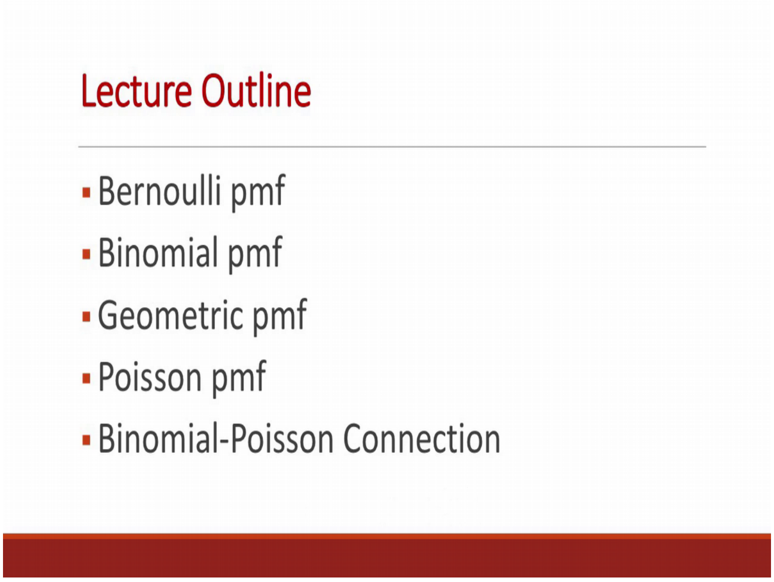
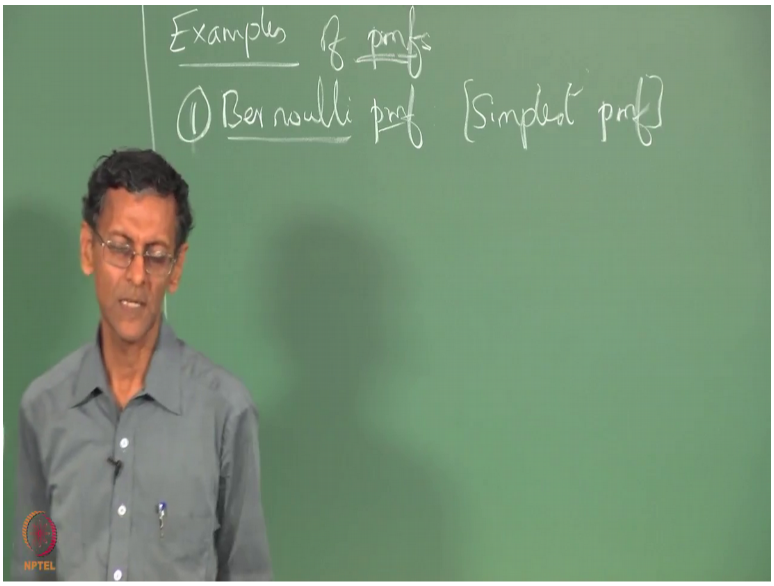
Probability Foundations for Electrical Engineers
Prof. Aravind R
Department of Electrical Engineering
Indian Institute of Technology, Madras

Lecture - 17
Important PMFs

(Refer Slide Time: 00:13)



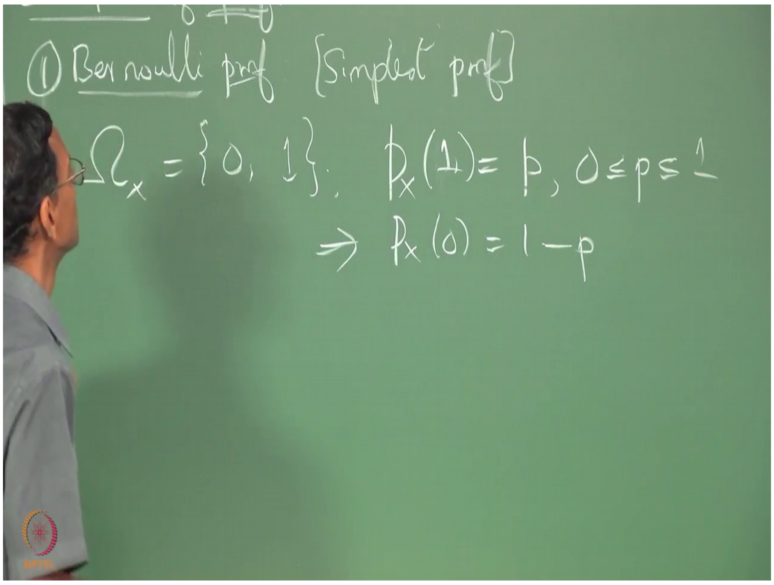
(Refer Slide Time: 00:18)



