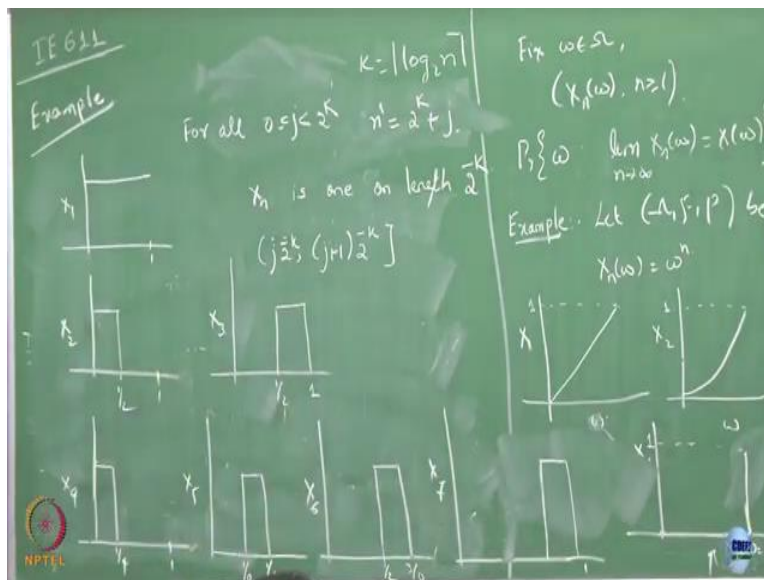


Introduction to Stochastic Processes
Professor Manjesh Hanawal
Industrial Engineering and Operations Research
Indian Institute of Technology, Bombay
Lecture No. 22

Convergence of Sequence of Random Variables (Part 2)

So, I am, now I am going to look at another example which will lead us to the second definition of convergence, second notion of convergence called convergence in probability. So, what is the example?

(Refer Slide Time: 00:32)



So, let us again try to look at a sequence of random variables I have, I am going to draw one random variable here x_1 , again these random variables are defined on my unit interval probability space, this x_1 is going to be 1 throughout and then I am going to define two random variables x_2 which is going to be, I want to split this half which is going to be like this. I am going to.

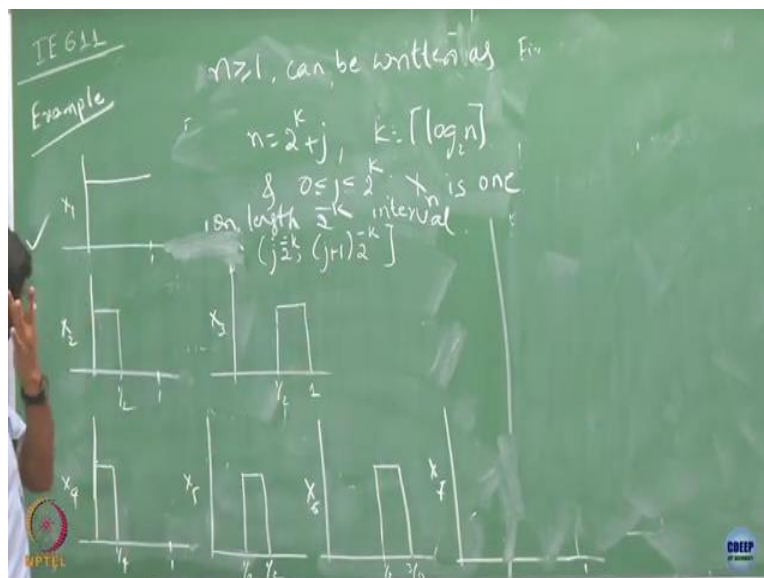
So, let us call this random variable 1, let us call this 2, let us call this 3, let us call this 4, 5, 6 and 7. So, this random variable x_1 is going to take 1 throughout, so x_2 is going to take 1 in the first half, x_3 is going to take 1 in the second half, now x_4 is such that it is going to take value 1 in the first quadrant, x_5 is going to take value 1 in the second quadrant, x_6 is going to take value 1 in third quadrant and x_7 is going to take value 1 in the last quadrant.

