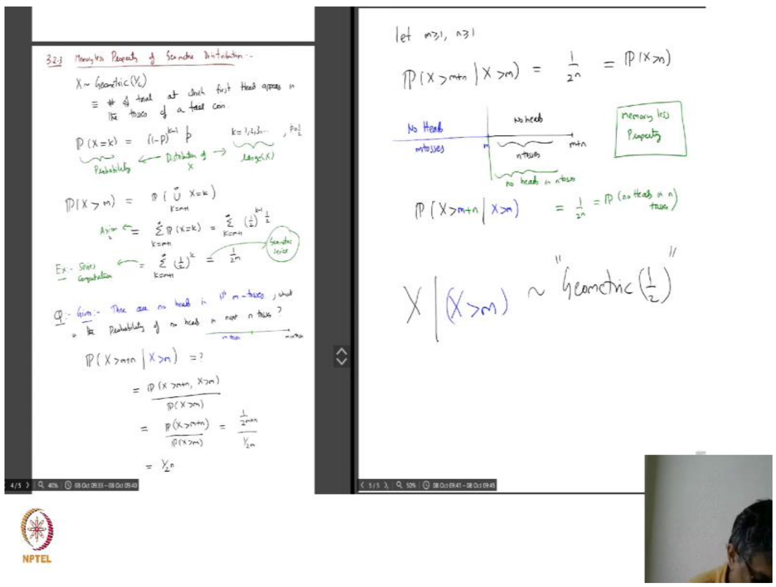
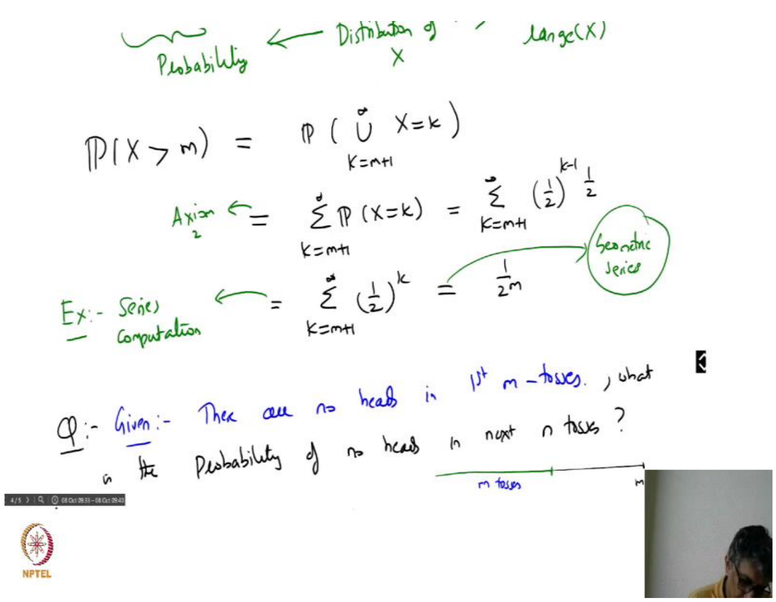
Introduction to Probability-With examples using R
Professor Siva Athreya
Theoretical Statistics and Mathematics Division
Indian Statistical Institute, Bangalore
Lecture - 18
Functions of Random Variables

(Refer Slide Time: 00:18)



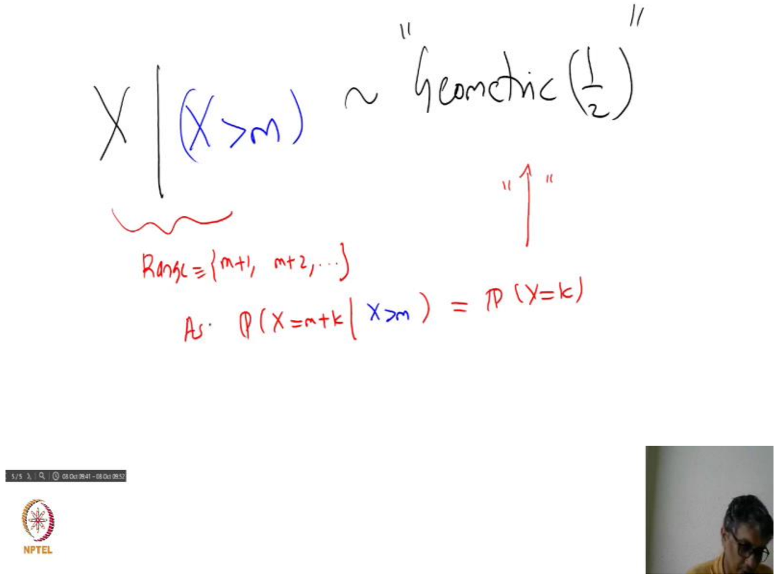
Of X. Okay here. Let me go back to my original view.

(Refer Slide Time: 00:31)



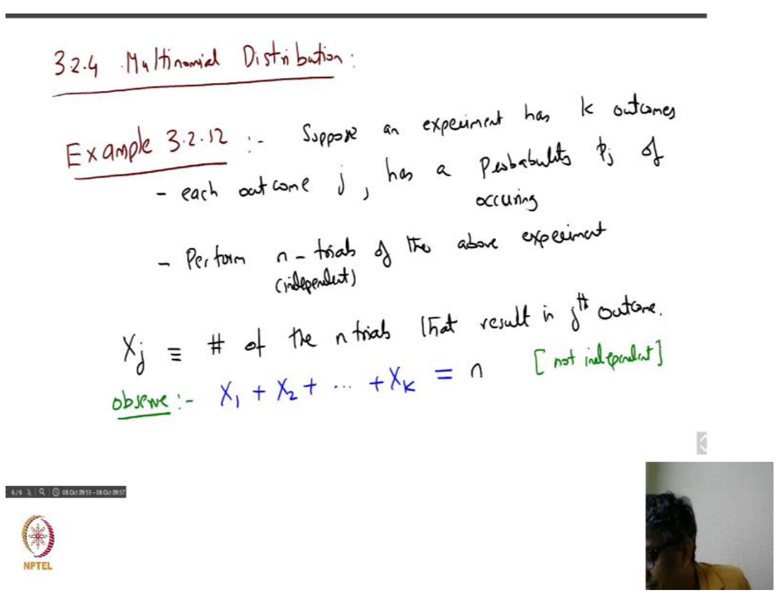
This is the geometric memory less property that we discussed so far. So, just to sort of be pedantic just make sure, that just to understand this quotation remark in the following.

