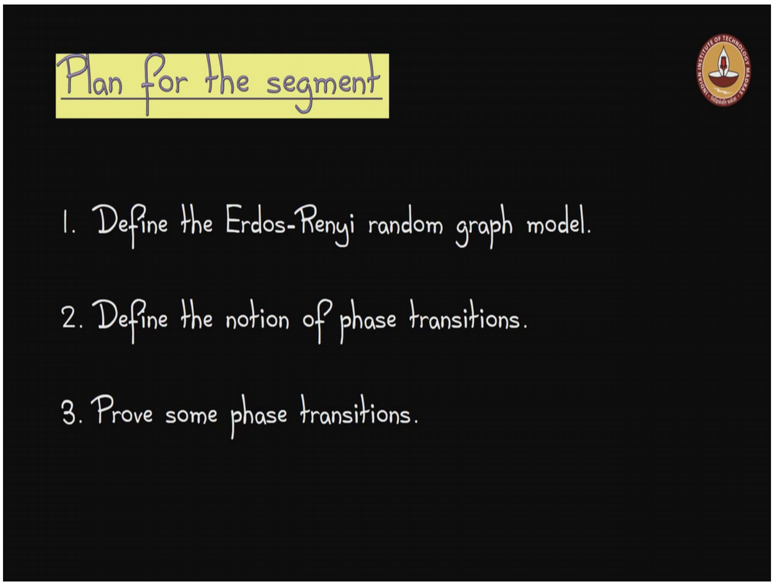
Probability and Computing
Prof. John Augustine
Department of Computer Science and Engineering
Indian Institute of Technology, Madras

Module - 04
Application of Tail Bounds
Lecture – 24
Segment 5: Random Graphs

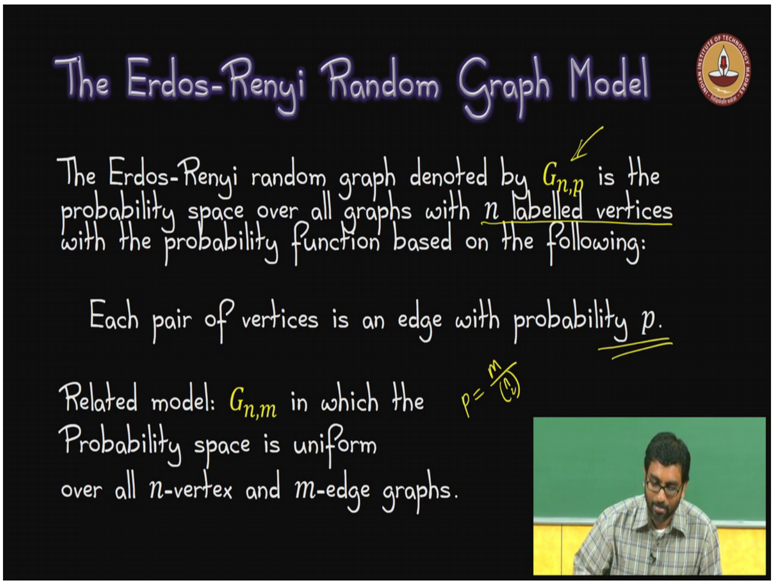
So, let us get start. So, it is a in segment 5 of module 4 in this module again because segment 4 was a little bit on the longer side I am going to just limit it to segment 5. So, hopefully it is a short.

(Refer Slide Time: 00:37)



So, what we are going to do today is we are going to introduce a very beautiful random graph model they just a new model it is very very simple and that is probably why it is so beautiful and leads to a lot of interesting things in particular this notion called phase transition will define that and we will show at least one phase transition ok.

(Refer Slide Time: 01:02)



And so let us without further a do let us get started with the definition of the Erdos-Renyi graph model. So, it is a little bit of history so this paper was published probably in the early sixties and since then this model has really gotten the attention of people lots of very pretty and results have improved.

Basically what is this? The it is a graph denoted by it is it is a random graph. So, essentially what it is a probability space it is we are denoting that by G n, p it is the probability space over all graphs with n labeled vertices and if it is a probability space then there must be a probability function associated with it.

