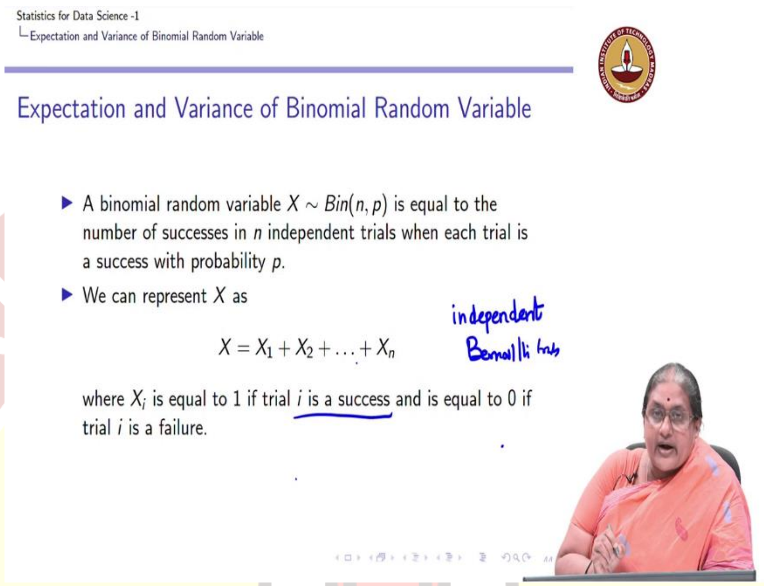
IIT Madras

ONLINE DEGREE

Statistics for Data Science - 1
Professor. Usha Mohan
Department of Management Studies
Indian Institute of Technology, Madras
Lecture No. 10.5
Binomial distribution - Expectation and variance of Binomial random variable

(Refer Slide Time: 0:16)

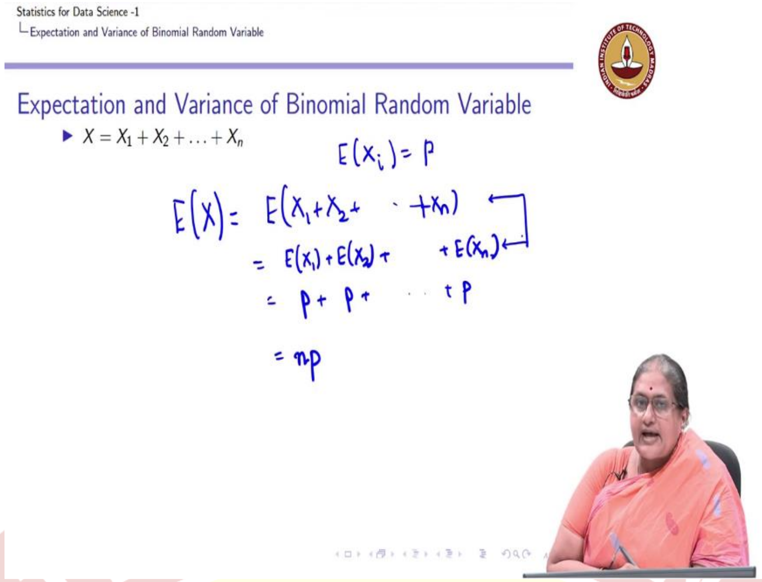


In this section we are going to see about how we can compute the expectation and variance of a binomial random variable and again apply those concepts to answer questions. Now, what is the expectation of a binomial random variable? Now, the expectation of a binomial random variable is equal to the number of successes, so you know a binomial random variable X is equal to the way we have defined a binomial random variable, we have defined it as the number of successes in n independent trial, where the probability of a success is p.

So, now I can represent, we also said the binomial random variable arises as the sum of n independent. Again I am emphasizing on independent and identically distributed Bernoulli trials. Now, I know that if $X_1, X_2, \ldots, X_n$ are independent Bernoulli trials, it takes the value 1 with the success p and $(1-p)$ that is what because I say the probability of a successes with probability p. Expectation of each X_i again we have computed the expectation of each X_i to be $(1 \times p)$ which is p.