(Refer Slide Time: 00:36)



This takes the range of this random variable. The range is m plus n, m plus 1, m plus 2 and so on so forth. And I should understand the fact that X equal to m plus k given, X is equal to m, X is at least m. This probability was the same as the chance that X is equal to K. That is what its shows. Very good. And that is how I, I understood this quotation remark.

(Refer Slide Time: 01:43)



Very nice, so now,we do one more important example and then I will sort of move on. So, the example is kind of important. We have this, you may have seen this also in, class 12 or otherwise as well. Is the idea of multi-nominal distributions. So, lets quickly do this.

So, here are the, this is a very specific example that I would like to sort of do. So, this is example 3.2.12 in our book. So, the examples are following. So, before we did only the binomial case, there were two outcomes. So, now let us assume that there are k outcomes. Suppose an experiment has k outcomes. So, each time you, you do an experiment there are k possible outcomes. So, 1 or like so success and failure, you have many outcomes.

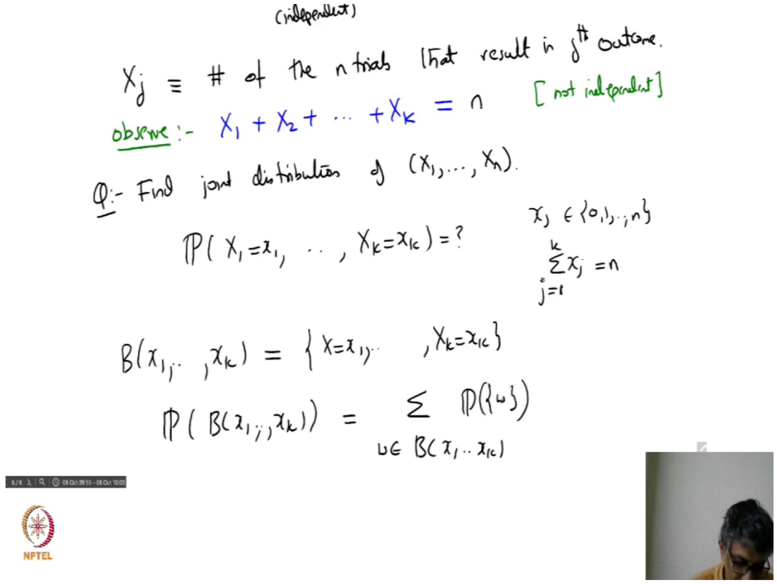
And let us say each outcome, has each outcome j, each outcome j, each outcome j has a probability Pj (())(03:15). Probablity Pj, so probability Pj of occurrence. So, each outcome has a probability. So, now I want to understand let us say how does this experiment behave. So, now this is again a place where random variables become very useful. So, if I were to write out this whole thing in, let us and we do let us say we perform n trials. We perform n trails of the above experiment and independent trails.

So, now if I do this, now if I would write down all the events and all the outcomes, in a sort of have n trials. Each trial has j outcomes possible. And write out all the sample space completely and I try to use that to compute the events and the probabilities of the events; it will be quite hard to do. So, what we do is, this is where random variable again come into use. So you let Xj, so you let Xj, be very careful, X sub j is the number of trials among the n, in which the jth outcome occurs. So number of trials, number of the the n trials, that result in the jth outcome. Is that clear.

So, that means, I am just doing this random variable. So, one thing is clear is that, so they are clearly dependent because the following idea. If I look at X1 is the number of trials in which outcome 1 appears. I look at X2 the number of trials in which the outcome 2 appears and I will go all the way uptil k and I get X sub k .

If I look at, if I add all these up; I will always get n. So, this is observation. So, they are clearly dependent, so they are clearly dependent. So they are not independent. Of course, so therefore not independent. So, the X's are not independent, because this sum adds up to n, so there is a clear relationship. So, if I know X1 through Xk minus 1, then Xk is uniquely determined. This is one important factor. Very nice.

(Refer Slide Time: 06:00)



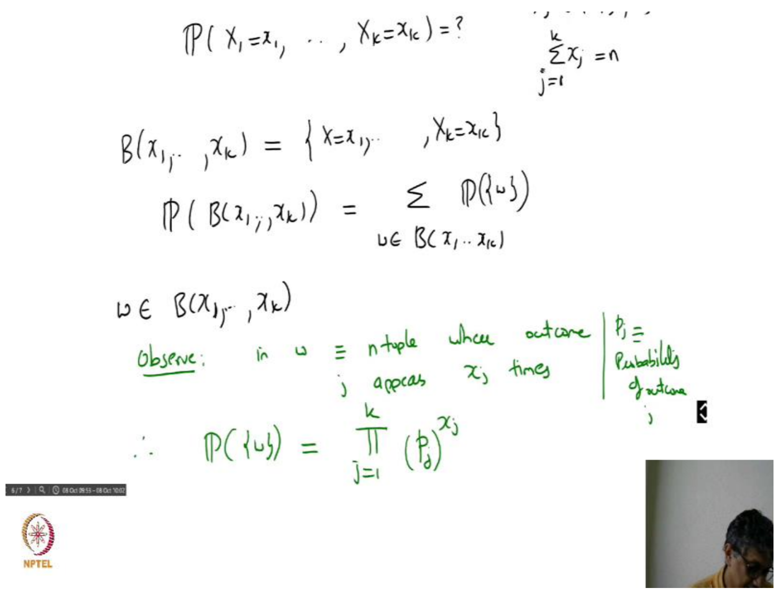
So, now how do I do this. How do I find, how do I find. So this is, this is the first example, interesting example, where we need to find. Let us say I want to find, the joint distribution of X1 to Xn, with the joint distribution of X1 to Xn. And that is the first question I have. How do I do this. So, now the, this is called the multi-nominal distribution. We will discuss that as we go.

