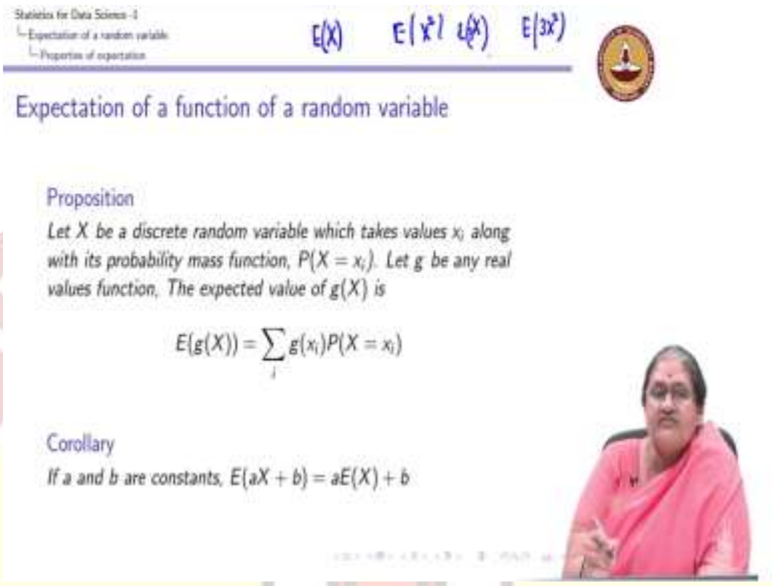
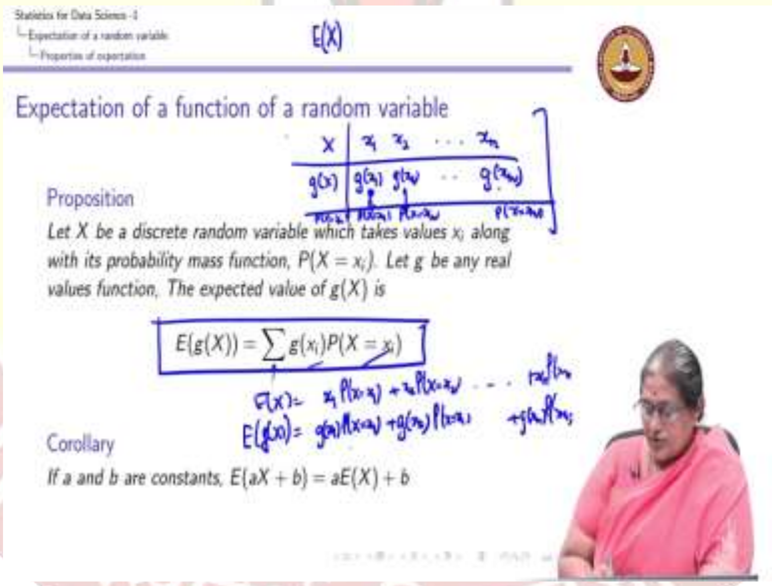
IIT Madras

ONLINE DEGREE

Statistics for Data Science - 1
Professor Usha Mohan
Department of Management Studies
Indian Institute of Technology, Madras
Lecture 9.3 - Expectation of a random variable

(Refer Slide Time: 00:15)



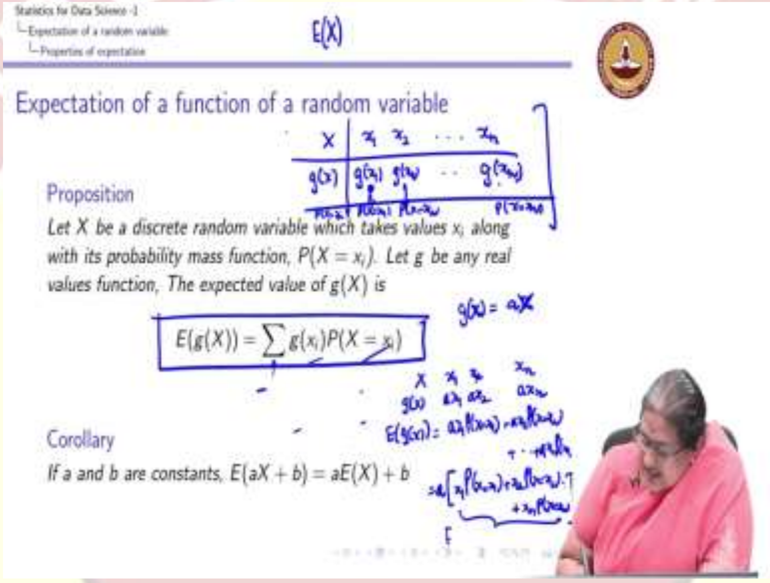


Now, suppose I am interested in knowing about, I have a random variable X, I have defined expectation of a random variable X. Now, suppose I am interested in knowing about expectation of, say, X^2, or expectation of 3^X, or expectation of $3X^2$. So, I am expect. In other words, we are interested in knowing about expectation of a function of X.

So, if I have X, which is a random variable, which I again takes, for now I am assume it takes finite number of values, then $g(X)$ is going to take $g(x_1)$, $g(x_2)$ with values $g(x_n)$. The probabilities would remain the same. So, I have $P(X = x_i)$ is $P(X = x_1)$, $P(X = x_2), \ldots\ P(X = x_n)$.

I know our expectation of X is going to be $x_1 P(X = x_1) + x_2 P(X = x_2) + \cdots x_n P(X = x_n)$. I can define expectation of $g(x)$ to be $g(x_1)P(X = x_1) + g(x_2)P(X = x_2)$, because it takes value $g(x_1)$ with $P(X = x_1)$, it takes the value $g(x_2)$ with $P(X = x_2)$, it takes the value $g(x_n)$ with $P(X = x_n)$. And that is what is summarised here, as expected value of $g(X)$ is $\sum_x g(x)P(X = x)$.

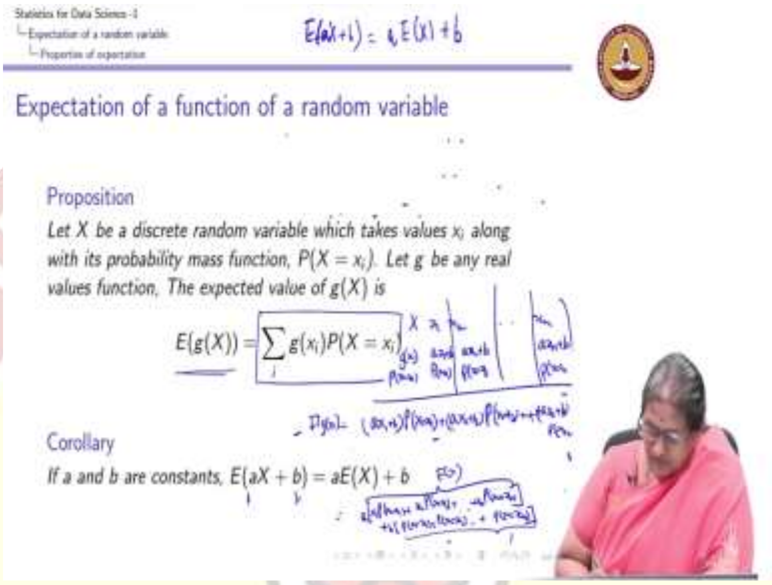
(Refer Slide Time: 02:34)



As a natural corollary for it, if $g(x)$, so suppose I have now $g(x)$ is a constant times x. Suppose, $g(x)$. Suppose $g(x)$ is just a constant times x, $g(x)$ is a constant time x, then I know that, again, let x take the value $x_1, x_2, \ldots\ x_n$, $g(x)$ is going to be $ax_1, ax_2, \ldots\ ax_n$, expectation of $g(x)$ is going to be $ax_1 P(X = x_1) + ax_2 P(X = x_2) + \cdots ax_n P(X = x_n)$. I can remove a outside, I get $x_1 P(X = x_1) + x_2 P(X = x_2) + \cdots x_n P(X = x_n)$. So, we can recognise this term inside the bracket is $E(X)$.

