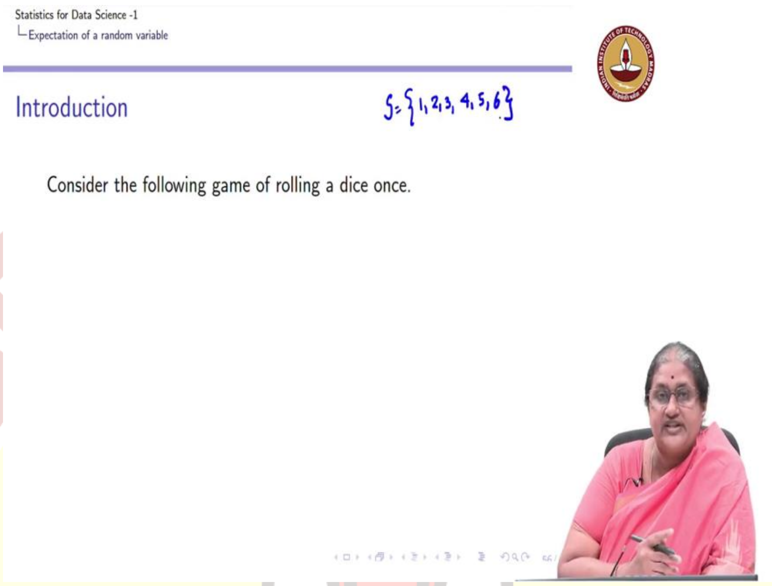
IIT Madras

ONLINE DEGREE

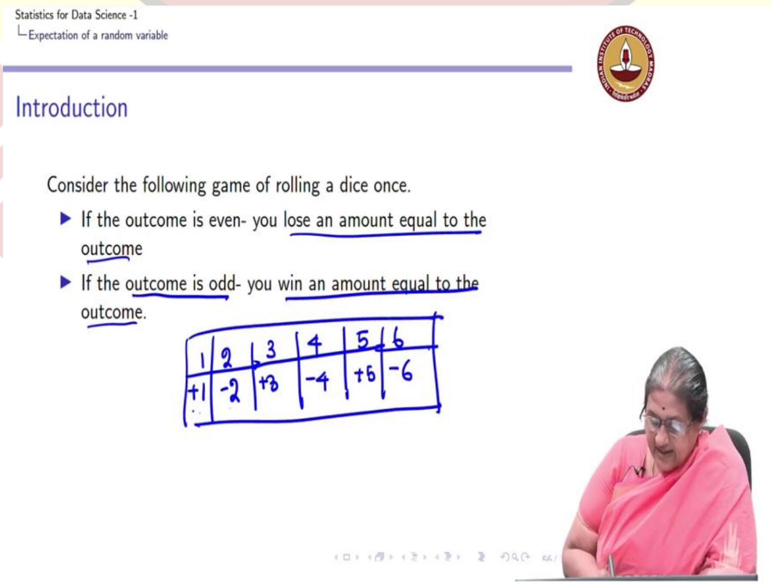
Statistics for Data Science - 1
Professor Usha Mohan
Department of Management Studies
Indian Institute of Technology, Madras
Lecture 9.2 - Expectation of a random variable

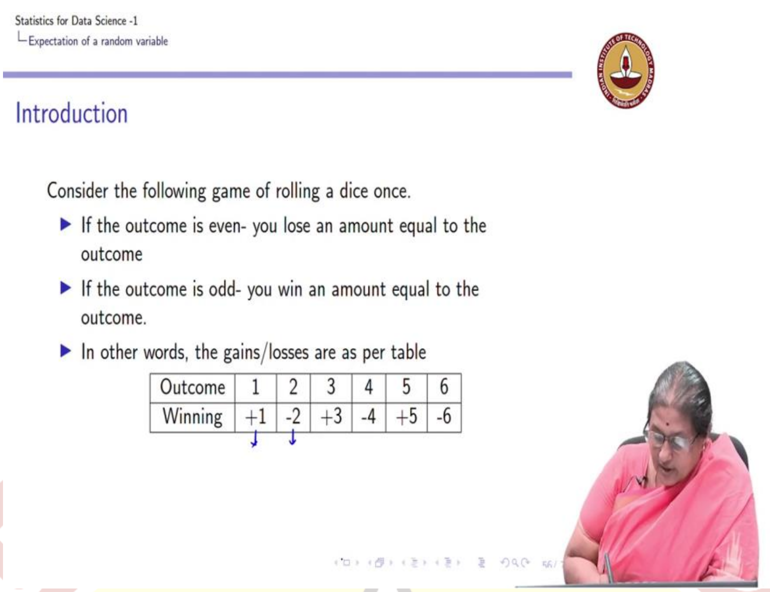
(Refer Slide Time: 00:14)



So, now we are going to introduce a very important notion that is the expectation of a random variable. Now, consider the following game of rolling a dice once, it is a fair dice so I am just rolling it once, I know my sample space here is going to be any one of the outcomes and the outcomes are 1, 2, 3, 4, 5, 6. I am talking about this 6 sided fair dice.

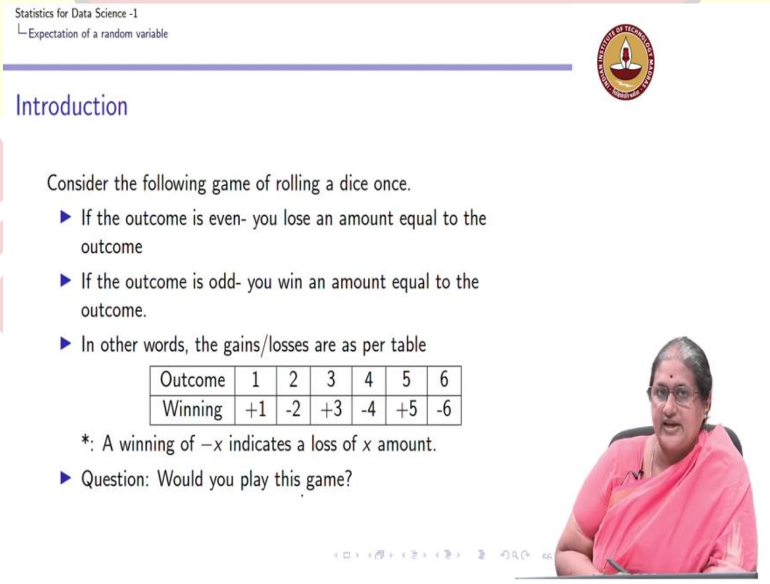
(Refer Slide Time: 00:45)





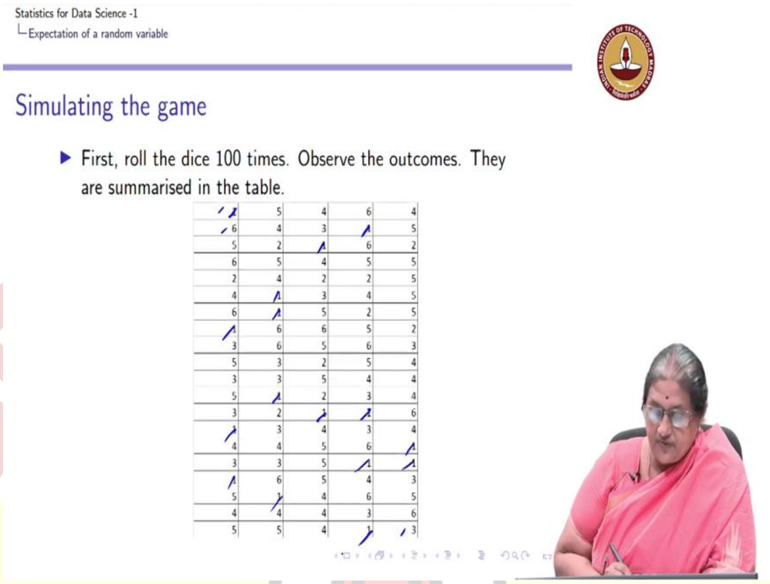
Now, let us play this game. If the outcome is even, you lose an amount equal to the outcome. If the outcome is odd, you win an amount equal to the outcome, my outcomes are 1, I lose, it is odd, so I win an amount of +1, outcome is 2, I lose an amount so if I win, I am going to put a +1, if I lose, I am going to put a minus, 3 I again, it is odd so I win +3, 4 is even I lose -4, 5 is odd I win +5, and 6 is even I lose -6. So, there is an element of chance. And I can tabulate or summarise my outcome, as I term my negative winning is a loss, positive winning is actually a gain. So, these are my outcomes.

