Introduction to Stochastic Processes

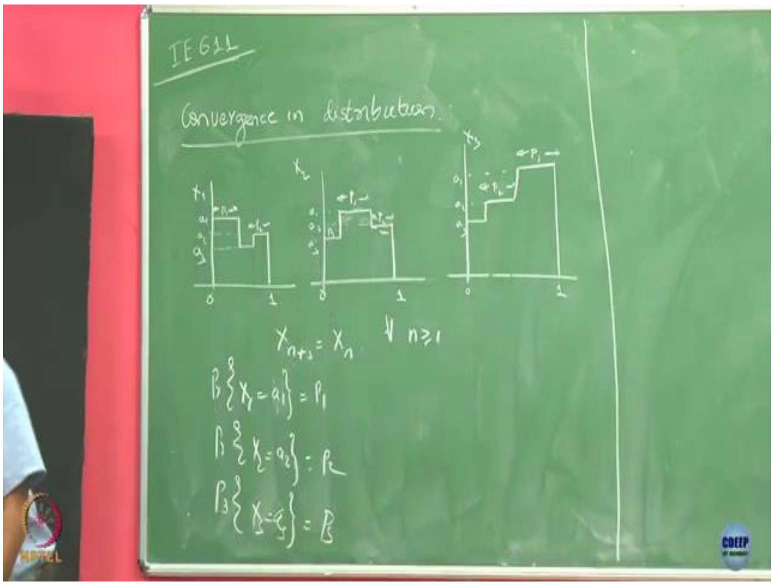
Professor Manjesh Hanawal
Department of mathematics
Indian Institute of Technology Bombay

Lecture 23

Relation between different Notions of Convergence

We started discussing about the convergence of sequence of random variables in the last class. So we defined different notions of convergence. We talked about what convergence in almost sure sense, we can talk about convergence in probability, we talked about convergence in mean squared sense, and at the end we defined convergence in distributions. So, Let us take an example here. Let us say I have 3 random variables in this form.

(Refer Slide Time: 0:53)



Let us say it takes some numbers here and then it has something, so this is all random variable x 1 and Let us say I have another 1. So let me call this, values here. Let us say this is a 1 this is a 2 and this is a 3. And similarly let me just drawn another one here it is something like this. So again, all these things are same, this is a 1 this is a 2, this a 3 and my scaling is not correct.

So let us recall the definition of unit interval probability we have defined earlier and let us say we have 3 random variables defined on this unit interval probability space like this, so this is x 1, this is x 2 and this is x 3. So, notice that the way I have drawn is here, all these random variables are going to take only these 3 possible values, a 1, a 2, a 3, and these widths are same, I mean and, my scale may not be correct, but assume that so this whatever this width is, let us say this is like p 1 whatever this is saying P 1 here and this is p 2 and let me, let us say this is p 2 width here.

And let us call this circle third part as p 3. So all of them have same taking 3 values, but it is just like the intervals on which they are taking these values is going to be different. So, here it is taking value p 2 on this interval and here it is taking the same value p 2 on a different interval here, so what I mean here is that is fine. Now, let us say that x 4 is now again this x 5 is this x 7 is this X 8 is this x 9 is this and x 4 is x 9, 10, 11 is this like it is now, my random variables are periodic questions of this. So, what we mean is x and plus 3 is simply going to be x n, for all n,

Now if you have a sequence for random variables like this they expect it to converge in almost sure sense or in probability.

Student: No.

Professor: No, right because this is periodic and it is fluctuating so much but if you look at the distributions. So, what is the distribution of x 1, x 1 is actually discrete value random variable here actually it is only taking 3 values, a 1, a 2 and a 3 and it is going to take that with probably to what p 1, p 2 and P 3.

And it is also again, the same is the distribution against same here, this x 2 is also taking 3 value 7 a 2 and a 3 with what probabilities? Again p 1, p 2, p 3. So you should look into these distributions, these random variables are identical, they are the same, it is just like that putting that mask on a different intervals otherwise, they are the same.

