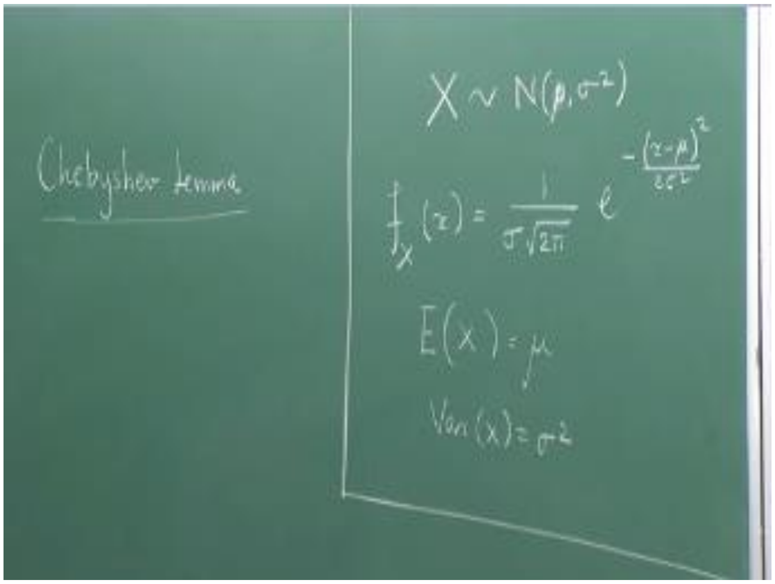
Lecture - 04
Cheybyshev Inequality, Borel-Cantelli lemmas and related issues

Yesterday we had spoken about random variable, their distribution function, their density functions. We have not gone into too much detail because we expect that you can look it up in any standard probability book. So, again just I want to recollect that distribution functions are essentially tools which allow you to view the range of the random variable, that is the whole idea and density functions specially for continuous random variable allows you to calculate the probability.

And let us now mention that certain random variables can behave in a certain typical way. That is they can follow a certain type of distribution function that is their distribution function would be actually known means you can actually know the form of the distribution function. Most of you who have done some basic probability in your basic engineering courses would know that there is a famous distribution for a random variable to follow is the normal distribution sometimes called the Gaussian distribution.

So one mathematician had told me that being a mathematician I should always call it to be a Gaussian variable, but I somehow want to, I always like to say that x follows a normal distribution with mean, median, and variant sigma square. So this distribution has a density function of this form.

(Refer Slide Time: 01:46)



So this is a very famous one and you know that the expectation of x or the mean value is meu and the variants, so please remember our aim is not to teach you very basic probability and statistics. I am sure there are a lot of courses in (()) 02:27 which many couple of courses which will do it. Our aim is to very quickly with the very basic introduction to take you to the mathematics or probabilistic tools that are required for finance. Our aim is not just knowing probability theory then we should not have reached this point in this few lectures.

So what is our next step. Our next step is to learn this very important lemma called Cheybyshev's lemma and then learn the Borel-Cantelli lemma. These 2 lemmas are useful throughout, they will be used throughout because they actually try to look at the behavior of sequences of events and sequences of events are very important when you talk about finance. So let us write down the Cheybyshev's lemma.

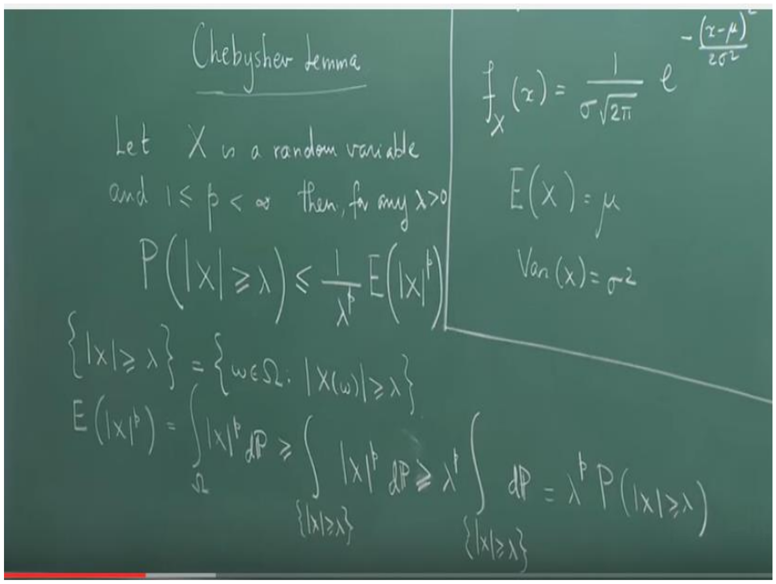
(Refer Slide Time: 03:27)