And remember what is important again, is only the pmf right. The name some people know in some books, you find these are examples of random variables. They are not really examples of random variables, there are really examples of pmfs, in some sense right. So, we are only looking at, we are giving the pmf importance say. The pmf is a model for as a probability model. And the X is just this the capital X is just something and the name we are giving to that random variable which takes values, just for convenience right. So, what are the important examples? We start with I am going to number them. The first one is the simplest one. The Bernoulli pmf is a simplest and more right simplest non-trivial pmf that we have. So what is it that the basic non I mean the most simplest pmf, it needs to have at least 2 possibilities just like head tail right.

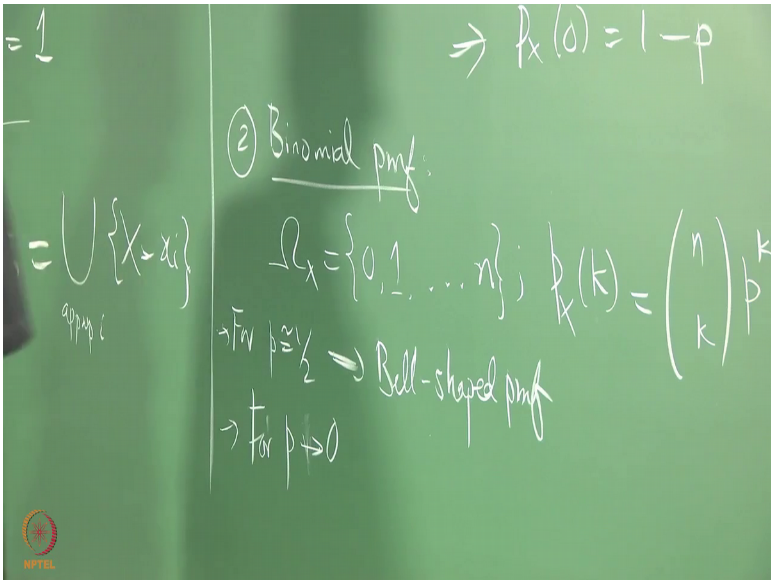
So, instead of head tail we are going to call it 0, 1.

(Refer Slide Time: 01:25)



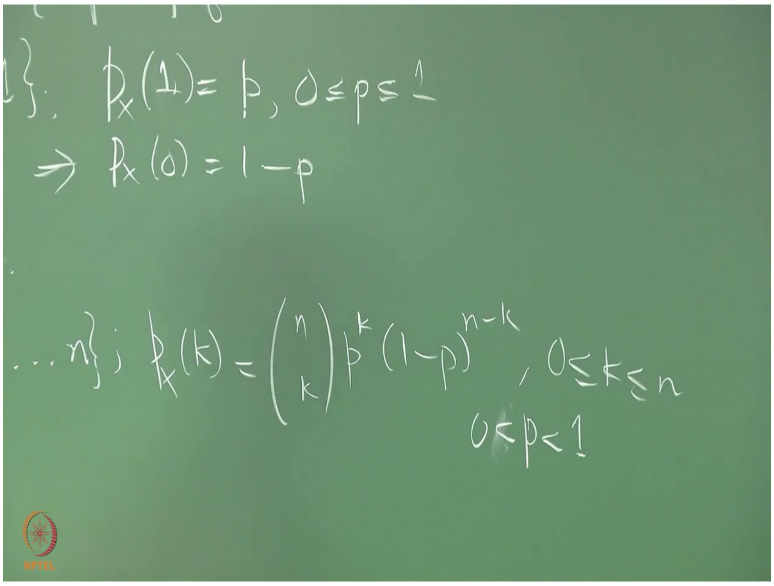
So, here omega x is basically 0 comma 1. And p x of 0 or p x of 1 we say is some number p. So, this pmf is parameterized entirely by this number p, which could be any number between actually in the region that case you are also allowing it to be to take the extreme values right, so obviously, p x of 0 must be what 1 minus p. So, note that right I am not going to keep writing this again and again right. The convention is to let this number small p or the; this the letter p denote the probability of 1 and 1 is almost always identified with either a head or with success right. 0 is identified with tails or failure right. So, that we will keep that convention for the rest of the course. So, this is right, the,you cannot get anything simpler than this. Anyway;

(Refer Slide Time: 02:30)



So, the next is, we have already seen the Binomial pmf. Here omega x is the set of integers. Let me see which number letter I used here I used did I use capital n or small n. I will use small n lowercase n. So, where all did I use capital n in the past. I guess I used it for the receiver number of endpoints. But here anyway, we will stick to small n right lowercase, all right.