So, this is where we want to understand convergence in the distribution. So we have to go beyond convergence in almost sure convergence of probability and convergence means square. And we are interested in convergence in distributions. Now what is the limiting distribution here? So if you have a sequence distribution what will be the limiting distribution, so the limiting distribution is something which is probability that x 1 is equals to a 1 is p 1. The probability that x 2 is equals to a 2 equals to p 2 and probably x 3 equals to p 3. This is going to be any random variable, which takes this values a 1 and a 2 ,a 3 with probability p 1, p 2, p 3, this is going to limit distribution here.

(Refer Slide Time: 6:32)

Now Let us look at another sequence of random variables. I am going to define a sequence of random variables here. What is u here u is uniform random variable. I am going to look at a sequence of scaled uniform randomly variable here like this, u I am going to fix scale by minus 1 to the power by to the power n by n. Now let us try to understand how the distribution of this looked like.

So let us take n to be odd. If n is odd, this random variable is going to be positive random variable or negative random variable.

Student: (())(7:24).

Professor: It is going to be negative value and then let us look at how it is CDF look like. So here value u takes value between what 0 and 1 right. Its smallest value is going to be 1u equals to 1, in that case, it is going to be minus 1 to the power minus 1 divided by n because n is odd here, okay and it is largest value is going to be 0.

So if you look into its CDF it is going to start from minus 1 by n here and how does it go? It goes all the way up to here linearly and hits 1 here right and then it is 1 here. Now if you look into the case when n is even how does the CDF look like. So in this case, the smallest value is going to be what it is going to be positive random variable when n is even, what is going to be a smallest value, 0, what is going to be the largest value, 1 by n and it is going to be and then it saturates here right x 1.

So, I have this CDF which is going to look like this depending on whether n is odd or even and as n tends to infinity, what is going to happen to this? It is going to be just a function and what is where is the jump happening at origin at x equals to 0. Okay, now let us understand. So we expect that in this case, the limiting distribution to be what? Which is the 1 which is going to take almost value 0 with probability 1. so that is why limiting distribution.

So my limiting distribution is just this. Now, let us try to like understand how my CDF converge at different points of x. So my CDF are function of x here, let us see for different values of x, how they converge. Suppose if I take a value of x here and the negative side of my real line, that is x is less than 0. If I let n go to infinity, what this value will converge to, it is going to converge to 0, right? Because this guy will keep shifting to the right side at some point, whatever x you have here, it is going to go beyond that and it is always 0. And this guy is anyway 0 for x less than 0.

Now if you are going to look at f of x n of x for x greater than 0, what is this going to happen? It is going to be 1, so this guy sequence converges to 1, that is fine. And that is matching with this when I going to take x less than 1. This is going to be 0, and when I am going to take x strictly positive, that is going to matching with 1. Now let us look at the case where x equals to 0.

So when x equals 0 so what is the value of this function at x equals 0? It is going to be 1 because we have right continuity property and what is the value of this at x equals 0?

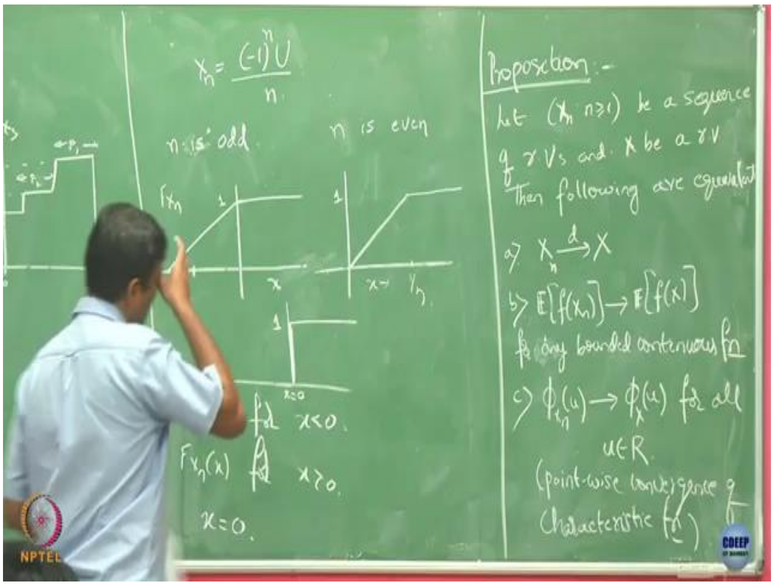
Student: 0.

