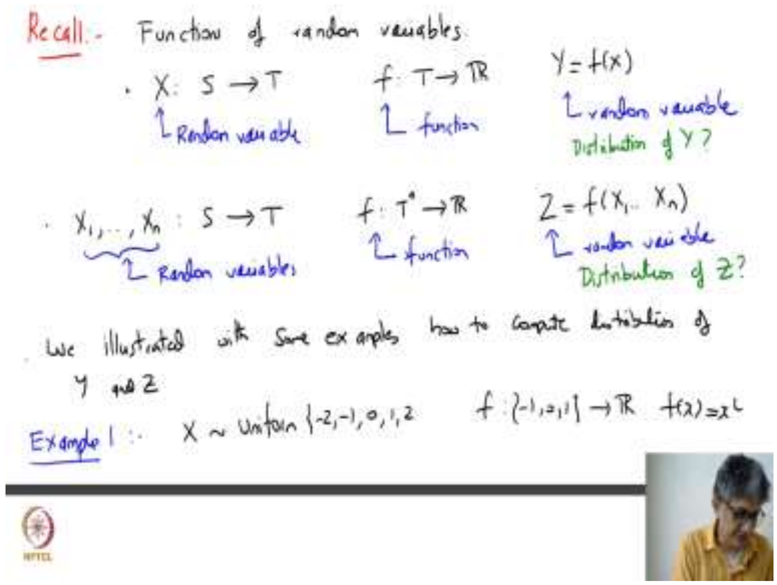
Introduction to Probability with examples using R
Professor. Siva Athreya
Theoretical Statistics and Mathematical
Indian Statistical Institute, Bangalore
Lecture No. 19
Sums of Independent Random Variables

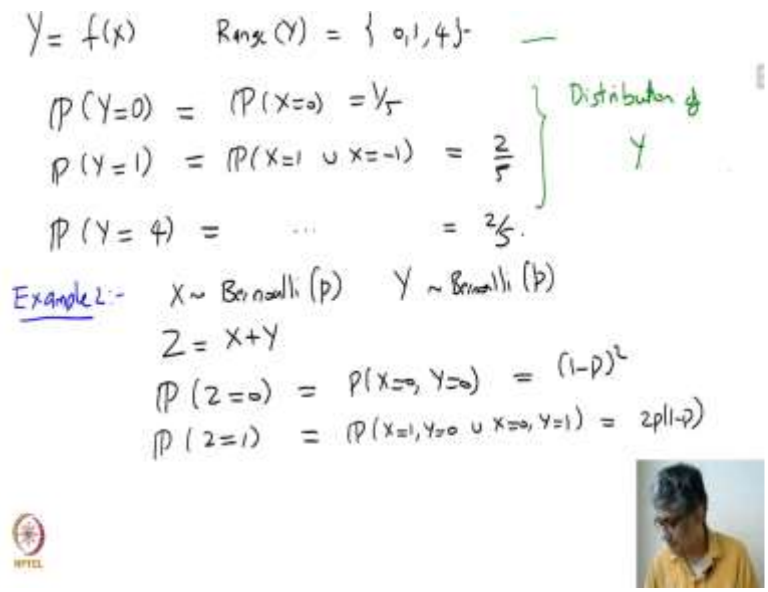
(Refer Slide Time: 00:18)



So, let me start again. So, we were doing this idea that, last time about functios random variables. I said a function a random variable could be X from S to T. And f is a function from T to R and Y is a function of X. So, X is a random variable.Y also a general random variable given that f is a nice function and we were interested in understanding the distribution of Y.

Then the second topic was they have had function of n random variables. I could take a function f from Tn to R. I could understand the random variable Z which is a function of X1 through Xn. X1 through Xn. So, Z also is a well identified random variable and depending on how nice the function f is, then one can also try to understand the distribution of Z. So, we illustrate this with some examples, how to compute the distribution of Y and Z. So, let me do example 1 was when X was uniform minus1, 2 and 0, 2. That means X took values minus 2, minus 1, 0, 1, 2 equally likely manner that the chance of X assuming any value was one-fifth.

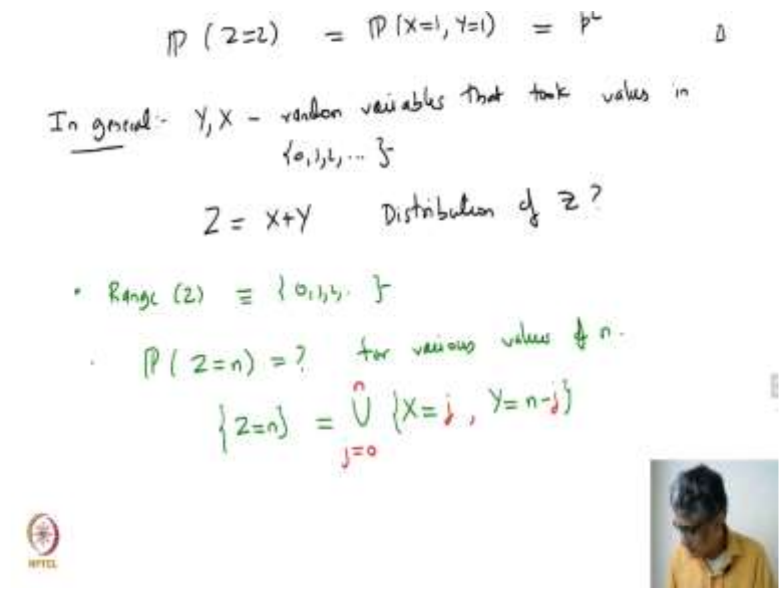
(Refer Slide Time: 01:49)



And if you took the function f of x equal to x squared, then we found out that the range of Y was 0 to 1 and 4; 3 values. And then we showed that the probability of Y equal to 0 was same as X equal to 0 was one-fifth. If Y was 1, then X could have been 1 or minus 1. And that gave you two-fifths and if Y was 4, then X could take the value minus 2 or 2 and that would give you again two-fifths. And that sort of computed the distribution of Y. So, this is just. This both these things the range of Y and the and the value with which takes the probabilities constitute the distribution of Y.