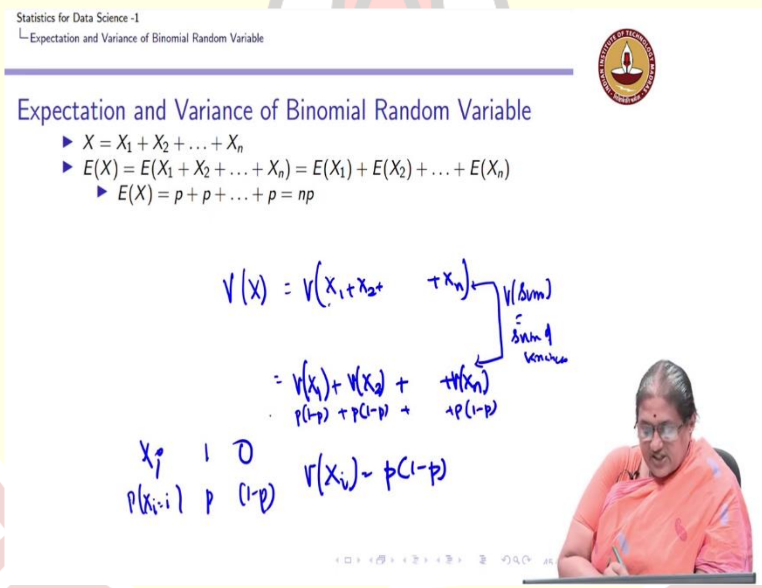
So, I have the expectation of each X_i ,so $E(X_1) = E(X_2) = \cdots = E(X_n) = p.$ And I am expressing my binomial random variable as a sum of n independent Bernoulli random variables.

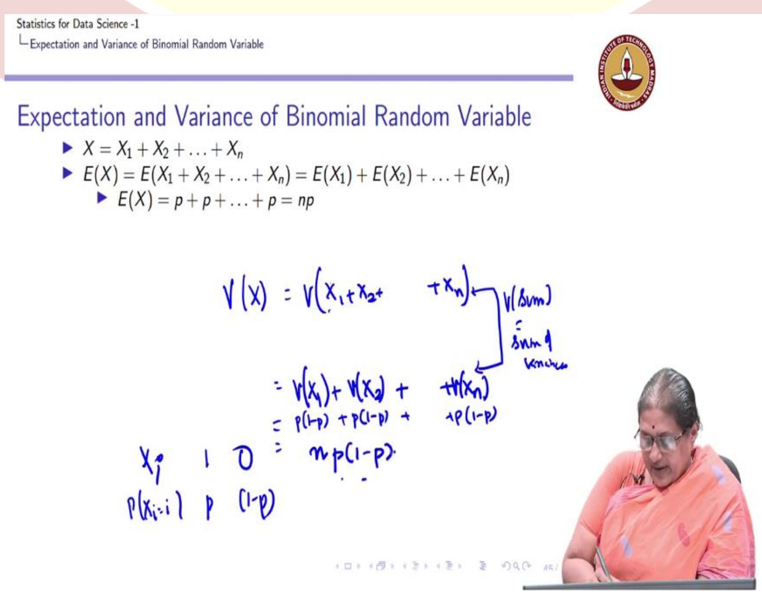
(Refer Slide Time: 2:14)



So, once I have this I have my X which is expressed as a sum of n independent Bernoulli random variables, expectation of each X_i equal to p again from properties of expectation, remember $E(X) = E(X_1 + X_2 + \cdots + X_n)$, which is the expectation of sum is sum of $E(X_1) + E(X_2) + \cdots + E(x_n)$ each one of them is p, so I add p, n times to get $E(X) = np$. The key thing we have used here is to see the expectation of sum is sum of expectations.

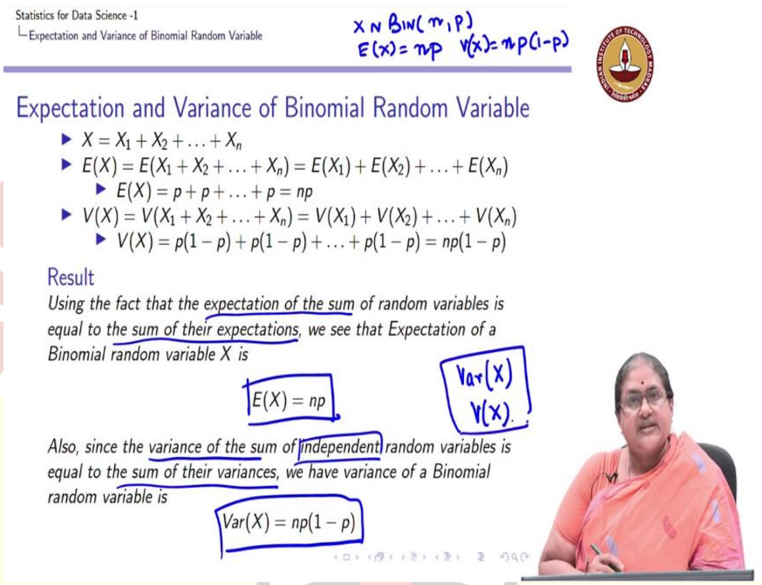
(Refer Slide Time: 3:13)