So X is a random variable. So we obviously distinguish random variable from a random vector. We had just given a definition of random vector in the last class, but we will just essentially consider random variable that is this X is from omega to R. So let this be a random variable and p is a number which is lying between plus 1 and plus infinity.

Then the probability, see a lot of things in mathematics is actually you are estimating certain quantities which you cannot calculate. So essentially this tells you how to estimate probability if you know if you can find an average, right. So it tells you that so if you take the modulus of the random variable that is basically you are essentially talking about non-negativity here and this is bigger than equal to lambda, a lambda and for any lambda greater than 0.

The fact the set of all omegas for which the modulus or absolute value of x omega is bigger than or equal to this positive quantity lambda, this probability is always less than whatever p you choose between 1 to infinity does not matter, you will always have this to be true. This is called the Cheybyshev's lemma. Now how do you prove it. Proof is very simple.

(Refer Slide Time: 05:21)



So let me denote in short this set which is the set of all such that, actually greater than equal to, so this whenever I write this I am actually meaning this that has to be kept in mind. Now how do I formalize the steps. So again if I go by the standard definitions which we have already given in the last class and this is nothing but integral.

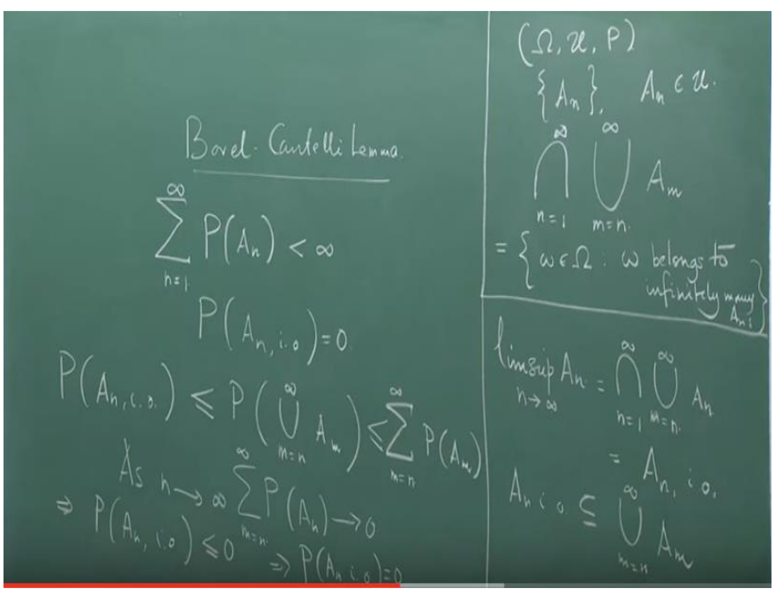
Of course, everything can be written in terms of the distribution function which you know, but here we are writing it in the more compact integral or measured theoretic form. This a whole is obviously bigger than the part so this is so in general either if I consider the omega is just for this set obviously these are subset of this so the integral would be smaller because you are talking about non-negative functions.

So this would be observe that what does this mean on this particular set mod of x omega is always bigger than lambda so it will be greater than equal to lambda so greater than equal to lambda to the power of p because here I have the raised the modulus of X to the power of p so what I would finally get from here is again I have to put the greater than equal to notation. So I have lambda to the power of p.

So what does this mean? This simply means the probability of this particular event, probability of this event so this is nothing but lambda to the power of p probability of and you see you have proved the Cheybyshev's inequality. Of course, Cheybyshev's inequality will have applications which we will come very soon.