So, I have, so we can recognise that this value inside this expectation of X. Hence, I know I can write that $E(aX)$ is nothing but $aE(X)$, a is a constant. So similarly, let us look at the case where $g(X) = aX + b$, where a and b are constants. So again, let us look at that case where I have both a and b are constants.

(Refer Slide Time: 04:16)

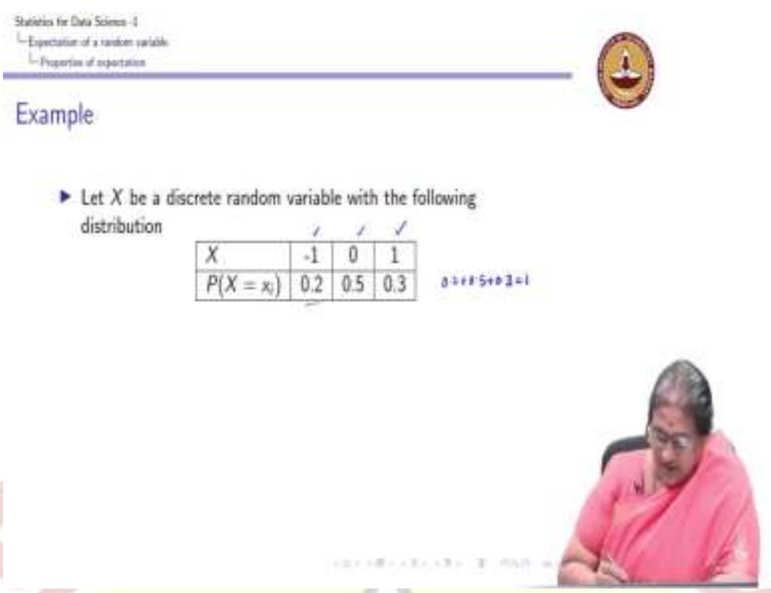


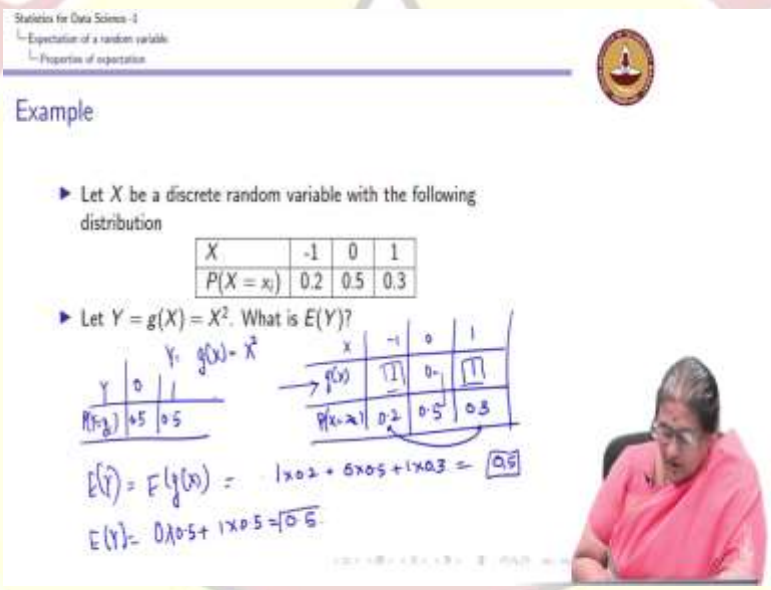
So, what do I have I have my X. So, I have my X, which is taking the value $x_1, x_2, \ldots\ x_n$, my $g(X)$ now is going to be $ax_1 + b, ax_2 + b, \ldots\ ax_n + b$ with the same probabilities, this is $P(X = x_1)$, this is going to be $P(X = x_2)$, this is going to be $P(X = x_n)$.

So, my $E(g(X))$ is going to be $(ax_1 + b)P(X = x_1) + (ax_2 + b)P(X = x_2) + \cdots (ax_n + b)P(X = x_n)$. Now, we can see that this is nothing but I can write this as $a(x_1P(X = x_1) + x_2P(X = x_2) + \cdots x_nP(X = x_n))$ the first portion, and the second I can write it as $b(P(X = x_1) + P(X = x_2) + \cdots P(X = x_n))$.

Now, the first portion that is this is nothing but $E(X)$. Now, since I have the second portion, this is going to be a probability mass function. So, this adds up to 1. Hence, I can check and verify that $E(aX + b)$, where both a and b are constant is the same as $aE(X) + b$. And this arises as a general corollary to $E(g(X))$ is a same as $\sum_{x_i} g(x_i)P(X = x_i)$.

(Refer Slide Time: 06:22)





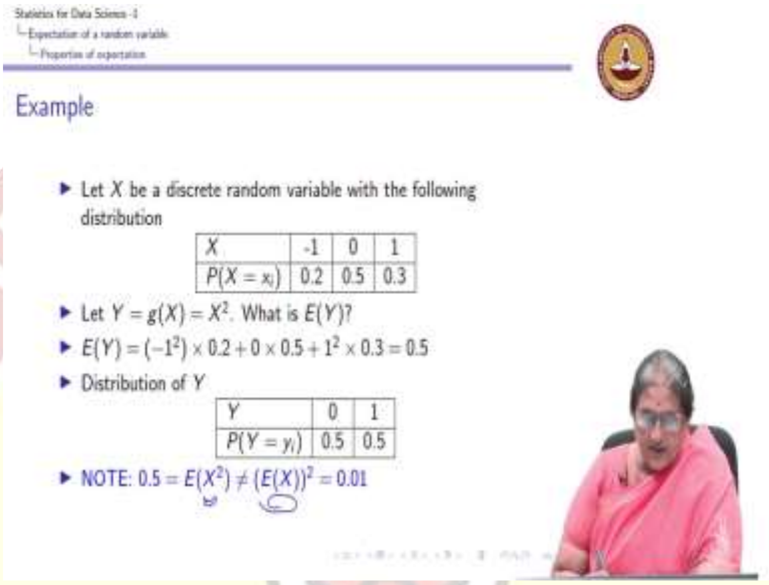
So, now let us look at a simple example where X is a discrete random variable with the following distribution, again is this a probability mass function X takes the value -1, 0 and 1 yes, this is a probability mass function, because 0.2 + 0.5 + 0.3 equals to 1. So, it is a probability mass function, I have each probability greater or equal to 0.