(Refer Slide Time: 03:07)



And what is p x of k; what is p x of k here, k for any integer between 0 and n is what?

Student: n choose k.

N choose k, p to the power of k, 1 minus p to the power n minus k. So, this is valid for 0 to n. We have already seen this model derived it in the context of k successes in n trials right; except now that we are giving, we are giving it the, giving that number of successes a name x. x refers, now, you can now directly say what does x stand for, number of successes right.

So, that, it is complete now. And of course, I think we have already seen the fact that they that this is satisfied right. The probability is obviously to add up to 1. So, what is the shape of this pmf by the way? We have to now, no here there is no big issue in looking at the shape because there are only 2 points. But here, the shape is important to understand right. What do you think is the shape of this pmf? If you plot those numbers as a sequence, now we have a discrete time sequence here if you think of it as in that sense right. We have all studied dsp so we know what sequences are like. So, if you plot this what is envelope of this pmf look like?

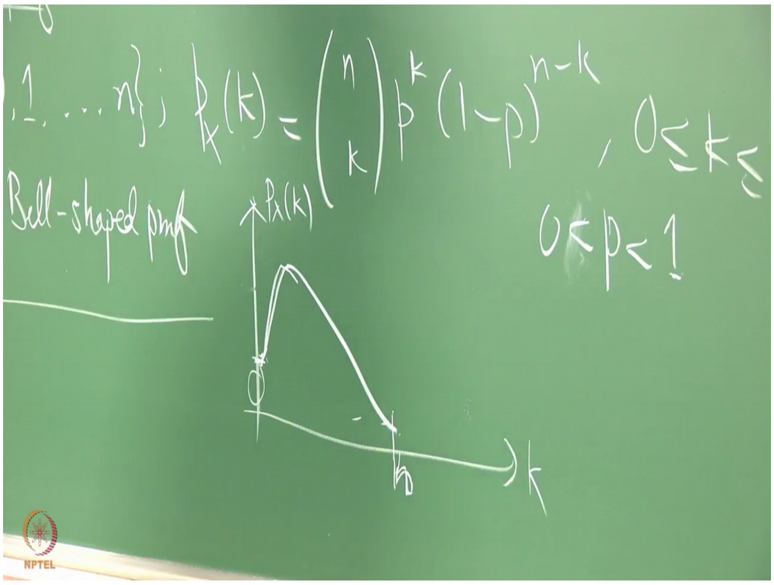
Student: (Refer Time: 04:33)

Student: It goes close to a bell wise.

It is not necessarily always close to a bell. It is close to a bell, the answer is it is close; it is a bell shaped, but it is not really bell shaped for small p or large p, p close to 0 or p close to 1. I have to make again, note that+ this p is typically between it is not right, it is not exactly 0 or 1 because, you have some degeneracy if you push it exactly to 0 or exactly to 1 right. So, right if you say it is exactly 0 then that means, the success cannot happen at all right. If you look at this so I mean, this p you know what this p to be exactly, this factor to be either 0 or this factor to be exactly 0. So, therefore, we put p between 0 or no, but you can make this be arbitrarily close to 0 or arbitrary close to 1 here right. So, at those extremes, it will not really look like a bell shape. So, it looks like a bell shape only if only for p around half, right.

So, for p equal to half we get bell shape. And as we assume that n is fairly large right, to a more than 20, 30, 40 whatever right; p around half all right. So, if p is very close to 0, for example, what do you what kinds of a shape will you get? A p tending to 0. I think let me draw it here.

(Refer Slide Time: 06:01)