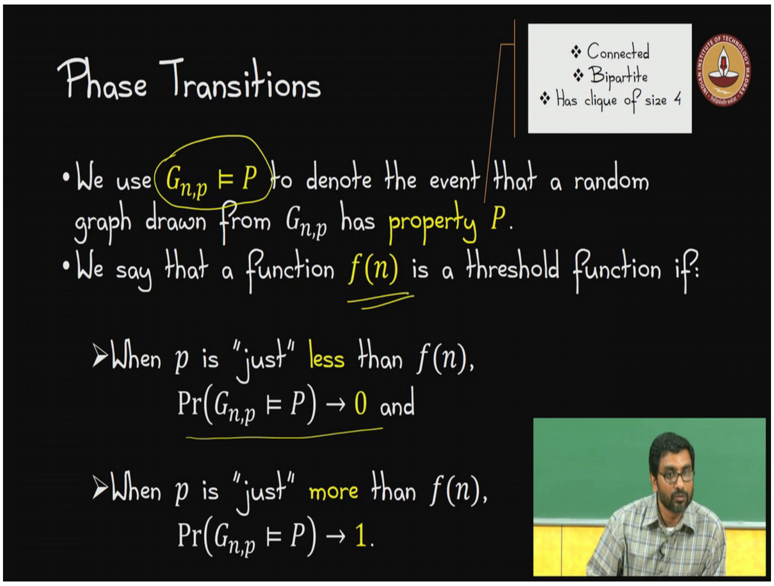
How do you get the probability function in some sense implicitly for each pair of vertices in you create an edge with probability p so that accounts for the two parameters that show up in the notation G n, p n is the number of vertices p is the probability with which any pair of vertices you see will have an edge ok.

So, now, one way to think about it is you create the n vertices and for each pair you toss a biased point with probability p and if it shows up heads put an edge between them otherwise move on to the next pair and these are all independently chosen of course. So it is a very elegant simple way to define a random graph ok.

And actually closer to what Erdos-Renyi originally defined is this notion called G n m ok. In this case this is actually very close these two are very closely related G n m is again a probability space again over all n vertex graphs, but now you choose m edges at random from uniformly at random from all n n choose two possible vertices edges ok.

And these are essentially similar if you set up the probability p to be well if you set up the probability p to be equal to m over n choose 2 then essentially both will be about the same they are not exactly the same, but about the same our focus of course, will be only on G n, p.

(Refer Slide Time: 03:50)



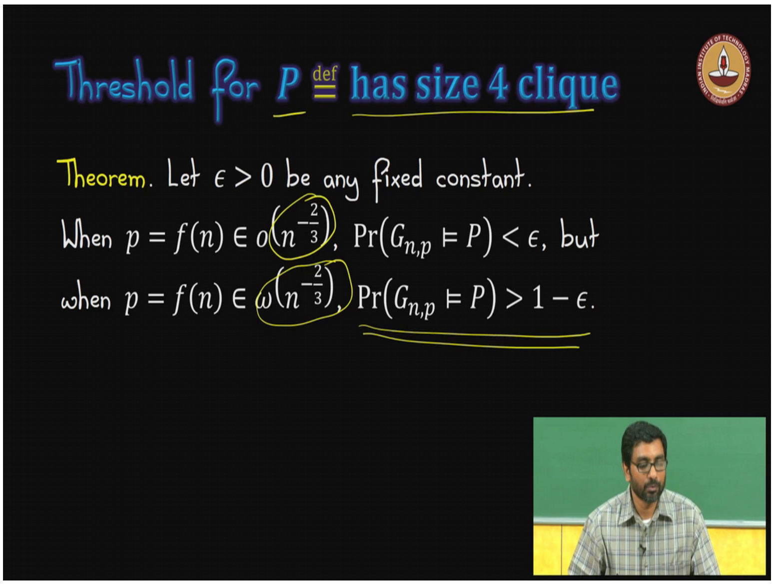
And we use we introduce a notation G n, p models p or has the property p it it denotes the event at a random graph drawn from this distribution this property space has the property p and what is the property P it could be any graph theoretic property graph is connected or bipartite or has a clique of size 4 or whatever.

You need property you know we can ask thus a graph drawn from this probability space have that property or not that is in the whether it has the property or not is the event and then you can talk about the probability of that event ok. And what do we mean by threshold I mean phase transitions if a property has this nice threshold function some function such that when p the probability with which edges are chosen.

If p is just less than a threshold function then the probability that you see that property tends towards 0 which means most likely you do not have that property when the p is just less than the threshold, but then we you when you raise the probability just above the threshold then immediately the probability that you see that property skyrockets towards the almost always there probability that is probably one all right.

So, this is phase transition so when you can come up with such a threshold function then you we say that you experience a phase transition at that threshold function ok.