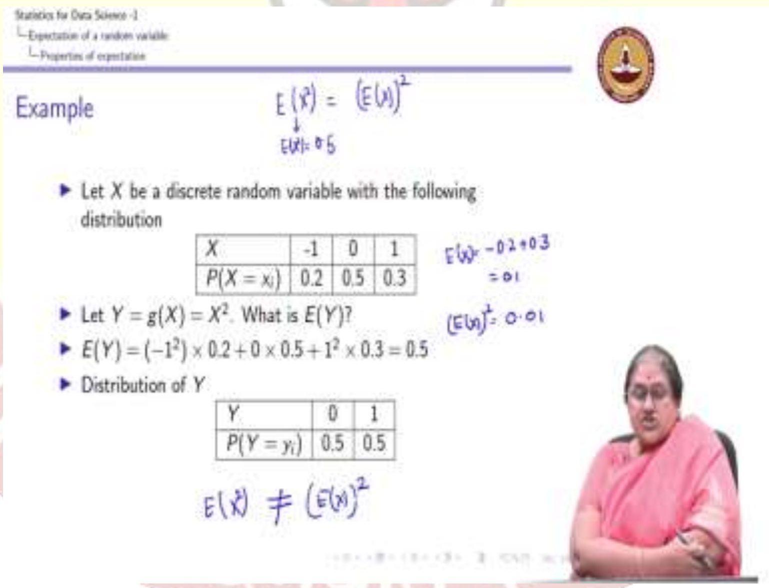
Now, let us define $g(X)$. So, let Y be another random variable, which is equal to $g(X)$, which is equal to X^2. So again, X takes the value -1, 0, 1 $g(X)$ takes the value X^2 which is going to be 1, 0, 1, with probability 0.2, 0.5 and 0.3, this is $P(X = x_i)$ the same probability.

So, my expected value of $g(X)$, which is my expectation of Y, is going to be $1 \times 0.2 + 0 \times 0.5 + 1 \times 0.3$, which is equal to 0.5. So, I have expectation of $g(X)$ 0.5. Now, when you notice $g(X)$, I can also write this as Y, which is $g(X)$ takes 2 values and those values are 0 and 1.

Now, what is the probability with which Y takes the value 0, the probability with which Y takes the value 0 is 0.5, and the probability with which Y takes the value 1 is 0.2 + 0.3, which is again a 0.5. Hence, if I looked at from the first definition $E(Y)$, it is going to be $0 \times 0.5 + 1 \times 0.5$ which is equal to 0.5. This matches with what we got earlier.

(Refer Slide Time: 08:39)





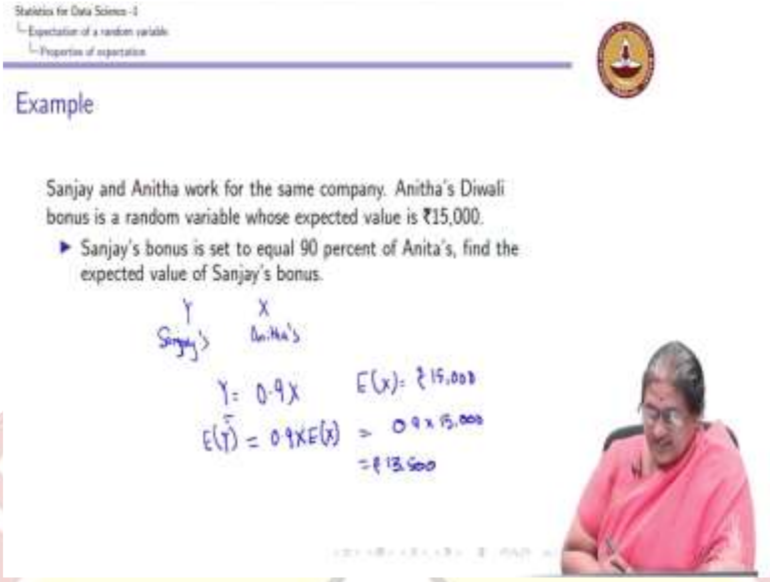
Hence, given a random variable, I can find out what is the expectation of the function of a random variable. A very common mistake people do is to think $E(X^2)$ is the same as $(E(X))^2$, what we are seeing just now is $E(X^2)$ was 0.5, what is $E(X)$?

$E(X)$ is -0.2 + 0.3, which is a 0.1. So, the $E(X^2)$ is 0.01. So, $E(X^2) \neq (E(X))^2$. And this you should be very careful about understanding this that expectation of square of a random variable is not the same as square of expectation. So, expectation of square is not same as square of expectation.

(Refer Slide Time: 09:55)

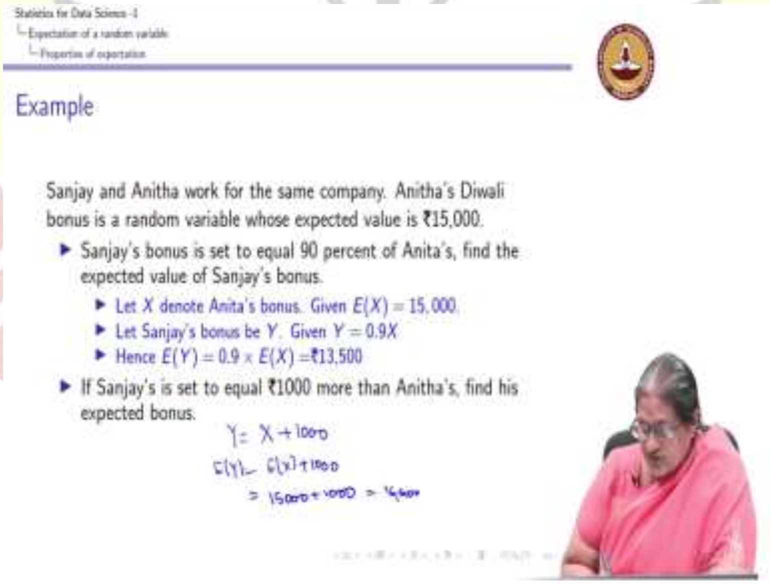
The next example, where we go for an application of what we have learned is the following application. So, Sanjay and Anita work for the same company, Anita’s Diwali bonus is a random variable whose expected value is 15,000. So, let X be the random variable, which represents Anita's bonus, what is given to me is $E(X)$ is 15,000 rupees. So, that is what is given to us.