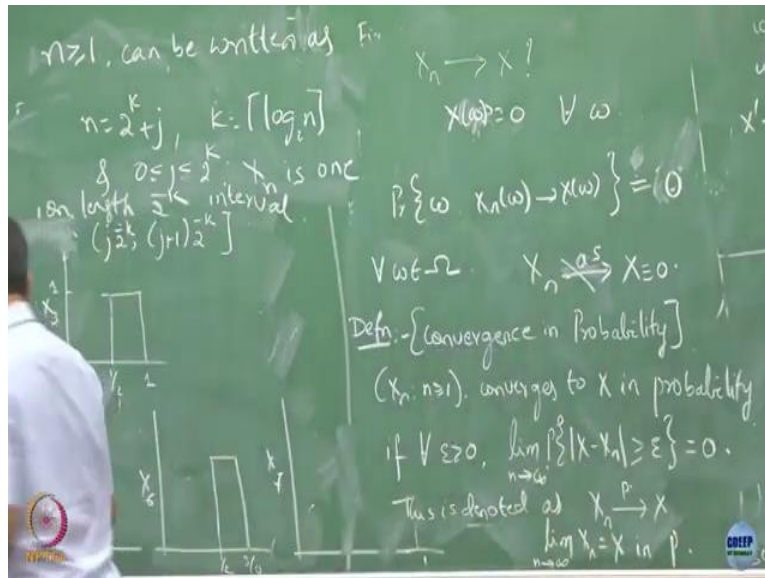
Now, I can keep on grading sequence like this, right, okay so you see that like I have doubled up from this is like a going in a geometric sequence and I can generalize this and write it as these equations as n prime equals to 2^k plus j , okay. So, let us take from n and then let us, let k to be the logarithm of that and then if you let j go to 0 to 2^k for all the indices which are between 0 to 2^k , 0 included the n prime index will have, it is going to take length 2^k on this interval.

So, just like this is generalization but you can, just let, just try to understand for the case n equals to 1. So, n equals to 1 what happens, this is going to be 0, right and in that case what is the possible j 's I have? It is only 0, right? Then I have 2^0 , 1 and j is 0. So, I have n prime equals to 1. So, that corresponds to this case and what is this value? Is going to be of length, k inverse what? k was 0, right? So, this is 1 and it is in the interval 0 to 1.

And now if you go and take further for these two things I mean you can continue to do this and you will see that in, in this first half when k equals to 1 this is going to be 0 and half, you are going to get the value of 1 and when you go the next index, you are going to get this is going to take the value 1 in the next interval. So, you can like this you can generalize this description of your x_n 's and then you have a sequence of x_n . And now let us try to understand where this, if I have a sequence of random variables like this, where do they converge?

(Refer Slide Time: 06:16)





So, let us say n equals to 1 then I can write it as, define like this. So, take any n that can be expressed as in this form for some k and some j , right. So, and for that it is going to be of length 2 to the power minus k over this interval. For example, if you are going to take n equals to 3 , what is the k and j corresponding to this?

K is going to be 1 and j is going to be, (1) so then 3 can be expressed in this form and for this for is the value of this? k is equal to be 1 , this is going to be in the half and what is that interval?

Student: 1 by 2 to 1 .

Professor: So this is going to be, this j is going to be 1 , this is going to be 1 by 2 to 1 . So, this is covered the entire thing. And now for this kind of x_n I have, where does it converge? For first of all any case what should be the limiting random variable here? So, as you go down-down like this what you expect? You expect this random variables to put mass on a small-small intervals. It is going to take value 1 every time but it is going to take that value 1 on a smaller-smaller intervals.

So, what do you expect as n goes to infinity, what? It is shifting, it is like putting mass on a narrower intervals and it is not putting, so here it has put mass on the first quadrant but rest of the quadrant it has put 0 mass, here it has put mass on the second quadrant and on the other part it is going to be 0 . So, like that it is always continue to happen. So, due to which okay let us assume you know that as n is going to tend to infinity right, this interval is shrinking.