Professor: 0 Right? So I have the sequence which is alternating between 1 and 0 with such a sequence converge, No. So it is not converging at when you are going to look at x equals 0, so it is converging for x less than 0, converging x greater than 0, but for x equals to 0, it is not converging.

So, as you see this, this limiting distribution has a jump or a discontinuity at the point x equals 2 0. But on all the other points, it is continuous. But now, you see that at the point of discontinuity, this convergence is not happening. So, in general that is why, when we said that in the definition of convergence and distribution, we said that a sequence of random variable convergence in distribution if the CDF convergence to the limiting CDF at all points of continuity.

So, that was our definition of convergence in distribution. So why we ignore that because, Let us say if you look at the sequence of this CDF they are converging at every point except for the point of discontinuity here, but still like we it is valid that this distribution we can assume, I mean we can interpret it that is converge to this point, but we have to just ignore this point of discontinuity here. So, just like to account for this in our definition of the convergence and distribution, which we stated last time, we have explicitly stated that, we look for convergence of CDF only at the point of this continuity of the limiting distributions.

(Refer Slide Time: 13:48)

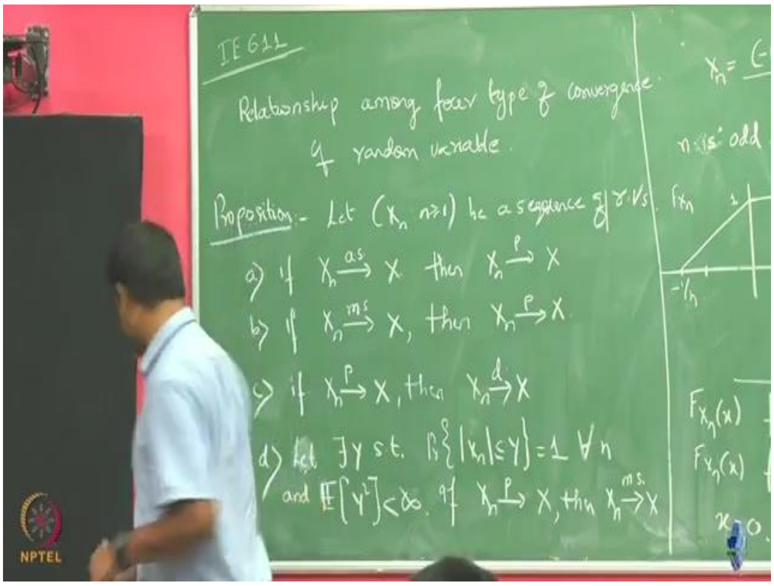


We have already discussed that distributions are somehow associated uniquely with their characteristics function. Because character function clear defines distributions and vice-versa. So, then based on this we can directly state the following result, which I am just going to state without proof. So, if we have a sequence of random variables and then another random variable x, then we are going to say all these 3 statements are equivalent either you say that, xn converges in x in distribution. Recall this notation, this is what we mean by converges in distribution.

This is equivalent to say that if you are going to take the expectation of your random variable, but not directly the random variable, expectation of some function of this random variable where this function is continuous and bounded, if this sequence of expectation converges to the expectation of function of that limiting random variable, then this is also implies that they converge in distribution and alternative characterization of same is. You look at the convergence of the characteristic function.

So phi of x n is the characteristic function of my random variable x. So, if you compute at some point u then it should convergence to the characteristic function of x, again computed at the same point u and this should happen for all u. If this happens, then we can again say that my sequence of random variable converges in distribution. So, I am just going to skip this proof, but we will take this result for granted. Please look into the proof in the book. Fine, so, we have now characterized all the 4 kind of convergence notions we studied. Now the question is what is the relation between them?