(Refer Slide Time: 01:47)



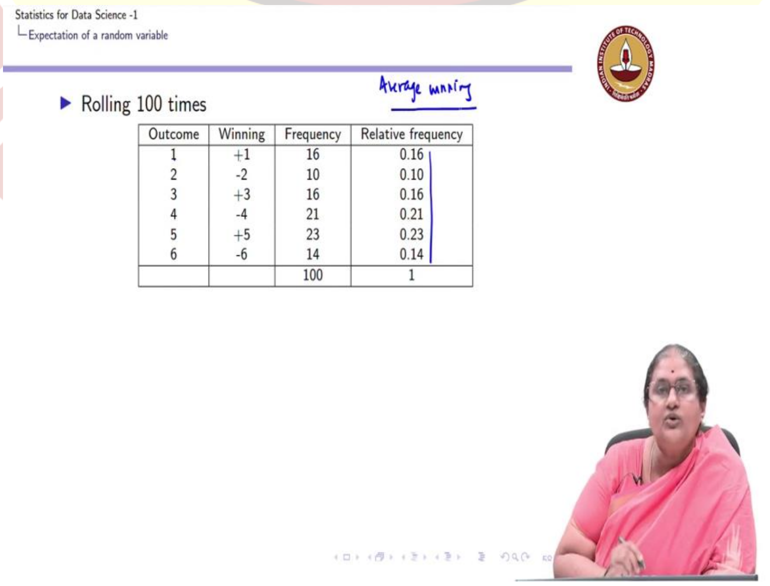
So, the question we are asking here is, would you play this game? And intuitive answer to this is I will play if I am a rational thinker, if I am a rationally thinking person, I will play a game only if I am expected to win this game. So, I am again telling something about expectation. So, now let us go and analyse this problem from a probability point of view.

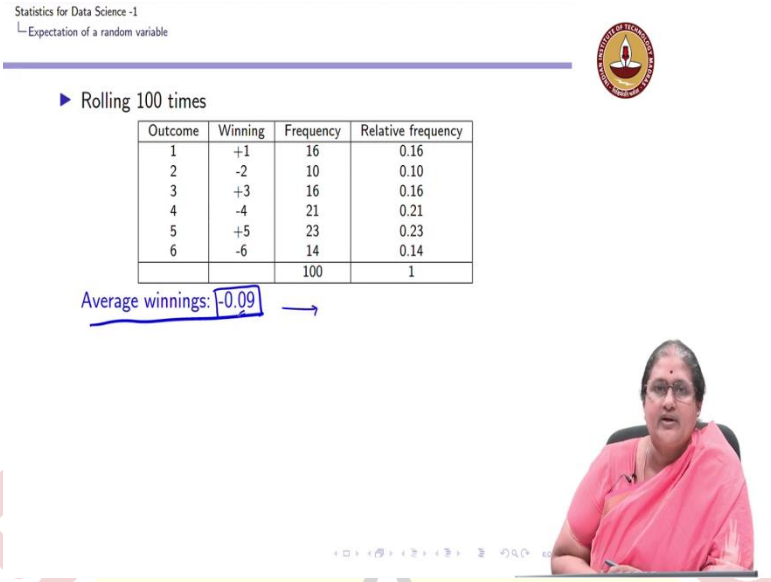
(Refer Slide Time: 02:24)



So, I want to know whether I should play this game. So, I am going to just roll the dice 100 times, once I role a dice 100 times and observe the outcomes. So, you can see that I have 100 outcomes which I have summarised here. So, this is the first outcome was a 1, second outcome was a 6, so forth my 100th outcome was a 3. So here, what I do next is I observe how many times are the frequency distribution of these outcomes, I note down 1 appears 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16 times, my 1 appears 16 times.

(Refer Slide Time: 03:17)

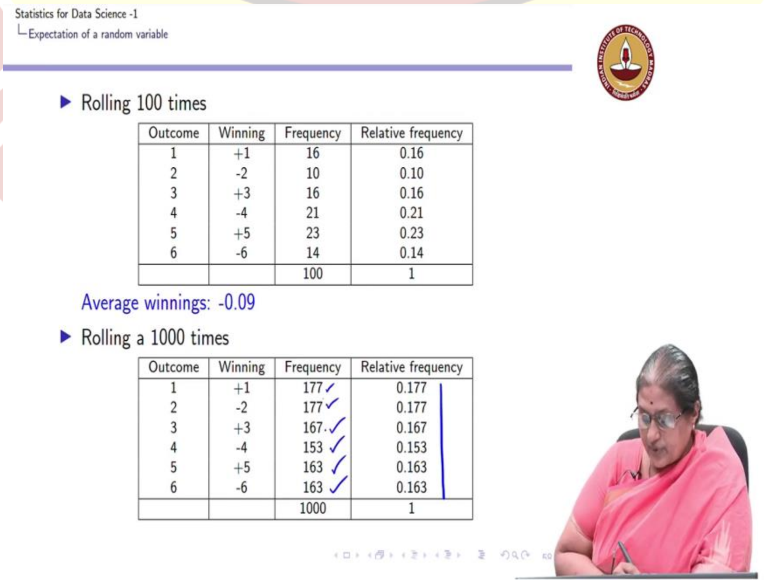


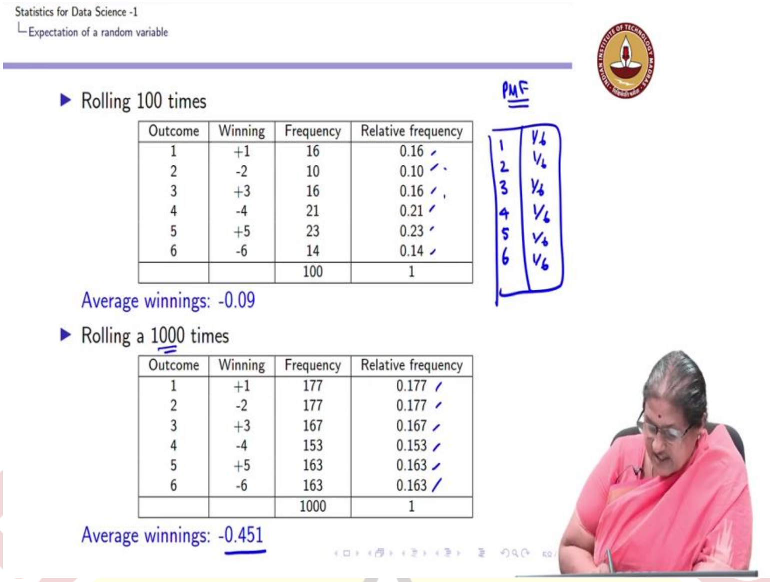


So similarly, I plot the frequency distribution for each one of these outcomes, 1 appears 16 times, 2 appears 10 times, from the frequency I can obtain the relative frequency, this is something which we have already seen how to get it. And in addition, I am going to see that from the frequency and relative frequency, I can compute what is my average winning, that is, I know that plus 1 my relative frequency is 0.16, I find out what is the average, we also know how to compute the averages. And you can check that the average winning is -0.09.

So, I am actually the average winning is a negative. So, this might say that, you do not play the game because it is negative. But again, it is not that negative, so I might be tempted to play the game.

(Refer Slide Time: 04:19)

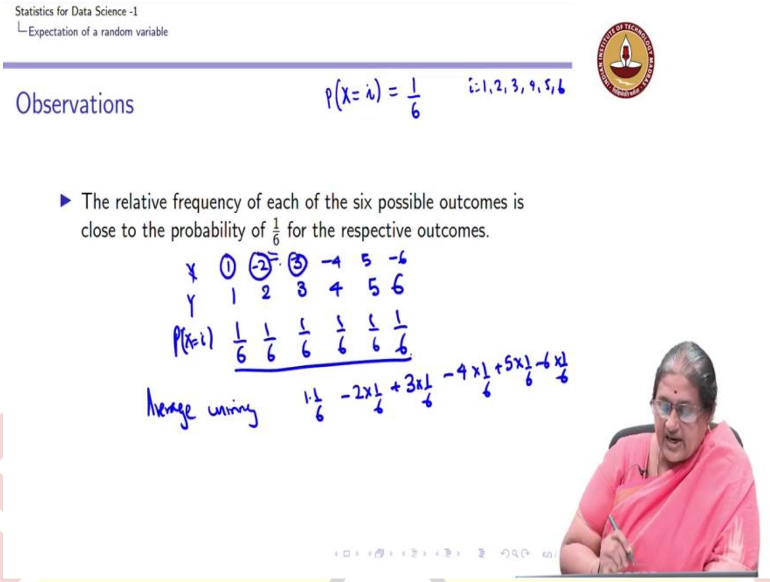




So, I continue the experiment, I roll the dice for 1000 times. Now, when I roll the dice for 1000 times and I count the frequency. Again, I repeat the experiment. Now, you can see that 177 times I had a 1 appearing, 177 times a 2, 167 times a 3, 153 times of a 4, 163 times a 5 and 163 times a 6, I count the number of times each outcome appears. Again, I compute, you can see the relative to frequency which I have computed here.