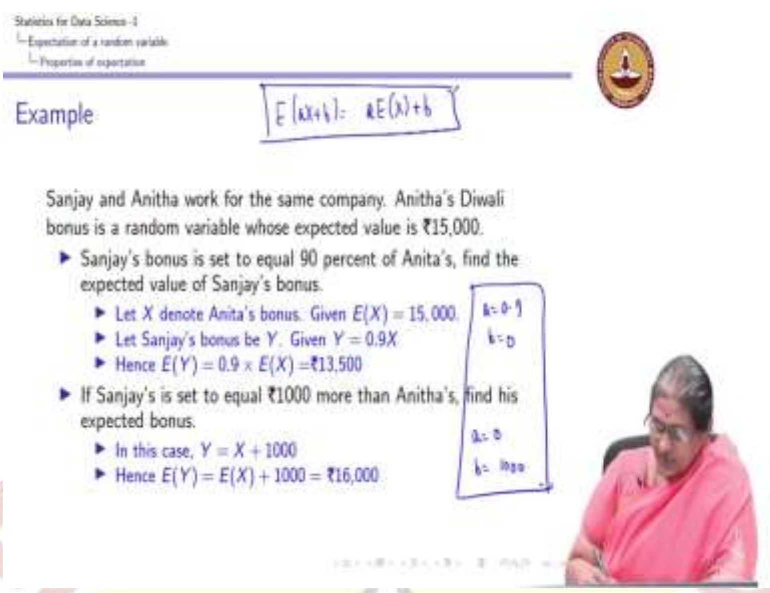
(Refer Slide Time: 10:28)



Now, the next thing that is given to us is Sanjay's bonus is equal to 90% of Anita’s bonus. So, let Y be Sanjay's bonus X is Anita’s bonus Y is Sanjay's bonus what is given to is Y is $0.9X$, I know what is $E(X)$ this is given to be 15,000 rupees. So, the question is what is expected value of Sanjay’s bonus that is $E(Y)$. From our earlier result I know that if Y is $0.9X$ then $E(Y)$ is 0. $E(X)$ which is 0.9 ×15,000 rupees which is 13,500 rupees.

(Refer Slide Time: 11:32)

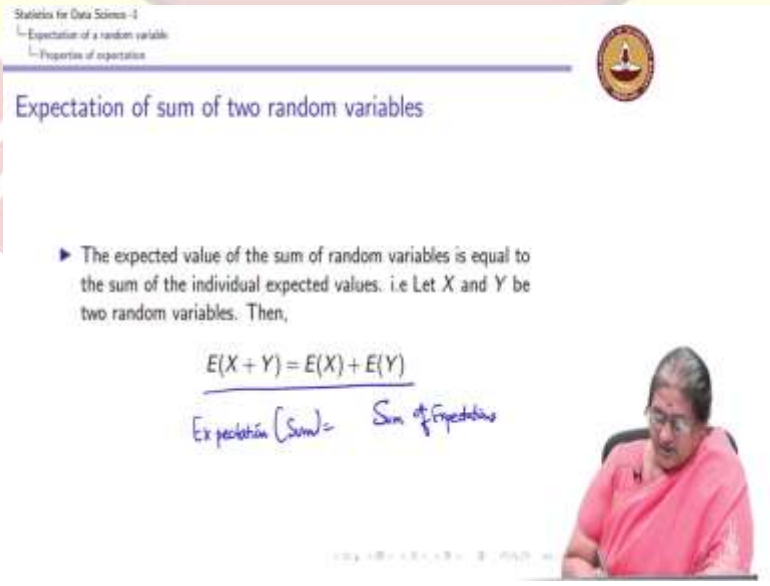




So, I can see that Sanjay's bonus expected values of Sanjay's bonus is 13,500. But suppose Sanjay's bonus is set to be equal to 1000 rupees more than that of Anita's bonus, then find the expected bonus. So, in this case, I have Y is $X + 1000$. So, $E(Y)$ is going to be $E(X)$+1000. I already know $E(X)$ is 15,000. I add 1000 hence, the expected value of Sanjay's bonus is going to be 16,000 rupees.

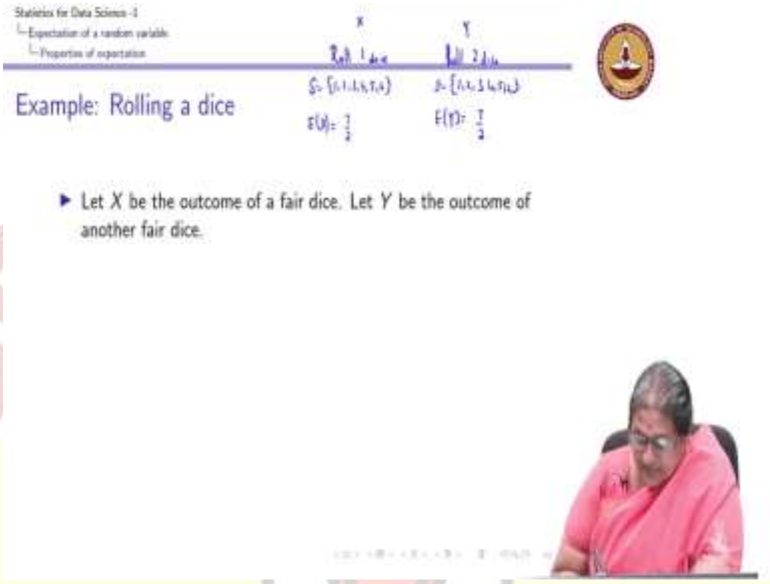
So, you can see how we have just applied the expectation property of $E(aX + b)$= $aE(X) + b$. In this case, my a was 0.9 and b was 0. In this case a was 0 and b was 1000. So, just you need to identify what are your constants to apply this property.

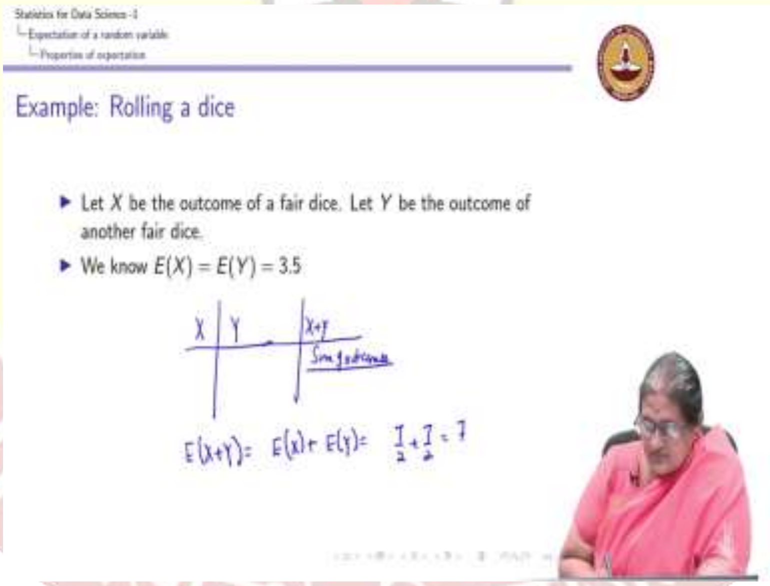
(Refer Slide Time: 12:51)