Similarly we did this other calculation when I said if Z is equal to X plus Y, how do I find the distribution of Z. So, one was I took Z equal to 0, that can only happen if X is 0 or Y is 0. And the other thing is that Z could be equal to 1. In this case what would happen? This case X could be 0 or X could be 1 or the other way around. So, you could decompose the event as X is 0, X is 1, Y is 0 or X is 0, Y is 1 and this came out to be 2 p times 1 minus p.

(Refer Slide Time: 03:11)



And similarly the chance that Z was equal to 2, was given by X equal to 1 and and Y equal to 1 That is the only way it can occur and that happens with chance p square. And we use independence of X and Y. So, X and Y were independent to do this. This is two examples we. So, what is happening here, I just want to sort of generalize this idea. So, in the two pre-examples, what it was they were known quantities in subsets. They were uniform, we understood it as an experiment and when X plus Y were two random variables. But in general, how would I do this.

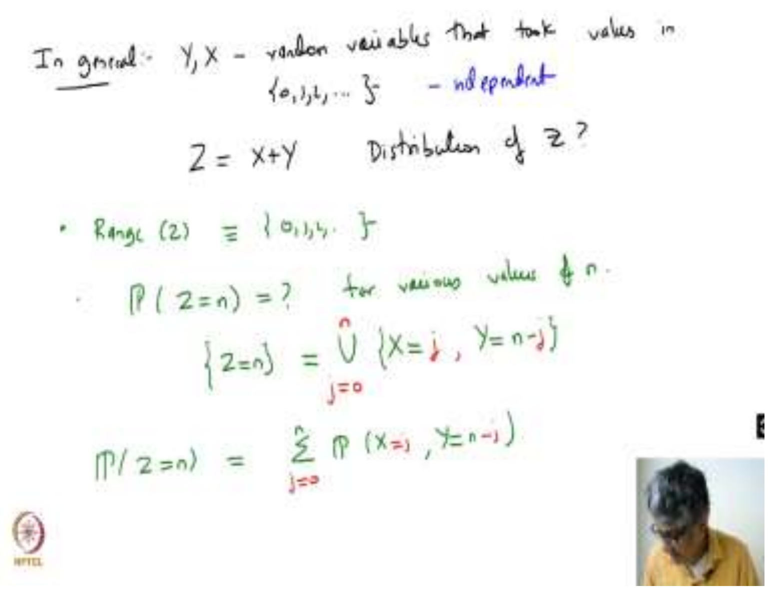
So, in general if X was a random variable, let us say X random variable that took values on the real, on that is your natural numbers and union 0. Let us say 0, 1, 2 and so on so forth. And let us say even Y did the same thing. So, X and Y, Y comma X are random variables, that took values on in 0,1 up to n and I had Z is equal to X plus Y.

In general, how would I try to find the distribution of Z. So, how do I find the distribution of Z. So, let us try and observe what we did in the previous example, in example 2. So, the idea is, the key idea is the following. So ,what you do is first thing, your first observation is, you see that range of Z is also the same as 0,1,2 up to 1. So, that is an observation you can easily make , because you take two non- negative numbers, you add them up, you will always land up in the non-negative number.

So, then all you have to do is understand how to compute the probability of Z is equal to number n. For is what for various values of n. This is what you would understand. So, how does one do this. So, one do one does this by decomposing the event Z equal to n. So, you take the event Z equal to n. This is an event. Now how can Z be n? Z can be n , if X can be, let us say a number, number j and Y correspondingly has to be the number n minus j. For the sum to be n.

Let us go back here, let us say n and then all you have to do is you have to take this thing and then, let me use curly brackets for events. So, I have z equal to n is the same as this amd that is just the union of j equal to 0 to n so that will, that is exactly the way in which Z can achieve the value n, when X is j and Y is n minus j. And note that n minus j is always non-negative and j is non-negative. So, it is well-defined events and these are the only ways in which Z can be equal to n .

(Refer Slide Time: 06:45)

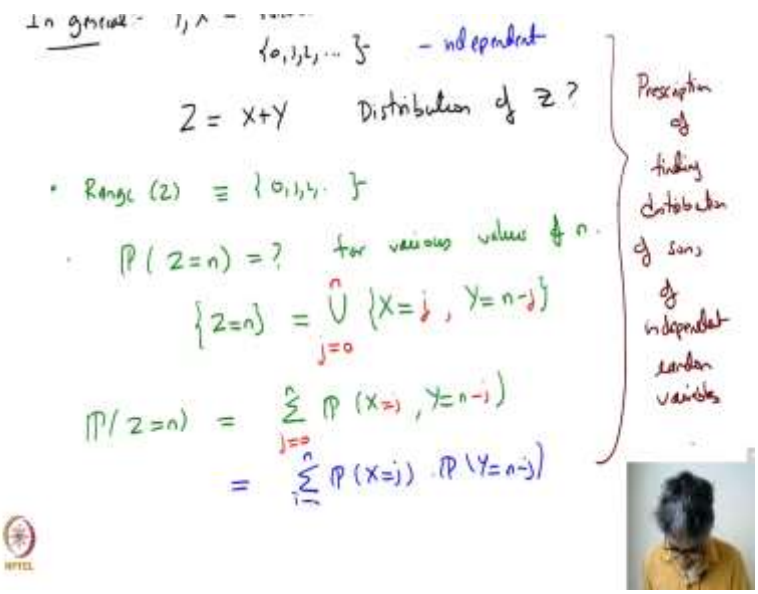


And of course now the observation is that they are all disjoint events, because if X is j and Y is n minus j for one j it cannot be anything else. So, then you can use the axioms of probability to show that the chance that, let me go back to my green. The chance that Z is equal to n, Z is equal to n is the same as the probability of this person but they all disjoint. So, that is the same as let me write down. That is the same as they are all disjoint events.