And if you compute the average winning, it is -0.451, which again, tells me that my in total or an average my winnings is negative. So, it does not help me or does not give me an indication of a positive gain. So, I might not decide to go ahead and play the game. Now, you can notice something here, this 0.16, 0.1, 0.21 all these relate to frequencies, which I have here, the relative frequency when a role for a longer number of time, and you can recall what was the probability mass function of the outcome of a die I had the following 1, 2, 3, 4, 5, 6 with probability $\frac{1}{6}$, $\frac{1}{6}$, $\frac{1}{6}$, $\frac{1}{6}$, $\frac{1}{6}$ and $\frac{1}{6}$.

(Refer Slide Time: 06:11)

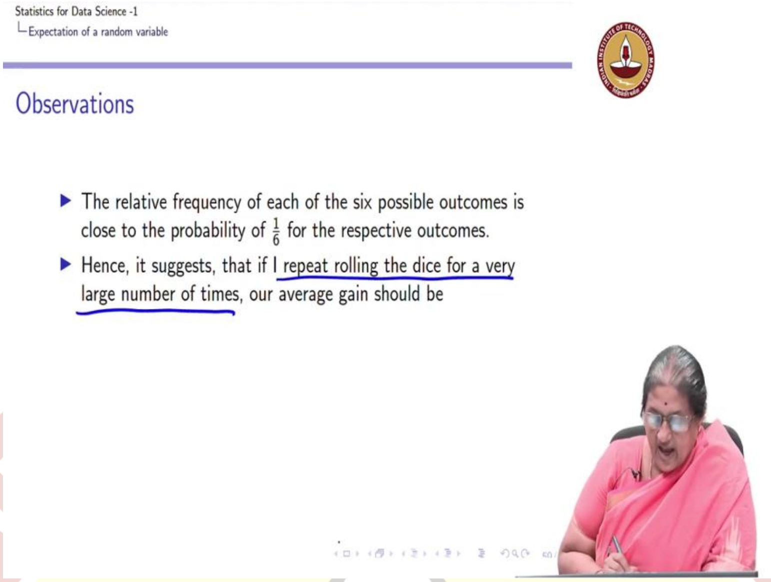


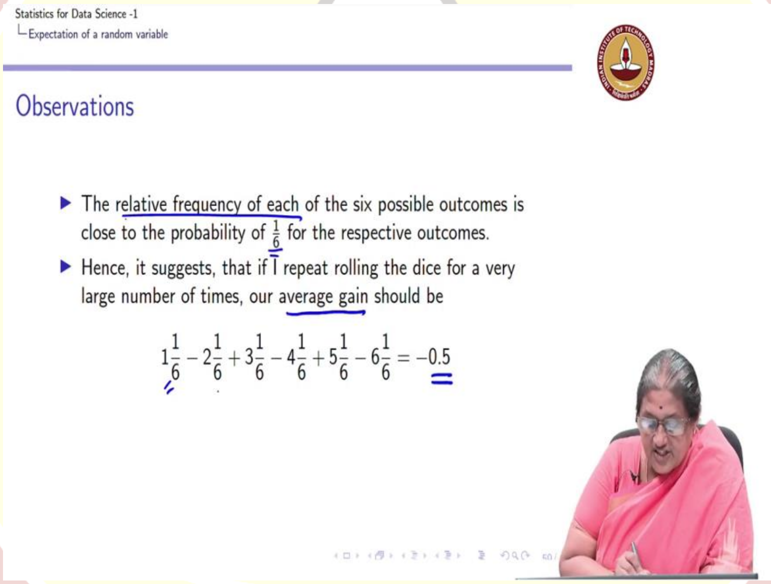
So, you can recognise that what is stated as the relative to frequency is precisely this $\frac{1}{6}$, or it is very close to this $\frac{1}{6}$. In other words, the probability of the random variable which is the outcome, taking a value i, which is equal to $\frac{1}{6}$, i going from 1, 2, 3, 4, 5, 6 is the same as the relative frequency, this is something which we observe.

So, if I have x, so I can stake, x takes my value 1, 2, 3, 4, 5, 6. With the probability x equal to i, $\frac{1}{6}$, $\frac{1}{6}$, $\frac{1}{6}$, $\frac{1}{6}$, $\frac{1}{6}$, and $\frac{1}{6}$. Now, if I am looking at this as winning, so let me take it as Y takes the value 1, -2, +3, -4, 5, and -6, these are my winnings.

So, I can see that in expectation with the same probability or let me call this X and I can call this Y, it does not matter to me. So, I can see that this relative or my average winning can be rewritten as $1 \times \frac{1}{6}$ - $2\times\frac{1}{6} + 3 \times \frac{1}{6} - 4 \times \frac{1}{6} + 5 \times \frac{1}{6} - 6 \times \frac{1}{6}$, where these 1 by 6 are the weights associated with the values the random variable is taking.

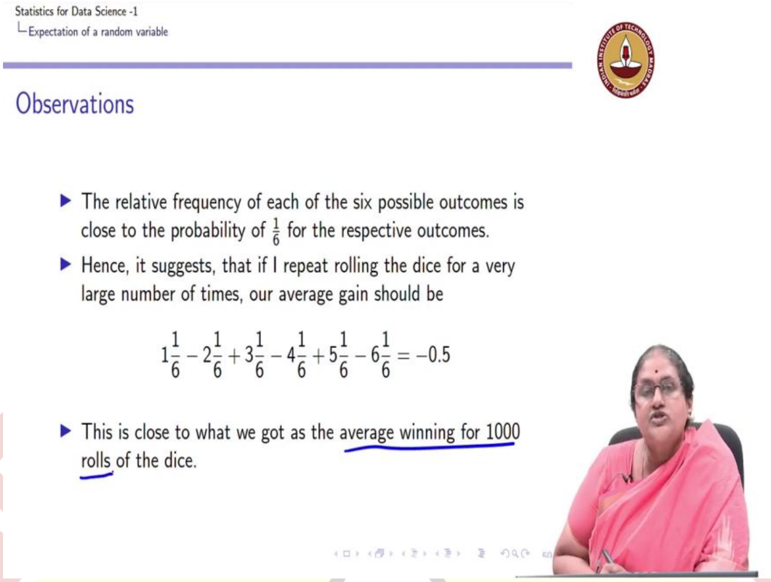
(Refer Slide Time: 08:00)





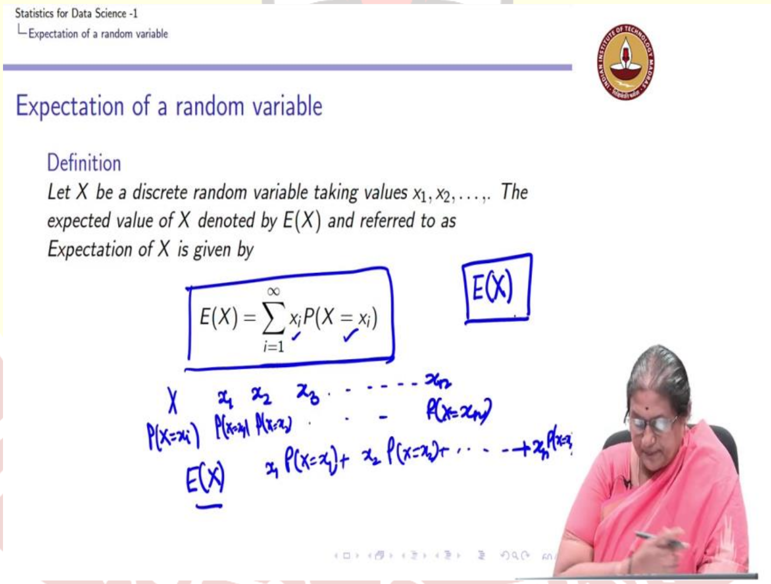
So, it suggests, what is it suggests that if I repeat rolling the dice for a very large number of time, then my average gain would be the following, which is -0.5 I repeat. So, it suggests that if the relative frequency is close to $\frac{1}{6}$, that is what we see. And if I repeat rolling the dice for a very large number of times, then the average gain would be -0.5.