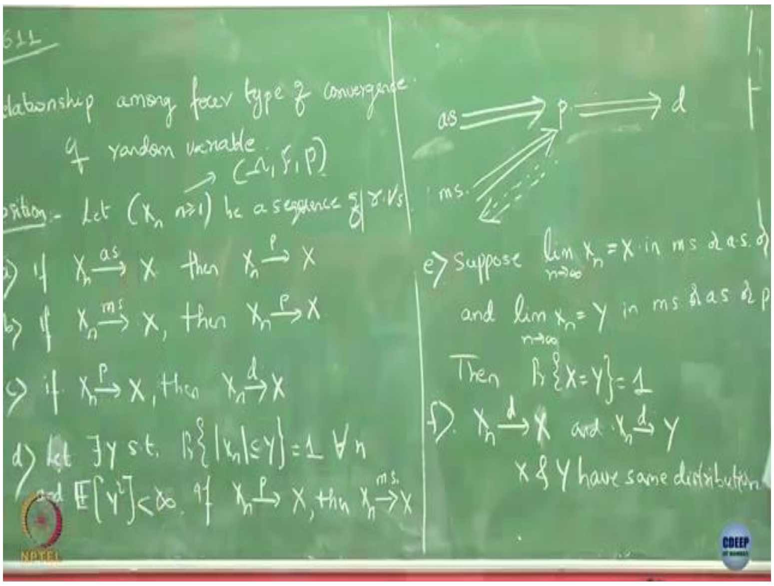
(Refer Slide Time: 16:08)



So I am going to treat it as a proposition. So, let xn be the sequence of random variable be a. Now suppose, if I know that my sequence xn converges to x in almost sure sense then this implies that xn converges in probability, the same random variable x. So this we already said and now if Xn converges to x in the means squared sense to some random variable x then xn converges in p again to the same random variable x.

If xn converges to x in probability then xn converges to xn distribution and d, such that for all and…

(Refer Slide Time: 19:08)



So what this result says is suppose if I have a sequence of random variable that convergence in mean squared in almost sure sense then that implies that it convergence in probability also, second point says that if our sequence of random variable that convergence means squared sense then that also means that it convergence in probability.

And the third point says that if our sequence of random variable that convergence in probability that means that it also converges in distribution further. Now the question is we had shown implication in 1 direction, What about the other directions? Is it true that P implies as and P implies means squared sense?

So, this part d answers that question partly it says that convergence in P implies convergence in means squared sense that is this direction provided something happens under some condition not always true. It says that, if there exist a random variable y such that all my 6 of random variables are dominated by that random variable y with probability 1. And this random variable is such that further it has finite second moment.

If I can find a, such a random variable y then this is true that my convergence in probability also implies convergence in means squared sense. So, this implication that converges and P implies convergence and almost sure that is not true in most of the cases, even if you look at the example we studied, you already had an example what we discussed where it converges in probability, but not necessarily in the means squared sense. So, also it is not so, easy to come up with conditions like when convergence in P implies convergence in means squared sense.

Student: If more (())(21:38) less than equal to y it is true for all Omega.

Professor: That is the meaning of that.

Student: It should be true for all Omega?

Professor: What is the meaning of this?

Student: (())(21:49)

Professor: So, this means that if I take set of all this omega where xn of omega the absolute value of that is appointed y of omega that set of omega should have probability 1.

It may happen that this condition is violated on some omegas where that mass is 0. For example, when we have a continuous value it may happened that 1 point, this condition may not hold but that 1 point may have 0 mass, I do not care about such points.

Student: Then what about the relation between (())(22:24)?

Professor: So that we are going to come through P, So if we have almost sure, first we will check whether is that implies P and once I know it implies P, then this condition will come to my rescue.

So it is not like I do not have a direct route here. I have to go here and if after going here further, if this condition holds then, I have a clear. I can then say something about mean squared sense. So this also, in general, we do not know I mean, we do not have a proper condition, like this when this is going to be true.

So, we only whatever we know that we have stated here under his conditions this holds, but if you are going to for some reason want to use that distribution convergence distribution implies convergence P you need to provide a proof for that, it may happen that for a specific example, you have in your hand convergence distribution may be also implying converges in distribution, but you need to establish that, but if you have a case where you have convergence means squared sense you can just say that okay using this proposition we state that already implies convergence P. Once we have in this hard lines this continuous lines we know and when you are using this dashed lines, you have to first establish this hypothesis that exists such a y and for the lines which we do not have you need to provide a proof before if you want to use it all.

There is 1 more important property that will just write down it says that suppose xn converges to x in x let say wait a minute xn converges to x in ms squared, or (())(24:36) square or in P and I have limited xn equal to y, again in mean squared sense or almost sure sense or P then it must be the case that probability that x equals to y equals to 1.