It is only going to put a mass on a very-very small but right now we are unable to imagine what is that small interval where it is going to put mass. But everywhere other, other path may be it is going to put a 0 mass, let us say whether x_n converges to some x where x is 0 for all, x of ω is 0 for all ω . Okay, now let us try to understand, so for this what we need to do? We have to understand whether probability that ω x_n of ω converges to x of ω is going to be 1.

Let us see, so take any value of ω , so if I want to include this ω in this set that x_n of ω should converge to 0, because that is what I have taken x to be 0. So, now let us say that ω is here, let say for time being let us take ω to be 0.3, so here ω 0.3 got value 1 when it is x_1 is giving value 1 to this and what is x_2 is giving it? 1, but what is x_3 giving it? And here what is x_4 is giving? And what is this x_5 giving? 1, so if you are going to construct such a binary thing.

So, here it 1 element, 2 element, 4 element then I have 8, 16 like that. In each of this rows that ω will have 1 at one of the graphs. In all the other graphs it is going to take 0 value, right. So, however, deep you go you will see that after some number, is it possible that x of ω is going to always get 0? Or if you go further it is going to take 1 again. So, for example, if you start it took 1, it took 1 here, it took 0 here, 0 here, it took 1 here and if you come back and when we went in the this row, it took 1 again.

If you go to the next row below it again took some time 1, right, so however down you go you will end up with some n where there it is going to take value 1, right. And if that is happening what is whether x of ω , x of 0.3 is going to converge in this case at all. No, right that violates our definition that however far I go I will find an n such that after that my x of ω is going to be 1, so that is why it cannot converge to 0.

And this is true for any ω , right, do you understand this point? So, there is no ω here such that x_n of ω converges to 0, so then what is this set if my x is 0? This is a null set, right and if this is null set then this probability is actually 0, not 1. So, because of that my x_n is in this case x_n is not converging to x almost surely and where x is 0, identically 0. Okay, but as you see from here as my x_n goes most of the time I am getting the value 0 for each of the

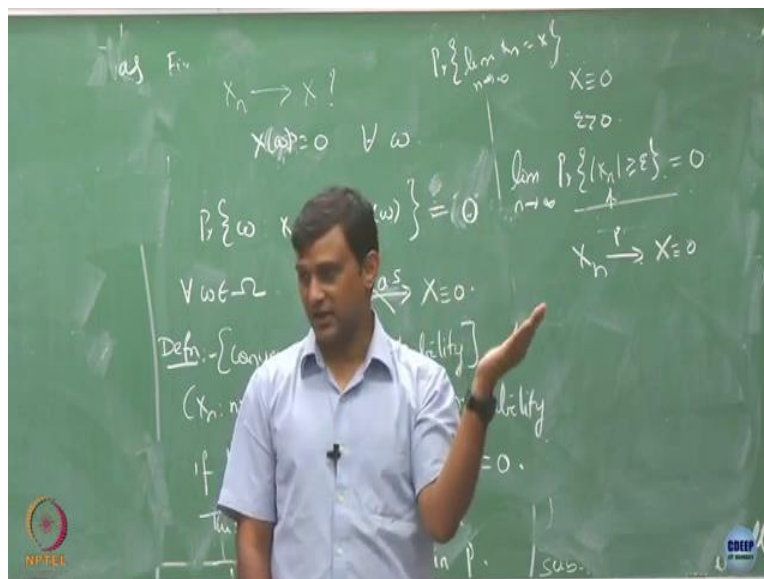
omegas but it only happens that but after some time 1 happens but rest of the time it is going to remain 0.

So, it looks like most of the time my random variable is going to take value 0 for omega, I mean as I have proceeding n but it is for some n further I go I gets violated that is why it is not true. But still you want to, like you will see that most of the time for different values of n it is going to take value 0, then it looks like may be like I could think of this sequence of random variables convergence to a random variable x which is identically 0 but this definition of almost sure convergence is not capturing that notion.