So, when I am having the sum of expectation, I get $E(x) = np$. Now, if I am looking at the variance of X, I again know $Var(X) = Var(X_1 + X_2 + \cdots + X_n)$. Now, I am going to use the property of variance which says variance of sum equal to sum of variances if my $X_1, X_2, \dots X_n$ are independent random variables, recall this from our properties of variance, if I have n independent random variables then variance of sum is the sum of variances.

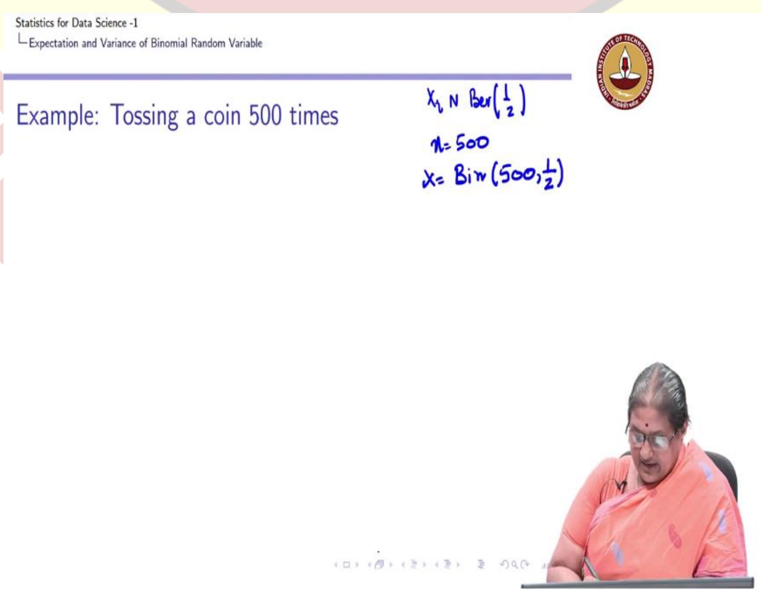
I use this then I get that $Var(X_1 + X_2 + \cdots + X_n) = Var(X_1) + Var(X_2) + \cdots + Var(X_n)$. Now, X_1 is taking the value 1 and 0, X_i take the value 1 and 0 with probability p and $(1-p)$, we have already seen that $Var(X_i)$ is going to be $p(1-p)$ that is the variance of a Bernoulli random variable. So, $Var(X_1) = p(1-p)$, $Var(X_2) = p(1-p)$,…, $Var(X_n) = p(1-p)$, giving the $Var(X) = np(1-p)$, (Refer Slide Time: 4:50)



So, I find $Var(X) = np(1-p)$, $E(X) = np$. So, the key result is we use the fact that expectation of sum is the sum of expectation we get $E(X) = np$ and using the fact that the variance of sum of independent random variables is equal to sum of variances we get $Var(X) = np(1-p)$.

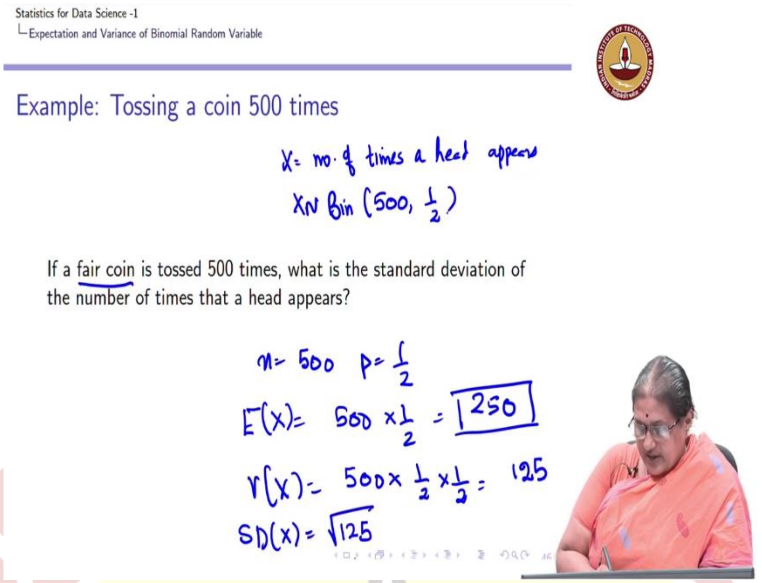
So, in summary if I have X which is a binomial random variable with parameters n and p, the $E(X) = np$ and the $Var(X) = np(1-p)$. Remember I am using $Var(X)$, I use $Var(X)$, or $V(X)$, both of them represent the same.