So, what it is saying is the limits are unique up to the probability unique up to the set on which we have on the set with probability 1. So, you remember in the first class, in the previous class, we had a simple example, where we defined xn of omega is equal to omega to the power n they recall that example. So, for that example, we said that the limiting x is 0 on all omega or we came up with another possibility with x is 0 for all omega strictly less than 1 and x at omega 1 is 1 right?

So, these 2 were both possible limits according to that definitions, but then they have their equivalent in the sense that they, they have they put they are the they put they take the same values on the space which has probability 1.

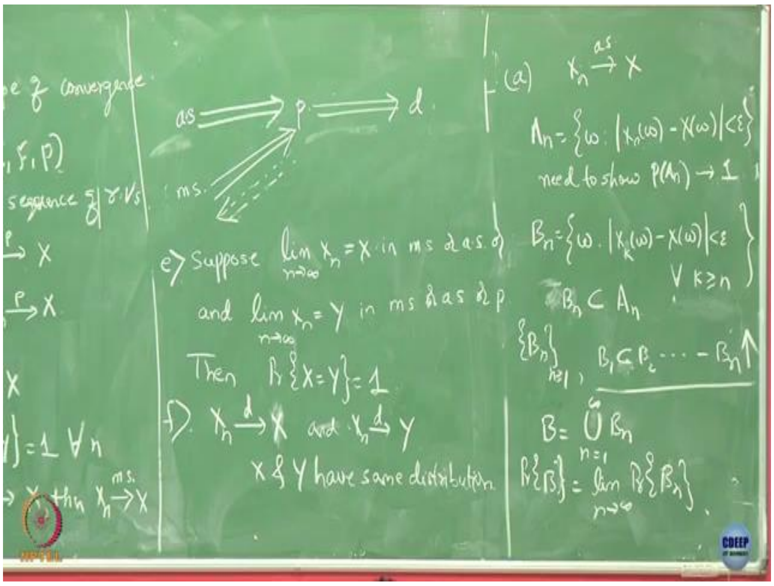
Student: For that example omega raise to n we are not defining a limit for every omega, right?

Professor: No because this is already said right we are interested in unit probability space unit interval probability space we have restricted our probabilities on our sample space to be unit interval.

Student: So even in that case then mod xn and then we will have to restrict our sample space according to the problem.

Professor: No, here I have not stated so those are specific examples. You can take the sequence of distributions, Let us say they are all under some probability space. On this probability space we are defining everything this, this xn are all defined on this. Fine, so the last point is let us say, I have and xn equals to y in distribution. So let us say I have I established that xn convergence to some random variable X in distribution and y xn converges another random variable in again distribution, then it must be the case that X and Y has same distribution.

(Refer Slide Time: 28:48)



Let us try to x, at least through the first a and b parts, which will also highlight some of the, or like, help us revisit some of the concepts we defined before. For example, continuity of probability and all. So, let us try to prove it. We have let us assume that xn goes to x, almost surely. And now we need to show that, that implies xn convergence in P.

So let us define a set An equals 2. I had defined it as An. Now, if I want to show my xn converges in probability what I need to show, I need to show that probability of this An as n goes to infinity is 1 that is the definition of convergence in probability. we need to show let us try to see if we can do that, so for that we need to construct some specific sets.

So, let us define Bn to be all omega xn of sorry k, x of omega to be less than omega for all k greater than or equals to n. See what I have done is I had interested in set An which for a given n wanted included all this omegas which satisfies this condition.

Now I have slightly refined it and defined a new set where I want this condition to hold not only for n but for all k greater than or equals to n. So, this was a particular n now, I wanted to hold not only for n also for n plus 2 and all the way up to infinity. So, because of this, is Bn is a set which is contained in An or other way around. So, An is contained in which is most in general where I am asking for more, this is Bn, So Bn for not only n I want is to happen for everybody.

So, it may happen for some omega this may not happens they may drop out from this. So because of that which 1 is correct? This 1 or other way around? Other way around, this one is correct. Now, let us look into this sequence of Bn. Now. I want to look at so, is this sequence of Bn they are monotonically increasing or decreasing? So, if I increase n, let us say I increase n from 10 to 15. Earlier I wanted everything beyond 10 to satisfy now, I only want everything beyond 15 to satisfy.