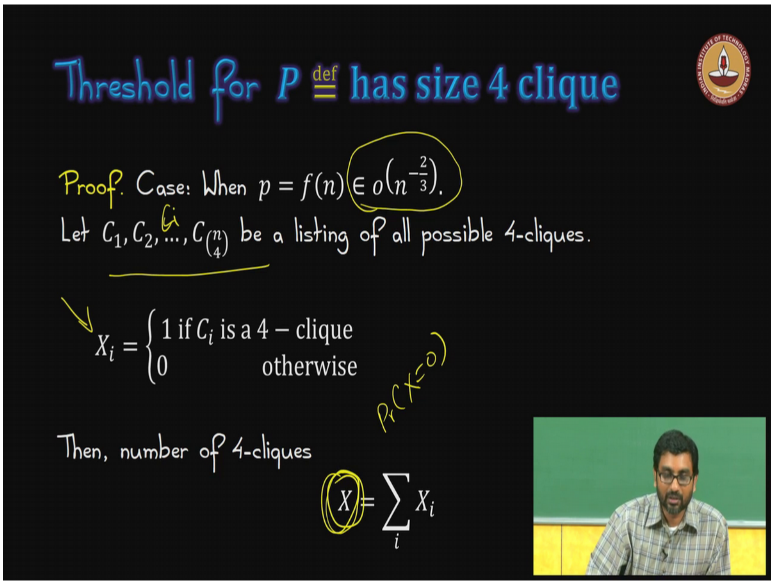
(Refer Slide Time: 05:50)



So today what we are going to do is study the one such phase transition. So, we are going to look at the phase transition for the property that property p that says that the graph has a clique of size 4 ok. In other words way to think about it is when you that is when you when your probabilities are very very small and you create this graph on n vertices the when the p value is very very small you do not see cliques of size 4 you might see cliques of smaller size, but it is very difficult to see cliques of size 4.

But then when you just increase it beyond the threshold you suddenly start to see cliques of size 4, and soon as it n particular here the threshold function is n to the minus 2 by 3 when you set the probability to be little o of n to the minus 2 by 3 then the probability of seeing cliques of size 4 tends towards 0 and otherwise the moment you say that the probability of having edges is little omega of n to the minus 2 by 3 then immediately the probability that you see that property tends towards 1 ok. So, let us see how we can prove this theorem the first claim is reasonably straight forward.

(Refer Slide Time: 07:26)

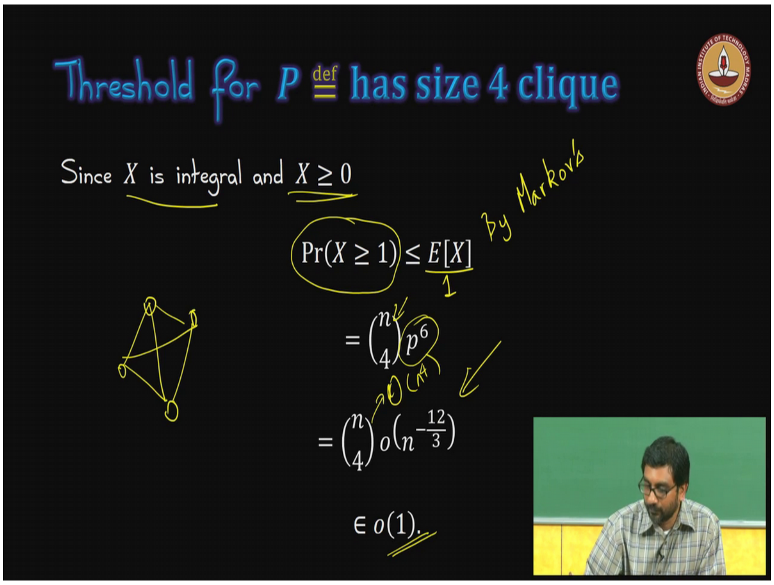


So, this first claim is when the probability is less than or not technically little o of n to the minus 2 by 3 we won't be able to say that you do not see clique so size 4.

So, how do we do that well let us start by listing the cliques all possible cliques there are n choose 4 possible cliques ok. So, you list them one by one and some arbitrary. And you we are finally, going to shoot for the number of such cliques that get formed in particular we want what we are finally, shooting for is the probability that X equal to 0.

We want to understand what is the probability that you do not see cliques right but towards that what we are going to do is we are going to set up the variable so that this X can be broken into these X i's X i's one if the i-th c i is a 4 clique and 0 otherwise it is an indicator random variable. So, clearly now then the number of 4 cliques is given by uppercase X which is a summation of all these individual X i's.

(Refer Slide Time: 08:46)



And you will notice that X is integrals it is an integer and X is always going to be greater than 0. So, now we can ask the question what is the probability that X is greater than or equal to 1 we want to argue that this probability becomes very small little o of 1.

So, what is the probability that X is greater than or equal to 1 first of all we claim that this is less than or equal to the expectation of X. Why is that so that I was thinking of a slightly different way to do this but yes.

Student: (Refer Time: 09:24).