May be I need to capture, I need to define something weaker than that, so for that we have this another notion of convergence called convergence in, so what it says? Okay, so now we are going to say that a sequence of random variables converges to x in probability if for any epsilon greater than 0 now if you look at this sequence of probabilities then what is that this limit is now outside.

If you are going to look at the probability of this event being larger than epsilon if this sequence converges to 0 then we are going to call this as convergence in probability. And we are going to denote this x_n converges to x put a p here or limit as n tends to infinity x_n is equal to x in p. So, now let us see whether this example we had here satisfies this definition.

(Refer Slide Time: 15:06)



So, let us again take the same thing this example here let us take x to be 0, so to do this, to verify this I need to verify it for any ϵ greater than 0, right. So, let us take a $(0)(15:20)$ and ϵ greater than 0. Now, look at this, so x is anyway 0 for all the points. Now, I am asking what is the probability that $\text{mod of } x_n$ is going to be greater than or equals to ϵ .

So, $\text{mod of } x_n$ I can just take it as x_n because x_n 's are positive in my case, x_n being greater than or ϵ equals to 0. If I look at the limit of this where this limit is going for this example? So, recall that this x_i 's here I have defined it on unit probability space, right. So, the probability that it takes a value in some interval is nothing but the length of their interval, okay.

So, now what is the probability that x_n is going to take a value greater than or equals to ϵ . If you take n , any n , what will be the probability of just $\text{mod } x_n$ itself? What is the probability of $\text{mod } x_n$ itself? It is just $(0)(16:48)$ to the length of that interval where it is taking value 1, right and as n increases these values are shrinking, right. So, the probability where it is going put a mass, positive mass is also shrinking very fast.

And this is ϵ here is strictly positive, right. So, if it is strictly positive you should be able to find n sufficiently large such that after that you are going to put a value in a smaller-smaller intervals whose probability is going to come down below ϵ . So, then this event will never hold and this probability will be 0, so because of that is it true that this is going to be 0 for any ϵ greater than 0. So, then does it satisfy my definition of converges in probability?

That is right, so x_n converges to x here in probability, so what is the difference between them, convergence in almost sure sense and the convergence in probability? So, in a way look at this when we are looking at this sequence at any n you are looking at the joint distribution of your sequence, sorry your limit and a given x , it is looking at these distributions at any point.

But whereas if you looked into the, recall the definition of almost sure, there you looked at probability, so here you are computing a probability on the entire sequence, so entire joint distribution of the entire sequence, is affecting the definition of almost sure. Whereas, the definition of the probability convergence in probability is only looking at a pairwise distribution, the joint distribution of this pair x_n and x at any time but whereas this is looking at the joint distribution of the entire sequence.

In a way this almost sure, so that is why this almost sure convergence is a much-much demanding the requirement then convergence in probability. So, later we will see that, that indeed the case almost sure convergence is as much stronger property than convergence in probability because almost sure convergence implies convergence in probability but not the way around. Okay, so now, okay so before I do that if you look at this definition what I asked is, the probability that x_n going to take a value larger than epsilon to be converging to 0.