So, it should be increasing function right because as n increases, I only want smaller number of sense conditions to be satisfied. So, I will have B 1 contained in B 2 like this or Bn is an increasing sequence

Student: Sir in Bn you are saying that this condition should be satisfied for all k greater than or equal to n.

Professor: Right.

Student: And An should be satisfied for only n.

Professor: So this is like I am looking for all this omegas with satisfied this at only n.

Student: Yes.

Professor: Now here I am looking for all this omega not only satisfied at n but also every point after n.

Student: Then n, so how are we saying Bn is a subset of An.

Professor: So, let us say something satisfied n is it necessary that it satisfies at n plus 1 also, it may not satisfy at this point right. So that point may drop out of this. So because of that, a point which may be belongs to An that need not necessarily belong to Bn.

Student: (())(33:54) belong to Bn.

Professor: What you are saying here Bn is going to be for all this, what is the issue here?

Student: He is saying An belong to A.

Professor: We want this to be correct. Let us take a punch. Let us say some omega belongs to An. Can you guarantee me that that also belongs to Bn? just think about it. Let us say… are you convince or not Bn is monotonically increasing. I just said let us take n equals to 1. Now you want this to be satisfied k greater than equals to 1, like all the points this should be satisfied now let us take k n equals to 10 when you have 10 you want to be satisfied only after 10. So here compared to the first case in the second case you have lesser condition to check.

Student: (()) (35:24).

Professor: Yeah, I mean that is what like it could be the same. That is what our convention right we have said that unless it is a 6 subsets then I would have written like this. Now, when I have this, I know that what is my limiting B my limiting B is going to be union of B of n and what is my convention in that case, what is going to be probability of B is equals to limit as n tends to infinity probability of... So this way I just because I have a sequence, which is monotone, I could apply this continuity of probability and write it like this.

(Refer Slide Time: 36:25)

Now, Let us take this set. Now, let us look at all this points, for time being, which are going to satisfy this condition, now what is the doubt?

Student: I am still not clear about that Bn is a subset of A.

Professor: Convince yourself later, he is not going to convince you . So, let us say this I have set of this omega which satisfies this condition. Let us think of a particular omega in this case. Now, according to the limit of this limit definition here, we know that xn omega minus x of omega is going to be greater than epsilon for some for all n greater than n equals 2 n epsilon this is true, right? I just applied the definition of the limit.

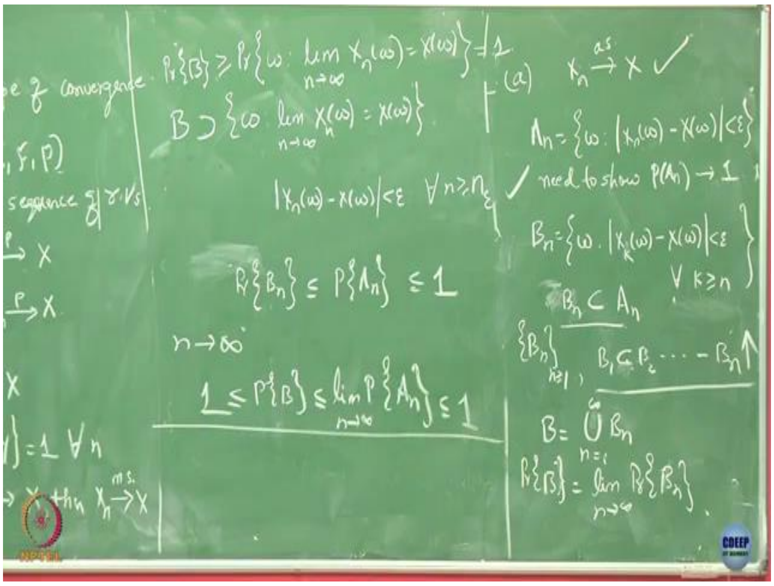
If for some omega this is true, I know that, that omega should also belong to this B. Is that true? Because I know that if this is the case, then for some B here for some Bn this is already satisfied, right and this is union so that omega should belong there. So because of that, I know that my B contains this set, so B is contains here.