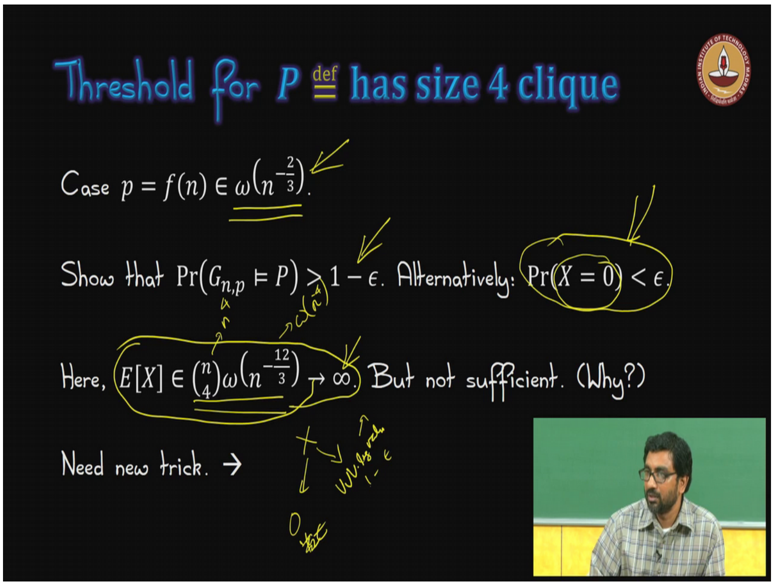
You can you can do this, this way I had a slightly different idea, but that is picked up another way to see this great so, by Marcos let us even put it that way alright ok. So, then what do we do so this is at most expectation of X. So, we need to bound the expectation of X actually we can get the exact value of expectation of X because first of all there are n choose 4 such indicator random variables so that is and each one of them is one with probability p to the 6 why is p to the 6 showing up here.

Student: (Refer Time: 10:13).

Yeah because you are talking about 4 vertices and you are asking whether all of them have all pairs are connected by edges and so you will have 6 edges and so that is why you have p to the 6 ok.

And we know in this case we are in the case where p value is little o of n to the minus 2 by 3 and p is raised to the power 6 so you get n choose 4 times little o of n to the minus 12 by 3 and n choose 4 is o of n to the 4 and then but this term is little o of n the this is what n to the minus 4 so overall it is little o of 1 and convinced about that so, so the first case was nice and easy. So, let us look at the second case.

(Refer Slide Time: 11:18)



So, this is the case where the threshold function I mean [vocalized- noise] this probability function is little omega n to the minus 2 by 3. So, it is more than n to the minus 2 by 3 and in this case we want to show that the probability with which the property will hold is tending towards 1 remember epsilon is a very small quantity.

In fact, for any small epsilon we want to be to prove that it is greater than 1 minus epsilon another way to prove this is just the opposite the probability that X equal to 0 is at most epsilon if this is the this is the case where you do not see any 4 clique some sense the (Refer Time: 12:16) in this context. The bad event is happening with less than epsilon probability ok. So, let us see.

So now, here let us look at the expectation of X in this case remember p value is little omega of n to the minus 2 by 3. So, when you plug it into that formula remember we said n choose 4 times the expectation of each individual X i's you get this quantity see this is n to the 4 and this is going to be little omega of n to the minus 4 ok.

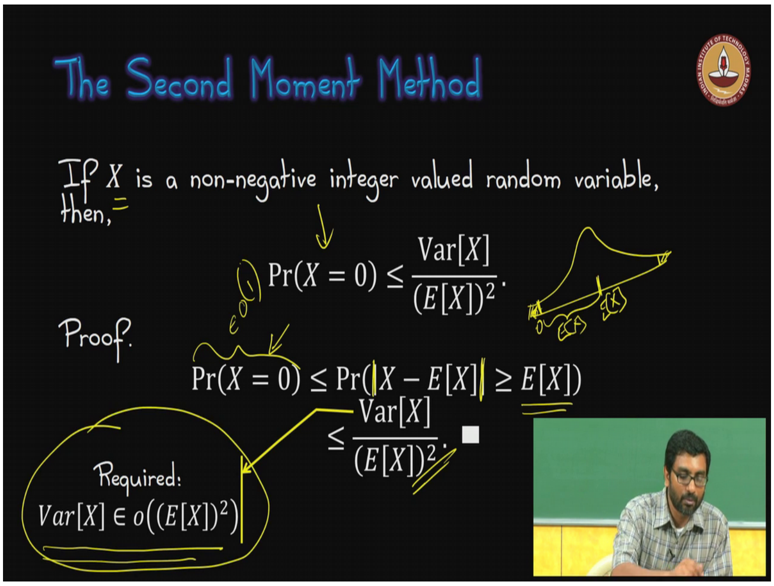
So, overall it is going to be tending towards infinity and so this seems promising. Why because now what we what we are able to see is that the expectation of X is tending towards infinity which means that this X equal to 0 intuitively is an unlikely event ok, but that is not sufficient technically. Because we won't be able to show that X equal to 0 with tech we won't be precise about this we will show that the probability that X equals 0 is at most epsilon.