(Refer Slide Time: 08:40)



This is very close to what was our average winnings of 1000 rolls of a die, you recall, the average winning of 1000 rolls of a die was -0.451.

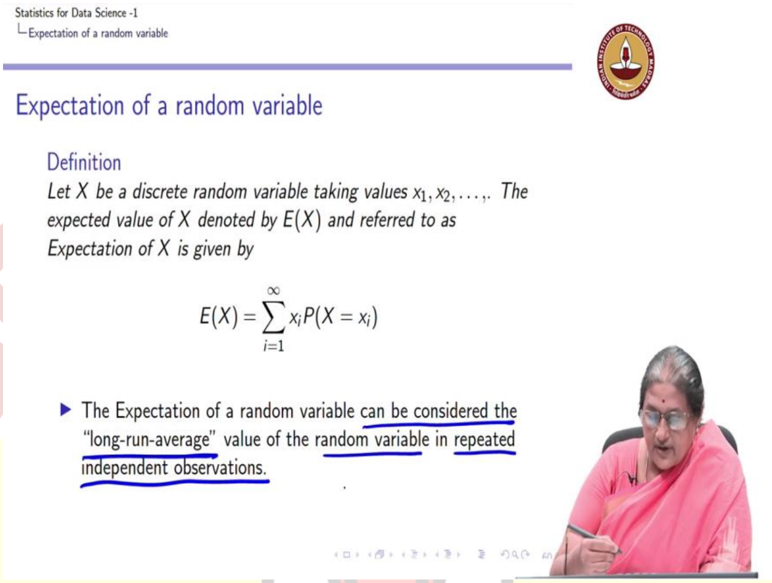
(Refer Slide Time: 08:56)

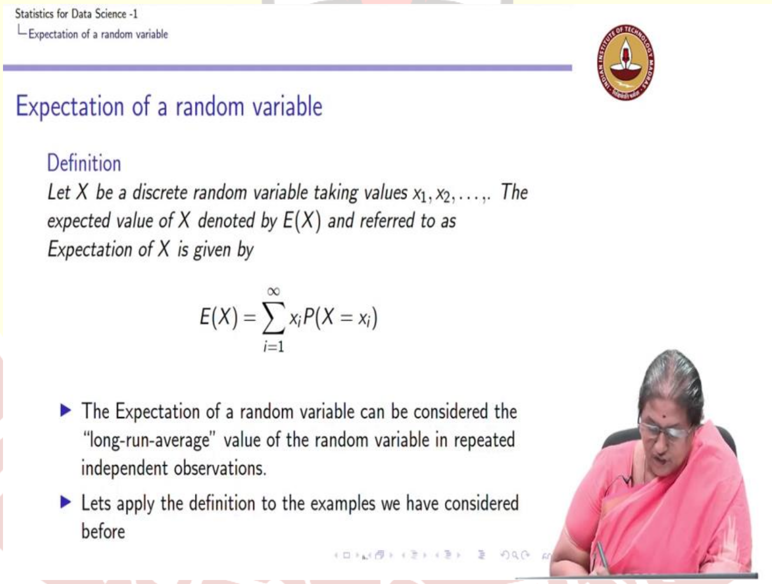


So here, you can see that the expectation of a random variable, I can define the expectation of a discrete random variable. Suppose X is a discrete random variable, which takes values x_1, x_2, x_3 and so forth uncountably many, countably infinite with respect to probability x_1, probability x equal to x_2, so forth, then the expectation of the random variable is x1 into probability x takes the value $x_1 + x_2$ into probability x takes the value x_2 and so forth, which can be abbreviated as summation i going from 1 to infinity x i probability x equal to x_i.

If x takes finite number of values with the same probability, then I have $E(X) = x_n \times P(x = x_n)$. So, this is what we refer to as the expected value of a random variable, we also refer to as expectation of X, and it is written or it is denoted by $E(X)$.

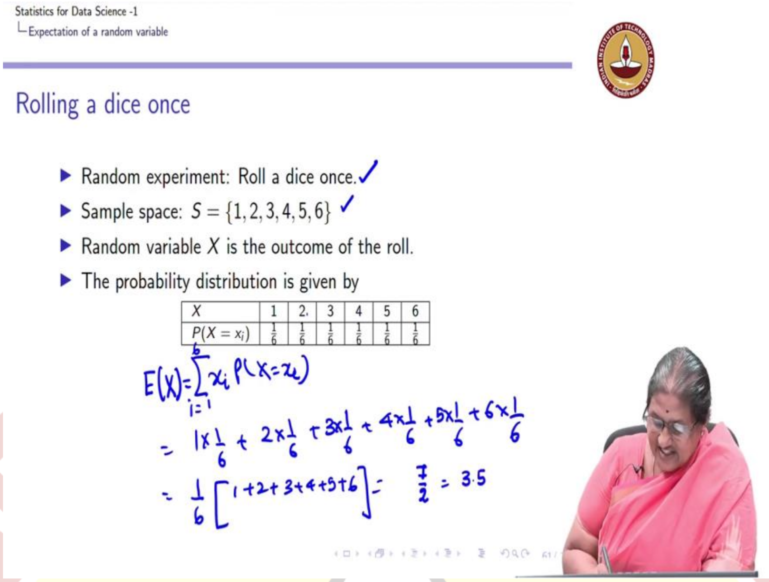
(Refer Slide Time: 10:25)

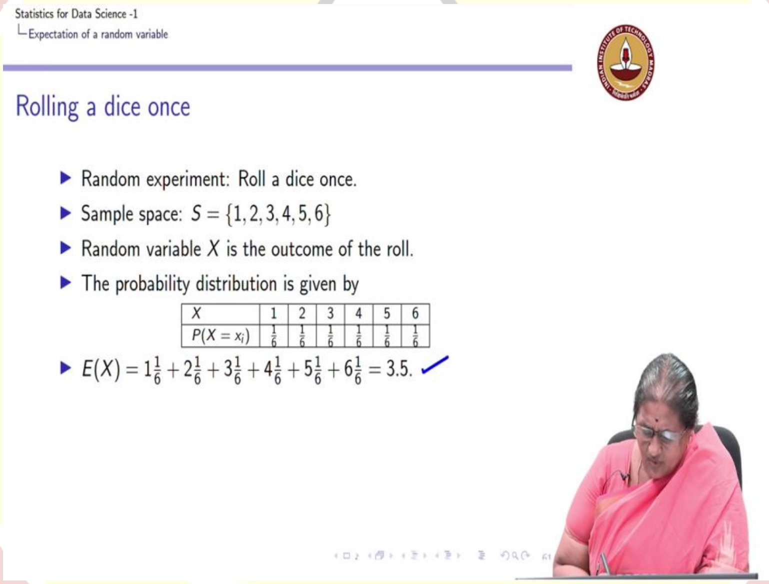




The expectation of a random variable is can be considered to be the long run average value of the random variable in repeated independent observations. So, if I am rolling a die, then I can expectation of the random variable can be viewed as a long run average.