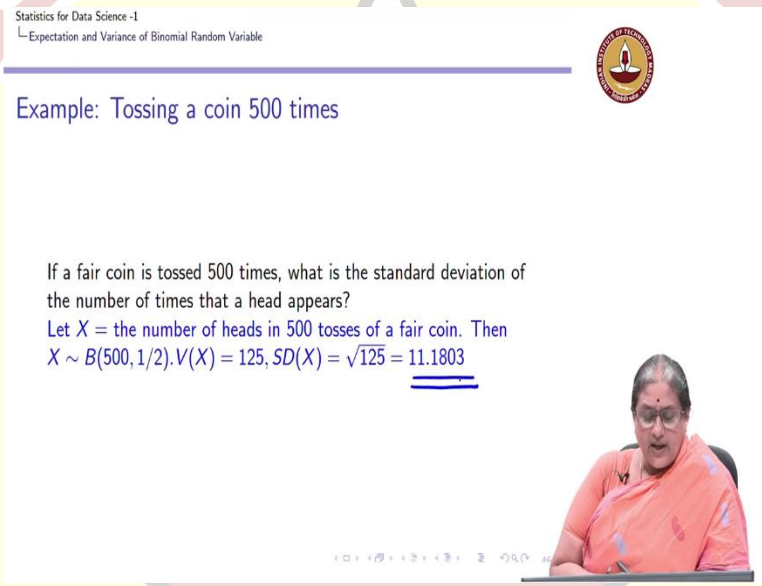
(Refer Slide Time: 5:52)



So, now let us look at a few examples and understand how to use this concept of expectation and variance. I am tossing a coin 500 time, again each time I toss a coin I know I have a Bernoulli trial, it is a fair coin then I have each of my times X_i is a Bernoulli trial with parameter $1/2$. So, if I am tossing a coin 500 time, my n is 500. So, if I am counting the number of successes then X is a binomial with parameter 500 and $1/2$.

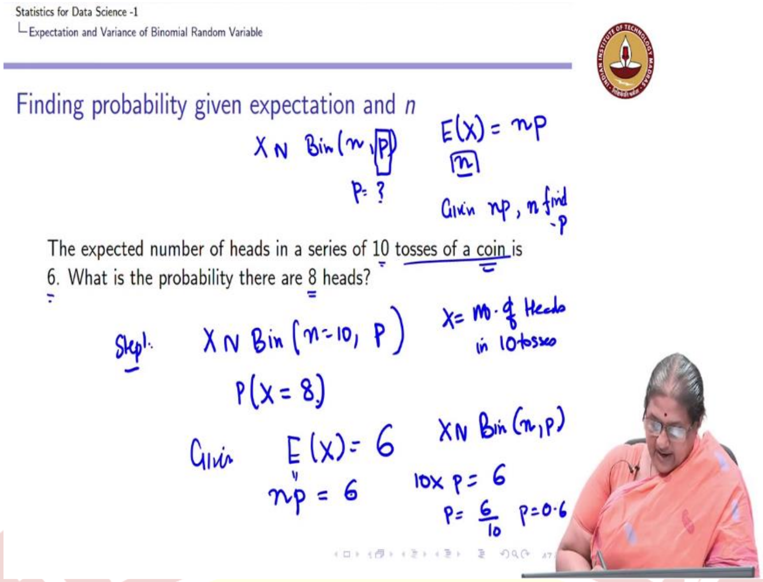
(Refer Slide Time: 6:33)





So, if I am tossing a fair coin then $p = 1/2$, 500 time the question is what is the standard deviation of the number of times a head appears. Again if X is the number of times a head appears, I know X is a binomial with parameter 500 and $1/2$, $1/2$ because I have a fair coin, $n = 500$, $p = 1/2$, my $E(X) = 500 \times \left(\frac{1}{2}\right) = 250$. The way I can interpret this if I keep tossing a coin 500 times I can expect 250 of the outcomes to be head. The variance of X is going to be $np(1-p)$ which is 125. The standard deviation of X is $\sqrt{125}$. So, I get the standard deviation of X is around 11.1803.