So, one thing the you know why is this not sufficient let me ask you this why is this not sufficient why is this not sufficient. You could have situations where the expectation is very large because let us say for example, a situation where X takes the value either 0 or a very very very very very large value 0 with some probability say half and this very very very large value with or rather.

Let us say 0 with probability epsilon and some very very large value with probability 1 minus epsilon and if this large value tends towards infinity then itself you will get expectation of X to going to infinity, but that won't satisfy this requirement that you have so that is why we need to be a bit careful here.

And this is where it connects to tail bounds remember we are in the context of tail bounds we are trying to understand applications of tail bounds. We need a new trick we need a tail bound based trick.

(Refer Slide Time: 15:00)



And what we are going to see is what's called the second moment method and where actually going to apply Chebyshevs, but in fact, we are just going to rewrite Chebyshevs in the in the manner that we want and then we will try to apply it.

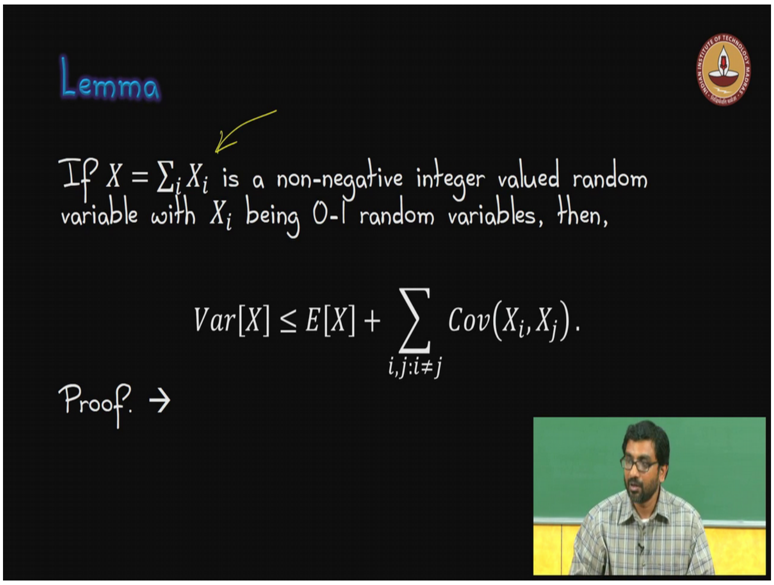
So, now, we are going to assume that X is a non negative integer integer random variable. Now we want probability that X equal to 0 and this is exactly what we wanted over here ok. And this probability that X equal to 0 is at most the second moment method uses this inequality that it is at most variance of X divided by expectation of X.

The whole squared and it is actually coming from a very very straightforward application of Chebyshevs. So, what is X equal to the property of X equal to 0 you can it write this so what is this so now, you want to know what is the probability that X equal to 0, and let us say this is the expectation of X what Chebyshevs tells you is it bounce both of them this is the 0 mark. So, it bounds X minus E of X greater than or equal to some parameter a right in this case that parameter is nothing, but E of X itself right. So, in here this is also E of X here ok.

So, you can basically write this X equal to 0, event as well it is basically X equal to 0 will be a subset of this event that says the absolute value of X minus E of X is greater than or equal to E of X and then you apply Chebyshevs equality and you get.

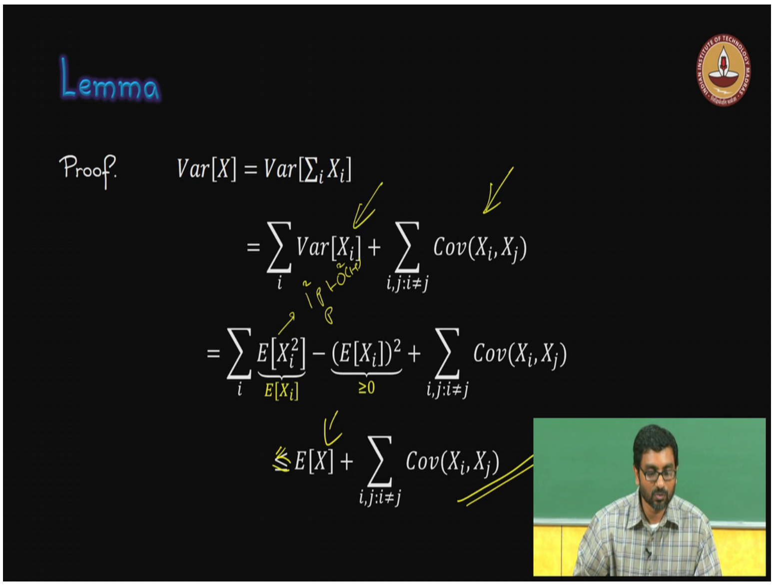
In our context so, going back to you know we want to be able to say that this is in little o of 1 we want to show that this is little o of 1 basically now what we are need to do what we need to show is that the variance of X is little o of expectation of X if we show that this whole thing will fall into place so that is going to this is going to be our focus from now on.