So, how does one do this. So now if you, what do you want to do you want to find the chance that X1 is equal to x1. Xk is equal to xk sorry and Xk is equal to xk. And what happens to the Xk, each of the xk's can be, take values from 0,1 up till n. And your one condition that the sum from i equal to 1 to k of the x, little xi’s; add up to it. Because there are n trials, so maybe I should use j because I have been using j so, so I will use j here. So, I will erase this. So I have x sub j and the j outcomes can come from 0 to 1.

So, I have n. Let me just do once yeah that is right. So, how does one do this. What does one do. So, here one should be little bit careful and one should see how does one exactly compute this, this problem. I will let you think a little bit on this. So, let me see, how do you want to do this. How do you compute.

If I give you a sequence of X1 and Xk . How would you do that. So, here the idea is the following it is very easy. So, now let us look at the event. Let us look at the event, B of X1, Xk, that means the outcome 1 appeared X1 times, outcome 2 appeared X2 times. So, let us define that as the event X equal to X1, X equal to Xk. Lets define this. So, then you know that, the chance that of the event B of X1 through Xk, is the same as the sum over all omegas, that is all the sample points in the event BX1 through Xk times the and each sum you add the probability of omega (())(09:15).

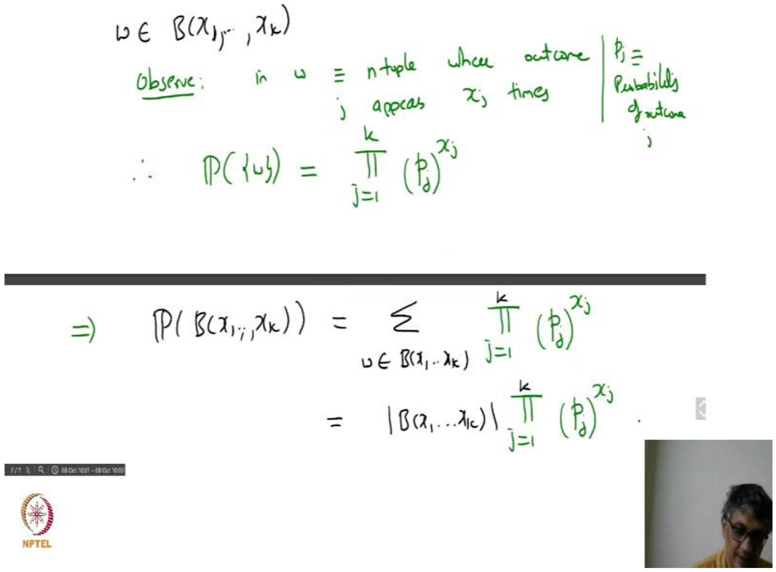
(Refer Slide Time: 09:17)



This is how you would compute the probability of B. But look at any element, look at any element in B of X1 through Xk. So, this is a sequence of n trials and look at each outcome. You want outcome 1 to appear X1 times. Outcome k to appear Xk times. So, then this is observation. So, this is observe. In omega the sequence is a, is an n tuple where outcome j appears Xj times and that is how omega looks (())(10:06).

So now if you want to compute the probability of omega, they are all independent trials and you know outcome j appears Xj times and you know outcome j comes with probability P sub j. So, this is just going to be the product of j equal to 1 to k. Outcome j occurred probability P j and yeah it occurs X j times (())(10:37).

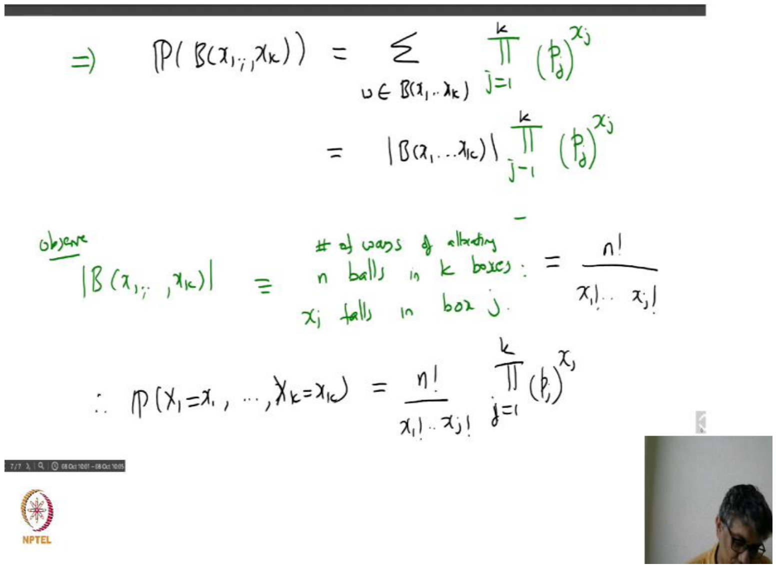
(Refer Slide Time: 10:37)



Therefore this, because this happens and Pj is the probability of outcome j. So, once you have this, then if you want to compute the probability of B. So, the same that, this will immediately imply, let me go back to black; the probability of B of x1 through xk is the same as the sum over all omegas in the event B of x1 through xk and each of them has the same probability.

So, the same probability is, is this probability. Let me just start copy this. Let me duplicate it. Put this down here for you. So, same problem. But now they are all the same probability because result all depends on what sequence it was. So, this is just going to be the same as the the size of B of x1 to xk times this problem. That is what is it going to be.