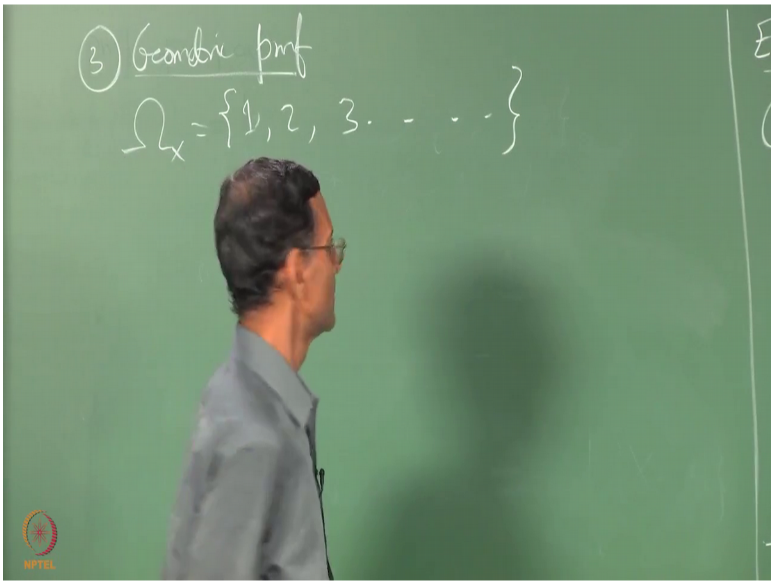
Where do you think the probability will be concentrated? For p, for b, for small values of p where will the probability being concentrated?

Student: near 0.

Near 0 right. Because you are saying that that the probability of success is on any trial is not very high right. And you are counting the probability of case success; obviously, no success or very few successes are going to be a lot more likely than k close to n right. So, you will get something like. And remember you are not ever going to go exactly to 0 right. So, it will be like this. So, I am exaggerating this gap out here, for especially, for large n, but it will be like this. You note that right I mean if I sub substitute k equal to 0 or n I do not get 0 right. I get some finite number if p is not you know it is somewhere in between right.

So, the fact is this peak will occur close to k equal to 0. And the complementary situation for p close to 1 right. It is, that has to happen because of common sense right. So, whatever we get from the maths has to agree with our intuition here.

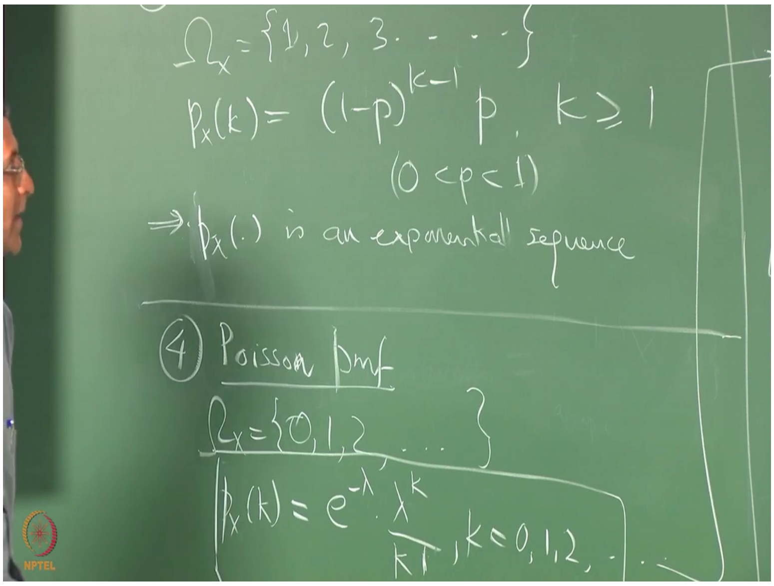
(Refer Slide Time: 07:37)



Then, the Geometric pmf. This is again something that we have seen right.

So, here omega x is un-count countably infinite. So, we are saying 1, 2, 3.

(Refer Slide Time: 07:53)



What did we say? got this; we wrote it as 1 minus p power k minus 1 times p. Again this parameter p is between 0 and 1 and right. The expression that we have is for these values of k right. So, these are all you know defined to be pmfs defined on integer sets for an integer arguments right. But as I said yeah that does not have definitely always a case if you take, for example; the number of the average number of successes if you define something like that; k by n that would not be an integer. But you will soon get a pmf right. Each of these values will get divided by cap divided by small n and you can you will inherit all these probabilities without making them sit hard on any digit right.

So, you should not think that omega x is always some collection of integers, no, right. I think we have said it right from the beginning that they can be any real numbers right. So, we have you know in the geometric case we have this very important pmf, which again as we said earlier is represents the number of times you keep trying till you hit that first success right. In their base; you know in a repeated sequence of trials where each trial can have success probability p right we have sorted done this. So, once again I mean I am I will leave it to you to think. You know this is basically an exponential sequence right, in the language of dsp. So, p x by itself is an exponential sequence. It is an exponentially decaying sequence right. And I think we have all learned to love these sequences from our dsp days right; a power n, u n, except that you are starting from 1 and not from 0. But some textbooks also start from 0. If you start from zero, that means, what are you doing with? You are actually counting the number of failures before the first success. So, if you redefine it that way you can start from 0. But that is a little less common than this right. To start most commonly we start from 1.