(Refer Slide Time: 17:30)



So now we need to understand the variance for which let us try to bound the variance when we have a certain property we have this property that this X is the summation of 0 1 random variables if that is the case well it; obviously, looks very similar to the variance formula that that we already are familiar with so we just need to do some minor massaging ok.

(Refer Slide Time: 17:57)

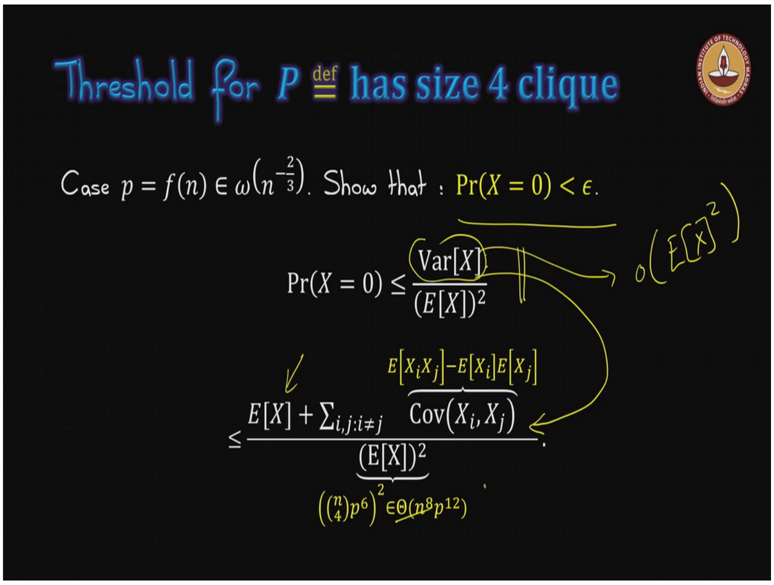


So, we know that variance of X is the summation of the individual plus these covariance terms ok. The covariance terms all is already matching what we have over here so, we do not worry about that so we only worry about the sum sum over all the individual variance terms and well then you can just apply the formula variance of X i's is expectation of X i square minus expectation of X i the whole square that is just a formula ok.

But now we are going to take advantage of the fact that X i is a 0 1 random variable it is a 0 1 random variable E of X i squared what is that it is going to be one squared with probability p and 0 0 with probab 0 squared with probability 1 minus p and that is nothing this 1 square does not it again shows up as one and so it is again just p right. So, this is nothing but, expectation of X i and here we want an inequality so and this is a positive quantity. So, we simply can ignore it you want the less than or equal to over here ok. So, it is nothing, but expect sum over.

All ixi which by linearity of expectation is expectation of this big random variable X plus the covariance plus the covariance terms ok. So, this is this is what we have. So, now, we have a way to bound the covariance so we will just go ahead and bound.

(Refer Slide Time: 19:45)

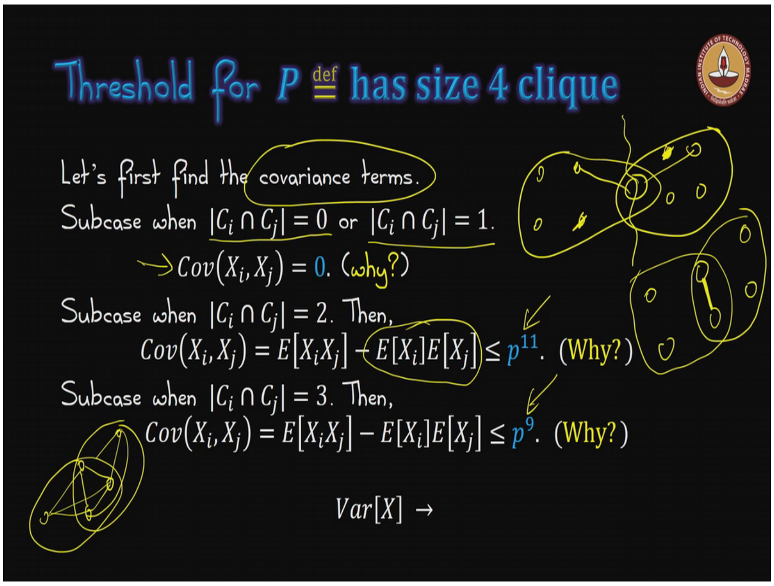


So, now we need to just to recall we want to show that the property of X is equal to 0 is at most epsilon and we are applying the second moment method and we have a way to bound the variance which is what is written over here expectation of X plus all the summation of all the covariance terms.