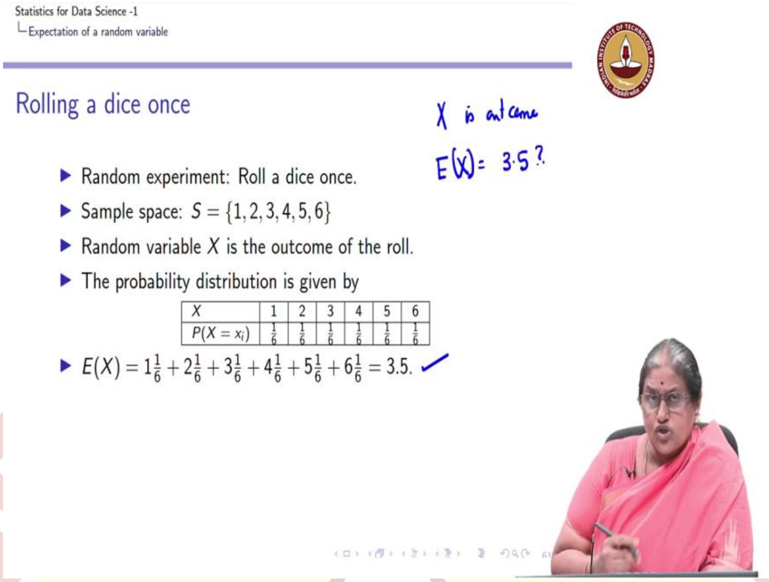
(Refer Slide Time: 10:55)

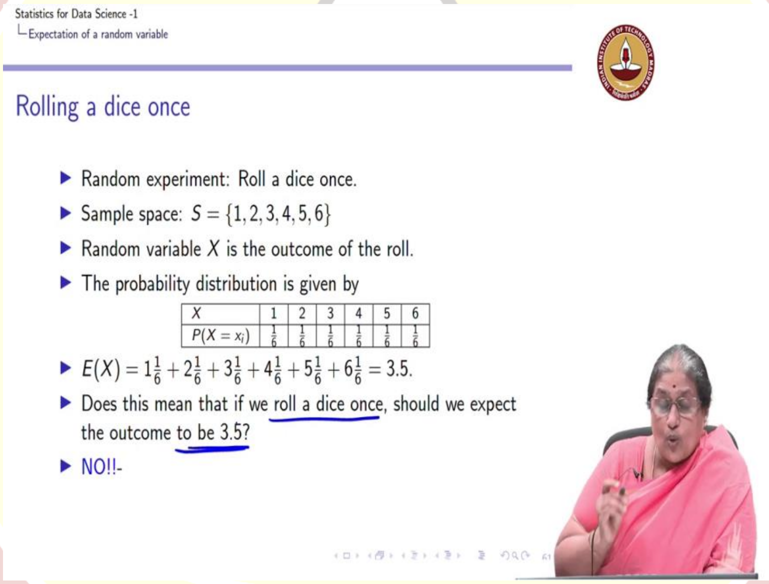


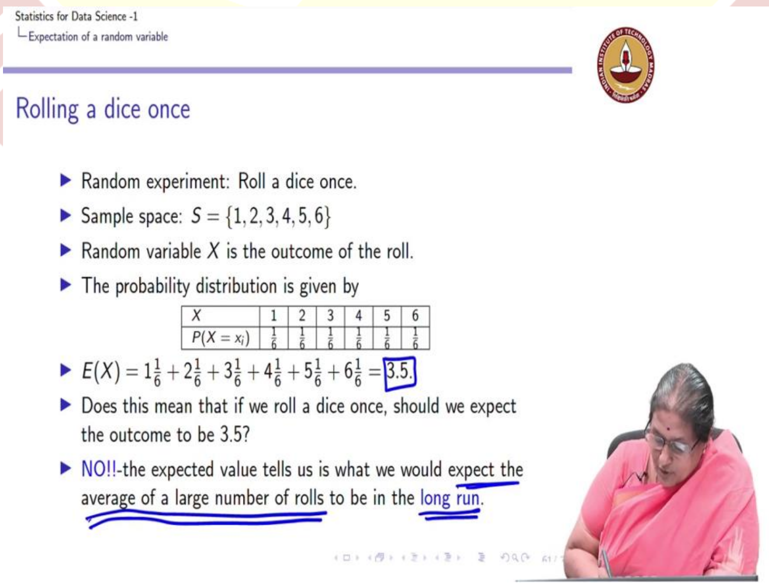


Let us look at a few examples. To understand this concept. Again, I roll a dice once I know the sample space is 1, 2, 3, 4, 5, 6, the outcome is 1, 2, 3, 4, 5, 6 equally likely with this is the probabilities. So, from my definition, my expectation of X is x_i probability x equal to xi summation i going from 1 to 6, which is going to be $1 \times \frac{1}{6} + 2 \times \frac{1}{6} + 3 \times \frac{1}{6} + 4 \times \frac{1}{6} + 5 \times \frac{1}{6} + 6 \times \frac{1}{6}$. So, I can take this $\frac{1}{6}$ out, $1 + 2 + 3 + 4 + 5 + 6$, which I can say is $\frac{7}{2}$, which is 3.5. So, the expectation of this random variable is 3.5.

(Refer Slide Time: 12:00)



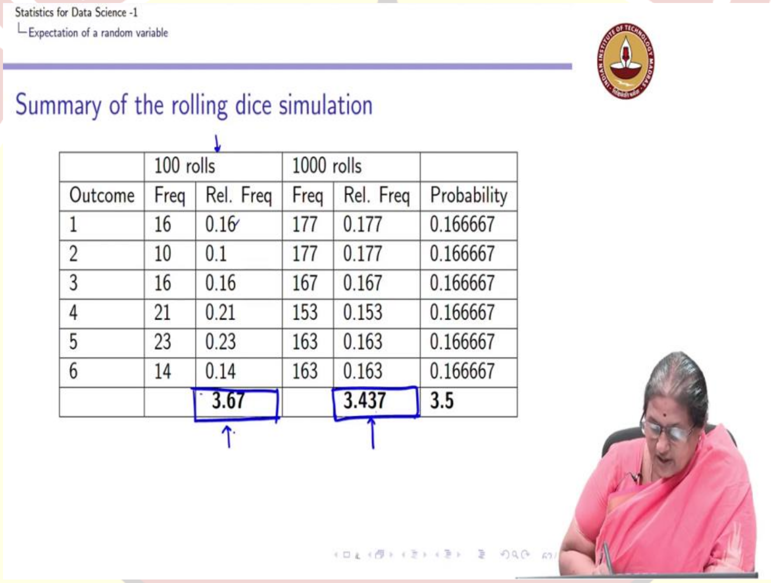


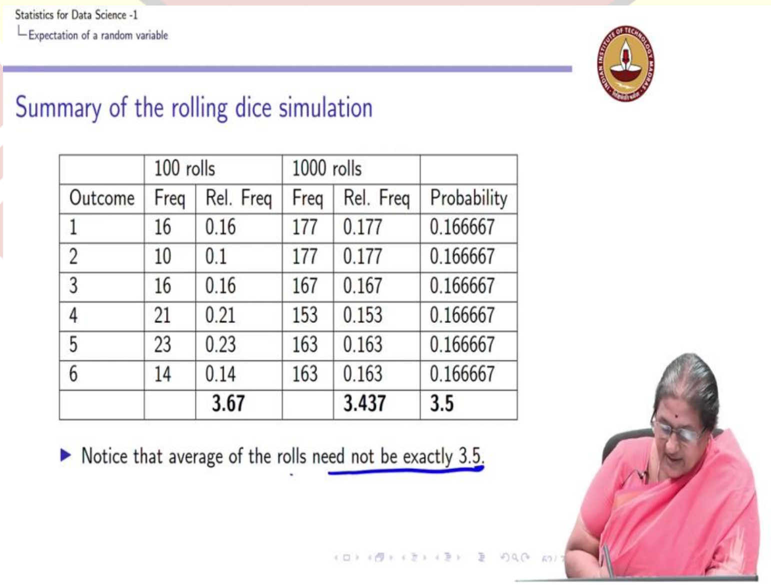


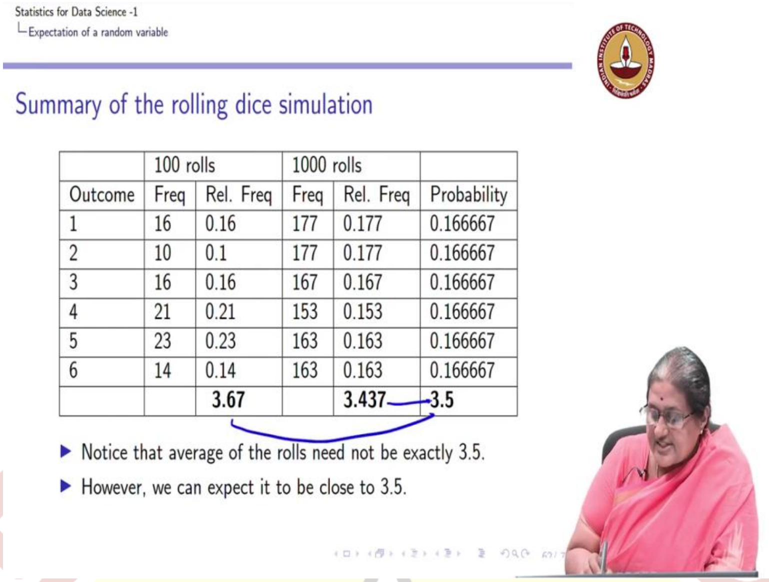
So now, let us interpret this random variable, X is the outcome of a role of a die, I am getting an expected value of X to be 3.5. So, would it be right for me to say that, if I roll a dice sufficiently enough number of times then I will get an outcome of 3.5? The answer is no. It does not mean that if you roll a die once you can expect the outcome to be 3.5. That is not what this 3.5 means.