So, same as summation j equal to 0 to n. The chance that X is equal to j and Y is equal to n minus j. If I want to use red, it means. Sort of earse this j part. So, let me go back, j equal to 0 to

n. And a X equal to j and Y is equal to n minus j. And in addition, if we were given that X and Y were independent. Let us go ahead, in addition if they were given they are independent.

(Refer Slide Time: 07:59)

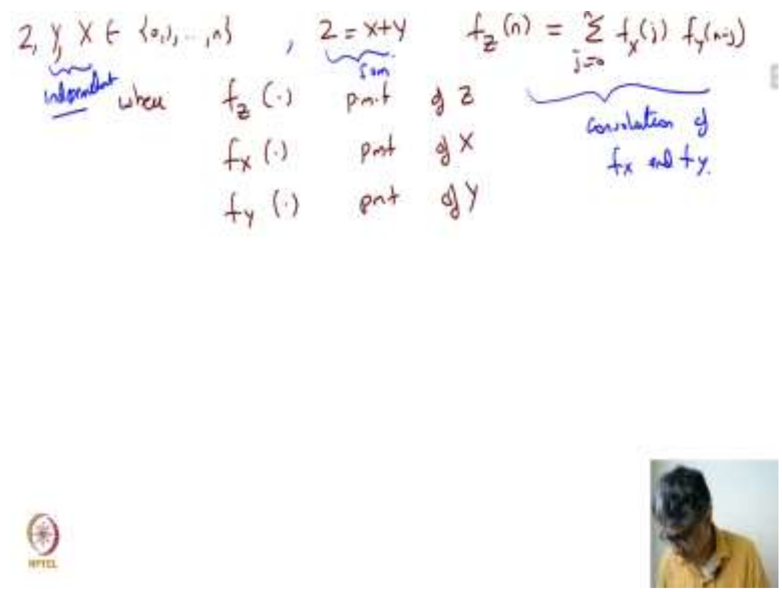


Let us write that in blue. In addition, if X and Y are independent as well; then what would happen, then you would come here and you would say fine. Now that is the same as the sum from j equal to 0 to n, the chance that X is equal to j times the chance that Y equal to n minus j. So, this step you can do. So, this would be one comprehensive way of understanding the distribution of Z.

So, if you have two random variables that took values in the space 0,1,2 all the way uptil n and they were independent, this is one prescription of finding distribution of Z. So, let me write this down. Let me write this down a little bit maroon color so it is clear. So, this is one prescription of finding. So, this is a prescription of finding distribution of sums of independent random variables.

That is one idea we have. This is one thing I already said. So, let us try and let us try and see if I can do this for one example, but I want to just spend a couple of seconds on this slide. So, this particular step is called this particular step is referred to as the convolution of the two functions or the two mass functions.

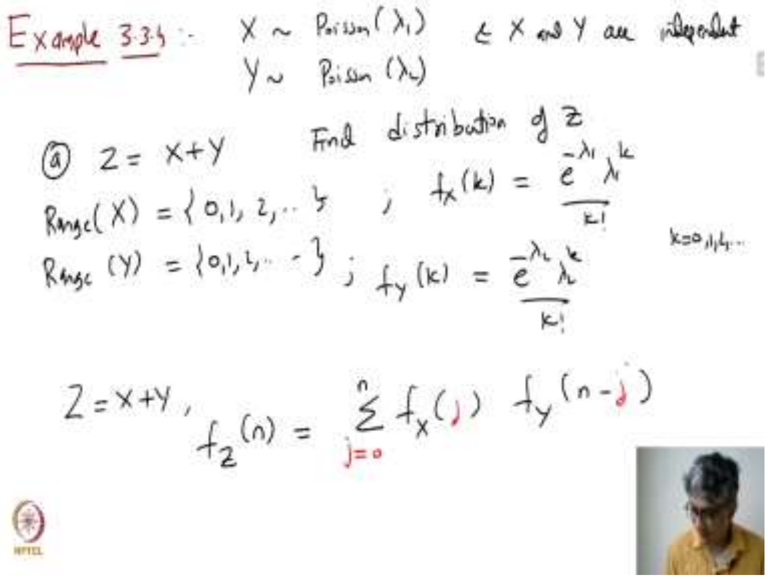
(Refer Slide Time: 09:50)



So, let me, let me maybe, I will just illustrate that a little bit more. So, in some sense what is happening is that, what we observed here is that, what we observed is the following that here X takes values in the fact X and Y, belong to 0,1 up to n; so does Z and then Z was equal to X plus Y. And then we found out that the probability mass function of Z of the point n, was the sum from j equal to 0 to n, of f sub x of j times, f sub y of n minus j; where f sub z was the probability mass function of Z and f sub x was the probability mass function of X and f sub y for the probability mass function of Y.