(Refer Slide Time: 8:00)

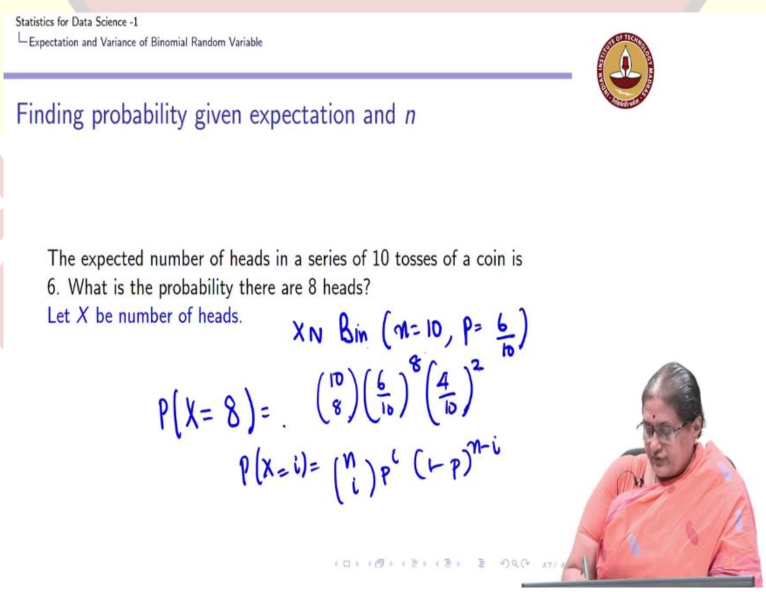


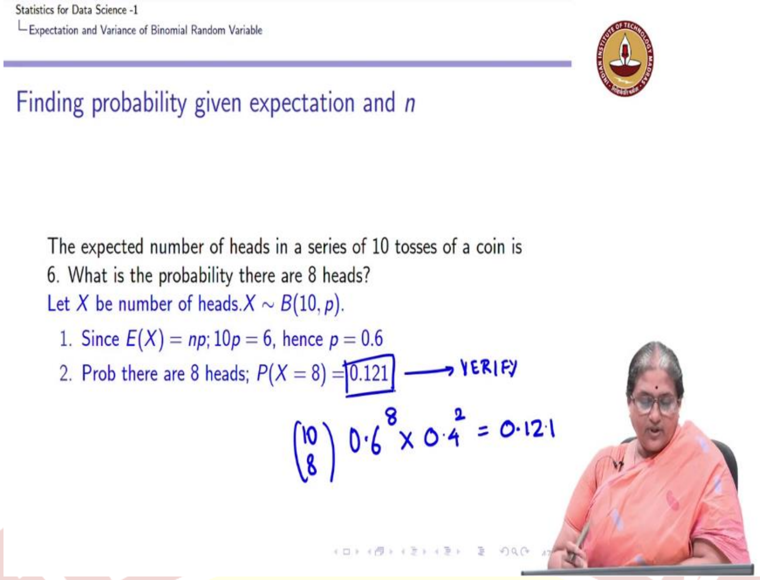
So, now the other thing is given the expectation and the value of n, can I find the probability? What does this mean? Again if I know X is a binomial random variable with parameter n and p. So, I am not given n and p, but I am given $E(X)$ and I know n. Can I determine p? The answer is yes, because I know, if I know $E(X)$ I know np, I know n, I can solve for p.

So, given np and n, I have to find p, that is what it is telling me. So, again the expected number of heads in a series of 10 tosses is 6, what is the probability that there are 8 heads? So, first step X is a binomial. Again I know it, the experiment is a binomial experiment because I am tossing a fair coin 10 times, am I tossing it 10 times, is it fair? We do not know, but I am just trying to, I need to identify what is my p.