Now let us apply probability on this Bn at An. If I going to take probability of Bn this is going to be…

Student: (()) (38:30).

Professor: The final 3 value (())(38:49) right probability has to be less than 1 and what I know… now let us try to invoke what is given to me. If I let this n I want to now show that this sequence takes the value of n. If I can somehow argue that my lower bond also goes to 1, then I as n goes to infinity, then I know that my probability as n goes to infinity is going to be 1 that is what I need to show.

(Refer Slide Time: 39:30)

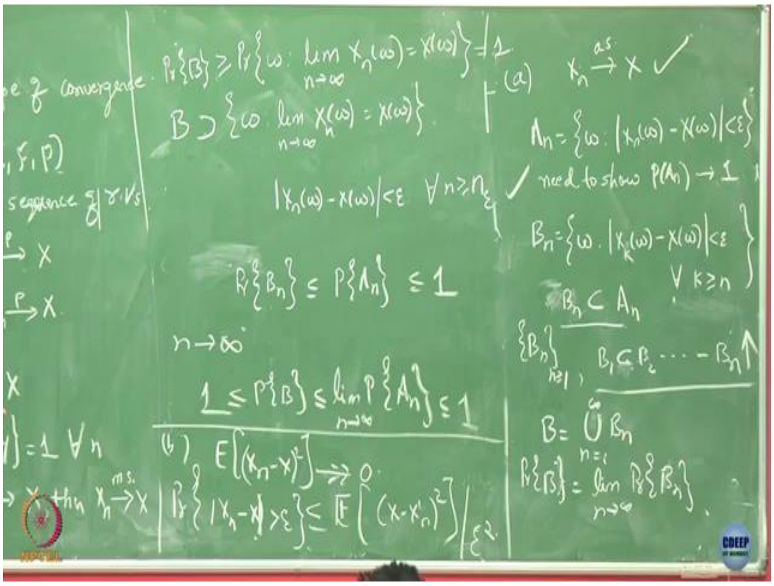


If I am going to apply probability on this, I know that probability of B because B is a largest set than this, it must be the case that probability of B is probability of omega limit n tends to infinity xn of omega because this set is contain, so this right hand side is contain in this, it means the case that the probability is less than this, but by my hypothesis, this quantity is going to be 1 because that is the definition of almost sure convergence.

Now, as I let n go to infinity, so, here if that is the case, I know that probability of B is lower pointed by 1 that means probability of B is 1, it has to be 1 and if I now let n go to infinity here. So, this quantity by definition as n goes to infinity this is going to be simply P of B which I have shown to be equal to 1. So, this is going to be a probability of B less than or equals to limit as n tends to infinity probability of An less than or equals to a 1. And this I have already shown you less than or equals to 1.

So that is why it must be the case that this sequence is equals to 1. So it implies that this quantity implies convergence in probability. Just also, let us quickly discuss this part also this is going to be slightly easier.

(Refer Slide Time: 41:35)



Now, suppose we assume Xn converges to x in the mean squared sense. What does that mean? I know that 0, goes to 0. Now, what I want? I want to say something about this because this quantity is related to convergence in probability.

So, how can I write this quantity in terms of this in terms of the expectation? So, do you recall any relations we studied so far, where probability of a random variable being larger than something I could express in terms of…

Student: Markov’s inequality.

Professor: Markov’s inequality, So, if I apply Markov’s inequality here, what is this going to be? What is this upper bound is? What did Markov’s inequality say? Probability that this quantity is greater than epsilon is upper bounded by expected value of square of this divided by.

Student: Epsilon square.

Professor: Epsilon square, epsilon is some constant but positive constant, however small it is and Now just apply if I let n go to infinity, by my assumption that xn is already convergence means squared sense, this quantity should go to what? It should go to 0, however, small your denominator is. If this quantity is 0, what is this quantity as n goes to infinity, it is going to be 0. And that is exactly what is the definition of convergence in probability, right? so fine.

So let us leave it here like this c and d, you can look into the book and you have to again go through construction of such sets, do some epsilon delta business to get the proofs right. So we will just leave it there. You can just skim through the proofs. But let us make sure that we understand these results further, so it is saying that convergence on P implies convergence D and you already talked about this.