So, this, you can understand that the, if you have if and of course these two independent was crucial, independent. So, if you have two independent random variables and take the sum of them then the distribution of the new guy is going to be given by the convolution. So, this is called the convolution of fx and fy. So, sums of independent random variables, the probability mass function is just given by the convolution of the two probability mass functions. So, this fact what I have shown in these examples. Very good very nice. So, let me go to the next page. Let me try and do another example. Let me do a example, example lets say, example is a.

(Refer Slide Time: 12:05)



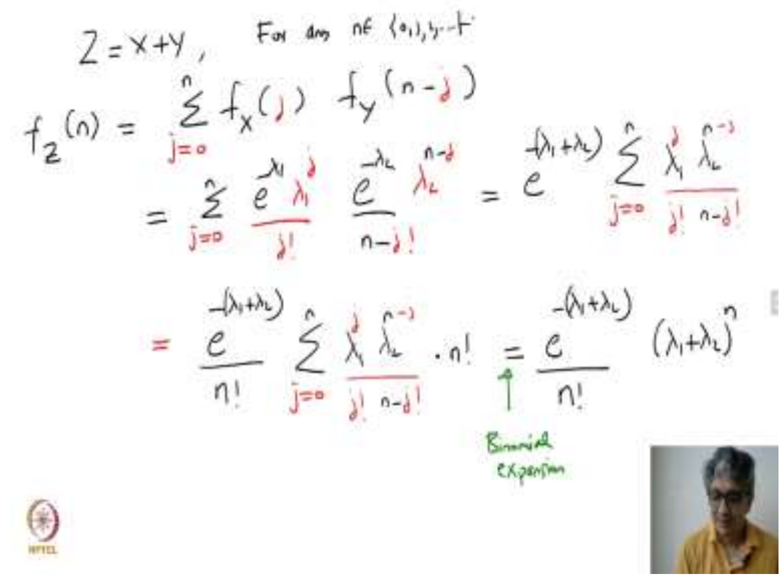
So, the example, 3.3.4 in the book. It is a good example. It is an illustrative example of how to do the convolution ,but let us just do it step by step. So, let me just let us take X to be Poisson with parameter Lambda 1. And let us say Y is Poisson with parameter Lambda 2. So, and let us say X and Y are independent. So, now I let Z equal to X plus Y. Let us do that first computation. Let us try and find the distribution of Z. Let us find the distribution of Z. So, how would one do this. So, again you do the same thing like you know.

So, X takes values , the range of X we already know is 0,1,2 and so on and so forth, that is Poisson. And we also know range of Y is 0,1 and 2 and so on so forth. And we also know that the probability mass function of X is given by e to the minus Lambda 1. Lambda 1 to the power k by k factorial and the probability mass function of Y is given by e the minus Lambda 2, Lambda 2 to the power k by k factorial. This for k equal to 0,1,2 so on and so forth.

Let me just maybe make the whole thing a little bit below so its same thing for both. So, this we know. So, 03 for the previous example previous calculation I know, that if Z is equal to X plus Y, from the previous page I showed this. I showed this if Z is X plus Y the distribution of Z is given by the convolution of the two probability mass. It means I know that f sub Z of n is going to be given by the sum from j equal to 0 to n. Let me write the color and I am writing a little bit red.

Let us see that I understand what is going on. Let me see j equal to 0 to n, go to n. So, let us say n is fixed and that is again in black and I have f sub x at the point j, f sub y at the point n minus j . So, let me write that in j. So, it is easier.

(Refer Slide Time: 15:07)



That is what it is. So, now I know the expressions for f sub x and f sub y. So, what will I get. I will just get, this is let me move to the side, so I can start again, nope. This will come here, we will erase this. Very nice. I will come back here, very good. So, this is the same as, so now I have to just plug in for f sub x and f sub j. So, let me write it down so because summation from j equal to 0 to n, I have that. I have f sub x at j this will be j equal to 0 to n. f sub x at j so what will that be. Either the minus Lambda 1. Lambda 1 to the power j over j factorial, I will have that.

And here I will have again, let me write that in a little black so it is a little bit distinguishable. So, e to the minus Lambda 2 over n minus j factorial and then I have Lambda 2 to the power j, n minus j. So, let me write this also in black, so all the terms that do not depend on j will be black. So, e to the minus Lambda 1. Lambda 1 to the power j. So, this is again, I will write this in red, so its clear. So, I have lambda 2 to the power n minus j, n minus j and then n minus j factorial. This is what have happened. This is what exactly happened.

But now you can observe that what all terms do not depend on j. So, let us see so I will, either minus Lambda 1 will come out. Lambda 2 will come out and (())(16.56). So, either Lambda minus 1 plus Lambda 2 at the top, that will come out and then I have the sum from j equal to 0 to n, j equal to 0 to n, I will have that.

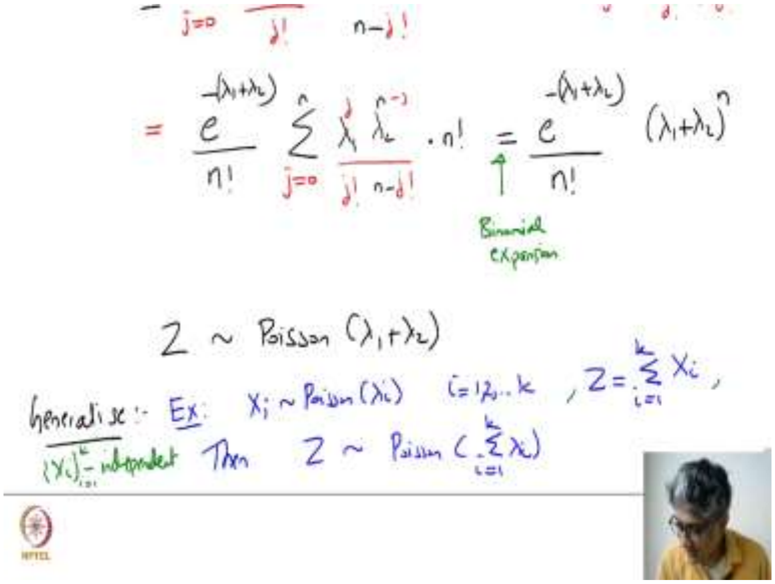
And what is inside what is left is Lambda 1 to the power j. Let me add that in, j and then I have Lambda 2 to the power n minus j. So, I will leave that as it is. I will not take that out, so I will leave that as n minus j and then I have divided by, I have j factorial and then I have a n minus j. This is what I have .