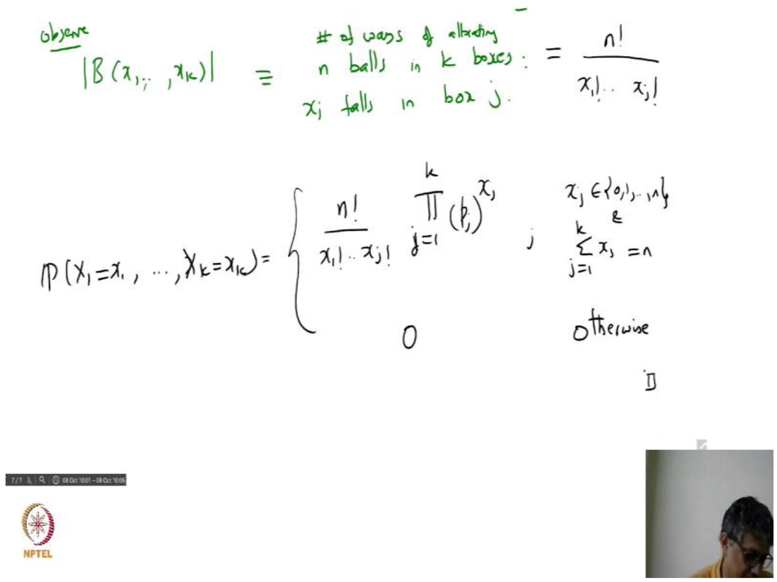
(Refer Slide Time: 12:26)



How do you find the size now. So size how do you find? So, size what you would define. You look at the number of outcomes, such that your your k boxes, your n, your sorry, your n boxes, so n balls and k boxes and you put, put the j outcomes in. So, the same as this computation let me write this down in a bit more precise manner. So this is the same, so now you observe, that B of x1 through xk, the size of this is the same as saying your n balls, your k boxes, your k boxes and then you have to do that Xj falls in, Xj of them fall in box k, box J.

And you want to know the number of ways. So, number of ways of allocating n balls in k boxes such that Xj falls in Xj of their (())(13:10). So, let me just cut and paste a little bit. Too close this whole thing. Let us bring it down, so whatever... So once I do this, this is,this we know from from either high school or whatever; it is the same as you look at n factorial at the top and you have x1factorial all the way to x j. So, therefore the joint distribution of X1 equal to X1, Xk equal to Xk. This joint distribution, is the same as you put the size of B as n factorial by X1 factorial, Xj factorial and then the product from j equal to 1 to k of p power p sub j to the power x sub j. So, that is the joint distribution.

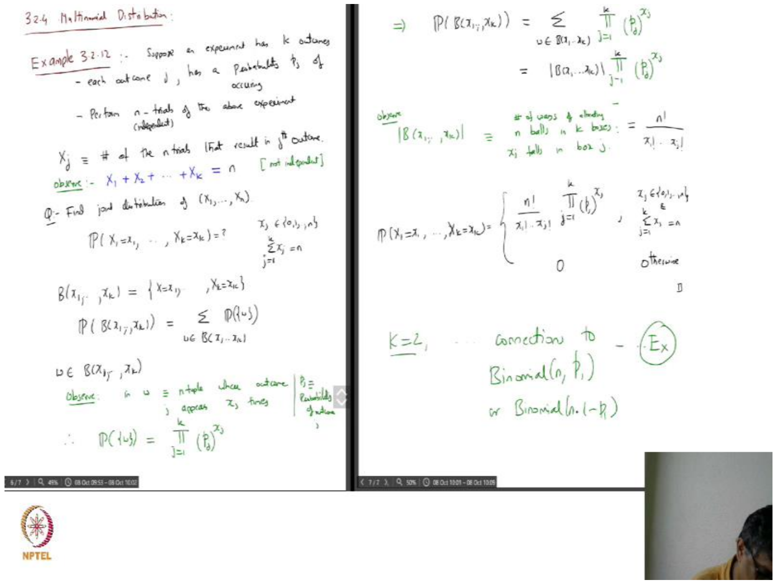
(Refer Slide Time: 14:13)



Of course I assume several things. I assume that so this will come out. So, therefore this will happen only in the two specific conditions, have to be satisfied. Let me just move this down this side, therefore this happens. Let me erase the k, this k. So, when does this happen. So, let me erase this part as well. (())(14.35) So, when does this. Let me move this down a little bit. Very good. So, let me erase this notes.

So, this is equal to this under what conditions now. This happens with each of the Xj's are between 0 and n. And we use the fact that the sum of j equal to 1 to n of the Xj's, j equal to 1 to k of j sorry;1 to k of the Xj's was equal to n. Under this condition this happened otherwise the probability is here. So, this is the kind of the joint distribution of X1 through X.

(Refer Slide Time: 15:45)



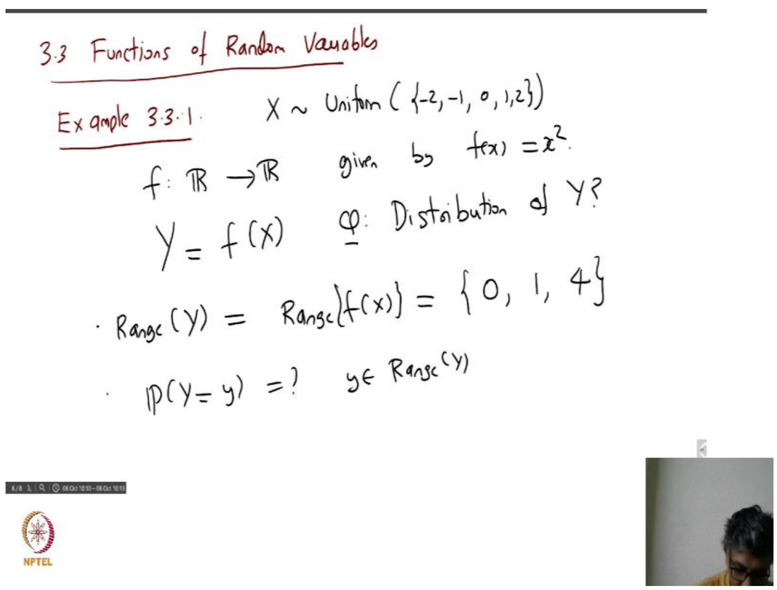
So, let me just do a split screen, so you can see everything in one shot. Just try and see if I can get that now, here and again. Very nice. I go down a little bit before and I have to go to this previous page here and here I have to go to the last page.Very nice. Very good. So this gives you a complete view of the multi-nominal distribution, that is I start off with the experiment, that have k outcomes. I perform n independent trials. I am interested in how each outcome, the type of each outcome is each trial.