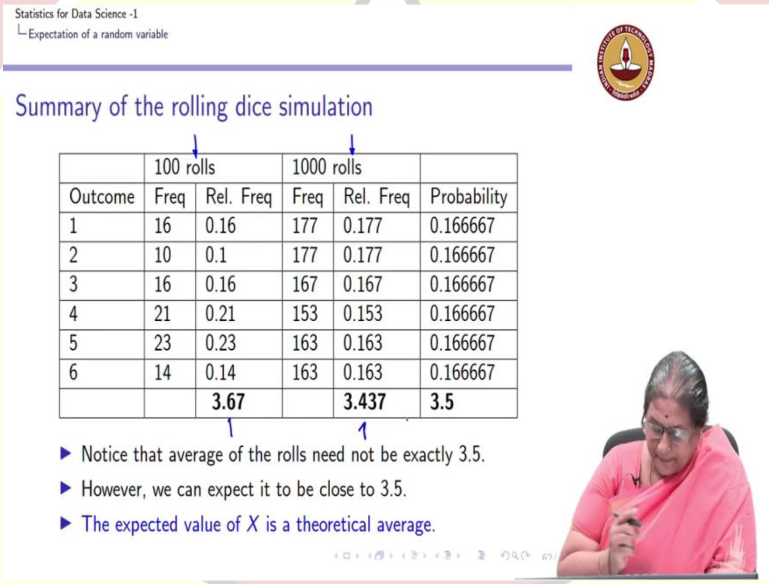
What it means is, if I roll a dice, sufficient number of times, that is I keep rolling a dice, then the average of a large number of rolls, the average of a large number of rolls, you can expect that to be 3.5 in the long run, it does not mean that you will have an outcome which is 3.5.

(Refer Slide Time: 13:05)





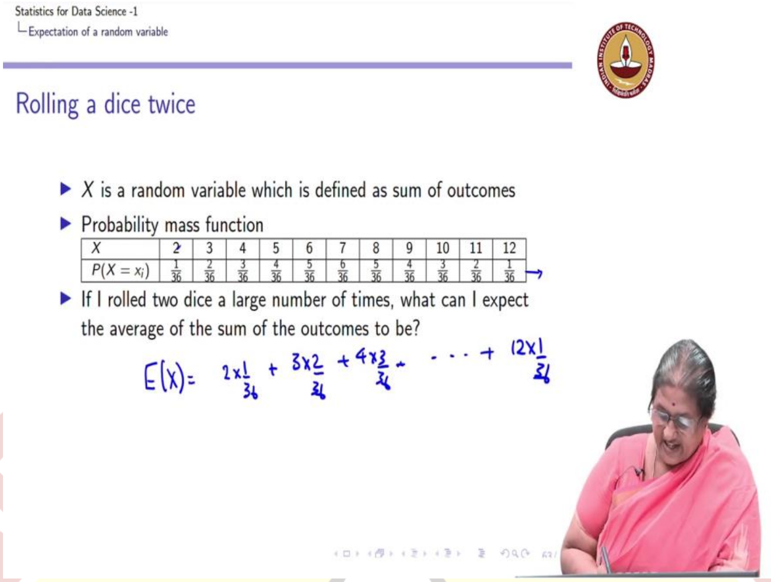


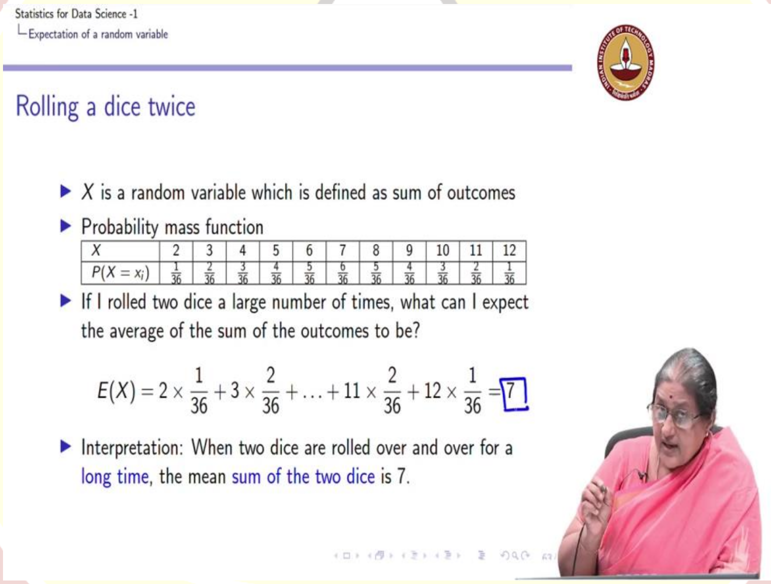


Let us look at the probability mass function. So, you can see that when you are looking at so now, if I summarise the rolling dice simulation, you can see in the relative and the relative frequency for a 100 roll the average was about 3.67. Here it was 3.437. So, can the 100 rolls there will the average value be exactly equal to 3.5 need not be here you can see it is 3.67 here you can see 3.47.

So, you can see that it need not exactly be 3.5, it would be close to 3.5, both 3.67 and 3.437 are close to 3.5 an expectation, or the expected value of X is a notion of a theoretical average. And you can see that when you are actually conducting the experiment of rolling a die, I could get values which are not exactly 3.5, but close to 3.5.

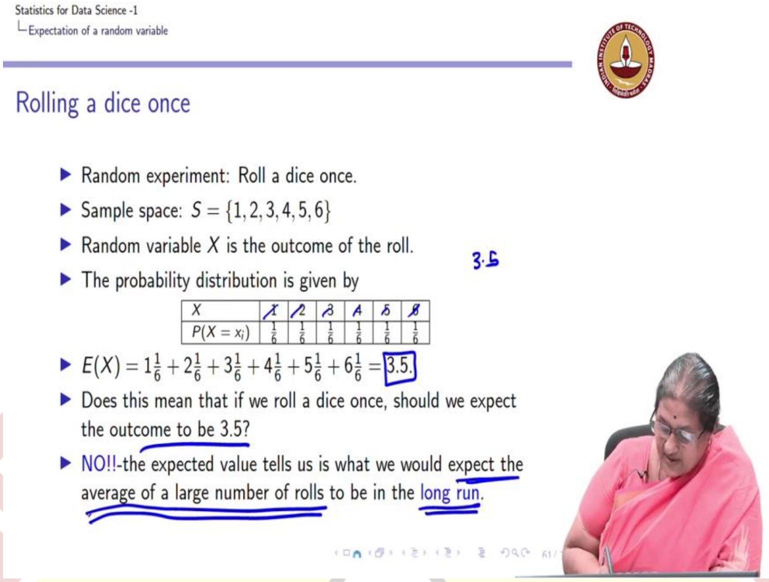
(Refer Slide Time: 14:18)





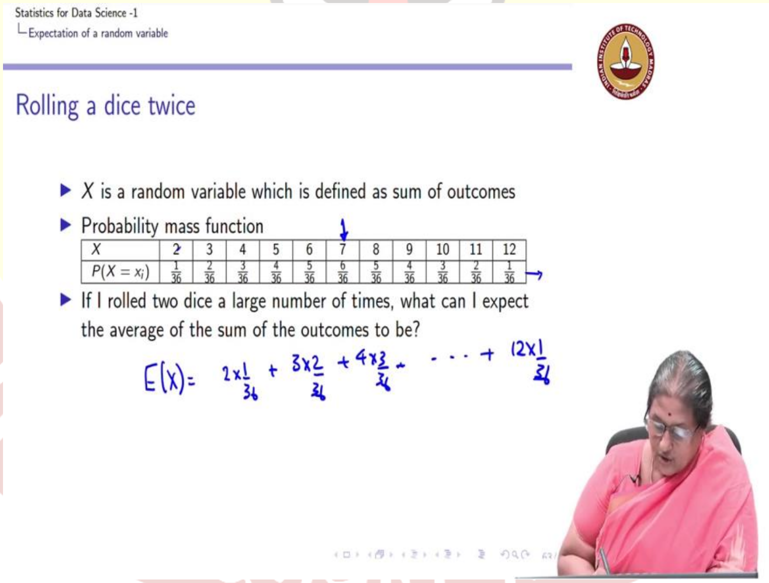
So, let us look at a random variable which I am rolling it twice. Again, recall we defined this random variable as the sum of outcomes. We saw that this was the distribution of this random variable. It takes values 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 with these probabilities. So, now the question is if I continue to roll a 2 dice for sufficiently long period of time, what can I expect the average of the sum of outcomes to be again expected X gives me the answer. So, I look at it, it would be 2 into 1 by 36 plus 3 into 2 by 36 plus 4 into 3 by 36 12 into 1 by 36 . And I can verify that this expected value of x is equal to 7.