So, now we are all in familiar territory now. This looks like a sum of two things you can understand the expansion little bit. So, what I do. I do a little trick, I just write this as; e to the minus Lambda 1 plus Lambda 2. That is at the top I leave it like this. I multiply and divide by an n factorial, so the n factorial outside the sum, then I take this sum that I have here. I am just making a mess of this. Very nice. I have to copy this and bring it here. Then I have multiplied an n factorial there.

So, I have multiplied by an n factorial inside as well, so nothing much changes, but then this person here is just like the binomial expansion. So, we will keep either minus lambda 1 plus lambda 2 outside, the whole thing divided by n factorial and if you just stare at a little bit, this is just the fact of Lambda 1 plus Lambda 2 to the power n. That is just a fact using binomial expansion.This is binomial expansion.In this step that is all I have used. So, that is a very nice crucial way of understanding it.

But now we are here, so that means the chance of Z equal to n, is given by Lambda 1 plus Lambda 2 to power n by n factorial. So, this is true for any n for any n in, let us say 0,1,2 up to 4 I have f z of n equal to Lambda 1 plus Lambda 2 power n into e to the power minus Lambda (())(19:22) divide by and factorial.

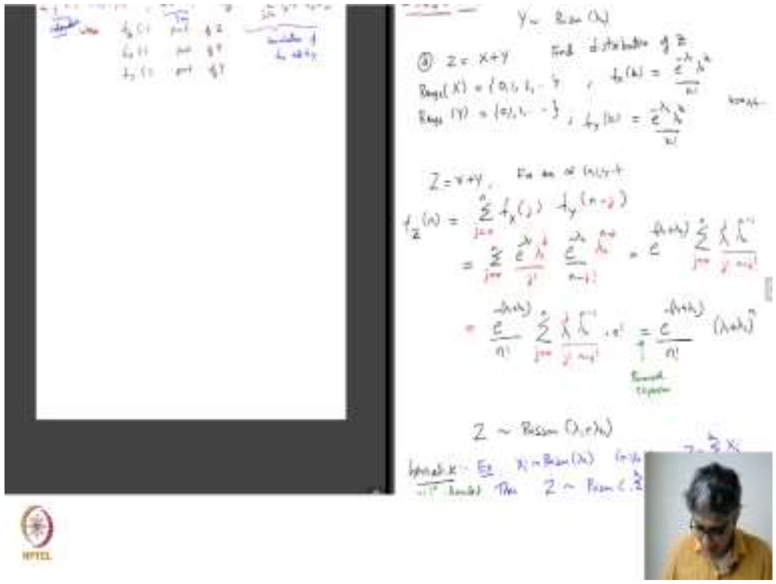
(Refer Slide Time: 19:26)



So, what does this mean. This just means that Z is distributed as Poisson of Lambda 1 plus Lambda 2. We have shown an interesting fact that, if X and Y are independent random variables; each of them is Poisson Lambda 1 and Lambda 2. The sum of poisson is again Poisson. So, one can also generalize the calculation. One can generalize quite easily. This is an exercise 1. I will leave it as exercise that if X1, Xi are all Poisson Lambda i, i equal to 1 up to k and I had Z is the summation i equal to 1 to k that is the Xi; then it just adds up. Then the distribution of Z is just going to be Poisson, of the sum from i equal to 1 to k of lambda i.

Of course I need Xi the crucial factor this for Poisson, Z is equal to this and I also have to take X i's are independent. So, this collection is independent. So, if I have a collection of independent random variables they add upto (())(21:01).

(Refer Slide Time: 21:03)

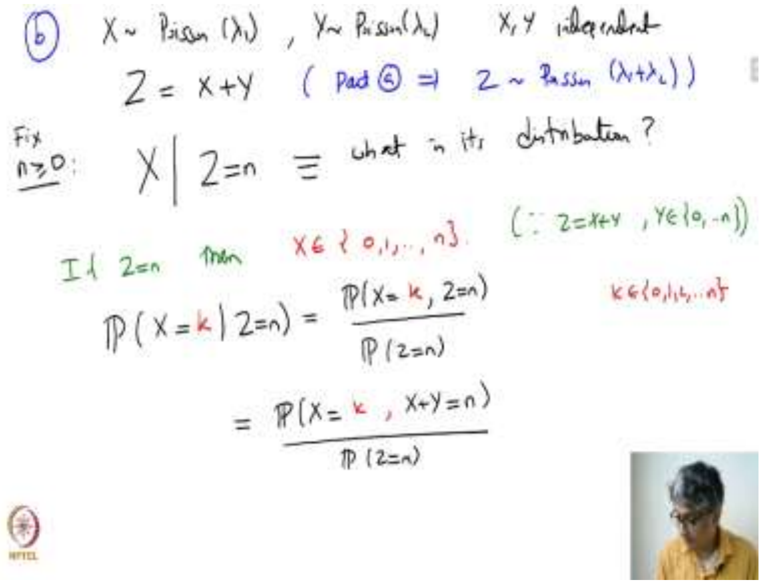


So, let us just recap what we did so far. Let me see if I can get that split view for you. Very nice. Let us go here, here again, split view. So, let me make this a little smaller. You put the previous page spread this page here it is a little bit smaller and we can recap a little bit. Not big, not smaller, I want smaller. So, the idea was the following, if I had independent random variables X and Y, I sum them, I called 1 over Z, then the probability mass function became the convolution of the two individual mass functions and that we implemented in this example of Poisson.