Now if you look at the so a few things that we need to worry about here. So, we need to worry about the covariance terms now and in the denominator we have an expectation of X the whole squared ok. So, so if you think about it if you recall we want to show that variance is what do we want to show we want to show that variance is little o of E of X squared right. So, this is not going to be too much of a problem so covariance terms that we need to be careful about now ok.

But anyway let us go through all these terms and just to recall the covariance has this formula covariance of X i and X j is expectation of X i X j minus the product of the individual expectations you have that formula and we already know the expectation of X so this expectation of X the whole squared is it is going to work out to n to the 8 p to the 12 that is just by applying formula.

(Refer Slide Time: 21:20)



So let us now focus on the covariance term so now we are going to focus on the covariance terms. And for the covariance we need to be a bit careful because remember this covariance is you know is over all pairs i j pairs right and funny things can happen. So let us break it down into cases the first are sub cases rather sub case what let us see the first sub case is this C i intersecting C j is either the cardinalities either 0 or 1 and I claim that this is the easy case with all the covariance terms are zeroes why is that?

Student: (Refer Time: 22:03).

Independent or saying no we need to be there just independent why are they independent. So, let us look at them one by one what is C i intersecting C j the cardinality equal to 0 what is that situation?

Student: (Refer Time: 22:22).

Yeah the cliques are completely disjoint there is one and then the other one what happens in one set of 4 vertices will have no bearing on the other set of 4 vertices what about cardinality of C i intersecting C j equal to 1. What does the picture look like what we will have to do we have to do to this picture yeah one of the vertices is common right.

So basically let us remove each. So, now we are asking these for forming a clique and are these 4 forming up clique and why are these still independent.

Student: (Refer Time: 23:12).

Yeah the.

Student: (Refer Time: 23:14).

So, the vertex one vertex might be common, but the edges are all chosen independently. So, because this edge whether this edge is chosen or not has no bearing on whether this edge is chosen or not ok. So, even though there is a common vertex the whether the cliques are formed or not are independent events.

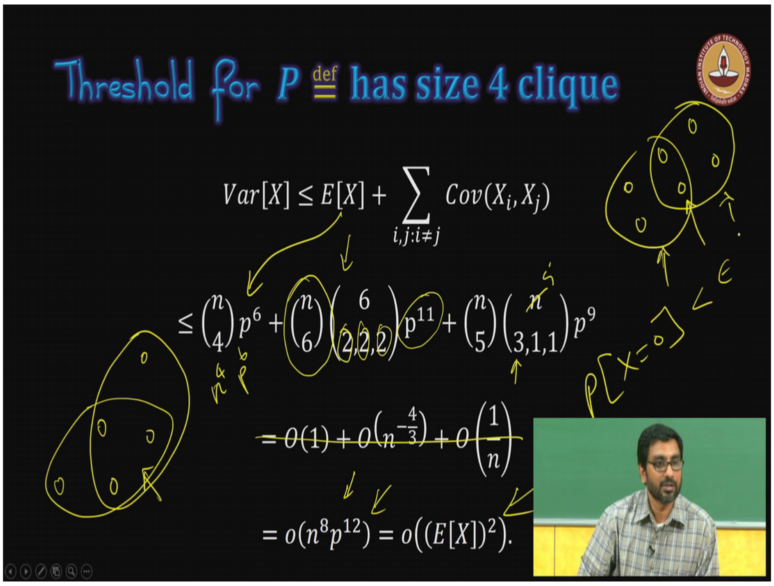
So, the covariance terms for those cases are 0. what about the case where C i intersecting C j equal to 2 so, what is let us draw the picture for it so we have two common vertices and then the rest of the vertices will look like this.

So, we are asking are these going to form a clique and are these going to form of clique ok. So, now there is essentially just one edge that is common. So, now, you start to see some amount of independence showing up sorry dependence. So, now, we are going to claim covariance is upper bounded by p to the 11 why is that well for one thing we want an upper bound.

So, we are going to ignore this this term we only have to worry about E of X i times X j. So now what we want what is that that is the expectation I mean that what is X i times X j it will be 1 when both cliques are formed and 0 otherwise. How do you get to form both the cliques? How many edges should actually get formed if you look at the two cliques are separate how many it is a total of 12, but one of those edges is common. So, you will need to form an eleven edges that is p to the 11 so that is so those that that will be the covariance term for the clique this case where this one common edge a very similar argument can be used for the case where 3 common vertices when there are 3 common vertices. Let us let us draw the picture.