So, the way I do is, I just count the number of times and outcome j appears right and then I compute the joint distribution of x1 through xk. So, even though the trials are independent, if you were to count the outcomes of each type, they become a very, very dependent sequence random variables.

And this gives rise to what is called the multi-nominal distribution and it is a very powerful tool and you can also check the following fact that; at n equal to 2, you will get, you can go back and get the binomial. So, in n equal to 2, you will count how many times 1 appears and how many times 0 appears. But if in n equal to 2, if you know 1, you know the other guy immediately. So there is no real (())(17.01).

Give me one second (())(17:07) Very good. So now I want to sort of move on to. So just one small observation, before I sort of move on, is that just check n equals sorry, k equal to 2 naught n equal to 2. I said n equal to by mistake. So just take k equal to 2, then just see connections to binomials (())(17:56). Something should just check and see, observed. I am observing both X1, X2 here. So, in binomial case if X1 plus X2 is n, if X1 is, I am only counting the number of successes means look at X1not X2. So, Binomial P1 or Binomial 1 minus P1. So this is, should just put here n and n. Very nice. Then let me go back to my full screen view (())(18:45) Very nice. Very good. We are good…

(Refer Slide Time:19:29)

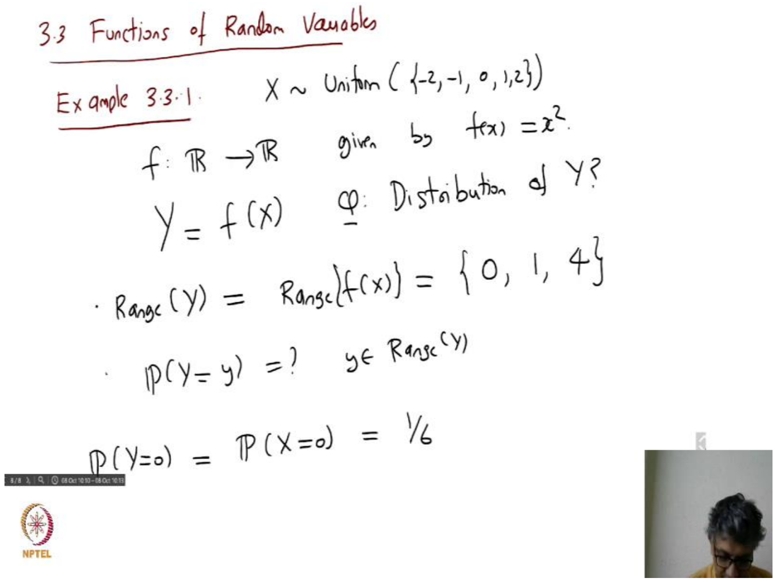


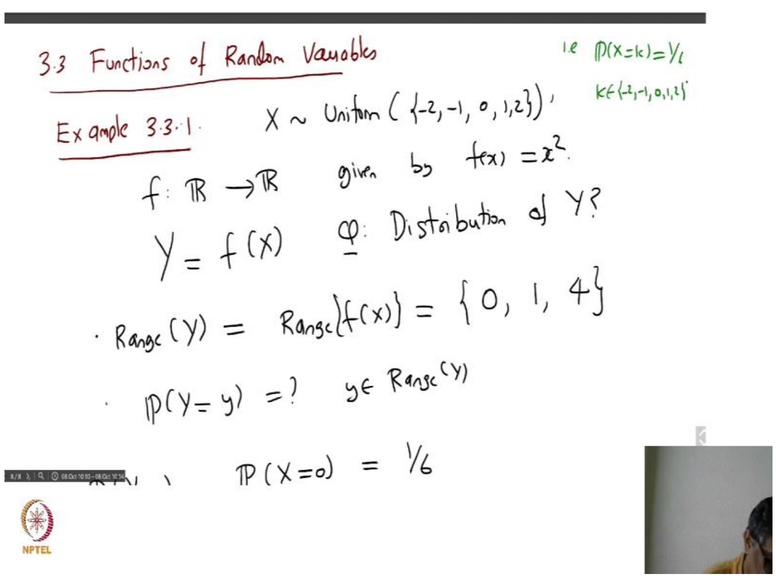
So, now I will close today's class with another idea, that it is kind of useful. It is like functions random variables. So, here is the, this section is 3.3. So I will do only the examples. I would not do the, I will come define it next time. So, here you go, 3.3 is a functions of random variables. So, let us do a bunch of examples. Let us try and let us , I will come to the formal definition after I do the examples. So let us do a simple example first. Let us say X example. Example 3.3.1. So, let us say X is a random variable that takes values uniform random variable, that takes values. Let us say in minus 2. So, it takes uniform value choose a, number uniformly between minus 2 and 2. So 0,1 that is what X does.

And let me now take a function, this function is a real line of the real line (())(20.32). And given by let us say f of x is equal to x square. So, now I will, let I just call my random variable Y as f at x and I ask you the question, what is the distribution of Y. So, now the idea is that, so first what I do; I first I find the range of Y. So, what is the range of Y.

The range of Y is X is minus 2 minus 1, 0, 1 and 2. So, range of f of x, f at x, that is just the set what all it will take. It will just take the values, when you square 0 you get 0, when you square minus 2 or 2 you get 4, when you square 1 or minus 1 you get 1. So, this is the range of X. Then our next thing I will do is to find the distribution of Y. I have to find the chance that Y is equal to Y for every Y in the range of, Y is equal to Y, is what; for a Y in the range of Y.