So, we took two Poisson Lambda 1 and Lambda 2, both independent.You took Z equal to X plus Y and then I found out that the random variable Z had the following distributions. You could do Z equal X plus Y here and then what you would get is that Z was Poison Lambda 1 plus Lambda 2. And one can generalize this if you have k random variables, all Poisson and they are mutually independent, then the following factors is true.

This is one simple calculation of doing this(())(22:31). Now suppose I do the I do a calculation like I want to do before. Let me go back to my whole screen. Very good, next page, very nice. Now I ask you a different question. Let us say I ask you a different question, I say, again I this I am in the same example. Same example, let us say I am going to part b.

(Refer Slide Time: 22:57)



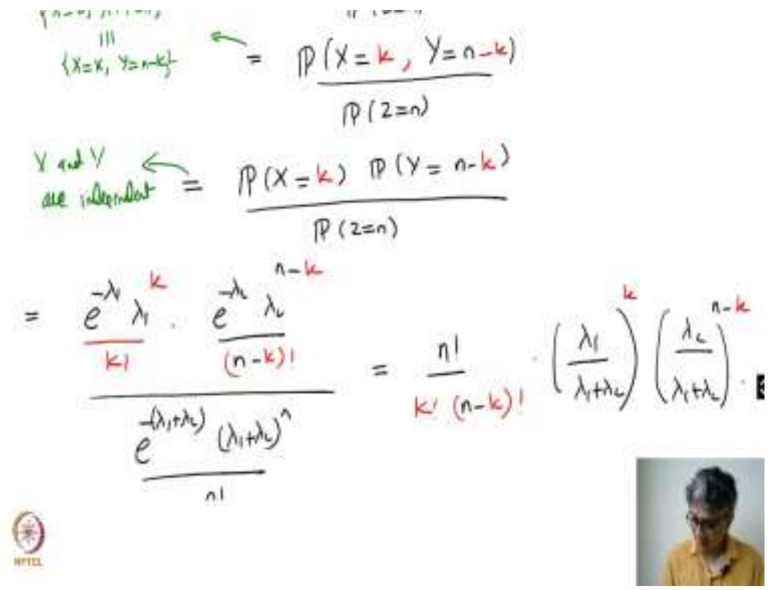
I will go back to my blue. Part b of the same example and I ask you the following question. So, X is Poisson Lambda 1. Y is a Poisson Lambda 2. X and Y are independent. X and Y are independent and I have Z is equal to X plus Y. This I know how to compute the distribution of Z from part a. So, from part a, part a implies that Z is Poisson of Lambda 1 plus Lambda 2. Now this I want to ask you a question from an experiment. Let us say I am given the event Z is equal to n. So, I am gonna ask you what can X be. So, X given is Z equal to n, what is the distribution of Z? So, let us say n is equal to greater than or equal to 0. I give the (())(24:09). What is the distribution of X?

So, what is the conditional distribution of X given Z. A little tricky one let us see how to do this. We know how to do Z equal to n. We know how to do X , treat that. Let us see so now one thing to observe is the following if, so the first observation is the following, so how do I find it out. The first observation is that if Z is equal to n, then X is Z is sum of X plus Y. Y is non-negative, then X can only take the values, X has to belong to 0,1 up to n. X can not take values more than that because this because, this is because Z is equal to X plus Y and Y is also in already in 0,1 up to n. That is good.

Let us get back to this, very nice. This is where we are. So, now, how do I do this now. Let us see, so now, say suppose I want to calculate the chance that I am given an n. n is fixed. So then the chance that X is equal to k given Z is equal to n. This is where am I. Let me go back to my old notation of making k as an X, so if you understand what k looks like. So, the k is in the set. So, k is in 0,1 up to n.

So, now how this is going to be. This is going to be equal to. So, how do you do this. If the chance that Z is equal to, I will add at the bottom, by definition. At the top you will have X equal to k and Z equal to n. This is what we have by definition of conditional probability. But now we know that that is the same as X equal to k and here you will have, what is that Z, Z is X plus Y. X plus Y is n the whole thing divided by the chances Z equal to n.

(Refer Slide Time: 26:43)



So, now how does one finish the problem off. So, the bottom we know how to do, because it is just Poisson Lambda 1 plus Lambda 2. The top how do you do, you look at it this way. So, Z is equal to n, the top you do a little trick. So, you know X is k, so you know X is k. So, that is easy to do. So, your X is k, that is k. Then what you do is you interpret the event X plus Y equal to n, because X is k now. So, I can replace the value of X by k above.

So, I will get Y is equal to n minus k. Y is equal to n minus k. Let me just do this in black, so that is clear. So, y is n minus k, that means all I am doing is that inside this event this whole thing it immense, this event all I have used the fact is that, if X equal to k and Y is equal to X plus Y is equal to n; that is the same event as X equal to k and Y equal to n minus k. You have two same events. That is what I have used here.