So, these are simple pmfs that we have already sort of seen right. And let me look at a very interesting pmf for those of you that may not have seen it explicitly, that is, Poisson pmf right. Let me write it here and we will continue there. Now, in practice right, certain counts of electron emissions, telephone calls in some interval basically right, the way in which this pmf has been observed to arise in practice is exact[ly] the number of times something happens right, in a certain time interval. Again I do not want to get too much into the history of this pmf and where all, but let me just briefly summarize right. A right from the 1920s people when people studying radioactivity when there is looked at a emission of, actually even earlier I guess right, but they started counting them using geiger counters about hundred years ago right.

So, what they did was, they observed the number of emissions that were happening over a long period of time and they partitioned the time interval into smaller time intervals and looked at the number of emissions that happened in any particular time interval. And they found that there was a statistical regularity about this. So each interval had a random number of emissions. But overall, there was right, you know if you looked at the; you know you could say you meaningfully talked of the probability that there would be 0 emissions or 1 emission or 2 emissions like that right. So, so that was the birth of the Poisson pmf right. And then the same model was also again I do not have, obviously, any first hand evidence of this, but the number of telephone calls in a particular interval that is in the pre cell phone year. Now I am not sure right if it still Poisson anymore right, but anyway.

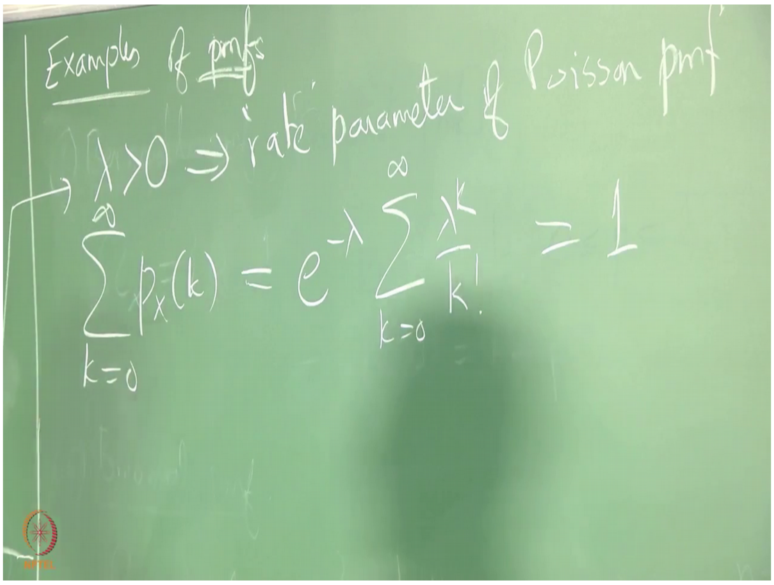
So, but it is, but there is no doubt that the Poisson pmf continues to be a very practically very important model. And I hope that you people, if you study this you know queuing theory and so on, you will definitely encounter this model again. So, at the; in this course we will of course make a lot of use of this model without any reference to what it actually means right the like any other most other things. So, what do we have what is omega x in this case? Again the bunch of integers, but now we start from; where do we start from?

Student: x 0.

We start from 0 right. Because, not the it is also 0 also possible. Only negative things are not possible. So, we start from 0 and we let it let the numbers right go without bound. There is no upper bound on this. Then, what is p x of k? Again, since, we have only now integers here I can just write p x of k. There is now a very important parameter called non negative parameter called lambda right, which is called the rate parameter of the Poisson distribution right. And the value of that rate parameter lambda is very is crucial right and it can be any non negative number again going all the way from very close to 0 to very high. And the nature of the Poisson pmf is fundamentally different whether you are looking at small values of lambda or large values of lambda right.

What is it in terms of lambda? It is e power minus lambda; lamda to the power of k by k factorial right; for k equal to 0, 1, 2 like that. So, these 2 in fact, Poission and Geometric are our first introduction to countably infinite sets right.