So, it is not telling whether, it is just telling a coin, it is not specifying whether it is a fair coin. But I know $n = 10$, let me keep p as it is, I am interested in knowing what is the probability that there are 8 heads where X is the number of heads or successes in 10 tosses? What else is given to us? $E(X) = 6$, that is what we are given. I know if X is a binomial with parameter n and p, np is the expected value of X which is given to be equal to 6, I have n equal to 10 so $10p = 6$, which gives $p = \frac{6}{10} = 0.6$.

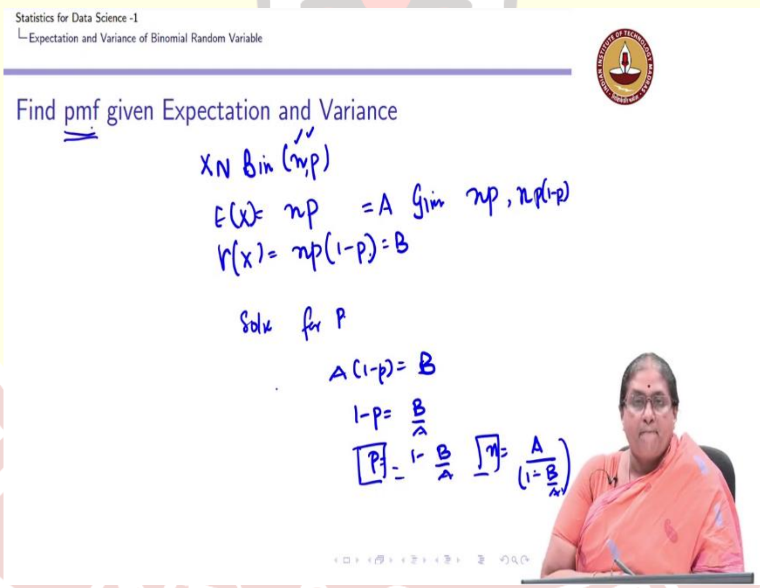
(Refer Slide Time: 10:27)





So, I can identify the distribution of the number of heads as X is a binomial with parameter $n = 10$, p=6/10. I can identify my $X \sim Binomial(n = 10, p = \frac{6}{10})$. Now, I want to know what is the probability that there are 8 heads, so I am interested in knowing this, I can see that $P(X = i)$, I know that the $P(X = i) = \binom{n}{i} p^i (1-p)^{n-i}$, $n = 10, i = 8$, so I get $P(X = 8) = \binom{n}{8}(\frac{6}{10})^8 \left(1 - \frac{6}{10}\right)^{n-8}$ which you can verify is equal to 0.121, I leave this as an exercise to verify that $P(X = 8) = \binom{10}{8} 0.6^8 (0.4)^2 = 0.121$.

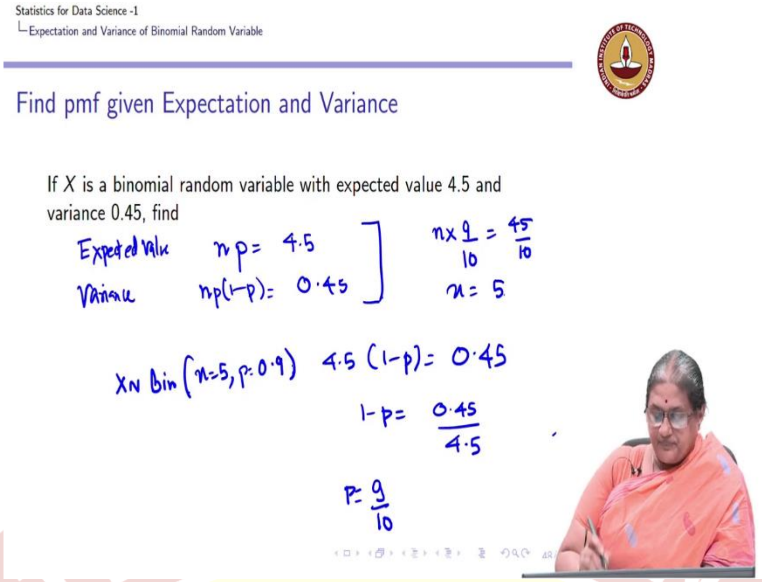
(Refer Slide Time: 12:00)



Now, sometimes we are given expectation and variance of a random variable. So, again if I know $X \sim Binomial(n, p)$, $E(X) = np$, $Var(X) = np(1-p)$. So, given np and $np(1-p)$. So, suppose I am given this is A and this is B, I can solve for p which is very simple, I know that $np = A$, so I know $A(1-p) = B$, which gives $1 - p = \frac{B}{A}$ or $p = 1 - \frac{B}{A}$.