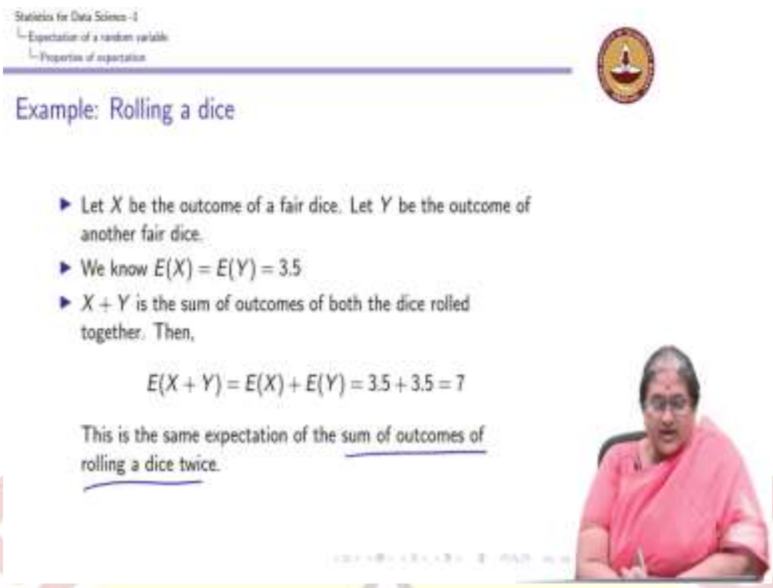


Now, let us look at what can I tell about expectation of a sum of random variables. If I am given two random variables with and I know that X relative or respective expectations, then $E(X+Y)$ that is expectation of sum is sum of expectations.

(Refer Slide Time: 13:30)





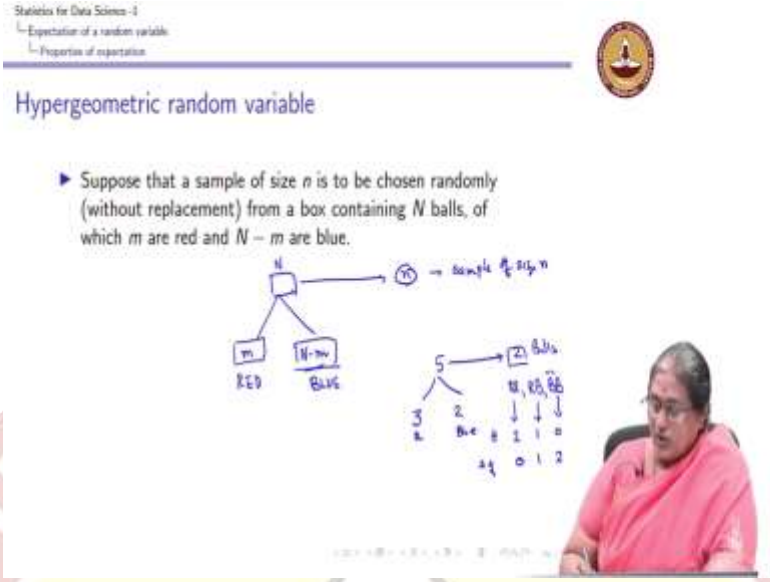


We are not proving these but let us look at example, again roll a dice. If I am rolling a dice and I say, I roll 1 dice, I roll another dice. The sample space in both these experiments are going to remain the same because I am just rolling a dice. So, if X is denoting the outcome here, we have already seen $E(X)$ in this case is going to be I have already seen this expectation was $\frac{7}{2}$ which is 3.5.

Again, a Y is representing the outcome in rolling the second dice, I also know $E(Y)$ is again $\frac{7}{2}$. So, if X and Y are outcomes of rolling dice separately, I know $E(X) = E(Y)$ which is equal to $\frac{7}{2}$. Now, what is X, X is the outcome of the first dice, Y is the outcome of the second die.

So, $X + Y$ would denote the outcome or sum of outcomes of both the dice. Sum of outcomes of both the dice is equivalent to so rolling a die twice. So, you can see that $E(X + Y)$ from my formula is $E(X) + E(Y)$ which is $\frac{7}{2} + \frac{7}{2}$, which is equal to 7. And this matches with our $X + Y$, which was same as the expectation or sum of outcomes of rolling a dice.

(Refer Slide Time: 15:18)



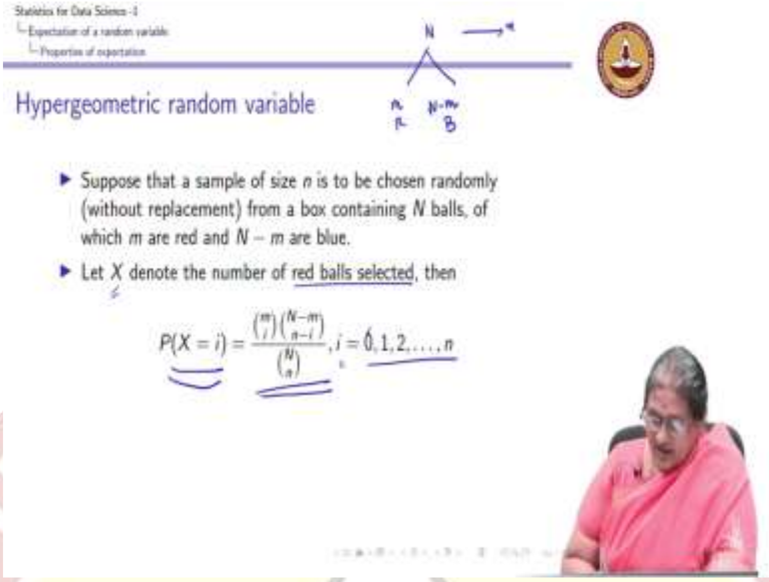
So, now let us look at a example of another important random variable which arises in a lot of application. And this is referred to as a hyper geometric random variable, a hyper geometric random variable. Suppose I have m of 1 type, and I have $N-m$ of another type, I have total of N, I have total of N balls, I can or N people of which m are of one type and $N-m$ are of another type.

Here, I am assuming m are red balls, and $N-m$ are blue balls. From this N, I am choosing a sample of size n. So, size this is my sample of size n. Again, this could have, again, this sample also could have red balls and blue balls, the question we are interested in knowing is suppose for example, I have 5 balls, of these 5 balls, I have 3 red balls, and I have 2 blue balls.

And I am choosing from here, 2 balls at random, then the possible choices for these 2 balls is I could have both are red, I could have 1 red and 1 blue, I could have both are blue balls, these are the possible chances I can have. So, if I look at the number of red balls in my sample, the number of red balls here are 2 number of red balls here is 1, the number of red balls here is 0.