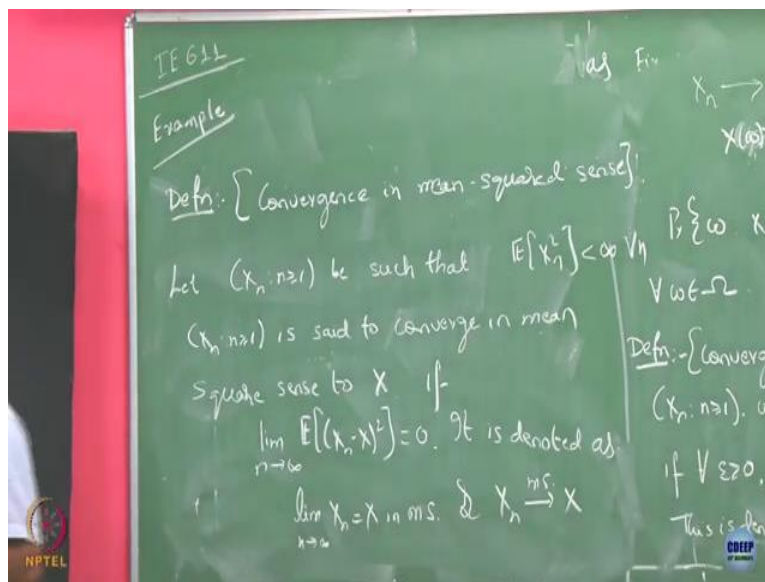
But it may happen that so here the value that it is going to take greater than epsilon is on a very small mass but it may happen that this value it takes on this small mass itself could be very-very large. For example, in all the examples I just erased now, instead of letting my x_i only taking up to value 1, I could let it take very large value, instead of 1 can make it 100, 200 whatever. So, in those case, yes it is happening that they are taking this value on a smaller intervals, but still the value they are taking and the smaller interval could be very-very large.

Now, when you are interested in finding the expectation of a random variable, you are just not interested in the probability. You will be also interested in the value taken with that probability. So, probability will linear product of the value of the random variable and the probability term. So, because of that it may happen that some sequence of random variables satisfies that but on the region where they are taking positive values small probability, that positive value could be very-very large.

So, in some applications this may not be desirable for you, you want it to be take small values even in that small intervals. So, this is like some cases like okay fine, it may happen that when I am going to put money in game, the probability that I am going to lose could be very small or like let us say lottery, the probability that I win is going to be very small, but if I win the amount is huge. So, the product is going to be very large.

So, when you think about this in the risk sense, something failuring is very small but if it fails it is going to be too expensive for me. So, in that case you are interested in both the value taken, it may be taking some huge value with small probability but that is of a concern for me. So, when that is the case instead of this, you may be interested in knowing whether it converges in some expected sense, okay for that notion.

(Refer Slide Time: 23:08)

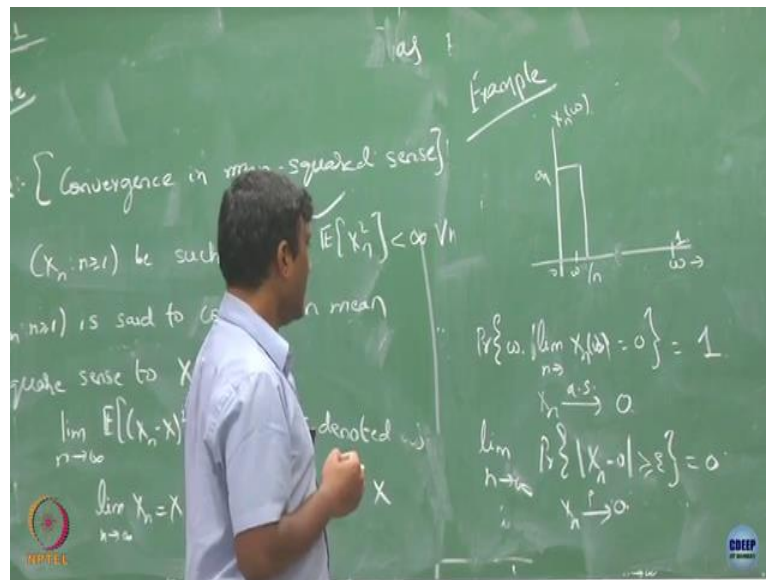


So, to capture those things we are going to define another notion called Convergence in Mean Square Sense. So, now instead of looking at the value of the random variable itself what I am now looking at is the expected value and not just the expected value but the mean square error from the random variable x . So, we are going to say that this sequence of random variables going to converge to x in mean square sense if there, if I look into difference, square difference and take their expectation that goes to 0 as n tends to infinity.

Okay, so notice that how could this condition that each of this random variable should be such that their second moment is finite okay. So, why is this? Now, the natural question comes in okay fine we are looking at different notions of convergence is it true that one already implies the other.