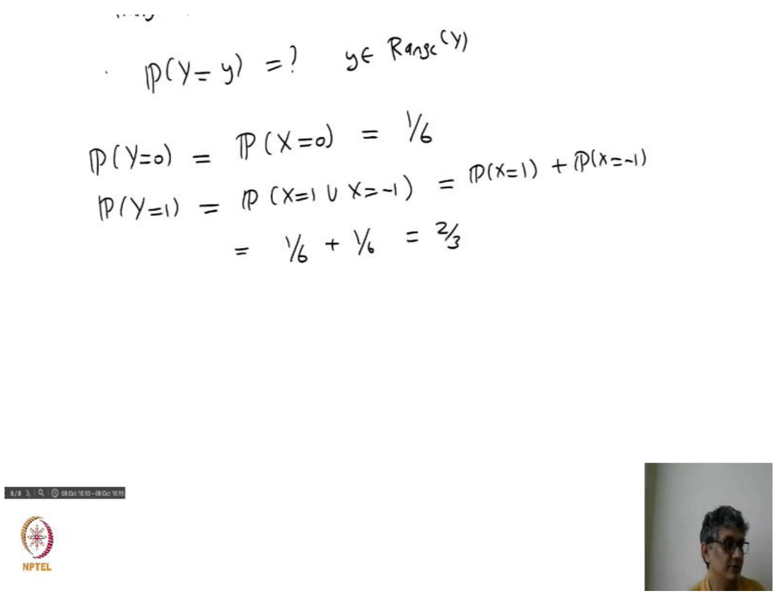
(Refer Slide Time: 22:02)

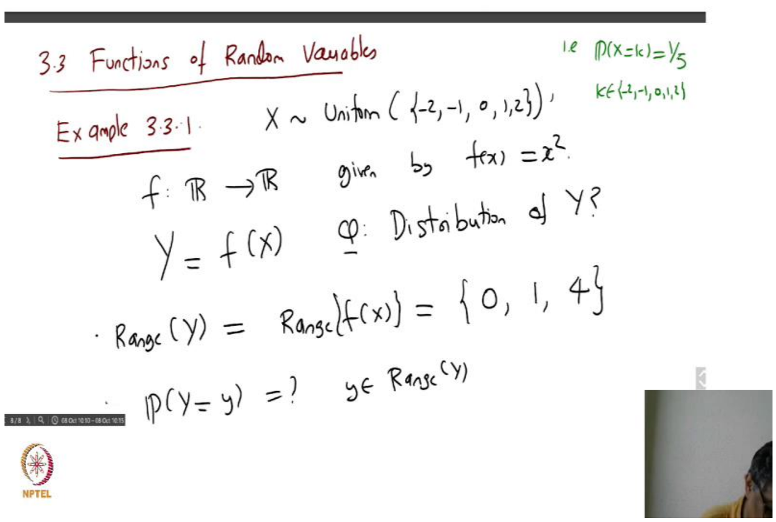




So, how do I do that. So, Y has three values. So chance that Y is equal to 0, so that happens when only in one case, that is when Y is 0 means X squared is 0 and that is the same as saying X is 0. . Then nothing else happens. So that is the same as, there are six possible outcomes in the uniform such as 1 over 6. So distribution of X, we know that the chance of X equal to k is equal to one sixth for k and minus 2, minus 1,0,1,2. So I will re-write this in green, so it is clear (())(22:51). So, that is uniform means the chance that X is equal to k is one sixth for k in minus 2, minus 0, 1 , 2.

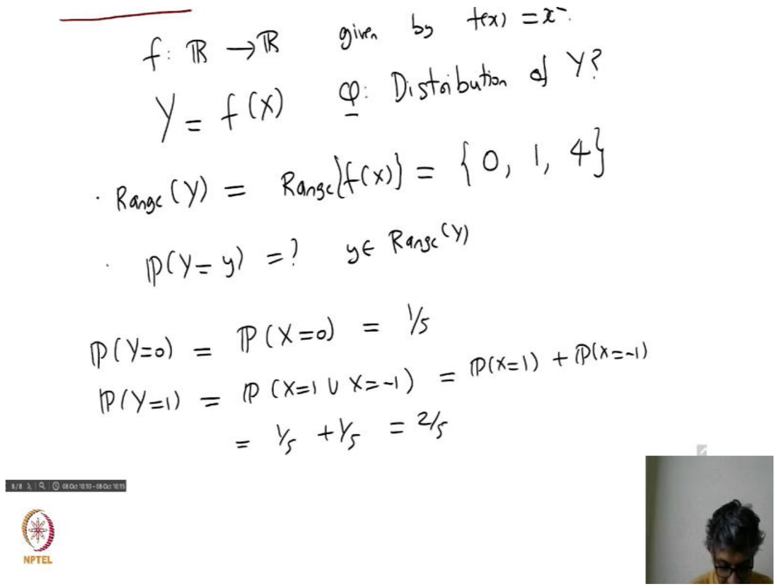
(Refer Slide Time: 23:16)





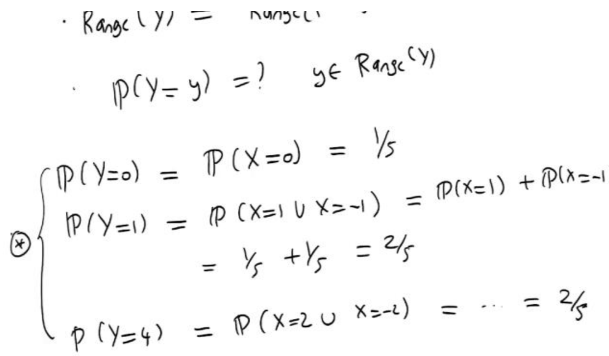
So, this example, if you do Y equal to 0, that means X has to be 0. This cannot be anything else. So, that means (())(23:13). So, let us do X equal to 1, so chance that X is equal to 1. (())(23:18)Y is equal to 1. So, the chance that Y is equal to 1. We do Y equal to 1. What do I get, that is the same as X is equal to 1 or X equal to minus 1, because if X is 1 also Y is 1. If x is minus 1, that also implies Y is 1. So, that is the same as the chance that X is equal to 1 plus the chance that X is equal to minus 1 and if you add the two up, I get one sixth plus one sixth I get 2 by 3. Oh sorry so it is a one sixth or one fifth. I am sorry how many numbers are there. There are 1, 2, 3, 4, 5. I am sorry about that. Yeah. So, here is five number. So, I should just correct this a little bit. This is five,that is one fifth.