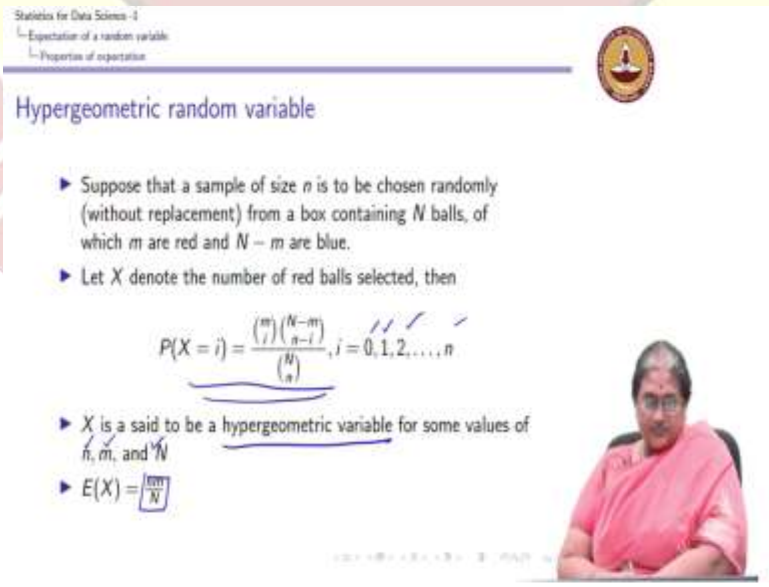
So, in this, if I am choosing both blue balls, I do not have any red ball, if I have 1 blue ball or and 1 red ball, I have 1 red ball and here I have 2. Similarly, if I am counting the number of blue balls, here it is a 0, here it is a 1, here it is a 2. So, the random variable of interest in this case could be the counting the number of red balls, or blue balls, the probability mass function, for example.

(Refer Slide Time: 17:56)



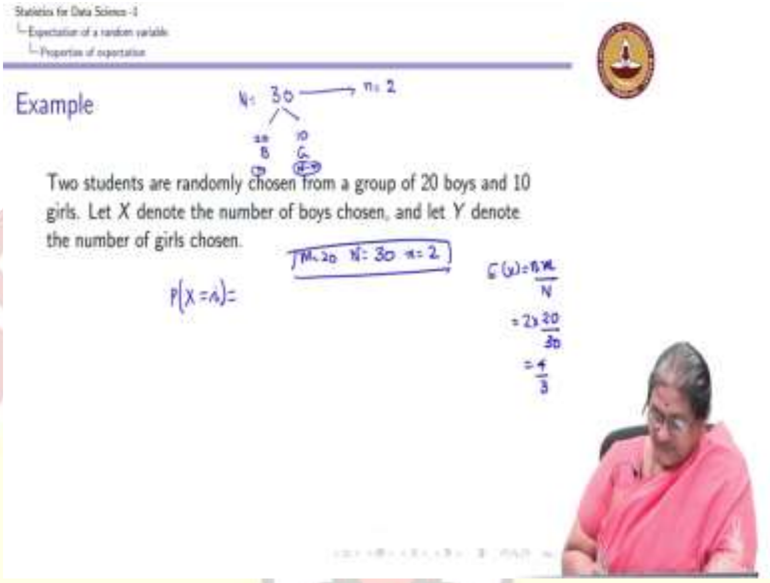
If I am looking at this example here, so here the probability mass function, where X is the number of red balls that are chosen. And I have said that this i can take value 0 to n. Again, recall, I have N I am choosing n, in this I have m red, I have $N - m$ blue. So, I could have both red balls, or I could have all the n red balls, or I could have all the n blue balls. So, I could have 0 red balls, the probability mass function or the probability distribution is given by this expression. How we get this expression would be discussed in tutorials, but this is the expression.

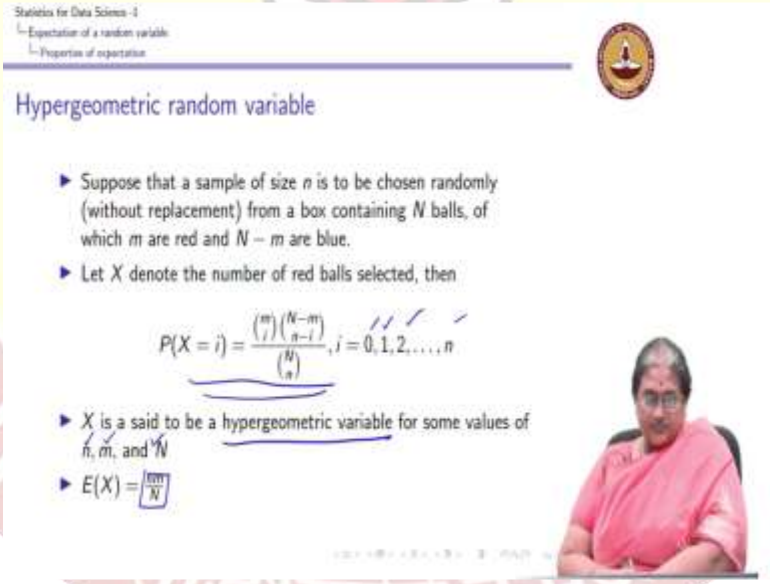
(Refer Slide Time: 18:44)



And you can also verify that this $E(X)$is $\frac{nm}{N}$, a variable of this kind, a discrete variable of this kind, which takes value 0, 1 to N is referred to as a hyper geometric variable for values of n, m, and N.

(Refer Slide Time: 19:11)