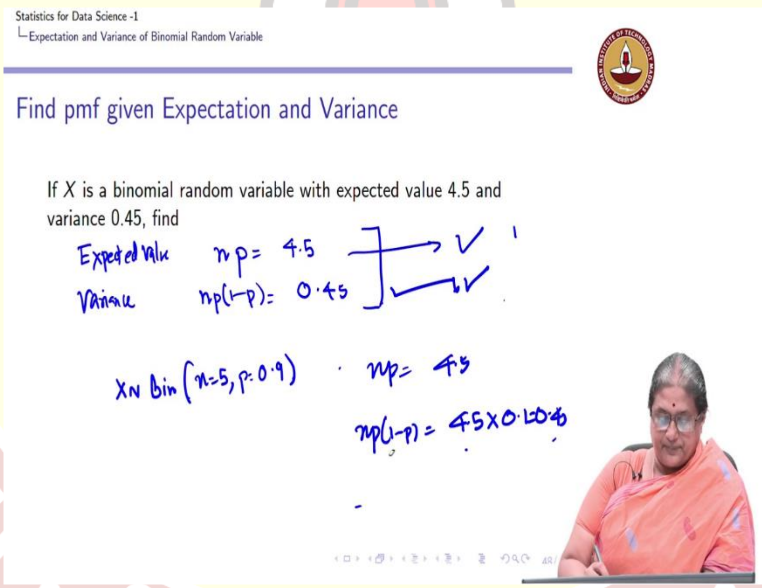
So, we can solve for p once I know p, I can solve for A, $A = np \Rightarrow p = \frac{A}{n}$, I know already $p = 1 - \frac{B}{A}$. So, given the expectation and variance I can solve for both p and n. I can solve for p and n. Now, if I am given p and n, I can find out what is the probability mass function of the binomial random variable.

(Refer Slide Time: 13:28)



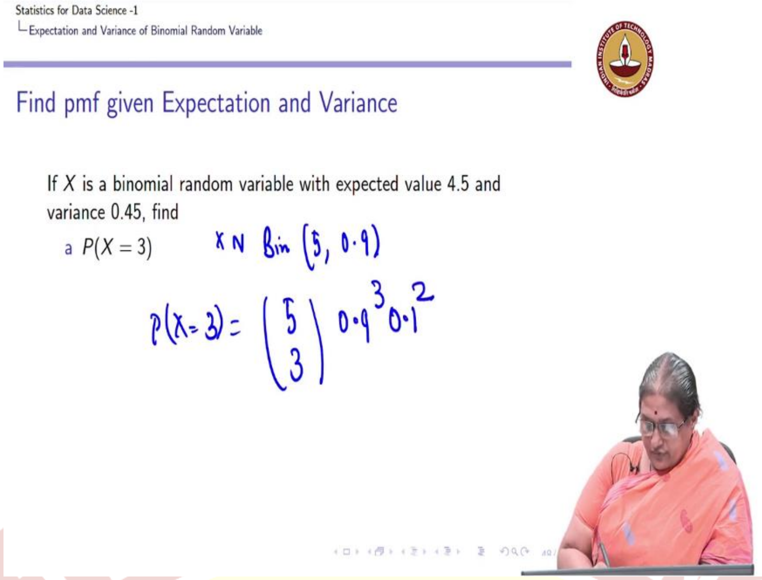
Now, let us look at an example, I know that X is a binomial random variable, $np = 4.5$, $np(1-p) = 0.45$, this is the expected value and this is my variance. So, from these two, I can obtain that $4.5p(1-p) = 0.45$ which gives me, $1-p = \frac{0.45}{4.5}$ which gives me that p equals to, which gives me $1-p = \frac{0.45}{4.5}$, so $p = \frac{9}{10}$. If $p = \frac{9}{10}$, I know $n \times \frac{9}{10} = 4.5$, which gives me $n = 5$. So, I have X is a binomial random variable with parameter $n = 5, p = 0.9$.