Another interesting idea that we will come to is the Borel-Cantelli lemmas. So we are now going to talk about what is the meaning of the notion of a sample point occurring infinitely often in a given sequence of events. So if there is a sequence of events say 1 dot dot dot n the question is that if I have say omega element of the sample space what is the probability or what does it mean by telling that this omega occurs infinitely often in this sequence, right.

(Refer Slide Time: 08:59)



So what we will first do is to consider a probability space and then I will consider a sequence A n where each A n belongs to U and then I will write this I will construct make this construction. I will fix up a value of n and then take m is equal to n infinity and then I will change and then I will shift the n so I will say I will take n equal to 1 then I will take 1 to infinity. So I will take the union of all the events, all the events here, then sorry I made a mistake this is intersection.

So what I am going to do, first I will take say n equal to 1 so take the union of all things union of all the sets here then I will take off take out A 1 from there so I would start from A 2 to A m, I will take the union of that and keep on doing it and then try to find a common intersection, right. This thing is often written as follows.

So what sort of final set it consists of. Now this is an element of the, question is, is this an element of the Sigma-algebra. So once you take the union the union is in the Sigma-algebra although unions are basically what you are taking you are taking the unions. So each of the union are in the Sigma-algebra. If you have countable union in the Sigma-algebra, then you know that the countable intersection are also in the Sigma-algebra, right because you can take the (())11:35 law.

So this set is actually an element of the Sigma-algebra. So this is, what sort of sample points are here. So this is all those things where omega belongs to infinitely many A ns. So what does this mean, belonging to infinitely many A ns. So it might belong to A1 then A3 then A5, then A6, then A7, then A8 and A9 and then say from A100 to A101 and just goes on, something like this. Suppose one omega belongs to A1 and A2 nothing else it does not come anywhere then we do not consider it here.

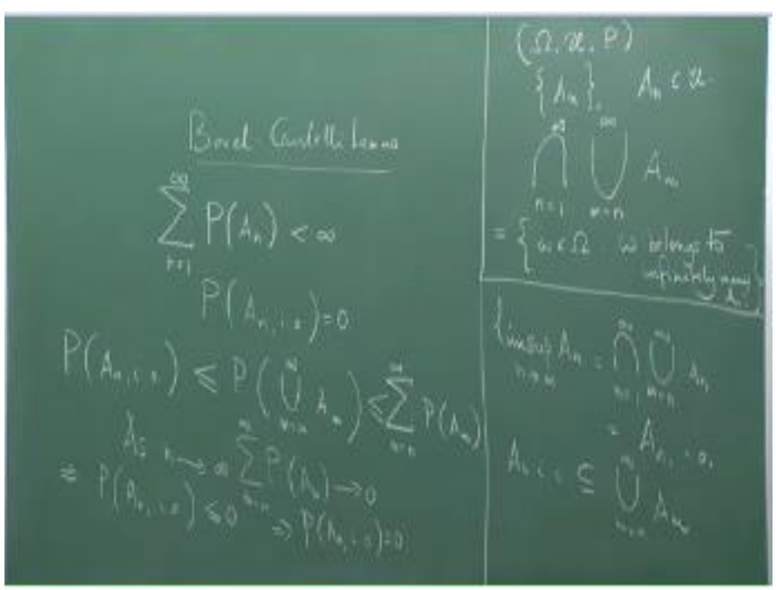
This is sometimes written as this set is also sometimes written this set is also sometimes written as because it satisfies the type of thing that we do with actual sequences of real numbers. It looks almost similar to that. I will allow you to ponder over what I am writing here whether it makes sense intuitively to call it (()) 13:06. So we also call it this.

But Borel-Cantelli lemma says that this gives you a more fascinating proposition. It says that now okay each of these P A ns if this is there if this is then this is a null event, this is a null event, that is beautiful. That it says if your probabilities have a finite sum this infinite series then this event must be a null event. It is a huge conclusion you are drawing. That is there is no omega which will belong to infinitely many A ns. If we (()) 13:55 all the omegas will belong to some finitely many A ns. That is a huge conclusion.

Sometimes you write this set also as some authors also write A n infinitely often. So, this one is just a symbolism of this, it does not mean anything. So I will just use this shortcut symbol so omega belongs to A n. So what would happen that there will be some A n in which there will be infinite number of omegas coming. That is the key idea here. Okay, do not bother about this point. Just you are essentially constructing a set that omega belongs to infinitely many A ns and we are trying to find the probability of this.