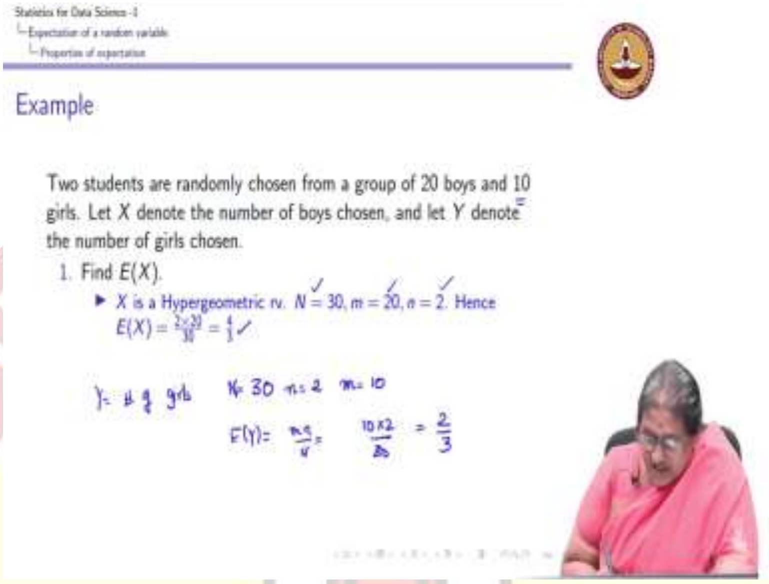


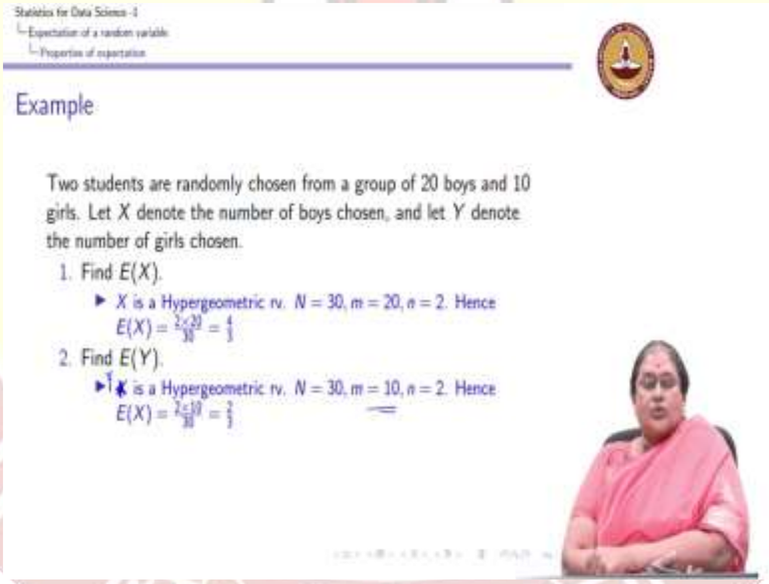


Let us look at an application of the hyper geometric random variable. I am choosing 2 students from a group of 20 boys and 10 girls. So, my N is 30. I have 30 people in these 30 people, I have 20 boys and 10 girls. So, this was my m, this is my $N - m$, which is 30 −20. This is of one type this is of another type. I am choosing n which is equal to 2.

I am choosing 2 students. Again, it could be both girls, both boys or 1 girl and 1 boy. So, my probability mass function of $X = i$ will follow what I have here, $\binom{m}{i}$ my m here is equal to 20, my N=30 and my $n = 2$, I can plug in these values in this equation of mine, what I am interested in knowing is the following. So, my $E(X)$ is going to be $\frac{nm}{N}$, which is $\frac{2.20}{30}$, which is $\frac{4}{3}$.

(Refer Slide Time: 20:38)

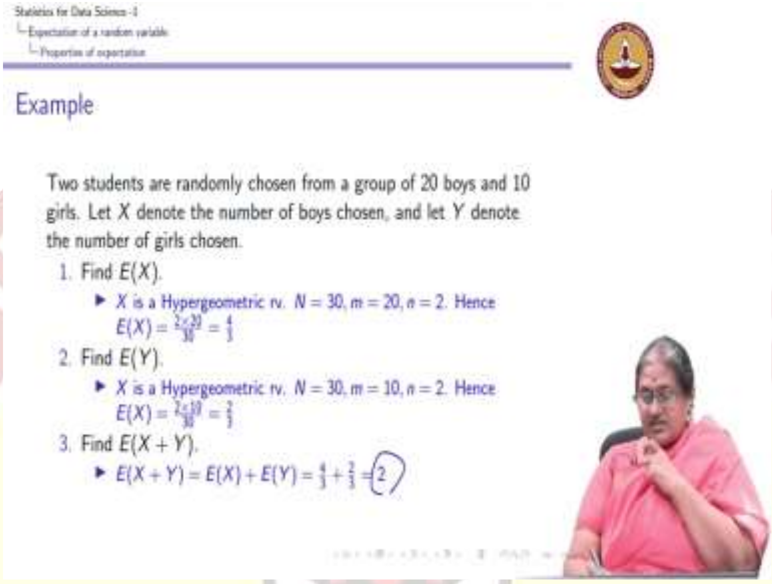


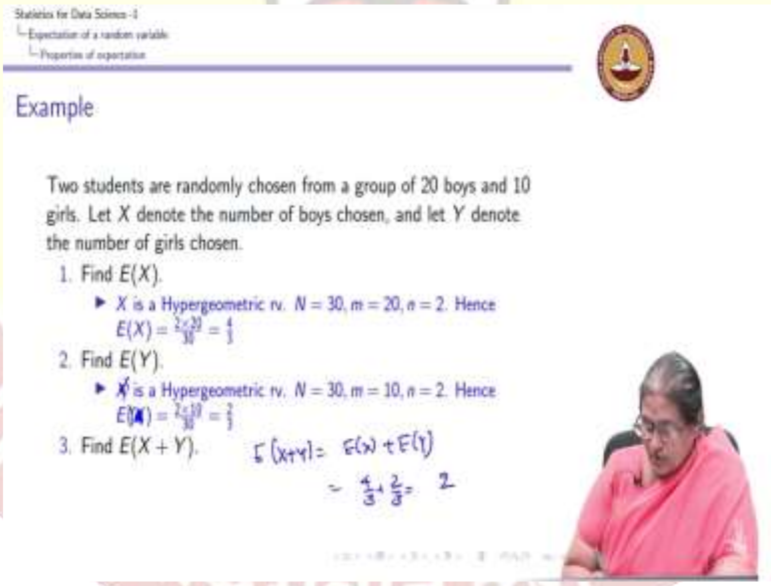


So, you can see that the $E(X)$ where I recognise N =30, $m = 20$ and n=2 is $\frac{4}{3}$. Now, Y is denoting the number of girls that is chosen. So, if Y is the number of girls that is chosen, now everything else capital N remains 30, small n also remains 2 but my m is now going to be equal to 10, because I have only 10 girls.

So, in this case, my $E(Y)$ is going to be $\frac{nm}{N}$ which is $\frac{10\times 2}{30}$, which is $\frac{2}{3}$ that is my $E(Y)$ can be shown to be equal to $\frac{2}{3}$ again by assuming or by seeing that Y is a hyper geometric variable with $m = 10$ and $n = 2$.

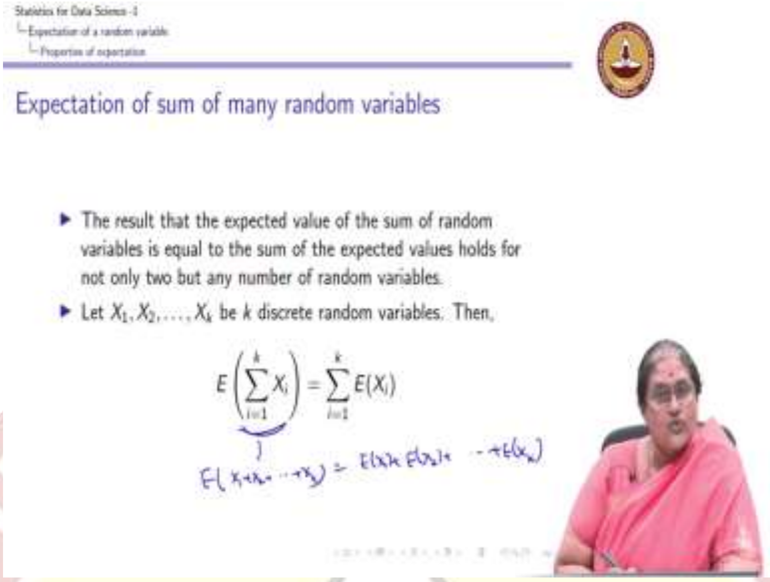
(Refer Slide Time: 21:44)