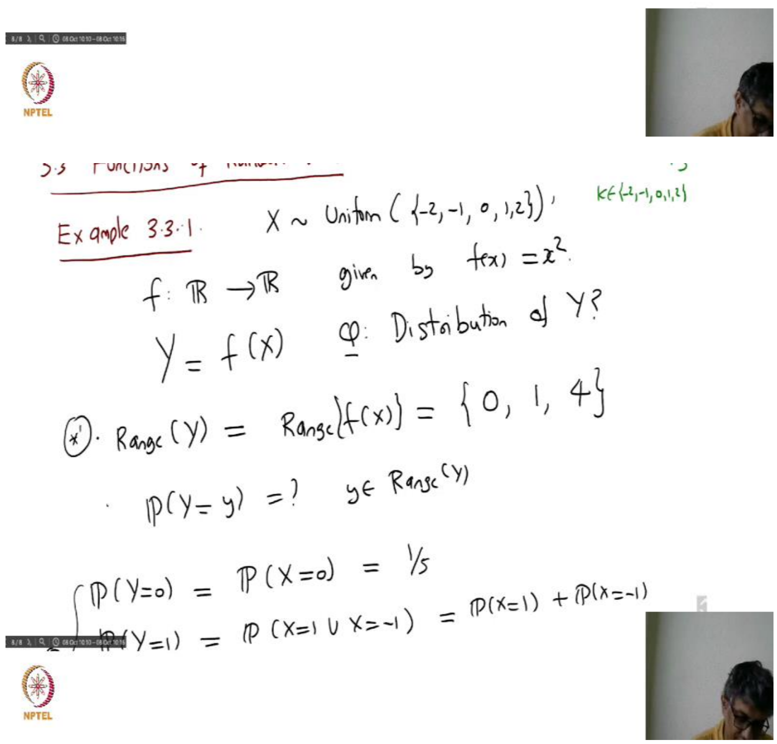
(Refer Slide Time:24:14)



And I come down here and this all one fifth. Sorry this is all one fifth. So I go back here like one fifth. And this is one fifth plus one fifth and that is this. Very nice.

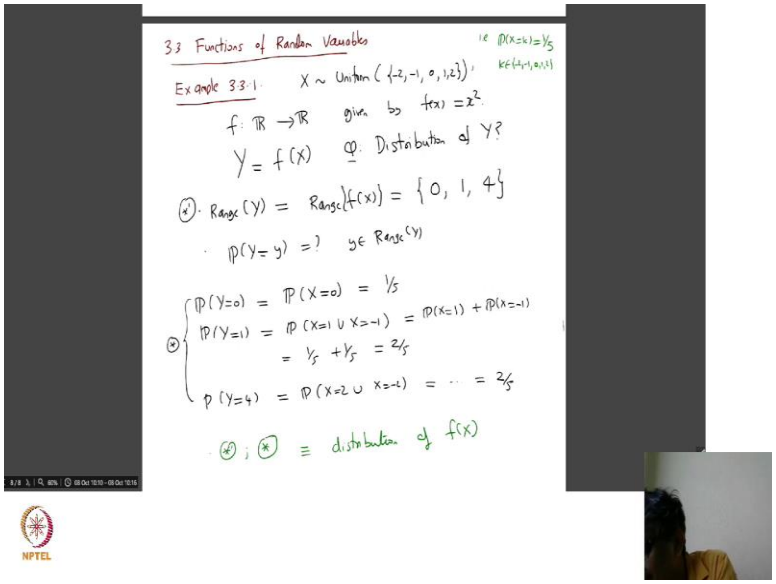
(Refer Slide Time: 24:34)





And then similarly the chance that Y is equal to 4, that is the same as the chance that X is equal to 2 or X equal to minus 2. And that if you compute the computations you just get again total this (())(24.51).

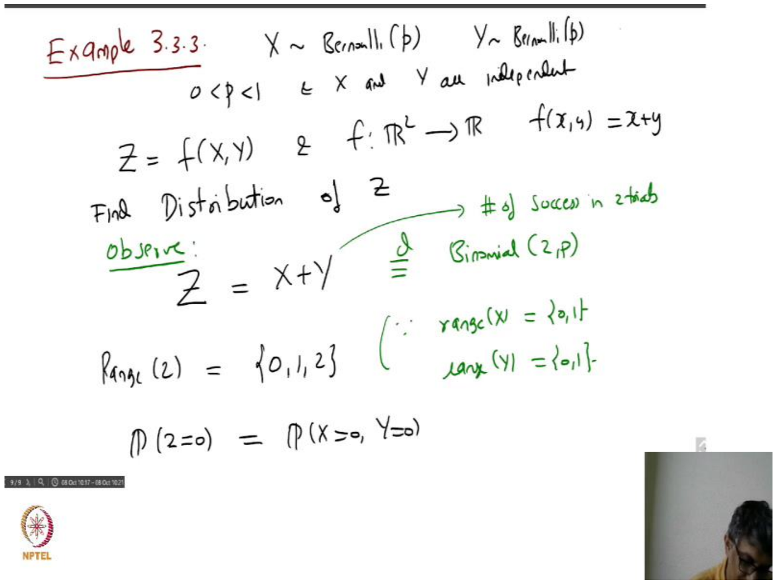
(Refer Slide Time: 25:00)