So, we already said that convergence in probability is a weaker notion of convergence in almost sure sense. And now we are talking about convergence in mean squared sense, is it already some weaker notion of earlier one or it is a total independent convergence we are talking about.

(Refer Slide Time: 24:52)



So, we look into that, so before that let us look into an example, okay let us take this random variable, again this is defined on a unit interval probability space. So, now each of this random variable I have defined such that if you fix an n then it is going to vary like this in between 0 to 1 by n it is going to take a constant value of $1/n$ and then it is going to take a value of 0 in the rest of the interval, okay. So, what do you think about it? Let us try to understand whether it convergence in almost sure sense. So, what I need to, so let say to understand almost sure sense I need to compute this probability.

So, first, what is the guess? What is your guess? What could be the limiting? So, if, so as n is tending to infinity right, it is going to shrink this interval and it is a kind of putting value, non-zero value on only a some small interval which is shrinking rapidly as n increases. So, most of the values are ω are only going to take values 0, right? Because this $1/n$ as n increases it kept on shifting to right, to my left and then most of the time.

So, you expect guesses like this will converge to 0 random variable, delta does not matter to me right like one point does not matter, what matters is the entire interval. If it is 0 everywhere I am going to just call it a 0 random variable. So, is it convergence in almost sure sense? Right, if I am going to take limit of as n tends to ∞ of $x_n(\omega)$ whichever ω I take, so let us say initially ω is here, it has some positive value $1/n$ but as n increases this guy comes to the left of ω and then it is going to get a 0 value, right?

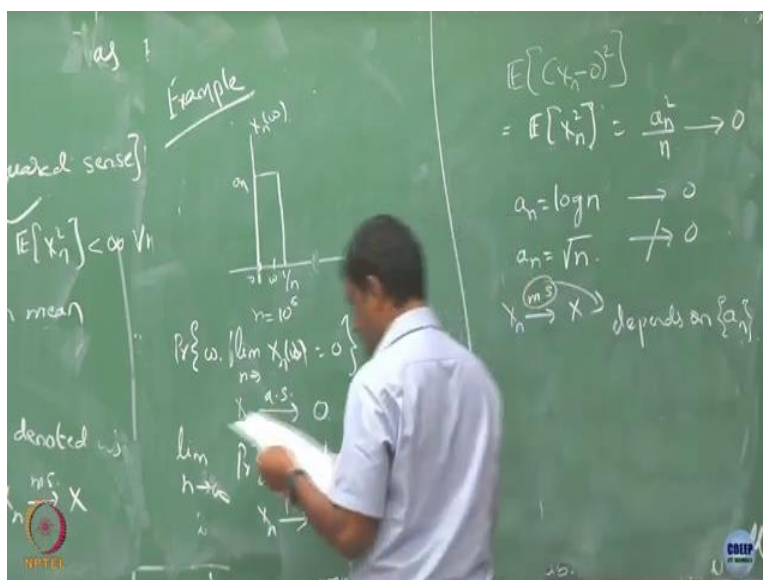
So, because of that every ω is going to be taking 0 value, so this is going to be 1, so x_n converges to 0 almost surely. Now, what about its convergence in probability sense? So, what I need to verify that, I need to verify whether limit as n tends to infinity probability that x_n minus 0 going to be epsilon is going to be 0. Is this true? So, again, so I have already substituted 0 for x here.

Now, if you look into the probability of this interval x_n that is nothing but this interval right and that interval is shrinking. So, after some time it cannot exceed epsilon, so after that point it is always going to be 0. So, it is also true, convergence in probability.

Student: (0)(29:19).

Professor: No, this is greater than epsilon, what you are talking about is probability of x_n mod x , in this case simply x_n is going to be simple what, 1 by n . So, what we are just saying that yes, it could take value greater than epsilon but on what regions? That is a region which is shrinking, right, small-small shrinking region that probability is shrinking so that equals to 0 as n goes to infinity. So, now let us compute convergence in mean squared sense, okay. So, what is this value then?