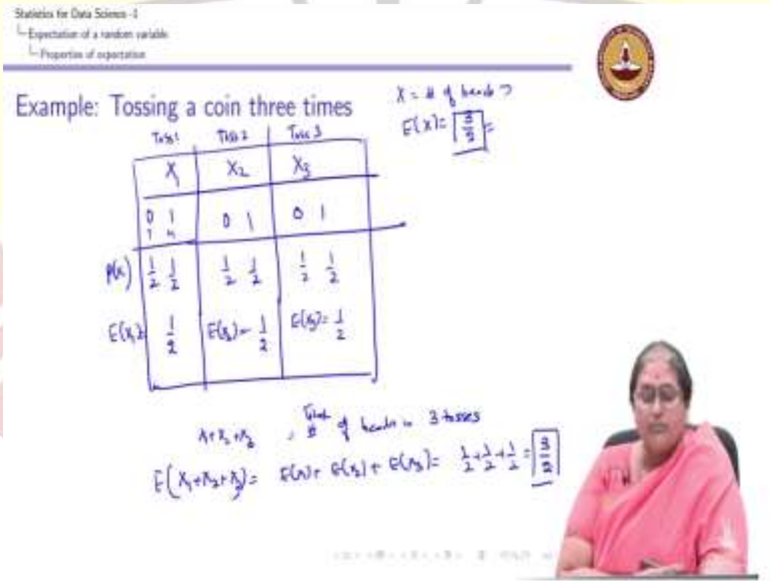
Now, what is the $E(X + Y)$, I can check that $E(X + Y)$ is $E(X) + E(Y)$, which is $\frac{4}{3} + \frac{2}{3}$ which is equal to 2 and I get $E(X) + E(Y)$ =2.

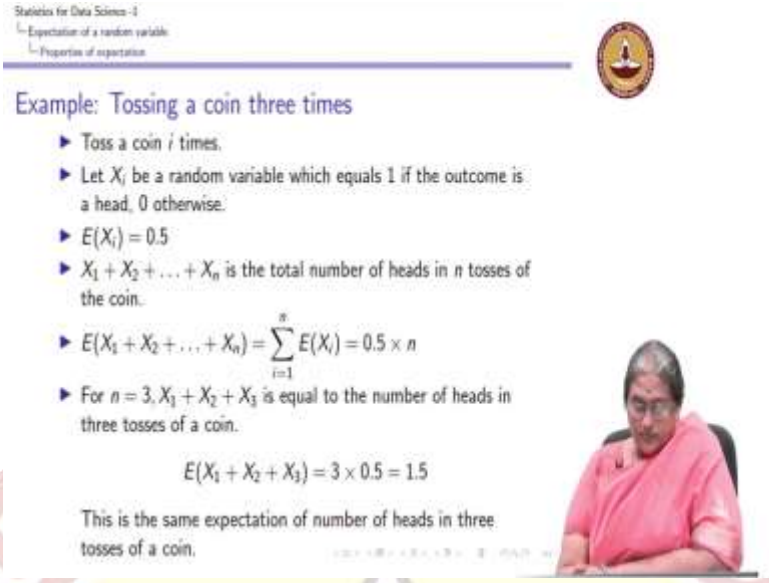
(Refer Slide Time: 22:07)



I can extend this property, what is the proper expectation of sum is sum of expectation to more than 2 variables. In other words, if I have $X_1,\ X_2, \ldots\ X_k$ discrete random variables, then the expectation of the sum which is equal to $X_1 + X_2 + \ldots X_k$ is equal to $E(X_1)$+ $E(X_2)$+… $E(X_k)$ when $k = 2$ it would reduce to what we have discussed earlier.

(Refer Slide Time: 22:50)



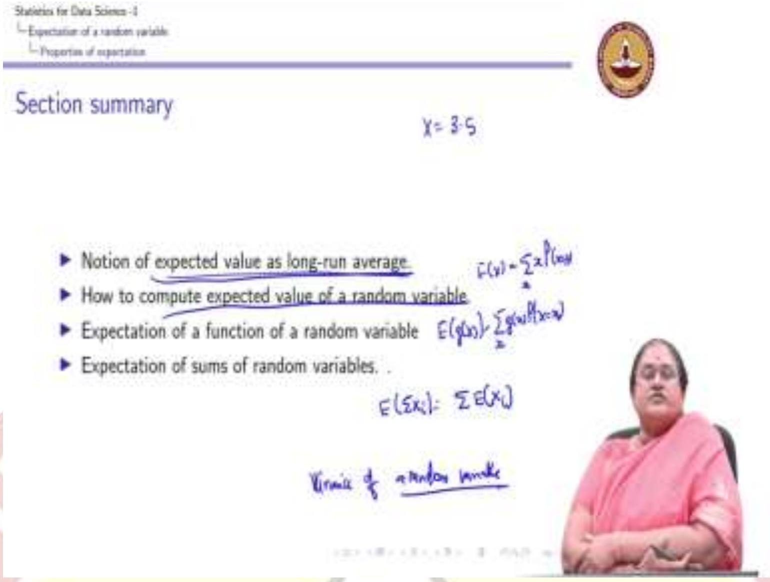


So, now let us look at the experiment of tossing a coin 3 times, we already know that if X is the number of heads, then we have seen that $E(X)$ is $\frac{3}{2}$ this is what we have seen already. Now, let us look at this experiment I find I toss a coin once I toss a coin second or toss 1 coin toss 2 coin toss third coin let X_1 be the random variable, which takes the value 0 for a tail and 1 for a head, with probability equal likely $\frac{1}{2}$ and $\frac{1}{2}$. I know $E(X_1) = \frac{1}{2}$.

Similarly, let X_2 be the random variable again takes the value tail and head with the same probabilities $E(X_2)$ is again a $\frac{1}{2}$. X_3 again takes the value 0 and 1 with probability $\frac{1}{2}$, $E(X_3)$ is again equal to $\frac{1}{2}$, now $X_1 + X_2 + X_3$, will is the same as number total number of heads in 3 tosses, or the 3 coins.

Which is same as number of heads, when I toss a coin 3 times, I can verify that $E(X_1 + X_2 + X_3)$ is $E(X_1) + E(X_2) + E(X_3)$ which is $\frac{1}{2} + \frac{1}{2} + \frac{1}{2}$, which is $\frac{3}{2}$. And this matches with what we already obtained for the expectation of number of heads, when you toss a coin thrice which is $\frac{3}{2}$.

(Refer Slide Time: 25:02)



So, in summary, what we have introduced so far and what you should be knowing now is the notion of the expected value as a long run average, word of caution is do not interpret it to be that that is the value the random variable will take. For example, we saw that X cannot take a value 3.5 when you roll a dice once, but rather it is the average of the outcomes.

The formula to compute an expectation of a random variable, which is $\sum_x xP(X = x)$, expectation of a function of a random variable which is $\sum_x g(x)P(X = x)$ and expectation of the sum, which is the same as sum of expectations.

And we looked at how to apply these concepts. So, the next thing which we are going to look at is what we mean by variance of a random variable. We are going to focus on the variance of a discrete random variable.