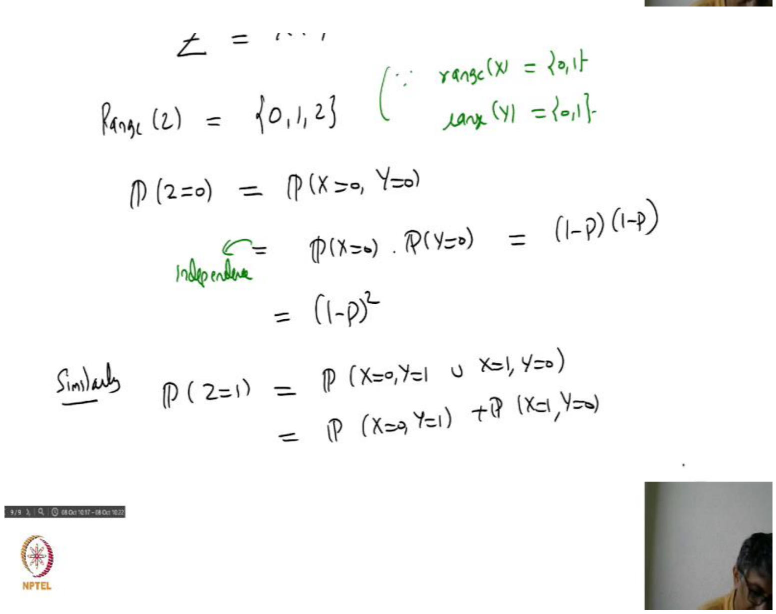


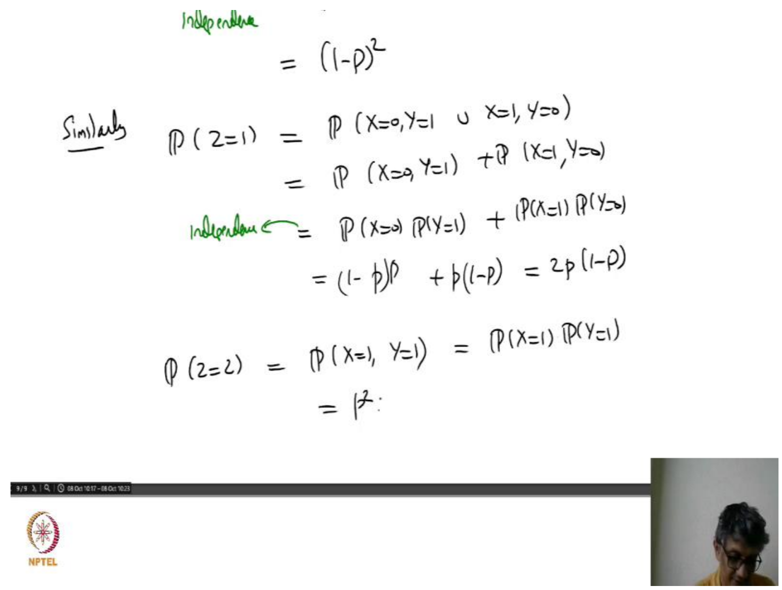
So, that is the distribution of Y. So, Y takes along with this. This is called star. Let us say a star and this let us say it is called star prime. So, star and star prime give you distribution of Y. Star and star prime. So range of X range of Y and the distribution of Y and the probability how it distributes the rates on the range gives you the distribution of Y. Y I will write Y as f of x.

So, this is one procedure by which you can compute. Let me make this little smaller. So, you can see the whole screen in one shot. A little bit more smaller. Very nice. So, this in one shot you can see how to compute the distribution of f of x. Given the distribution of X. So here the function was not one to one. But so to be a little careful. Here to go and sort of find the exact values of X for which match to the value Y and compute the problem.

(Refer Slide Time: 26:38)







Excuse me, so let us do a simpler. Another simple thing, we have seen before. So, here is another example. Let us say I take two random variables this time. So, let us say example I would say 3.3.3. So, let us say I do the following. Let us say I toss a coin once. I look at the number of successes.I toss a coin again. So, let us say I look at X is equal to Bernoulli p. It is a biased coin and then Y is also Bernoulli p.

So let us assume 0 is less than p and less than 1. And let us also assume x and y are independent. I do these two things. And then I tell you I give you a function of X and Y. I give you Z is a function of X and Y and I say f is a function from, I am writing a little bit complicated just sort of to make sure that I capture generality of it.

Of course the idea is very simple. So, I look at f of x comma y as x plus y. So, look at this and I ask you to find the distribution of Z. So, now you stare a little bit like this. What is going on, just nothing much is going on here, because all you want to do is, you want to find Z is equal to f of x y which is just X plus Y.

Now, observed this. Observe. This just look at X and Y are two independent variables. If you go back to the experiment idea, you already know the following that X is independent of Y. That means I am observing the number of successes in two tosses and that is the same as a Bernoulli binomial to come up with because what we observed is that this person right here is the number of such number of heads or number of successes in two trials and that is one way of understanding .

But in general if you want to sort of, if you do not want to go to experiments and you want to do it directly; you have to be a little bit more careful. So, how does one do that let us just check that. So, if you want to do it in a sort of methodical way, you have to be a little bit more careful. So, let us see how does one do that. One needs to be a bit more careful so one has to write out the events properly.

Let us do it for the example. Then it will be clear how to do it in general. So, here the you can do it, you can compare the experiments and do it. But the the correct way to do it is that first you observe. Like if you would like to do the previous example, you observe that the range of Z is going to be what, so X and Y are Bernoulli. So, X takes value zeroes and one. So, this is going to be 0, 1 and 2. So, this is because, this is because, range of X is 0 or 1 and range of Y is also 0 or 1. So, then what you would have to do is , if you want to compute the probability of Z equal to 0, when will this happen.

Will happen X and Y are 0, that is the same as saying the chance that X is equal to 0 and Y is equal to 0 and this will be the same as use independence. Here independence and you would get the chance that X is equal to 0 times Y is equal to 0 and that is the same as a P square or 1 minus P square. 1 minus P and that is the same as 1 minus P square.

Similarly, if you do a probability that Z is equal to 1.You can write out all the combinations. You can say that the same as happens when X is 0, Y is 1 or X is 1, Y is 0. You split it up into two sums X equal to 0. Y equal to 1 plus the chance that X equal to 1 and Y equal to 0 and then use independence now. Now, to say that X equal to 0 times the chance that Y is equal to 1 plus the chance that X equal to 1 and Y equal to 0.

So, here are using independence and then I would get (())(31.47). This would be just same as P times say 1 minus P times P and this would be the same as 1 minus P times p and I would just get the correct answer. I would get 2p times 1 minus P, which is what I wanted. And I can finally conclude that as Z equal to 2 will have, will have, all possible problems like X equal to 1, Y equal to 1, that is the only possibility and that is just by independence X equal to 1 times Y equal to 1 and that is the same as P square. So, I will build on this next time. I will take a two minute break and I will come back.