(Refer Slide Time: 30:18)



So, to compute this I need to compute expected value of x_n minus, so x I have taken to be 0, so can I compute this expectation? So, this is nothing but expectation of x_n square right, what is

this? So, x_n is taking what value? It is going to take value a_n and with what probability? $1/n$ and with other probability it is going to take value 0. So, this is going to be a_n^2 by n .

Now, I want this probability to go to 0, if I have to claim that x_n converges to 0 in mean squared sense, is this true? So, notice that n is a number now and I am dividing a_n^2 by n , will this go to 0 for any n ? For n tending to infinity a_n^2/n always goes to 0.

Student: () (31:38).

Professor: N, yes a_n is a function of n , so okay. So fine, let us say if I can choose like see I can define this a_n to be $\log n$. So, in this case it goes to 0, if I am going to, say a_n equals to square root n , it goes to 0?

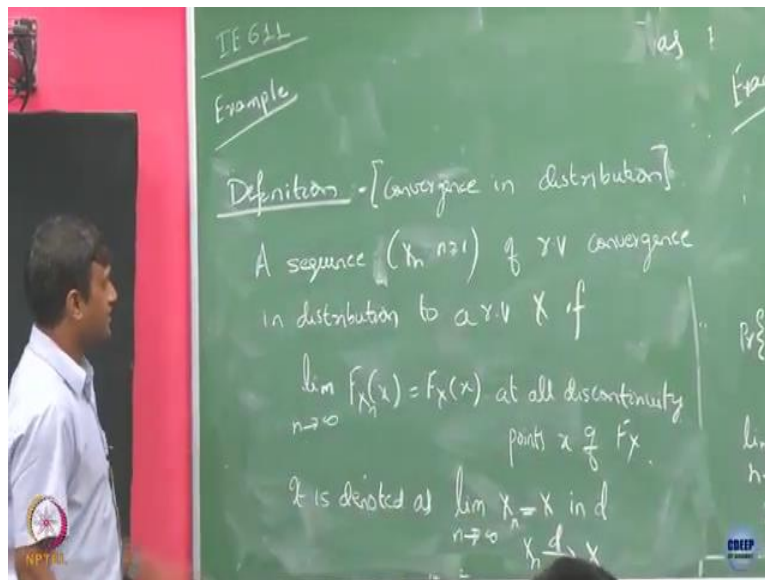
Student: () (32:11).

Professor: n^2 , so square root n is become ratio 1, right. So, this does not go to 0 in this case. So, in this case does go to 0, but in this case it does not go to 0. So, what is happening in this case? So, it looks like in this case the convergence in the mean squared sense here at least depends on how this amplitude is or how this value is. Like if this value of a_n is going only logarithmically in n and this guy is shrinking much faster than this guy, then this guy is going to go to 0, but if its amplitude the value taken is going to be much-much larger that means it is then this interval or sorry when this the whatever this value of n then this product could be still much-much larger, right.

So, for example, you could almost come very close to here but let us say n equals to some 10^6 to the power 6, you are like almost $1/n$ is like some 10^6 to the power minus 6, so you are putting positive value only in the small interval but you are going to take a value of also you are going to take the value 10^6 to the power 6.

Let us say like your chances of winning a lottery is 1 in million but if you win you are also going to get a million dollars or million rupees. So, in that case that value going to be still high, so here as you see whether it converges in mean squared sense or not depends on the sequence a_n , so depends on.

(Refer Slide Time: 34:35)



So, so far we looked at the sequence of random variable converges to some x , now we will look at the distribution itself. And if it is possible that like if I have a sequence of random variables which certain distribution they will converge to random variable which has some limiting distribution, okay. So, we are going to say that a sequence of random variable convergence and distribution to a random variable x if you are going to look at the CDFs.

So, f of x n denotes the CDF of random variable x n if you look at a point x and if it converges to a CDF f of x at that point and this point x are such that they are the continuity point of the CDF of the limiting distribution. Okay, in that case we are going to call it as converges in distribution, sorry continuity point of x . Okay fine then, so let us stop here.