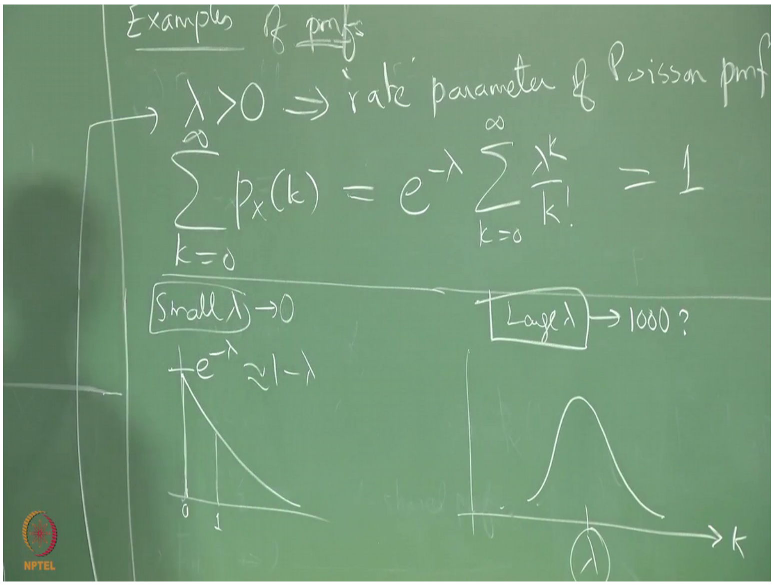
(Refer Slide Time: 14:36)



I mean, Ah when we are defining a pmf over countably infinite set of numbers. And in both cases right. You notice that x k goes to infinity, the p x of k must go to 0. Or the otherwise it cannot add to unity. So, what is a sanity check here? How do you ensure that this guy goes to 1? So, let me continue here. Lambda is strictly greater than 0 right, the so called rate parameter of the Poisson pmf right. All right. But how do we ensure that this is going to happen. Is this is this true? Why is it true? Because, you have the exponential series right. So, this is nothing but, this e power minus lambda turns out to be a normalizing term right. Because, this quantity here everybody knows from the study of the (Refer Time: 15:45) series that this must be e power plus lambda is it not. So, this is clearly unity.

So, once again we are using this right the factorial for the factorial notation k factorial is k with them by a bang right [explanation/exclamation] exclamation mark after it right. We will stick to this notation for factorial. It is the internationals notation. So, we do not want to use anything else right. So, please stick to that right. So, clearly this is valid right. So we have a valid pmf. And all these numbers are obviously, positive. There none of them is negative right. There is no way that this quantity here can ever be negative. So, that is also trivially satisfied. The important thing is the shape of this pmf, right. That will tell you exactly where the probabilities are concentrated. Where are they concentrated, for small lambda and for large lambda? For small lambda, again you get something that is what you saw for the binomial. Small values of k are much more where all the probability are in fact residing. In fact, you can actually simplify this right. If you say lambda so small, that you can actually ignore all the turn for k greater than 2. Because lambda power k, this if you take lambda equal to 0.01 or something that squared or that cubed is right people often ignore all those terms. They you just took, look at only 0 and 1, And then as you keep increasing lambda, what happens the whole character of the pmf changes right.

(Refer Slide Time: 17:31)



So, for small lambda you will get basically, just something like this p x of 0 is going to be the largest. What is the value of p x of 0? Lambda power 0 is always equal to 1 right. No matter what how small you make lambda. And 0 factorial is always 1. So, these terms are do not are always I mean they become 1 for k equal to 0. So, you have only e power minus lambda. For small lambda this e power minus lambda is what? You can sometimes replace it by 1 minus lambda right. And in fact, I should in fact, blow it up just to show you that I cannot even have, you know with this kind of a p k curve lambda is actually extremely small right. There is no point in saying this is 0, 1, 2, 3, 4, 5. It is like that. So, all right the pmf is going to 0 will go to 0 so fast right that you do not even want to look at it for even for k greater than 2 or 3 or whatever right. It is in fact very, so very heavily concentrated and on 0 and 1. Like I said, the limit is just taking the 2 values is 0 and 1 itself right. And for lambda equal 1 for I mean if you put k equal to 1 for example, you get lambda here and you get this e power minus lambda if you write as 1 minus lambda.

So, 1 minus lambda times lambda you approximate that by just lambda. Because you if you throw out lambda square, you guys just get that right.

So, 1 minus. So you can think of this 1 minus lambda for 0 and 1. That is also like a Bernoulli distribution. But it is a limiting case of Poisson distribution as lambda becomes very small. So, what happens when lambda is very large? So, this is small lambda going to 0. For large lambda going to a; you know 100 or 1000 or whatever. That is also possible physically right. There is no reason why you cannot; you do not encounter such large lambdas in practice. So, now what happens is what you think that happens.

Student: it (Refer Time: 19:48) it becomes centered at lambda.