These types of sets become very important when we will study central limit theorems so we will not do much proves there, but we will give you a hint of what can happen and so let us draw the conclusion. so if this happens it tells you that probability A n i o it is same as writing the (()) 15:30 0 that this is a null, set empty set. There is no omega which belongs to infinitely many A n s. So how do you prove this. Proof is very simple.

(Refer Slide Time: 15:51)



So if we look at this set, so let me take some n and look at this union m equal to n to infinity say n equal to 3 say, m equal to 3 to infinity. So that set is one of the set. So if you have A1 intersection A2 the intersection must be smaller must be a subset of either A1 or A2 is a subset of both. See if I take any A n it does not matter. This whole thing is a subset of union m equal to n to infinity because what I am trying to say is that.

This whole that is A n i o is a subset of for some n whatever n you choose not for some n all n actually whatever n you choose does not matter. This is always true. So if that is true which will immediately tell you that the probability of course the larger set would have a larger event but event which has more possibilities of occurrence will have a larger probability. But what does this mean, the second means by the very Kolmogorov axioms, but if I look at this.

What is this, this is a tail of a sequence tail of a rather a tail of a series summation P A and n equal to 1 to infinity is a tail of the series that is we are shutting off finite amount of elements cutting off finite number of terms of the series and looking at the remaining part. So the tail of the series will always go to 0 if this sequence right, if this sequence actually strictly less than infinity.

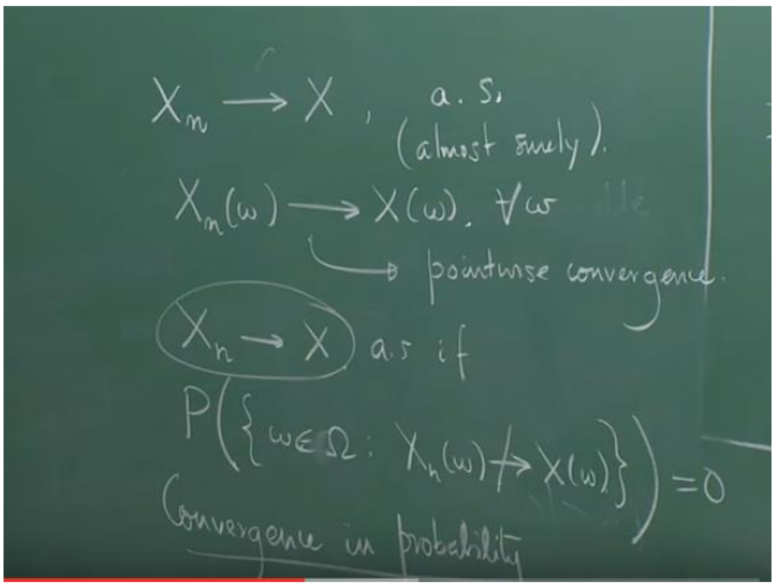
If this sequence is convergent then the tail always goes to so this is a notion from infinite series which I expect the viewers of this is supposed to know because that is the sort of requirement that we had that you have to have some understanding on mathematical analysis. Now if I look at this one., so what is happening as n tends to infinity this will go to 0. Now this is actually a fixed probability.

This is a fixed event. So which means as I take the limit from this side and this side this remains unaffected by the limit operation because this is already a fixed set, everything is done, so which from here, so we will have this, so we can conclude now that (())19:08 probabilities are always greater than equal to 0 and this will immediately tell me.

Now let me show you an application of all these things that we have just spoken, an application of the Borel-Cantelli lemma and application of the Cheybyshev's theorem and that comes out when we are talking about convergence of random variables. If I am using a vector valued random variable or random vector then I if I need to use it I will make an explicit mention of that.

Otherwise, please understand that we are always working with random variables that is functions which are mapping sample points to real numbers. Now there is something, there are 2 types of convergence which are very important; one is convergence in probability and convergence almost everywhere.