So, now I am in business now. I know how to do this now, because now I can use the independence of X and Y. So, I have X X equal to k. Our probability Y equal to n minus k and the whole thing divided by probability Z equal to n. k and Y equal to n minus k. So, let me do one small thing let me erase this light and in green. So, I will use observations in green. So, in this step what have I done, I have used the fact that the event X equal to k and Y equal to n, Y, X plus Y equal to n; is the same as, the same event as X equal to k and Y equal to n minus k.And here the, here I will use the fact that they are independent. Very nice.

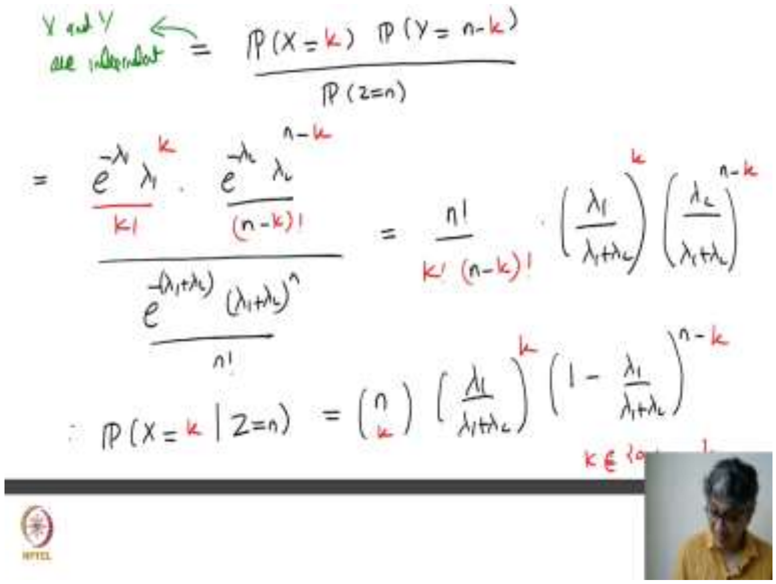
So, once I have these two facts, I have the following idea; that I know the answers now. So, this guy is going to be equal to e to the minus Lambda 1. Lambda 1 to the power k . I will highlight k. Whole thing divided by k factorial , that is the top. The next term is e to the minus Lambda 2. Lambda 2 to the power n minus k, the whole thing divided by n minus k factorial.

And then here the whole thing divided by Z equal to n, which you know is poisson Lambda 1 plus Lambda 2. So, e to the minus Lambda 1 plus Lambda 2. Lambda 1 plus Lambda 2 to the power n by n factorial. So, I do this whole jugglery I do up and down. So, I will get what n factorial at the top by k factorial n minus k factorial.

And then I have all the e's will cancel off. So and I, only Lambda’s will remain. So, I will get Lambda 1. So, may be I will do it next line. So, this is too short (())(30:07). So, maybe I will move it across. So, let us see this whole thing . So, this whole thing will move it across. This is equal to this and this is equal to like I said n factorial by n minus. So, by, let me write this in red by k factorial and then I have a n minus k factorial and then I have what do I have, I have Lambda 1 by Lambda 1 plus Lambda 2 to the power n minus k.

So, where did I start, I start off with this distribution X is Poisson Lambda 1, Y is Poisson Lambda 2. They are both independent. I looked at the fact I want to understand the fact that X is given the fact Z is equal to n. I want to understand as distribution and I came up with this formula that probability X equal to k given Z equal to n after this computation gave me this answer. This way.

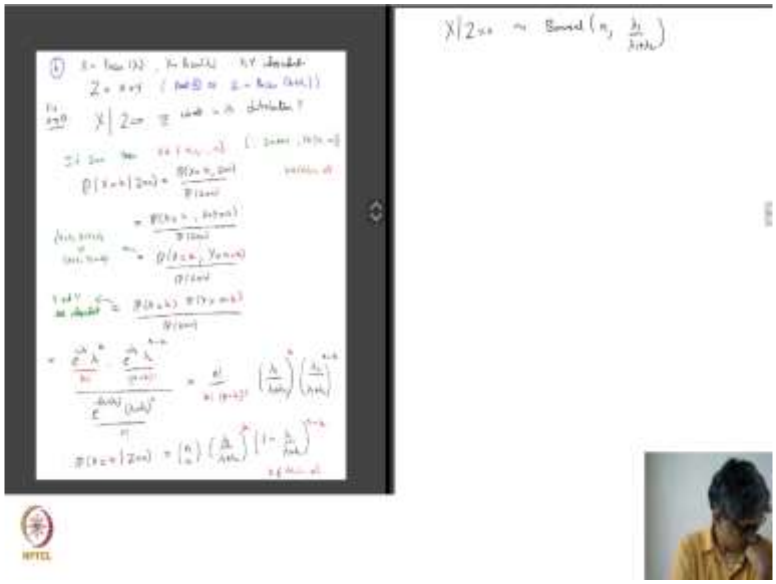
(Refer Slide Time: 31:34)



But this if you stare a little bit what do you get, this is the same as, how do you stare it, so now So, what have we shown, so we have shown, so we have shown, the chance that so therefore the chance that X is equal to k given Z is equal to n, Z equal to n is in black. about that. (())(32:00) Given Z is equal to n is the same as n choose k, n choose k; in my notation. And this I will write little differently. I will write it as Lambda 1 by Lambda 1 plus Lambda 2 to the power k and the inside the other part I will write it as 1 minus Lambda 1 by Lambda 1 plus Lambda 2.