It becomes, very much like a bell shaped curve centered at

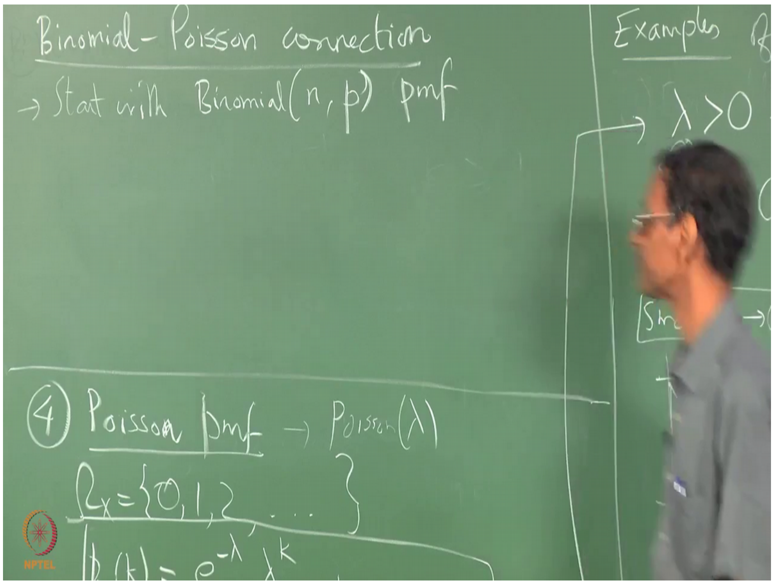
Student: lambda

Lambda.

Now, I cannot look at individual numbers anymore. I have to look at, you know I have to draw an envelope like this. Now lambda equal to 0, k equal to 0 will have very small probability. You go back this e power minus lambda will effectively kill all the probabilities close to 0. So, you will only get significant probabilities around k equal to lambda. As and you can, for the sake of argument you can say k lambda is an integer. There is no reason why you should be very you know should cut any that you know should be very whatever about it right general about it. You would say what happens if lambda is 1000, surely you will get k equal to 1000 here right.

So, that, so, this is a kind of pmf. That is a how that pmf transforms from small to large. In the binomial case, you do not have such as large spread simply because that number the parameter p for a fixed n is only going to vary between 0 and 1. Whereas, here you can get decades of variation. When I am saying decade I am talking of 0.1 to 1000. Lambda; you will that is so many decades right. What is a decade? A factor of 10 right. So, Poisson again is a very important practical model right. And this formula this p x of k e power minus. So please remember this right. I will expect you people, I am not going to give this 1 formula, I will not, I will expect you to remember along with the binomial expression. All of these things you should be able to, you should remember.

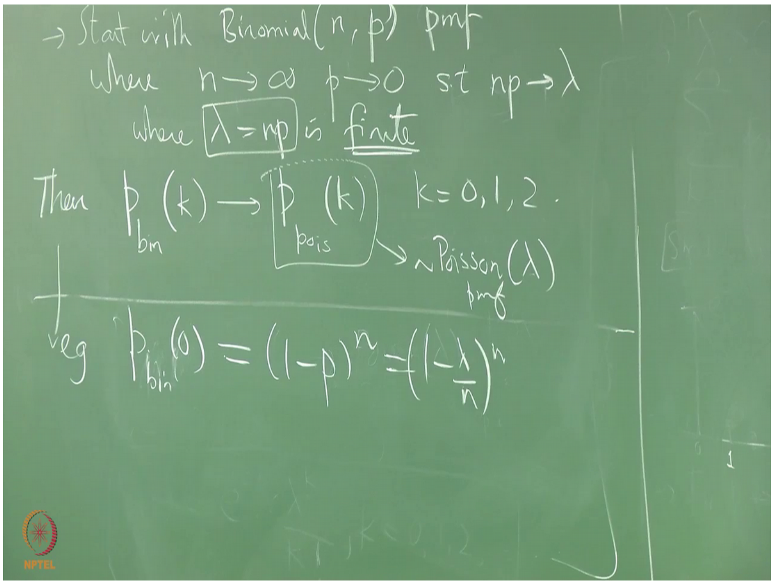
(Refer Slide Time: 21:44)



So, Binomial-Poisson Connection. So, this connection happens when start with Binomial pmf, right. In fact, this is a notation which is often used to describe this pmf. It just says binomial and you put n comma p. The 2 parameters of the binomial distribution right n and p.

You right there, and similarly this will be Poisson. How do you call this, you call this Poisson lambda right. So, I start with n comma p some right.

(Refer Slide Time: 22:41)



Where what, n is very large, n goes to infinity.

Student: p goes to small.