(Refer Slide Time: 20:18)

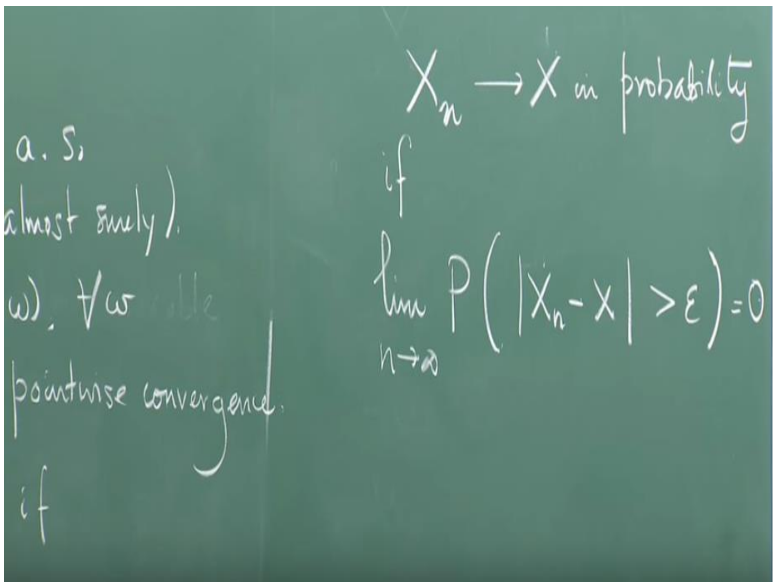


So, suppose I have a sequence of random variables X n and I say X n goes to another random variable X and we say it converges almost surely, almost everywhere convergence, this is called almost everywhere convergence, a.s. is a short form of the word almost surely. So when I am talking how, so these are random variables are functions, so a sequence of function converging to another function.

So when I would say when can I say that what is the meaning of a sequence of functions converging to another function. It means, technically it should mean that for whatever omega that I have in omega this is what it should mean if I am talking about pointwise convergence. So if this happens then this is what is called pointwise convergence, this is what is called pointwise convergence and then one needs to understand that in probability theory the tendency is always to throw out null events.

So we say X n tends to X almost surely if probability of the event it is the collection of all omegas such that this is not true. Probability that X n omega does not convergence to omega. The collection of all such omegas is a null event. There can be some omegas in the sample point for which this may not happen, but their collection would form a set whose probability would be 0. If that happens then we say X n converges to X almost surely and then there is a notion called convergence in probability. We will see how these are connected.

(Refer Slide Time: 22:36)



Convergence in probability, so what is the meaning of convergence in probability. Let us discuss this little fact. We will take 10 more minutes to finish everything so let me discuss this little fact, convergence in probability.

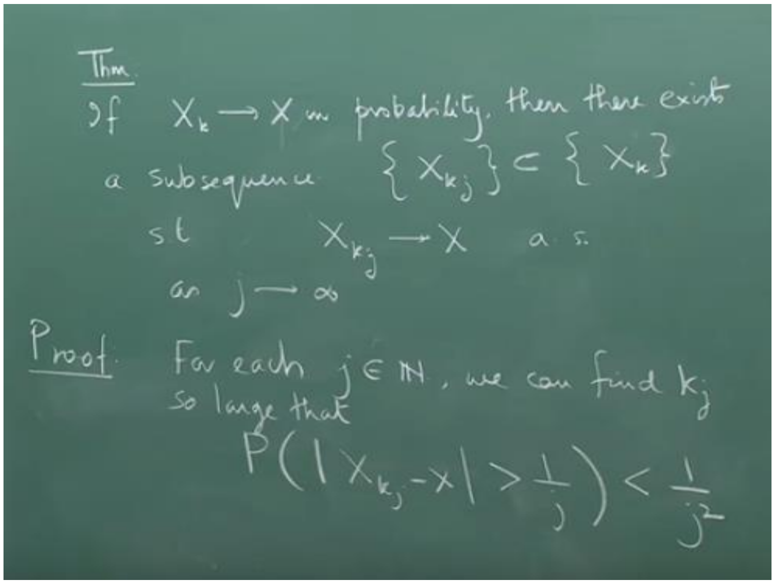
So convergence in probability means we say X n converges to X in probability if, so we are telling that a set of all omegas or any given epsilon take any epsilon, if the set of all omegas so this is the probability that the set of all omegas such that the distance between x and x is strictly greater than epsilon keep on increasing the value of n right, see suppose I say x on minus x strictly greater than omega.