So let me erase that again properly Lambda 1 plus Lambda 2, the whole to the power n minus k. So, what have, what is, what happens here. So, if you and this is all valid for k in 0,1 up to n. So, what is the punch line here. Let us tell it a little bit if X is poisson Lambda 1. Y is poisson Lambda 2. Z is X plus Y, then X given Z is this.

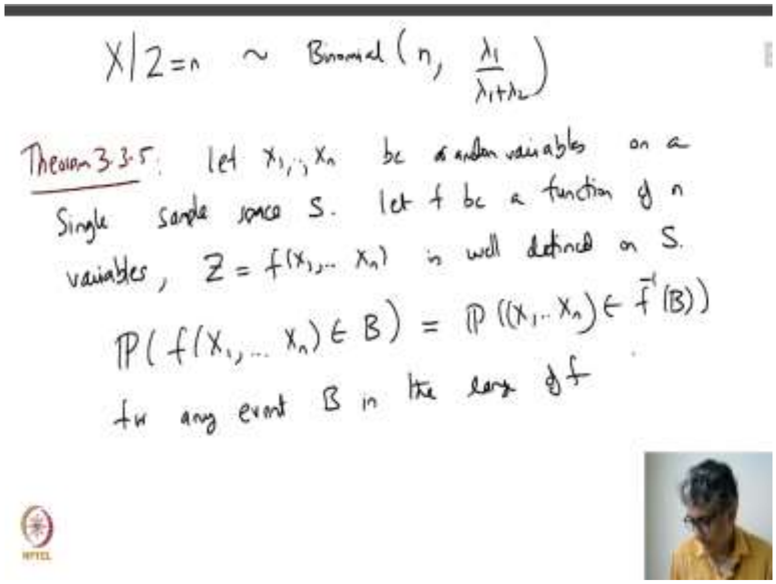
(Refer Slide Time:33:06)



So, what have we shown, we have shown a very interesting fact that X given Z equal to n is just binomial with Lambda 1 by Lambda 1 plus Lambda 2 as the p and (())(33:24). n comma Lambda 1 by Lambda 1 plus Lamda 2. So, let me just do a split view and show you what the whole calculation is, so you can, you can enjoy the calculation a little bit. So, the previous page, let me go here back to the small one smaller. This is what we have to show. It is a nice idea. So, you have two Poisson independent, you conditional the sum then the individual will become a binomial.

It is an interesting idea (())(34:09). Intitutively you can think about it. There is sort of an explanation for poisson, obedient by an approximation of, it came as an approximation to the binomial. But there is sort of an interesting idea of poisson in terms of a queuing model. Then this interpretation will become even more clear but as of now we will leave it like this, the calculation. So, let me just close off with the theorem and then I will, this ideas are widely applicable everywhere. Let me go back to the normal view. Next page you will know what is happened why is it is not coming.

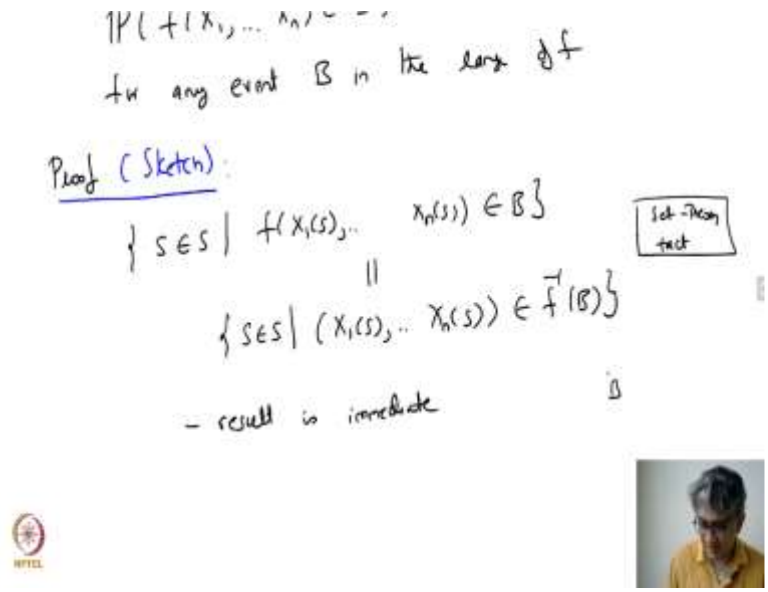
(Refer Slide Time: 35:08)



So, let me close off with the theorem, this class. So, the idea is that so this theorem is called Theorem 3.3.5 in the book and the idea is that let X1, X2, Xn be few random variables on a, let us see. For the time being I think this is the simplest way to do it on a single sample space S. Sample space S. Let f be a function from a function of n variables. Variables are on such that Z equal to f of X1, up to Xn is very defined. Well defined on S.

So, then the probability so the key thing we use in all the examples was we tried to understand the probability that f of X1 through Xn, is in an event b on the real line; that was the same as the chance that X1 through Xn was an event f inverse B. That is exactly what we use in this whole example. This is for any b which is subset and B an event in the range of f. Any event B in the range of f.

(Refer Slide Time: 37:13)



I will let you prove it yourself but I will just give you a sketch of the proof. Sketch, the key idea is that which you are using the examples as well which is the following idea that a sketch of the main part. The key idea we used was that the event, the set of all S in S such that f at the point X1 of s, f xn of s, Xn of S is in the event B. This if you use the definitional function, this event is the same as the set of S and S such that X1 of s, Xn of s is in the event f inverse.

So, this is a set theoretic inequality. So, once you understand this, then the problem the theorem is immediate. So, once the set theoretic notion, set theory fact. Once you understand this, then the result is immediate and this is what we have kind of used in the examples.