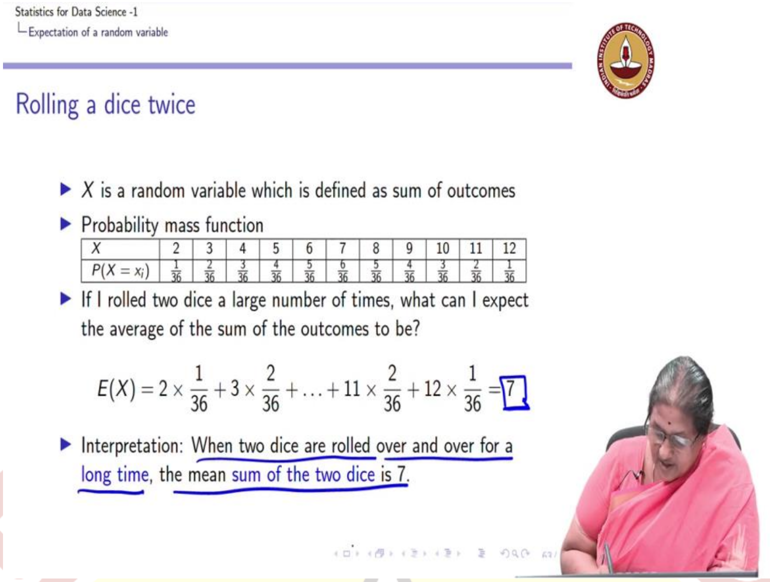
(Refer Slide Time: 15:28)



Notice that in the earlier case, the expected value of X need not be a value of X, it was 3.5 X does not take a value of 3.5. So, I am not saying that X cannot take a value of 3.5, we are looking at a average of these outcomes.

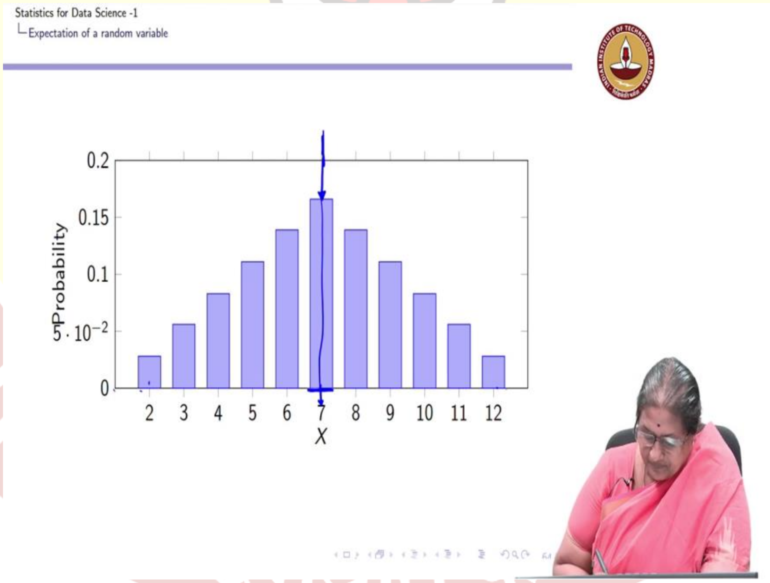
(Refer Slide Time: 15:49)





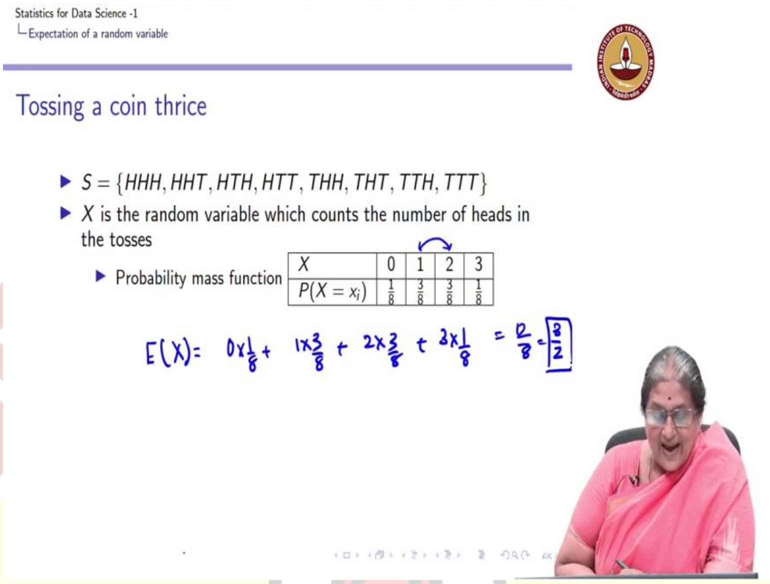
Whereas here X is taking a value, which it is already taking which was value of 7, it can happen, it need not happen also. So, expected value of X is 7. So, I can interpret or articulate it by saying that when 2 dice are rolled over and over again for a long time, then the mean sum is 7.

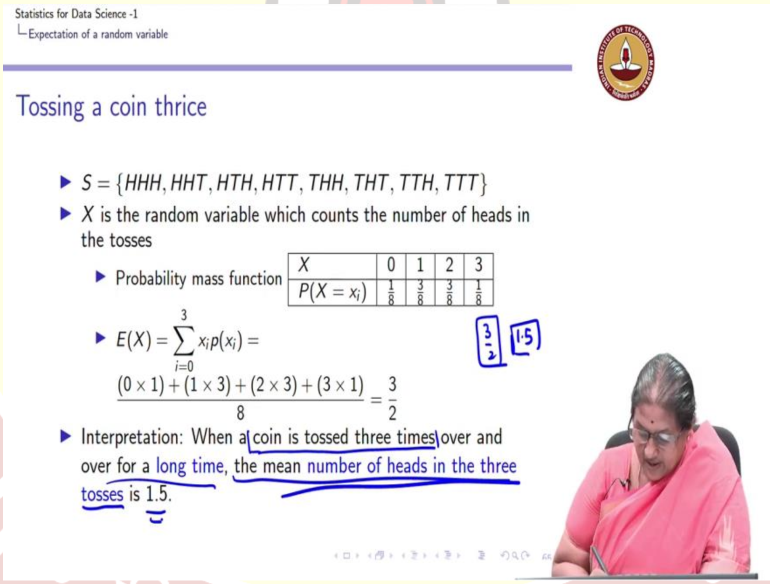
(Refer Slide Time: 16:14)



Now, interestingly you look at the probability mass function, you see that the distribution was, in a sense symmetric around this 0.7. And also, you can see that suppose you imagine this to be a rod with weights hanging, that this point in some sense balances the rod. So, of one of the physical interpretations of the notion of expectation is the centre of mass expectation, which comes from physics. This, I am not going to dwell more into it, but it just an intuition, you can think of these as weights being laid on a unitless scale. And you can see that this point actually balances the beat.