I collect all those omega that has a probability so you have x 2 minus x strictly greater than epsilon there could be some omegas where for which that is true so that will have a positive probability but then as I increase the value of n this should go to 0; that this distance cannot be bounded away from 0. There cannot be positive number such that this distance would be always greater than the positive number.

So the fact that that probability in the limit, the probability that this distance is bigger than some epsilon would be 0. So you are expected to, x n is expected to be coming towards x, that is the whole idea, that sort of explanation one might think.

Of course if you are satisfying this you are satisfying that but there is an interesting that okay if I am converging probability am I converging also almost surely. See if I am converging almost surely, I leave it to the listeners or viewers to decide whether they are converging probability, the answer is yes. Now the question is if I am converging in probability am I, is the same sequence also converging almost surely.

The answer is there is at least a subsequence which is converging surely to the same random variable. So this is a theorem if you want to write it like this.
(Refer Slide Time: 25:55).



If X k converges to X in probability, then there exists a subsequence X k j of X, X k j which is the subsequence of X k such that this subsequence converges almost surely to X. Now, this definition of probability you must understand that this should be true for all epsilon greater than 0. Whatever we have chosen epsilon this will always happen. It does not matter for some epsilon this will not happen.

This has to happen for all epsilon. So that is something you have to keep in mind when you are talking about convergence in probability. So what does it mean? So how do I prove this? So ideas that you take any integer j does not matter whatever integer you are taking, take an integer j.

We can always select k j or a index value so large that this difference, this difference the fact that the probability that the difference between X k j and X is bigger than some 1 by j square or 1 by j something like this can be made smaller than some quantity.

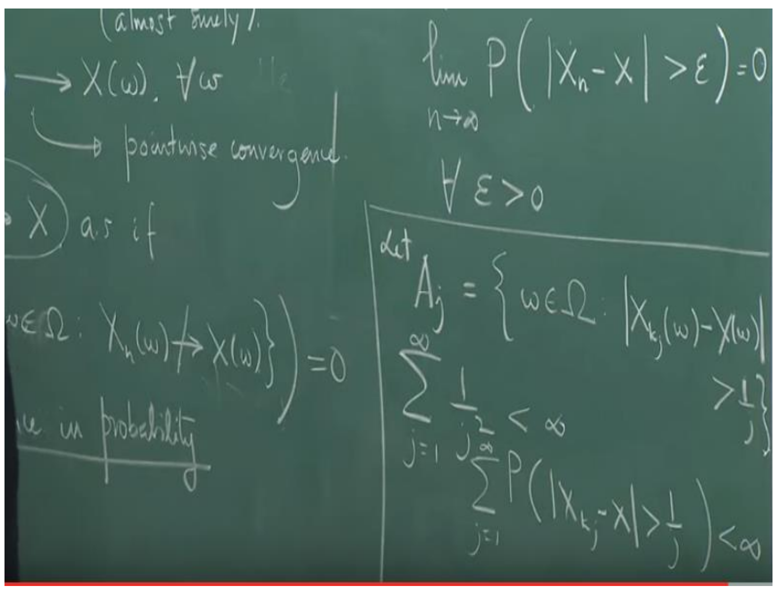
So the ideas that for each, why can we make it as smaller as I like, because I have this fact, because I know that it converges in probability that this what does this mean. This simply means that when n is large this probability can be made as small as I like, can be made smaller than some delta, right. So, given for each j in N we can find k j so large because this is where we are exactly applying the fact that we are talking about, the fact that x tends to x in probability that definition we are now applying.

So for each j we can find a k j so large that probability of X k j minus X this is bigger than 1 by j. This is my some sort of an epsilon. This is less than some delta, that delta may be chosen as 1 by j square. So I can find a k j if this is my delta. I can always find a k j for which this is bigger is lesser than 1 by j square. This is exactly I am applying this definition because this holds for whatever epsilon I want.