So, you have 3 common vertices and we are asking these forming a clique and are these forming a clique ok. So, now, there will be 1, 2, 3 3 common edges and then 12 3 4 5 6. So, there will be 6 other edges to a total of 9 edges. So, we have either so now, we have the covariance terms individually. So, we just need to put them into the summation ok.

(Refer Slide Time: 26:07)



So, recall that this is the covariance that we want to compute I mean sorry variance we want to compute. This expectation of X is this term n choose 4 p to the 6 we have already seen that whenever C i and C j have an intersection of either 0 or 1 elements we have just discarded them. So, we only have to worry about the case where they have 2 common vertices or 3 common vertices.

So, this is the case where they have 2 common vertices you may recall that when they when they had 2 common vertices the covariance itself first p to the 11, but how many such covariance terms will be there that is that is going to be the number of ways in which you can choose these 4 4 vertices for the first clique 4 vertices for the second clique such that they have 1 2 common vertices ok.

How do you count that first of all you need to remember the picture you need to first get 6 vertices after you get 6 vertices then you have to choose 2 of them to be the common vertices. So, get this is choosing the 6 vertices then out of those 6 vertices you want 2 to be in the middle.

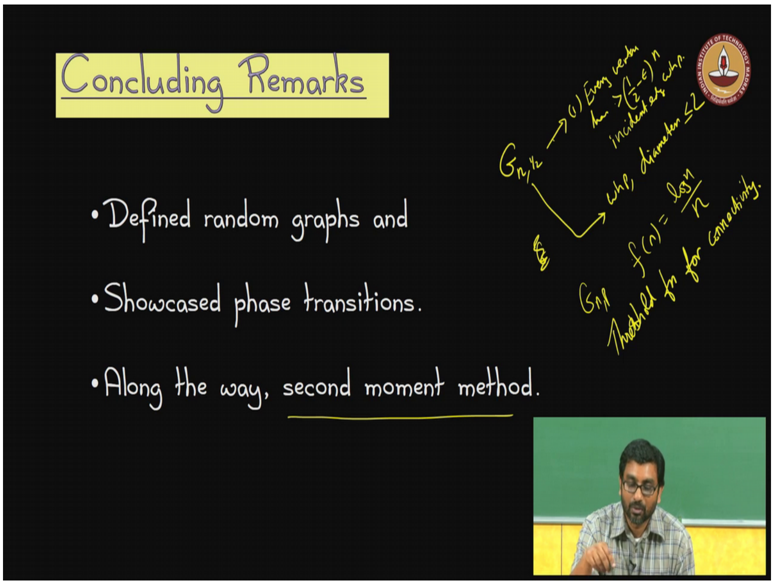
Then you want to choose 2 to be here and then you want of course, the remaining 2 will go here that is the number of such covariance terms that you will have so you get that and a similar argument it can be made for this 1. So, here again if you recall the picture 3 common vertices so this, the out of the out of the n vertices you choose 5 vertices and out of those 5 3 of them go in the middle and then one on the left one on the right alright.

So, now what we do is let us plug in the terms. So, in the interest of time what I am going to do is just point out. So, so remember that p is we need to be a bit careful here ok. So, let us for the moment just ignore this line ok.

So, let me see now let us just write these things in terms of n and p. So, here this is going to be an n to the 4 p to the 6 and that is clearly the claw of n to the 8 p to the 12 then this is going to be an n to the 6 p to the 11 again little o of n to the 8 p to the 12 this is going to be an n to the 5 p to the 9 again little o of n to the 8 p to the 12.

So, you get little o of expectation of X the whole square which is exactly what we wanted to show struct really brings us to the end of this segment. Because what we have shown with this we have shown that the probability is going back probability X equal to 0 is less than epsilon for any arbitrary small epsilon so that is exactly what we wanted to show and completes all the cases. Well let me conclude.

(Refer Slide Time: 29:49)



So, what we have introduced is random graphs define transitions and we have shown one phase transition we have shown along the way we we learned the nice method called the second moment method in the next module which we will do.

After and in this classroom setting we will be doing this after our first quiz will go into things like bowls into bins and some other computer science applications of them. So, thank you some easy questions first. So, look at this random graph G n comma half are ok.

So, here I want you to ask argue two things one is argue that every vertex has more than minus epsilon times n incident edges with high probability of course, should be very easy and the second thing I want you to prove in this is prove that with high probability diameter of this is at most two couple of things that you can you should be able to prove quite easily.

The other thing is that for this is for arbitrary p f of n equal to log n over n prove that this is a threshold function for connectivity for the property that the graph is connected there is some interesting problems to work on in the 10 minutes you should be able to do this.