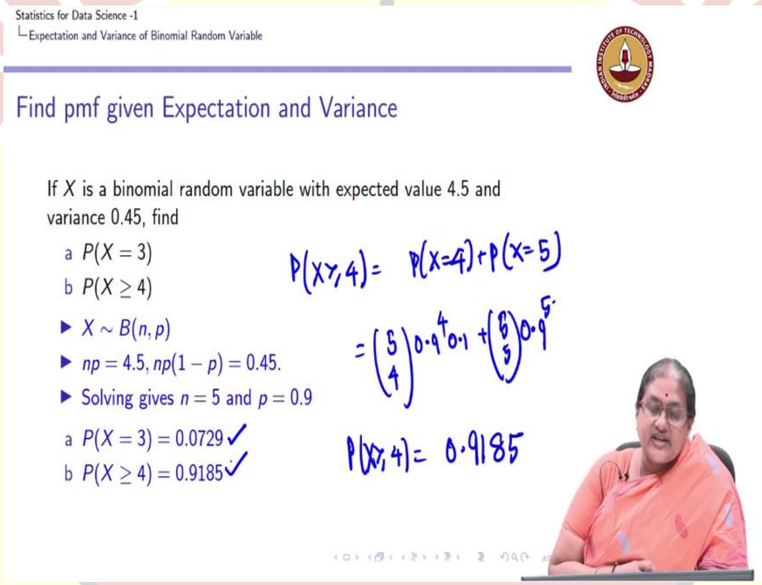
(Refer Slide Time: 14:59)



I can verify the expected value is 4.5 which is np, so I can verify np is 4.5 which is what I have here $np(1-p) = 4.5 \times 0.1 = 0.45$, which is what I have here.

(Refer Slide Time: 15:22)

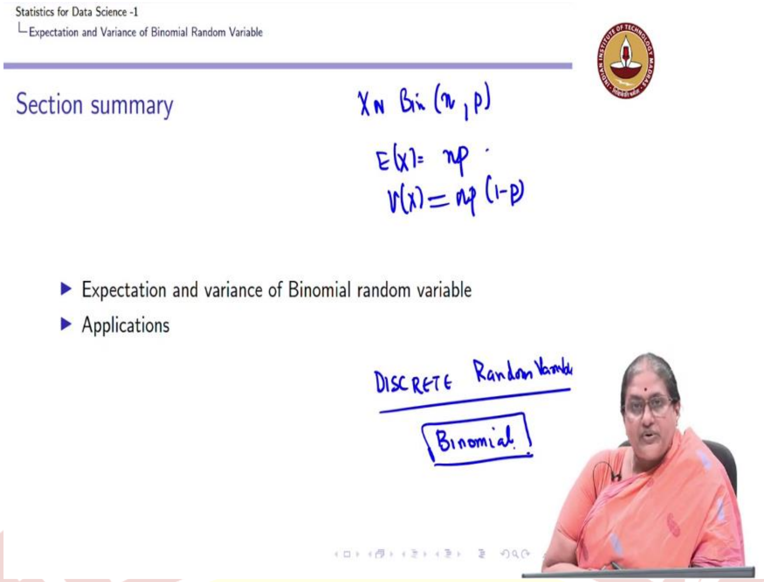




So, once I identify and solve for the unknowns n and p, I am interested in knowing what is $P(X = 3)$, I know X is a binomial with parameter 5 and 0.9, $P(X = 3) = \binom{5}{3}(0.9)^3(0.1)^2$ which I am going to get and I can verify that, that is 0.0729.

Similarly, I leave it as an exercise $P(X \leq 4) = P(X = 4) + P(X = 5)$, $P(X = 4) = \binom{5}{4}(0.9)^4(0.1)^1, P(X = 5) = \binom{5}{5}(0.9)^5$ and we can verify that $P(X \geq 4)$ is 0.9185, which is the answer I get here. So, given the expectation and variance I can obtain the probability distribution of my random variable.

(Refer Slide Time: 16:40)



So, in summary what we have learnt here is what are expectation and variance of a binomial random variable. Remember if X is a binomial random variable with parameter n and p, $E(X) = np$, $Var(X) = np(1 - p)$. So, given the expectation and n we saw how to obtain p, given expectation and variance you can solve for both n and p and obtain the probability distribution.

So, with this we complete the discussion on binomial distribution. So, the path we took was, we looked at discrete random variables and in the real, in the domain of discrete random variables, we focused on mainly certain applications which follow the binomial distribution. The next thing we are going to discuss is about continuous random variables.