P goes small, p goes to 0 such that np goes to a constant lambda. Independent of or np is lambda is like p equal to lambda by n type of a arrangement where, n times p is does not get it does not change with n or p right. Then it turns out that the binomial pmf instead of being a 2 parameter pmf p comma lambda, converges to a single parameter Poisson lambda pmf. Lambda is now finite, not right, it is not going to 0 or go into infinity. So, we have a finite, well let me write that also, where lambda is np is finite. Any, some real number, obviously positive. Then this binomial pmf actually right, the p if I say p binomial now I am not going to attack I call it a no put a subscript for a random variable, I just put the subscript corresponding to the pmf it iself. P binomial of k right tends to p Poisson of k, for all I mean for almost cases over k equal to 0, 1, 2. It does not matter right, because k can go as large as n. The upper, but in theory right when n is very large. So, therefore, there is no, usually there is no conflict in writing it like this.

This k, you know in theory of course, k cannot go larger than n, but n is very large anyhow right. So you just simply put 0, 1, 2. So, this pmf parameter, this is a what? This is Poisson lambda. These probabilities are taken from the Poisson lambda pmf where lambda is exactly given by np. Have you people seen the proof of this?

Student: yes.

So, I, we can skip that right. So, please go and I, you can look at. The demonstration of this is a very simple thing right. It just comes from basic exponential right. The properties of the exponential the numbers themselves right e power;

Student: (Refer Time: 25:20)

Yes.

Student: sir we are not seeing the board.

You are seeing it now. Is it ok?

Student: yes sir.

So, where does the property come in? Let me just say 1 small thing about it right. How do I, for example, show that p x of 0 or p of 0 converges. Take that simplest of a, the, that is the first birth probability itself. Eg; what is p binomial of 0 with np, it to be. What is it exactly equal to p. Let us come back to right. I hope that I mean I will write it and move away. So, you can see what is happening right.

Student: 1 minus p whole to the power n.

1 minus p whole to the power n. Now I want to, now bring a lambda right into this so what do I do. I write this as 1 minus, what is p;

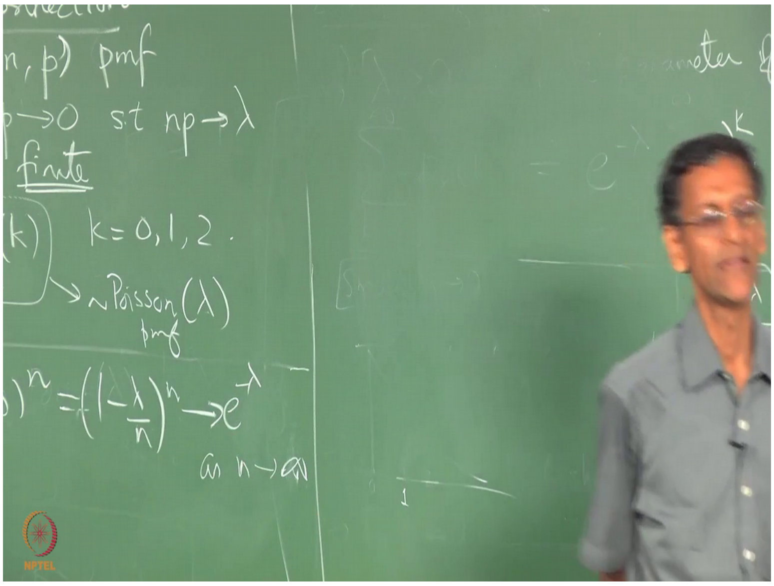
Student: lambda by n

Lambda by n power n. What is this? What happens to this is n goes to infinity, but fixed lambda.

Student: it goes to e (Refer Time: 26:36)

It goes to e power minus lambda right.

(Refer Slide Time: 26:45)



This tends to e to the e power minus lambda as n goes to infinity.

So, 1 minus lambda by n whole power n is basically why you get this convergence. And then you can prove it for or you can show it show the convergence for all other n also o[r] or other k right. So, the n goes away, the p goes away, all you are left with this lambda. So, it is like a point wise convergence right. When I say this converges to this, all the probabilities magically or whatever not the mathematical agency you shouldnt use your ma[ths] magically they are. Mathematically they arrange themselves so that the Poisson distribution x very closely describes this. What is the advantage. The Poisson distribution this is only 1 parameter lambda. It does not care about n. The n is not very important right. Because n is a large number and it can be 1,000 or 10,000 does not matter at all. And anyway those probabilities are very small. So, again here we are going to depending on the value of lambda, the probabilities will concentrate in some and lambda is typically not that large right in this in this situation right.

So I think you have seen right. So, I can move on here.

Thank you.