So once I have done that you might note that I can always chose k1 less than k2 less than k j minus 1 because we are essentially taking a larger and larger k j. So this can always be done. These are very simple techniques, mathematical analysis which mathematicians use, so I am essentially giving a small warning, even to make money you need to know some good mathematics though many mathematicians do not have money but.

Now let me try to show you what does it mean. Now let me construct, so what you are constructing. You are always taking an increasing k j. So this k j is actually going to infinity. So you are increasing j, as you increase j you increase the k j, right. So this is the increasing function of k j increasing function of j.

This is exactly the definition of a subsequence because in subsequence when you chose the n k or k j you are actually constructing and increasing function increasing the function of the indexes because you have to maintain the order. That is the meaning of a subsequence.
(Refer Slide Time: 30:58).



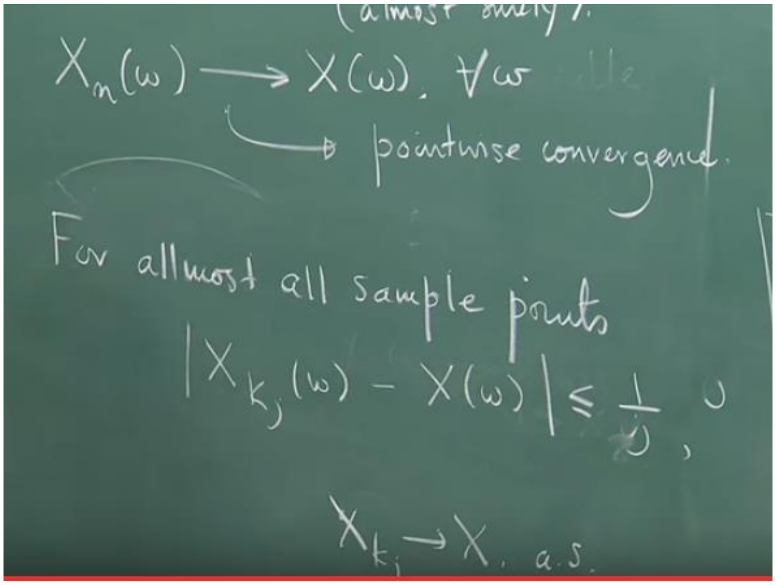
Now you construct, let us construct this sequences, this event. A set of all omega in omega such that X k j omega minus X omega is strictly bigger than 1/j. so I am looking at all such for every j j equal to 1, 2, 3, 4 I am just trying to figure out. So for given A j this should happen. So find, for every k, for take the k j and put that particular j as this A j index, index will be the same j. So here I have an infinite sequence of events.

Now I know that probability of all them is strictly less than 1 by j square so you can sum it up from j equal to 1 to infinity. The probability is strictly less than infinity because summation 1 by j square is a convergent series, right. So here I will get since summation 1 by j square is strictly less than infinity, it simply tells me that probability and this is a place where you apply the Borel-Cantelli lemma.

So what does the Borel-Cantelli lemma says that this thing this would imply that probability A j infinitely often is 0. Sample points omega do not occur infinitely often. So what does this mean? That is for almost all sample points except a few of them, for almost all sample points what do I have. So here is something we now have to look into.

So for almost all sample points, so for almost all sample points we have, so this is this probability is 0. So for almost all sample points this has to be true.

(Refer Slide Time: 34:25)



So omega cannot be occurring at infinite many A js. So for infinitely many omegas the opposite must occur. For infinitely many omegas this will not occur, only the opposite will occur.

So what does this mean where j is an integer greater than some particular j some index depending on the omega it does not matter. So what does this mean? This simply means that X k j goes to X in probability. This simply means that X k j omega minus X omega can be made as small as I like as j is increased. What does it mean but this is not true for every omega but almost all sample points omega. This simply means that X k j goes to X almost surely.

So with these ideas I would like to stop my discussion here and tomorrow we will see that how these ideas can be used for example to talk about central limit theorems and law of large numbers which are actually very useful while you study Brownian motion and Brownian motion is central to understanding the flow of stock markets and also the Ito calculus depends on Brownian motion.

The Ito calculus is the calculus of computing very (()) 35:55 pricing when the pricing varies. So let us stop here and let us see tomorrow what we can discuss about central limit theorems. Thank you very much.