(Refer Slide Time: 17:06)

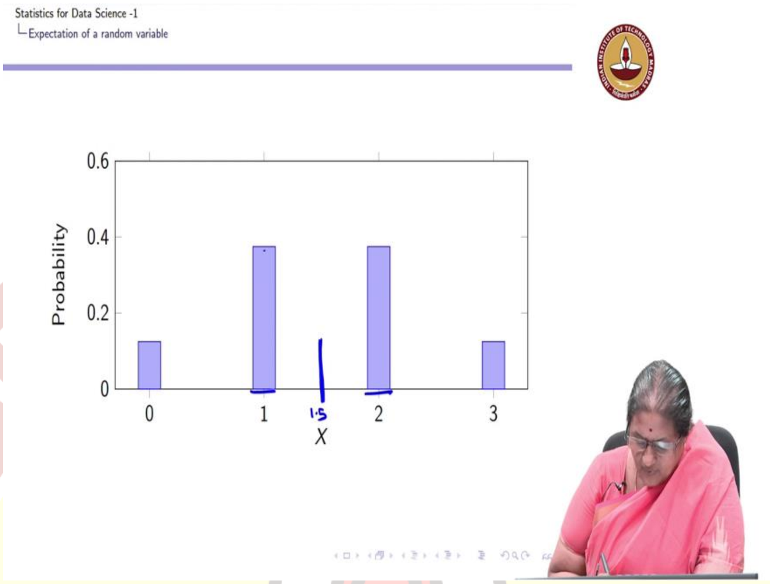




Now, let us look at the experiment of tossing a coin thrice again, we know, this is our sample experiment. And this is a distribution of the random variable. So, my expectation of x in this case is going to be $0 \times \frac{1}{8} + 1 \times \frac{3}{8} + 2 \times \frac{3}{8} + 3 \times \frac{1}{8}$, which is going to be 3 plus 3, 6 plus 6, which is 12 by 8, which I can see is 3 by 2.

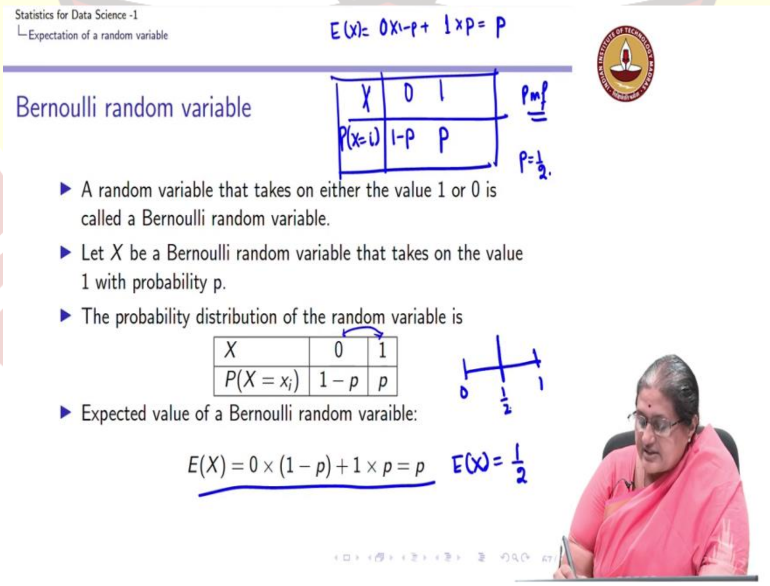
Again, I can see that expectation of X is 1.5. It is not a value, which X takes, but it is a value between 1 and 2. Again, the way I can interpret this value 3 by 2 is if I toss, if a coin is toss 3 times, and I repeat this experiment over a long period of time, the mean number of heads in 3 tosses is 3, 1.5 does not mean and I should not interpret it as the number of heads would be 3 by 2 or 1 and a half. This does not make sense. But the mean number of heads in 3 tosses is 1.5.

(Refer Slide Time: 18:26)



Now again, going to the probability mass function, I know that this was again symmetric, and you can see that the mean is somewhere here, which is 1.5. This is 1, this is 2, these two are same. And you can imagine that this acts like a fulcrum, which balances the mass.

(Refer Slide Time: 18:47)



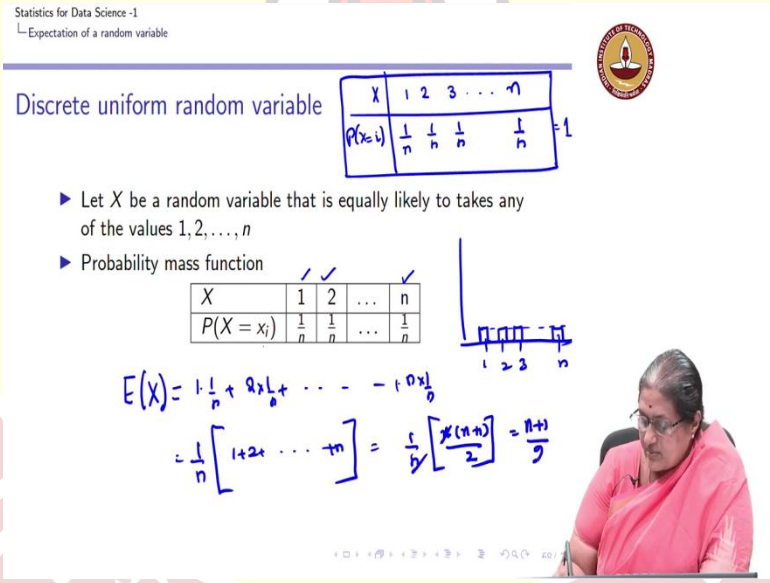
Now, let us look at a special random variable, which I refer to as a Bernoulli random variable. So, a Bernoulli random variable is a random variable, which takes 2 values. For ease of exposition, I just assumed that it takes the values 0 or 1, it takes only 2 values. And the probability with which it takes the values I am going to say it takes a value P, it takes a value

1 with probability P, since this should add up to 1, if it takes the value 1 with probability P, it should take the value 0 with probability (1 – P). So, this is the probability mass function of a Bernoulli random variable which takes 2 values 0 and 1.

Now, if this is the probability mass function, then the expectation of this random variable, it is taking the value 0 with probability (1 – P) and value 1 with probability P it is equal to P. So, the expected value of a Bernoulli random variable is equal to P. Now, if it is equally likely that X takes the value 0 and 1 with an equally likely probability, then I know that this P equal to half, which tells me that the expected value of X equal to half.

So, if P is equally likely than P equal to half, in which case the expected value of the Bernoulli random variable is equal to half, again you can see that this does not take either of these values, but it is balanced between so X takes 2 values, 0 and 1. So, the balance is at this point half.

(Refer Slide Time: 20:52)

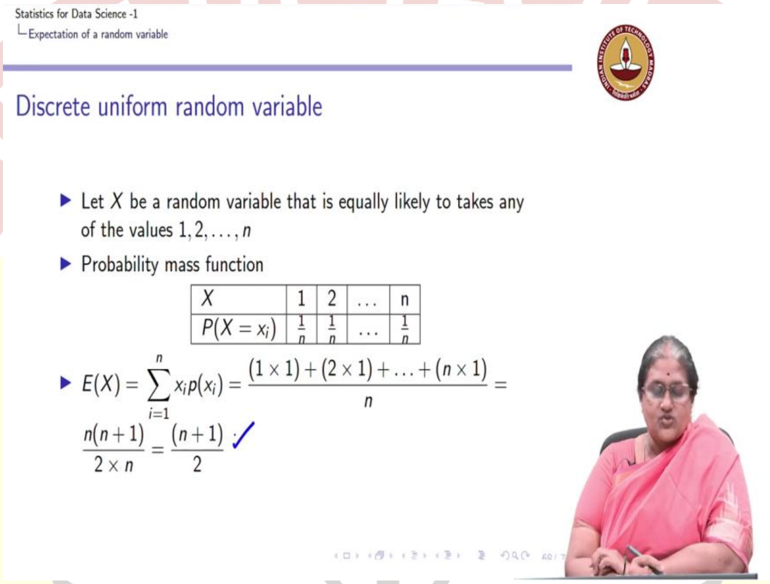


There is another random variable, which we refer to as a discrete uniform random variable. So, this again, suppose X takes the values 1, 2, 3 up to n. And the chance of a taking any value is the same, earlier Bernoulli took only two values, but here I am assuming it takes n value. So, the chance it takes the value 1 is equal to the chance it takes the value 2.

So, each 1 of them have a probability 1 by n, all of them add up to 1 we can verify that. So, the probability distribution is given by X takes the values 1 to n with a probability 1 by n, 1 by n, 1 by n, what is the expectation of a discrete uniform, so it takes discrete values 1 to n with equal probability, so my probability mass function if it takes the value 1, 2, 3 up to n, all of them are going to be so you can see that all the bars would be of the same height.

Hence, it is referred to as discrete uniform distribution, what is the expectation of this random variable? So again, it would be 1 into 1 by n plus 2 into 1 by n plus so forth, n into 1 by n, I can remove 1 by n is common 1 plus 2 plus n, which we can see is 1 by n. The sum of n numbers is n into n plus 1 by 2, I cancel these 2 out and I get a n plus 1 by 2.

(Refer Slide Time: 22:43)



Hence, you can see that the expectation of X is nothing but n plus 1 by 2.